IUBMB *Life*, 61(6): 627–643, June 2009

Critical Review

Protein Folding Simulations: From Coarse-Grained Model to All-Atom Model

Jian Zhang[1,2], Wenfei Li[1,2], Jun Wang[1,2], Meng Qin[1,2], Lei Wu[1,2], Zhiqiang Yan[1,2], Weixin Xu[1,2], Guanghong Zuo[1,2] and Wei Wang[1,2]

[1]*National Laboratory of Solid State Microstructure, Nanjing University, Nanjing 210093, China*
[2]*Department of Physics, Nanjing University, Nanjing 210093, China*

Summary

Protein folding is an important and challenging problem in molecular biology. During the last two decades, molecular dynamics (MD) simulation has proved to be a paramount tool and was widely used to study protein structures, folding kinetics and thermodynamics, and structure–stability–function relationship. It was also used to help engineering and designing new proteins, and to answer even more general questions such as the minimal number of amino acid or the evolution principle of protein families. Nowadays, the MD simulation is still undergoing rapid developments. The first trend is to toward developing new coarse-grained models and studying larger and more complex molecular systems such as protein–protein complex and their assembling process, amyloid related aggregations, and structure and motion of chaperons, motors, channels and virus capsides; the second trend is toward building high resolution models and explore more detailed and accurate pictures of protein folding and the associated processes, such as the coordination bond or disulfide bond involved folding, the polarization, charge transfer and protonate/deprotonate process involved in metal coupled folding, and the ion permeation and its coupling with the kinetics of channels. On these new territories, MD simulations have given many promising results and will continue to offer exciting views. Here, we review several new subjects investigated by using MD simulations as well as the corresponding developments of appropriate protein models. These include but are not limited to the attempt to go beyond the topology based Gō-like model and characterize the energetic factors in protein structures and dynamics, the study of the thermodynamics and kinetics of disulfide bond involved protein folding, the modeling of the interactions between chaperonin and the encapsulated protein and the protein folding under this circumstance, the effort to clarify the important yet still elusive folding mechanism of protein BBL, the development of discrete MD and its application in studying the α–β conformational conversion and oligomer assembling process, and the modeling of metal ion involved protein folding.
IUBMB *Life*, 61(6): 627–643, 2009

Keywords protein folding; protein structure; structural biology.

INTRODUCTION

Protein folding is among the most important and challenging problem in molecular and computational biology. In the early 1960s, Christian B. Anfinsen showed that proteins can fold reversibly under thermodynamic control and concluded that: (1) proteins organize themselves without assistant machinery into one of a myriad of possible conformations; (2) the native conformation is thermodynamically stable and, accordingly, the global minimum of the free energy landscape (*1, 2*). This thermodynamical point of view of protein folding was facing a famous paradox at the time—the Levinthal paradox, which states that if a protein has to explore the entire phase space to find its global minimum, it will take the age of the universe (*3*). Since the foundational work of Anfinsen, enormous efforts, including experimental and theoretical, have been done to solve the protein folding problem. Gradually, people are led to the new statistical view of protein folding which suggests that protein starts from an ensemble of unfolded states and folds through thousands of independent microscopic pathways, at last converging to the common native structure (*4–23*). This process is similar to the way the water flowing along different routes down mountains can ultimately reach the same lake at the bottom (*19*).

The new view of protein folding derives from the advances in both experimental and theory. Among many techniques, computer simulation stands for an important class of method in studying protein structures, folding kinetics, thermodynamics, free energy landscapes, and structure–stability–function relationship, as well as in engineering and designing new proteins and

Received 27 February 2009; accepted 3 April 2009
Address correspondence to: Wei Wang, National Laboratory of Solid State Microstructure and Department of Physics, Nanjing University, Nanjing 210093, China. Tel: +86-25-83594476. Fax: +86-25-83595535. E-mail: wangwei@nju.edu.cn

ISSN 1521-6543 print/ISSN 1521-6551 online
DOI: 10.1002/iub.223

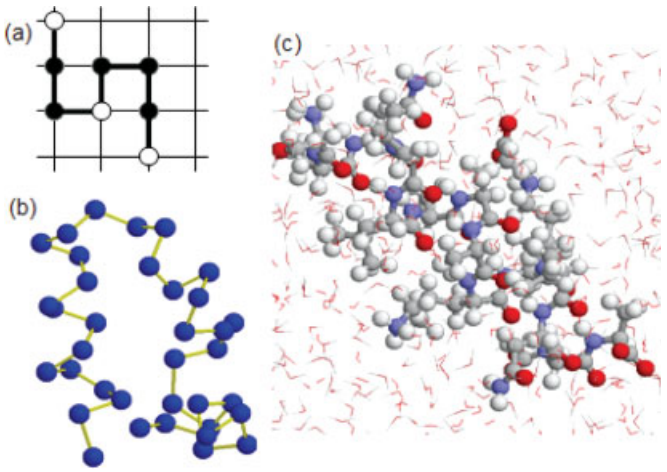

Figure 1. Three major levels of representation of proteins. (a) The lattice mode (a 2D lattice for example), where each residue is modeled as a bead and is restricted to the lattice grid. Two types of beads are shown here, hydrophobic (black) and hydrophilic (white). The potential function usually only includes contact energies. (b) The off-lattice model where each residue is represented as one (as shown here) or several beads (not shown). The beads are allowed to move freely in space. In most cases, the bond, angle, torsion angle and VDW interactions are considered at this level of representation. (c) The all-atom model of proteins where all atomic details, including all heavy atoms and hydrogen atoms in each residue, ion and water molecules, are explicitly modeled. All physical interactions, including the bond, angle, torsion angle, VDW, and electrostatic interactions, and sometimes the polarization effect, coordination bond, protonate/deprotonate process, and charge transfer, are modeled.

answering even more general questions, such as the minimal number of amino acid for proteins or the evolution principle of protein families (*5–47*). In general, the computer simulation of proteins can be divided into two classes, molecular dynamics (MD) and Monte Carlo simulations. Both of them require some levels of representation of the protein geometries and interactions, ranging from the highly simplified lattice model, to some intermediate levels of modeling, and to all-atom model with even water molecules explicitly considered.

The most schematic description of protein chain is self-avoiding walk on a cubic lattice with each amino acid represented by a bead and restricted to the lattice grid. In the so called HP model, two types of amino acids are considered, the hydrophobic and polar residues (Fig. 1a). Despite their simplicity, the HP and similar models have proved invaluable in shaping the new statistical view of protein folding (*5–16*). Here it is worth mentioning the pioneering work of Kolinski and Skolnick on high-resolution lattice models, which have the ability to accurately model protein structures and retain the computational efficiency of lattice models as well (*14–16*). The lattice models benefits greatly from the discretization of protein phase space; however, it also suffers from this strategy. The discrete nature of the model surely affects the folding behaviors, especially the dynamics of the system. The off-lattice models alleviate this problem by employing a continuous representation of the internal degree of freedom (*24–30*). One or more centroids can be used to represent an amino acid and a suitable interaction potential among them needs to be introduced (Fig. 1b). These simplified models allow people to simulate thousands of folding-unfolding events to obtain a detailed statistical description of the folding process. They have provided major insights into, for example, the enthalpy–entropy interplay, the structure of the transition states, the folding nucleus, folding pathways, cooperativities, folding rates, and free energy landscapes.

Although the simplified coarse-grained models have proved very useful in gaining insights to the general thermodynamic and kinetic features of folding process, they are obviously incapable of capturing the rich variety of physical and chemical behaviors of proteins. In such cases, the appropriate tool is the simulation based on all-atom representation of amino acids and all-atom potential functions. Though the time scale accessible to all-atom simulation is still hindered by its large computational costs, it has illustrated its power in, for example, the detailed characterization of the transition state structures, the complex folding pathways, the structure of the denature state ensemble and hydrodynamics (*31–47*). Most importantly, the all-atom model is necessary in the rational design of drugs and protein interactions (*48–50*).

Recently, the MD simulations are undergoing rapid developments and expanding their territories very quickly. The coarse-grained models are achieving more and more successes in large molecular systems, such as the ribosome-protein complex, the nucleosomes, molecular motors, and even the virus capsides. The all-atom level simulations, however, are providing more and more reliable details of the molecular process. Plus, the developments in the methodology, such as the multicanonical sampling, the replica exchange method, the transition path sampling strategy, and the Folding@Home project, allow people to deal with larger and larger biological systems (*37–45*). Currently, people are able to collect thousand of trajectories of length several ten to several hundreds nanoseconds; the overall computational length has exceeded the folding time of fast folders. It is believed by some optimists that the protein folding problem has been solved (*51*). At the same time, the protein models at the intermediate level have proved useful especially in protein structure prediction; Rosetta, for example, is such a familiar program. Nowadays, with the help of such level of protein models, high-resolution ($<$1.5 Å) structure prediction can be achieved for small protein domains ($<$85 residues) (*52*).

This review is by no means intended to cover all the advances in the area of computer simulation of protein folding; which has expanded to such an extent that it is impossible to do this job for any single review. The interested readers are referred to several reviews for a better understanding of this area (*16, 20,*

53–55). In this review, we would like to show a selection of topics and methods which illustrate some interesting new directions in the computational protein folding studies, mostly in our group. These includes but are not limited to the study of the delicate enthalpy-entropy interplay in ribosome protein S6, the exploration of the rules of disulfide bonds in the folding process of protein Tendamistat, the effort to clear debates in the folding mechanism of protein BBL, the works on chaperonin assisted protein folding and on amyloid formation, and the researches devoted to metal coupled protein folding.

THE ENTHALPY-ENTROPY INTERPLAY IN THE FOLDING PROCESS OF RIBOSOME PROTEIN S6

During the last two decades, the coarse-grained Gō-like models that incorporate only the structural information of native state have proved very successful in studying the protein folding process (*56–60*). This is because that the energy landscape of most small globule proteins is largely determined by their native topologies, which is supported by extensive experimental observations and simulation results. For example, the structures of the transition states (TS) are similar for proteins with similar topologies and are insensitive to site mutations; the folding rates exhibit a strong dependence on a topological parameter—the contact order, which measured the average effective distance of contacting residues along sequence. As a consequence of this fact, the folding pathway and the structures of the TS and intermediate states can usually be predicted by the simplistic Gō-like models. However, with the advancement of both experimental techniques and simulation methodologies, it was discovered that, in addition to the topological factor, the energetic factor also plays an important role, especially in some proteins such as the ribosomal protein S6.

The specialty of S6 was noticed in a circular permutation experiment (*61–63*), which kept the interactions between residues intact but changed the topology and loop entropy. The interesting observation was that the distribution of the Φ-value values of S6 changed from diffuse to polarized after circular permutation. As a contrast, the distribution remains diffuse or polarized for protein chymotrypsin inhibitor 2 (CI2) or src SH3 domain after circular permutation, respectively. To understand the interesting behaviors of S6, extensive experiments and theoretical studies had been done, and it was concluded that such behaviors are the consequence of the competition between biased contact energy and loop entropy in this protein (*64, 65*). For this reason, the failure of the simplistic Gō-like models in predicting its folding process is not surprising; such models treated all the native interactions equally thus suppressing the energetic factors to the lowest level. (Fig. 2)

There are several important theoretical works concerning this issue. For example, Shakhnovich and coworkers investigated the folding process of S6 using an all-atom Gō model in conjunction with restraints from experimental Φ-values (*64*). With those restraints, the biased contact energy was automatically introduced into their model, which was then able to reproduce most experimental observations. Following a similar idea, Clementi and coworkers developed a systematic strategy to incorporate experimental data into their coarse-grained model (*65*), which indeed exhibited an atypical distribution of contact energies. However, despite these successful works, the situation is not satisfactory since these previous models need experimental data to calibrate the parameters of contact interactions; they did not provide a transferable model with physically based potential energies.

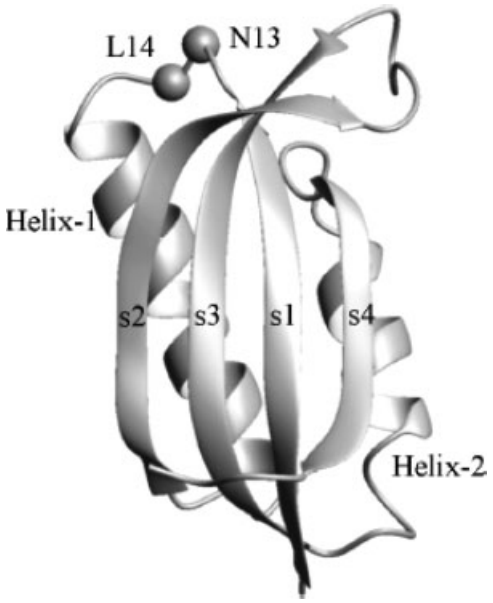

Figure 2. The native structure of the ribosomal protein S6. The figure is generated using the software MOLMOL. The figure is reproduced from ref. (*66*), with permission.

Toward building a transferable model, we developed a coarse-grained Gō-like model by introducing variable contact energies between residues based on their physical-chemical properties (*66*). For example, the Van der Waals interaction was modeled proportional to the number of heavy atom pairs within a cut-off distance between two residues, the hydrogen bond was defined based on the geometric structure of the native state, and the hydrophobic interaction was added to three major interfaces between four β-strands according to the contents of native contacts formed between two hydrophobic residues. To verify this model, the folding of the wild-type protein S6 and its circular permutant were simulated. The results showed a significant improvement regarding the degree of agreement with experiments. First, the folding pathway was correctly predicted by the modified model whereas failed by the original unmodified model (Fig. 3a). For example, according to the original unmodified model, the fastest event is the formation of the native contacts within the interface S2-S3 due to a shorter involved loop length and thus a lower conformational entropy change. How-

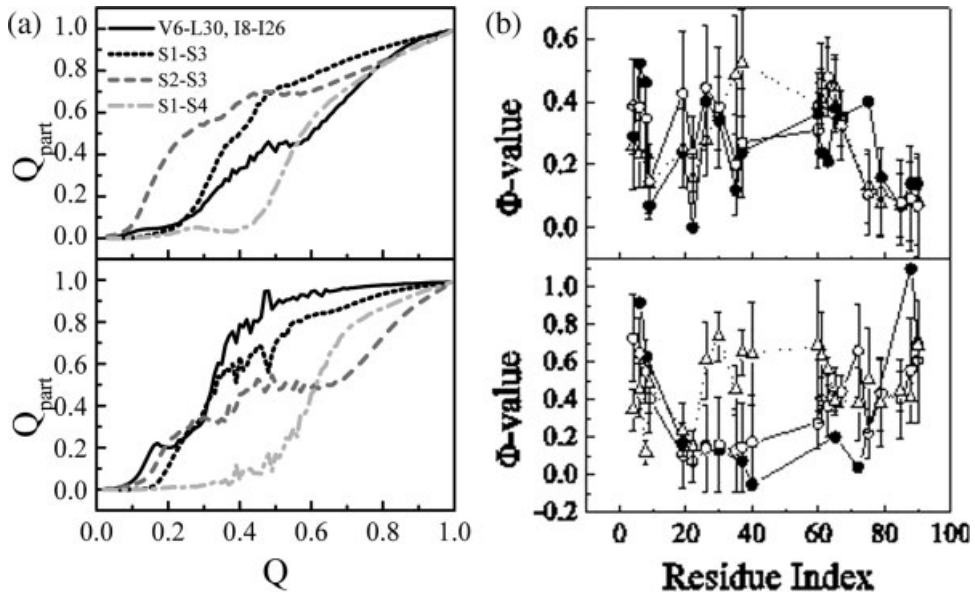

Figure 3. (a) The folding pathway of ribosomal protein S6 calculated based on the original Gō-model (upper left) and the modified Gō-model (lower left). Different curves correspond to different sub-structures. The *Q* value in the X-axis is the fraction of formed native contacts during folding process, which is often used as a reaction coordinate (RC) together with coarse-grained Gō-like models. Q_{part} is the fraction of native contacts within specific sub-structures indicated by the figure legends. This figure shows the relative formation speed of different sub-structures with the progress of folding reaction; in short, we called it the folding pathway. (b) The Φ-values of the wild-type S6 (upper right) and its permutant (lower right), demonstrating how well the structure of the transition states (TS) is predicted. The filled circles represent the experimental values (taken from refs. (*62, 63*)), the open circles are calculated from the modified model and the open triangular symbols are from the original Gō-model. The figures are reproduced from ref. (*66*), with permission.

ever, the modified model predicted that the formation of the interface S2-S3 and hydrophobic core are the fastest, due to the strong hydrophobic interactions in the new model. Second, the the Φ-values calculated by the new model showed a better correlation with that determined by experiments (Fig. 3b). Third, the folding nucleus identified by the new model was consistent with previous works (*64, 65*).

This work showed that the new model works well for protein S6 and highlighted the importance of modeling energetic factors, especially in this protein. Also, this work indicated a possible strategy used by nature to overcome the dilemma when functional requirement conflicts with folding cooperativity; the way is to use a delicate arrangement of interaction energy to balance the conformational entropy. By this way, nature fulfills the functional requirement and at the same time maintains the folding cooperativity and avoids the formation of folding intermediates, which may lead to harmful aggregation or amyloidosis.

THE RULES PLAYED BY DISULFIDE BONDS IN THE FOLDING PROCESS OF PROTEIN TENDAMISTAT

Many proteins, such as some membrane and secreted proteins in bacteria and eukaryotes, fold into their native structures requiring the formation of disulfide bonds. Disulfide bonds are covalent interactions between residues, and are vital for protein stabilities and activities (*67–69*). The folding process involving the formation of disulfide bonds is very complex. For example, in the refolding experiment of the ribonuclease *in vitro* by Anfinsen and coworkers, the different additive orders of oxidants and buffer led to completely different final structures and activities (*2*), indicating the delicate role of disulfide bonds in the folding process. Moreover, several experiments suggested that the folding process of some disulfide-contained proteins *in vivo* is even complex in that they require not only the participation of oxidants but also the help of many special enzymes (*70, 71*).

It is a challenging task to explore such a complex folding process for both experiments and theoretical analysis. So far, people mostly focused on the refolding process without the rupture and reformation of disulfide bonds. If the denatured conformations of a protein are obtained by unfolding from native state without presence of reductant, the disulfide bonds remain intact. It has been shown by many experimental and theoretical studies that the pre-formed disulfide bonds can significantly increase the protein stability. However, regarding the folding kinetics, there are two different opinions. One assumes that the pre-formed disulfide bonds will increase the folding rate because

Figure 4. The schematic representation (backbone model) of the native structure of wild-type tendamistat. The cysteines are displayed with ball-stick model and the disulfide bonds are shown by yellow sticks. The first disulfide bond (Cys-11–Cys-27) is located at the end of the first β-hairpin while the second one (C45–C73) connects the two outer strands-4 and strand-6. The figure is reproduced from ref. (*80*), with permission.

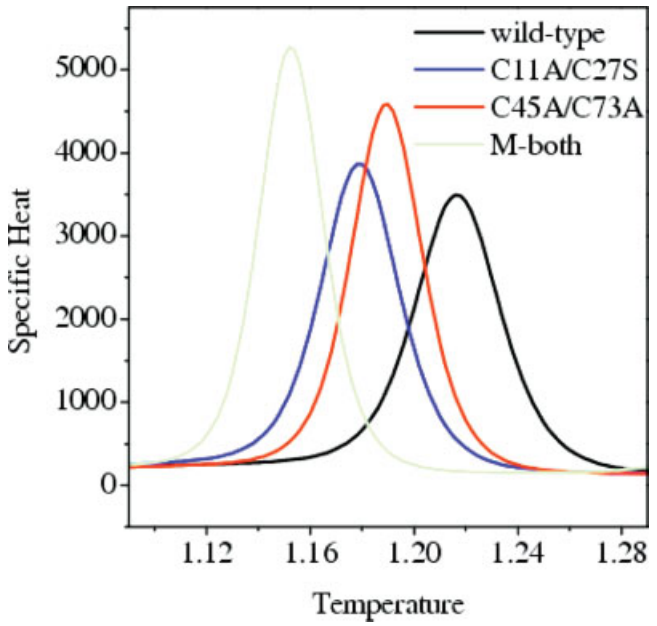

Figure 5. The specific heat Cv calculated as a function of temperature (in unit of ε/k_B, where ε is the energy and k_B is the Boltzmann constant). The transition temperatures (corresponding to the peak of the Cv curve) of the wild-type tendamistat (thick solid line), the C11A/C27S mutant M1 (dotted line), the C45A/C73A mutant M2 (thin solid line), and the mutant M3 with both disulfide bonds removed (dashed line) are 1.217, 1.179, 1.189, and 1.152, respectively. The figure is reproduced from ref. (*80*), with permission.

the size of phase space the protein has explored is decreased significantly by these preformed disulfide bonds (*72, 73*), whereas the other argues that they may lead to kinetically trapped intermediates and retard the folding process (*74*). It is still an open question which opinion describes the role of preformed disulfide bonds better or the answer is case dependent. Also, the folding process of proteins with disulfide bonds usually involves multiple pathways; the physical origin and biological behavior of these pathways also need to be clarified.

Protein tendamistat, a α-amylase inhibitor from *Streptomyces tenda*, is a good model system for investigating the effects of disulfide bonds on the stability and folding kinetics. It is an all-β-sheet protein with 74 amino acids and has two disulfide bonds, C11–C27 and C45–C73. There are extensive experimental studies on this protein (*72–79*). As a brief summary of these works, tendamistat exhibits a two-state folding behavior that is unique among all the disulfide bonded proteins discovered so far, and this behavior sustains if either disulfide bond is removed (*76*). A surprising feature is that the removal of the C11–C27 disulfide bond leads to a larger effect on the stability than the removal of the C45–C73, since the latter involves a longer sequence stretch thus is presumed to have a larger entropic effect (Fig. 4).

We developed a modified Gō-like model that incorporated disulfide bonds as well as hydrogen bond interactions to investigate the folding mechanism of tendamistat and the effects of removal of disulfide bonds on the folding process (*80*). Our model showed that the disulfide bonds as well as the three hydrogen bonds between the N-terminal loop-0 and strand-6 are of significant importance for the folding of tendamistat (*79*); without either interaction, the two-state behaviors become unstable or the folding pathway changes significantly. The simulation based on the modified model showed that tendamistat and its two single-disulfide mutants are all two-state folders, consistent with the experimental observations (*76, 79*). By comparing the folding behaviors of the wild-type tendamistat with its two mutants, it was found that the removal of either C11–C27 or C45–C73 disulfide bond leads to a large decrease in the thermodynamical stability and the loss of structure in the unfolded state; and the effect of the former is stronger than that of the latter, as shown in Fig. 5 (*76, 79*). We further studied the folding pathways of the wild-type tendamistat and its two mutants and got a detailed folding picture for these proteins, which complemented the experimental findings. The folding nuclei were also identified, consistent with those suggested by experimental studies (*81*). Therefore, a nucleation/growth folding mechanism that explains the two-state folding manner was clearly characterized, as shown in Fig. 6. Moreover, the effects of the removal of each disulfide bond on the thermodynamics and folding dynamics

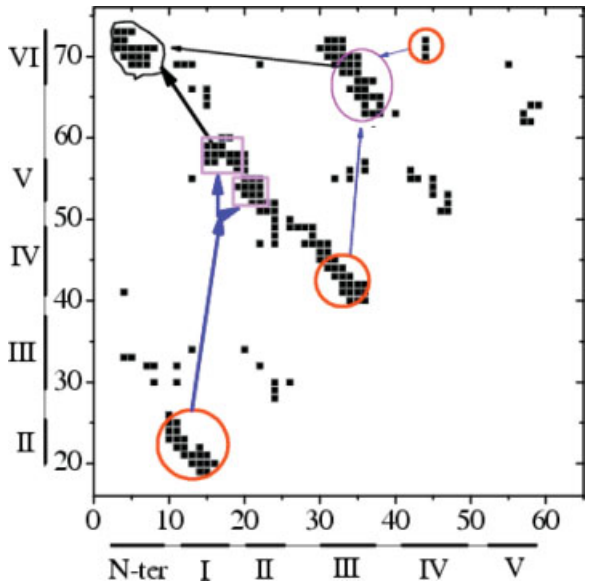

Figure 6. The folding pathway of the wild type tendamistat shown on a contact map. The X and Y axes are the residue indices. The black spot indicates there is a native contact formed between two residues. The contacts inside the circle regions with thick lines are formed before reaching the transition state (TS); these inside the rectangular regions are formed in the TS; these inside the ellipse region are formed in the later folding process except that several key contacts were found to be formed in the TS; and the contacts inside the irregular region are formed in the final stage. The figure is reproduced from ref. (*80*), with permission.

were also well interpreted from their influence on the nucleus. This work suggested that the modified Gō-like model is a workable model for describing the folding behavior of protein tendamistat and could be used to study other disulfide-bonded proteins.

THE FOLDING MECHANISM OF β-HAIRPIN AND ALL-ATOM SIMULATION

The coarse-grained model has made great contributions to our understanding of protein folding and is still undergoing rapid developments. At the same time, all-atom simulation with either implicit or explicit solvent model is attracting more and more interests of researches from both academic institutes and industry, which is attributed to the exponential increase in computational power as well as the recent significant advances in methodology. The computational cost of the all-atom simulation is enormous due to the large number of degree of freedom it has to deal with and the thousands or millions of local minima in the energy landscape, which may trap trajectories hence effectively slow down the sampling speed. Accordingly, many techniques have been developed to overcome this problem. Among them, the replica exchange method (REM) is perhaps the most familiar (*82, 83*). The REM method requires simultaneously running of several replicas of the same protein system, yet at different temperatures. The replicas will be subject to attempt to exchange after a fixed number of MD steps; the exchange probability is calculated according to a Metropolis-like criterion. The REM method is especially useful to overcome the multiple minimum problems because the replicas at low temperatures could be exchanged to high temperatures and get high probabilities there to transfer to a new region of phase space.

In the past, there are many works on the folding mechanism of β-hairpins (*84–91*). These works are important in that β-hairpin is one of the key secondary structural elements and many features of folding kinetics can be found in this structure; moreover, its folding time-scale is within microseconds, which is accessible for both fastest time-resolved experiments and advanced simulation techniques. Therefore, the β-hairpins provide a good model system to test our models and theories for protein.

Although extensive studies have been done previously, the folding mechanism of β-hairpins remains elusive. Toward better understanding of this problem, by combining all-atom explicit water simulation with the REM method, we studied the energy landscape, folding intermediates, structural features of the transition region, and the kinetics of the trpzip2 β-hairpin (*45*). The conclusions are (1) the very early folding process is initiated by the nonspecific hydrophobic collapse; (2) the structure of the transition region is characterized by a largely formed turn and the hydrophobic interaction between the inner pair of core residues. The turn formation and establishment of hydrophobic interactions are cooperative processes and occur almost simultaneously—it is not appropriate to say which one occurs earlier; (3) the roles played by turn and hydrophobic interactions in folding process are different. The turn sequence determines the size of conformation space of the denatured state that the peptide has to search within to find the bottleneck of the transition states, thus determines the folding rate. The hydrophobic interaction provides a compact ensemble to facilitate hydrogen-bond registration and stabilizes the nascent turn; (4) the folding mainly follows a zipper mechanism, however, the hydrophobic interactions are also critical. The folding picture is a blend of hydrogen bond-centric and hydrophobic core-centric mechanism.

BBL: THE GLOBAL DOWNHILL FOLDER?—STUDIED BY BOTH COARSE-GRAINED AND ALL-ATOM MODELS

The folding mechanism is one of the most interested topics in protein researches during the last decades. On the basis of the landscape theory, several possible folding mechanisms were proposed from the fundamental physics of protein systems, including the Type-0 (the downhill folding with no free-energy barriers) and Type-I (two-state folding with a high free-energy barrier sep-

arating the native state and the denatured state) (*92*). Among these, the two-state folding behaviors were carefully characterized through the widely cooperation between theoreticians and experimentalists, including the cooperativity (*93–97*), folding pathways (*98–100*), folding nucleus (*101, 102*), and many other interesting issues (*92, 103–105*). Meanwhile, other kinds of folding mechanisms were sometimes discussed (*106*). Some downhill behaviors for artificially designed proteins were observed in the condition with denatured state totally destabilized (*107–109*); it is just a nonequilibrium behavior and related to local regions of the whole landscape. In 2002, Munoz group proposed the first natural downhill folder (*110*), a small protein BBL, which is believed to have a barrierless behavior for a wide range of denaturant concentrations. This finding raised a lot of debates (*111–113*). Does the protein BBL really have the downhill behavior? What kind of factors produces such a special kinetic behavior? These questions have attracted much attention during the recent years (*114–123*). The molecular details are deserved to solve these problems.

As a computational approach, an off-lattice model with Gō-like interaction was used to simulate the folding of protein BBL (*122*). The specific capacity calculation showed that the model protein does have a low folding cooperativity compared with that of protein CI2—a most familiar two-state cooperative folder (Fig.7a). The free energy along the reaction coordinate Q, which is the fraction of native contacts, are generally unimodal at various temperatures, demonstrating a barrier-less folding feature (Fig. 7b). The free energy profiles at various temperatures all have a single valley, which gradually varies from the denatured state to the native state following the decrease of the temperature, suggesting a global downhill behavior with minimal frustration. To reveal the physical origin of such a special behavior, a comparison of the native structure with protein CI2 was made and it was found that BBL has a much lower ratio of nonlocal contacts in its native structure. Presumably, the ratio of nonlocal contacts may be used as an indicator of folding behaviors, global downhill via cooperative. This idea was supported by the statistics on a set of 17 proteins with different kinetics, where a strong correlation between the number of the nonlocal contacts per residue (N_N) and the cooperativity factor ($\kappa_2 = 2T_f\sqrt{k_B C_v(T_f)}/\Delta E_{cal}$, see the refs. (*95, 96*)) was observed (*122*). This correlation enabled us to semi-quantitatively predict the folding cooperativity solely from the native structure of proteins.

To further test the ability of the structural factor N_N in predicting the folding kinetics, a series of theoretical "mutations" on a two-state protein, the lambda repressor, were carried out (*122*). It was found that by modulating the ratio of the nonlocal contacts by "mutations," this original two-state protein was successfully changed to a downhill folder. Further, based on these results, it was proposed that a protein with cooperativity $\kappa_2 < 0.34$ or with $N_N < 0.90$ would be a good candidate of the global downhill folder. This criterion offered us an auxiliary tool to discover new global downhill folders.

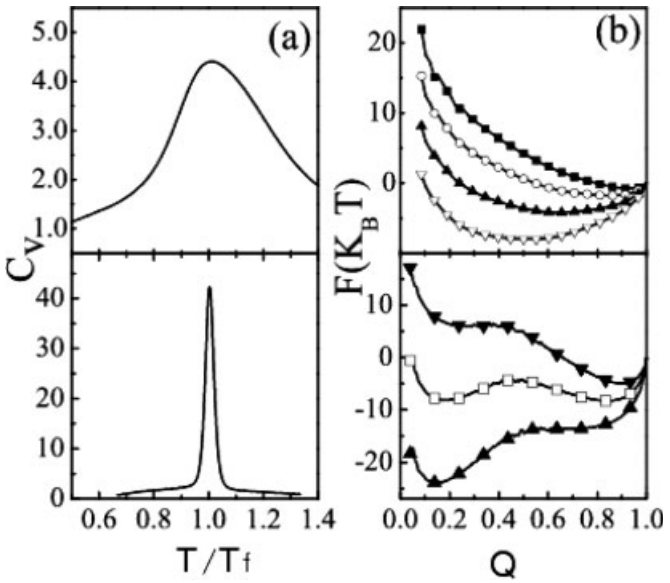

Figure 7. The heat capacity Cv as a function of the reduced temperature (a) and the free energy along the fraction of native contacts Q (b) calculated for protein BBL (the upper two figures) and CI2—a two-state cooperative folder (the bottom two figures), respectively. The temperature is normalized with respect to each protein's folding temperature T_f. In (b), different free energy curves were calculated at different temperatures; the upper curves correspond to lower temperatures. The figures are reproduced from ref. (*122*), with permission.

The stronger judgment of whether BBL is a global downhill folder may come from the all-atom simulations. Besides the coarse-grained model study of BBL described earlier, we also calculated the free energy landscape of BBL with replica-exchange all-atom MD simulations with water molecules explicitly modeled (*120*). Totally 64 replicas were used; the overall simulation time is 6.4 microseconds, comparable with the folding time (~7 microseconds at 310 K). To characterize the energy landscape faithfully, a total of four combinations of eight OPs (order parameters) were used to project the free energy surface. They are the fraction of native contacts Q, the radius of gyration of the whole protein R_g, that of the hydrophobic core R_g^{core}, the RMSD with respect to the native structure, the number of formed native hydrogen bonds nHB, the number of native helical residues nHelix, the number of formed native contacts in the secondary structure Q_2 and in the tertiary structure Q_3. The free energy landscape was found rather complex. As shown by Figs. 8a and 8b, only one broad populated state is populated, whereas Figs. 8c and 8d show two separated basins of attractions. However, even for Figs. 8c and 8d, the free energy barrier separating two basins is only $\sim 2k_BT$, indicating that the folding cooperativity is rather low. Figure 8 also suggests that among all OPs, the R_g^{core} and RMSD could be better parameters to detect the possible separation of states. Besides, it

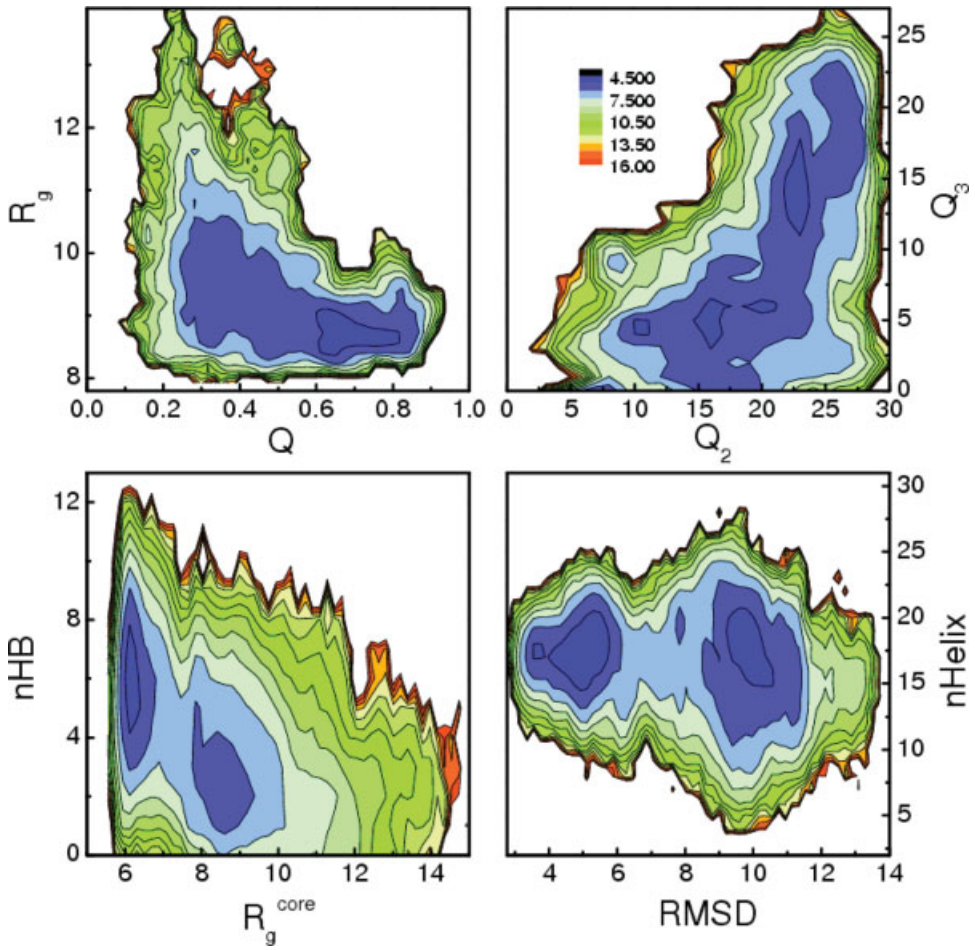

Figure 8. The two dimensional free energy projections calculated at temperature 330 K. The height between neighboring contour lines is 1 k_BT. Totally eight OPs (order parameters) are used, which are defined in the context. The figures are reproduced from ref. (*120*), with permission.

should be noted that Fig. 8 shows no evidence of existing of intermediate states, ruling out the multistate picture.

As the projection of energy landscape is dependent on the choice of OPs, we tried to circumvent this problem by constructing it in high dimensional space. The idea is that, (1) presumably, more OPs together can characterize the free energy landscape better; (2) while highly cooperative folding (indicating high free energy barrier) can be characterized well by only one reaction coordinate (RC), weakly or noncooperative folding (indicating flat free energy surface) may need more RC to describe the motions along the free energy surface; (3) since the free energy barrier in protein folding is entropic in nature, the increase in the dimensionality of the projecting space will decrease the barrier height, thus the dimensionality cannot be a large number; presumable, a dimension of 3–4 is a good choice. After constructing the energy landscape in a 4D space, we calculated the minimal pathway leading from the denatured state to the native state as a function of Q. The curves calculated at temperatures 300, 310, 320, 330, and 340 K are shown in Fig. 9. It can be seen that along these pathways, the free energy barriers are very low. For example, it is only 1 k_BT at 330 K and 1.3 k_BT at 300 K.

According to Figs. 8 and 9, the free energy barrier is rather low, 1–2 k_BT near the experimental melting temperature, depending on the way the energy landscape is projected. Moreover, taking consideration of the slower convergent speed at the barrier region, the barrier height may decrease slightly further for an even longer simulation. Therefore, 1–2 k_BT could be an upper limit of the barrier height near the melting temperature.

One importance difference between the cooperative and downhill folding is the conformational distribution of the transition states. To characterize this distribution for BBL, we collected all the conformations in the transition region, clustered them and then superimposed the representative structure of each cluster with the native structure (Fig. 5 in refs. (*120*)). The result showed that, roughly speaking, the relative positions of three helices are similar to that in the native state thus similar in all conformations. However, their spatial orientations showed

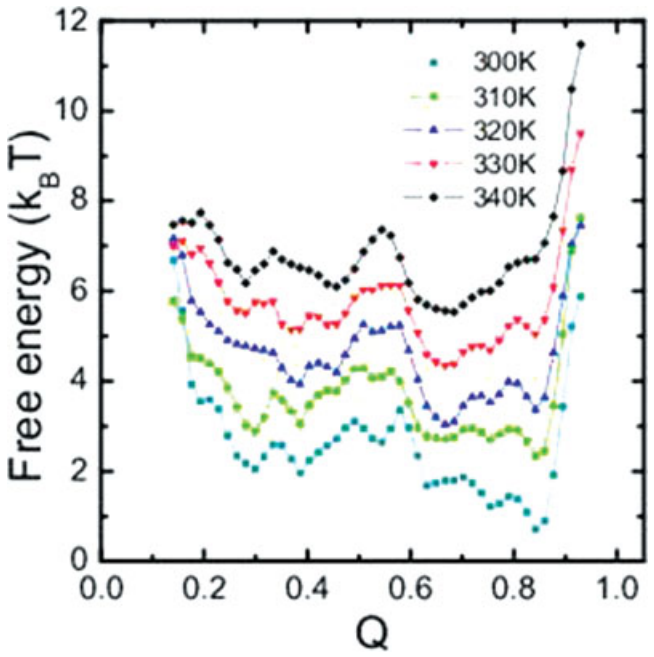

Figure 9. The minimal pathway along the 4D free energy surface as a function of the fraction of native contact Q, calculated at temperatures indicated by the legends. The curves have been shifted along the vertical axis for clear representation. The figure is reproduced from ref. (*120, 121*), with permission.

great variabilities, especially that of the second and the third one. The transition states of our construct of BBL showed a much higher degree of structural variability than that of cooperative folder NTL9, whereas a less degree than that of protein BBL predicted by coarse-grained model (Figs. 7b and 7e), Knott and Chan (*119*). Our specific construct of BBL indeed exhibited some cooperativity; however, it was very weak and exhibited a large conformational heterogeneity in its transition states.

On the experimental front, major emphasis had been placed on checking for the inconsistency in the melting temperature T_m between different probes (*110–112, 115*). We followed the same procedure as in experiments by monitoring the temperature dependence of the signals of five different structural features. They are the fraction of native contacts Q, the fraction of native tertiary contacts Q_3, the average lengths of helix-1 and helix-3, and the number of native hydrogen bonds nHB. After applying the standard two-state analysis using the Gibbs–Helmholtz equation for each *individual* signal, we found that their temperature dependence do not coincide with each other: the structural features that mostly define the tertiary structure melt at lower temperatures (For example, the Q_3 curve), whereas secondary structure sustains much higher temperatures. The melting temperatures range from 378 to 397 K, spanning a width of 19 K. A similar value, 17 K, was obtained if the T_m was calculated from direct derivative of the unfolding curves, which procedure was free of possible baseline problem. The large spread of T_m suggested that the five structural features unfold in a weakly or non-cooperative manner, consistent with the small free energy barriers shown in Figs. 8 and 9 and the structural feature of the transition states.

The spread of 19 K is significantly smaller than that of Naf-BBL (60 K) reported by Sadqi et al. by performing an atom-by-atom analysis in their NMR experiment (*115*). The most possible reason is the different pH values used; a neutral solvent was modeled in our simulation whereas pH 5.3 was used in the experiment. Physically, the low pH will change the electrostatic distribution of protein thus affects the enthalpy–entropy interplay in folding process, leading to a different free energy barrier and different cooperativity. If we *naively* relate the spread of T_m with the barrier height, since both of them are correlated with the folding cooperativity, the comparison of the spread of T_m indicates that the Naf-BBL at pH 5.3 may have an even low barrier height than 1–2 k_BT. In this way, our all-atom simulation supported the global downhill picture of BBL.

CHAPERONIN-ASSISTED PROTEIN FOLDING

The folding inside living cells (namely the *in vivo* folding) is a widely interested and difficult topic in protein researches. In the cellular environment, there are a large number of biological macromolecules and machineries (*123–130*), which have complex interactions with substrate proteins. Sometimes, the folding is coupled with synthesis process or with aggregations (as the formation of amyloid-shaped structures in central neuronal system) (*131–136*). The crowded and heterogeneous environment makes the *in vivo* folding largely different from the *in vitro* folding processes. During the recent decade, accompanied with the progress in both experiments and the simulations, there are more and more attentions on the *in vivo* folding (*137–154*). In these studies, the effect of geometric confinement is generally appreciated, which states that the inert confinement could accelerate the folding by reducing the entropy of the denatured states. However, in the realistic cellular environment, the influence exerted by the surroundings may not be simply inert (*149–151*). In fact, some experiments (*152, 153*) showed that the folding of several specific proteins can be retarded by the chaperonin GroEL, which generally encapsulates the substrate proteins inside its cavity and provides a typical crowded environment; the experimental results were different from the expectation based on the regular theory of geometric confinement. Therefore, a more comprehensive analysis is necessary to understand the folding processes in complex environments.

To study the chaperonin-related protein folding processes, we modeled the crowded environment as a space of shape of a cylinder and utilized a Gō-like potential to model the interactions inside proteins (*60*). Different from the previous studies, the interaction between hydrophobic residues and the cylinder wall was modeled as well (Fig. 10). It was found that the attractive interaction between hydrophobic residues and the wall can

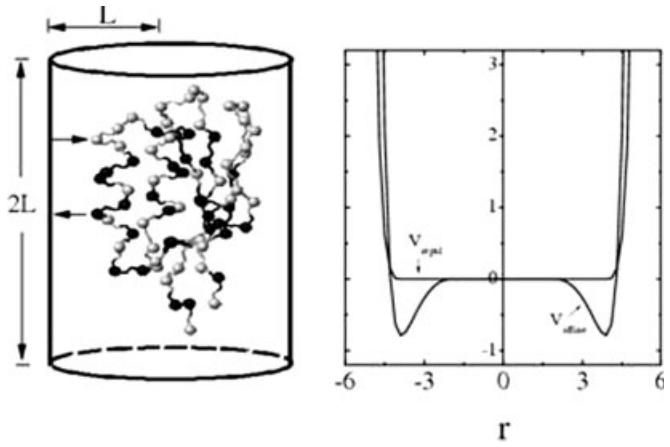

Figure 10. (a) The sketch map for the substrate protein encapsulated in the chaperonin cavity. The crowded environment was modeled as a space of shape of a cylinder. The hydrophobic residues are shown as black balls and the hydrophilic residues as gray. (b) The potential function between protein and the cylinder wall. The parameter r measures the distance in unit of angstrom from the center. The interaction between hydrophilic residues and the cylinder wall is repulsive, represented by the curve Vrepul; and that between hydrophobic residues and the wall is attractive, corresponding to Vaffine. The figures are taken from ref. (*60*), with permission.

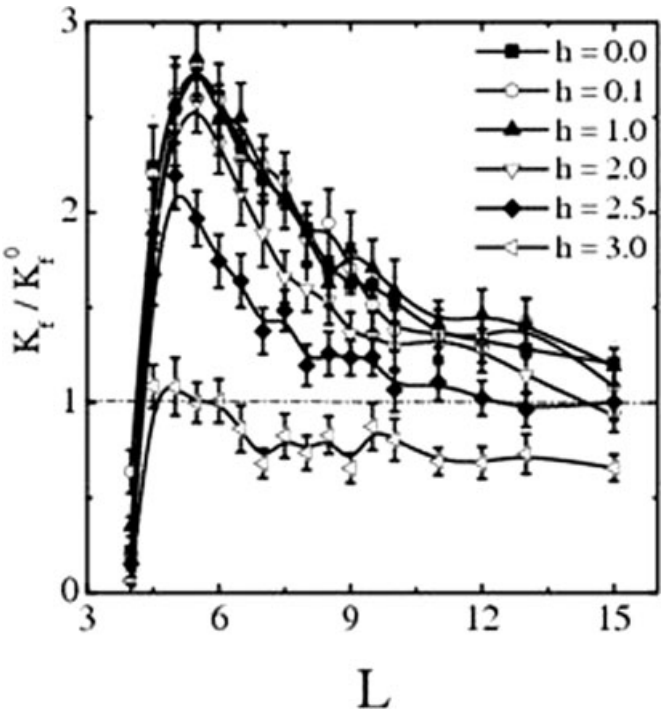

Figure 11. The folding rate as a function of the radius L (in unit of angstrom) of the chaperonin cavity at different affinity strengths h at temperature T_f^0 (the folding temperature in bulk). K_f is the folding rate inside the chaperon cavity and K_f^0 is that in bulk. The figure is taken from ref. (*60*), with permission.

weaken the acceleration effect of the confinement, though the optimal size of the cylinder to maximize the folding rate depends little on this affinity (Fig. 11). Also, it was observed that this affinity can induce unfolding, which stabilizes the unfolded structures and acts as a competitor of the intra-protein interactions. This result could explain the unfolding signals of substrate proteins observed in the T state of GroEL and provide a physical picture for the iterative annealing mechanism. Besides, a series of mutations were carried out for protein CI2 and studied by using this model. The experimental acceleration/retardation of the folding rates was successfully reproduced quantitatively (Fig. 12). This verified our modeling and suggested that the protein-environment interaction is important for a precise understanding of the protein folding process inside the crowded cellular condition.

THE DISCRETE MD SIMULATIONS AND PROTEIN SELF-ASSOCIATION

Biological cells are composed of very high concentration of macromolecules. In such a crowded environment, some ordered aggregates in the shape of amyloid fibrils are widely observed in biological systems. It was suggested that such an aggregation is a common feature of peptides and proteins. The design of new therapy for treating amyloid related diseases demands an extensive comprehension on the mechanism of the formation and propagation of such aggregates in complex cellular environment. To this end, discrete MD (DMD) simulation is a good theoretical tool since it is computational fast due to the discrete nature of its potential and the model is flexible enough to model chain dynamics. For DMD, the discontinuous feature of potentials makes all the motions propagate at constant speed; therefore the processing of events are algebra calculations rather than the integration of dynamic equations, greatly reducing the computational demands. The collisions and propagations of the beads are processed sequentially, and all these events build up the dynamics of systems. The DMD simulation has achieved many successes in the study of protein folding pathway, the structure of the transition states and denatured state ensemble, and especially protein–protein association and aggregation (*154–162*).

The EAK-series peptides are typical model systems in studying the self-assembling of peptide and formation of amyloid fibrils. These peptides are composed of three kinds of amino acids, the glutamic acid (E) carrying negative charge, the neutral alanine (A), and the lysine (K) carrying positive charge. The amino acids are often arranged with a certain order along the sequence. Among these peptides, the EAK16-IV is an ionic self-complementary polyalanine-based peptide chain and can form different kinds of aggregate structures at various conditions (*161*), which make it an interesting system to study the α

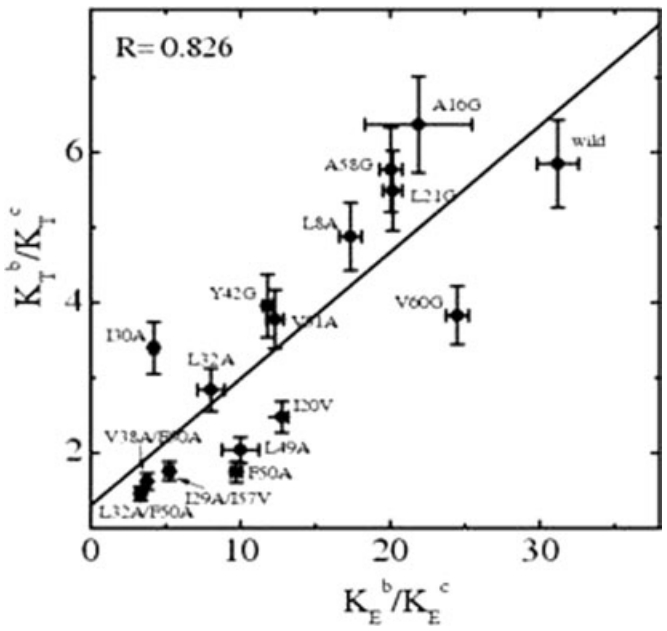

Figure 12. The correlation of the folding rate of protein CI2 and its mutants between the simulation results (indicated by the subscript index T) and those obtained from experiments (indicated by the subscript index E). The experimental values were taken from ref. (*153*). The superscript index c and b represent the folding rates inside the chaperonin cavity and in bulk, respectively. The correlation coefficient is $R = 0.826$. The figure is taken from ref. (*60*), with permission.

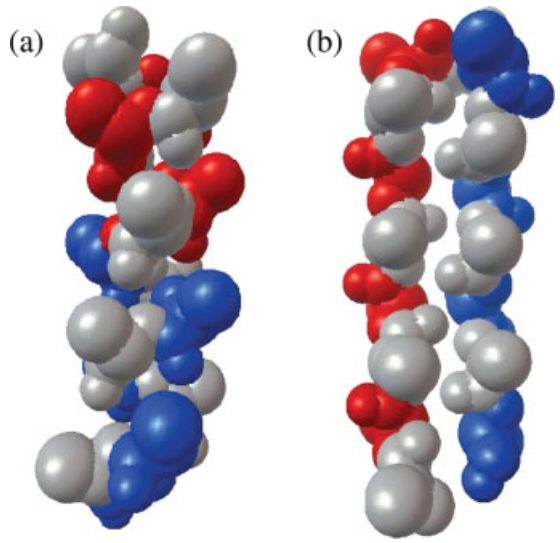

Figure 13. Snapshots of a full a-helix conformation (a) which has 12 a-helical hydrogen bonds with four a-helical turns, and a hairpin-like conformation (b) with a five residue turn located at the center. These conformations are colored based on the residue types, red for E, gray for A, and blue for K. The figure is taken from ref. (*161*), with permission.

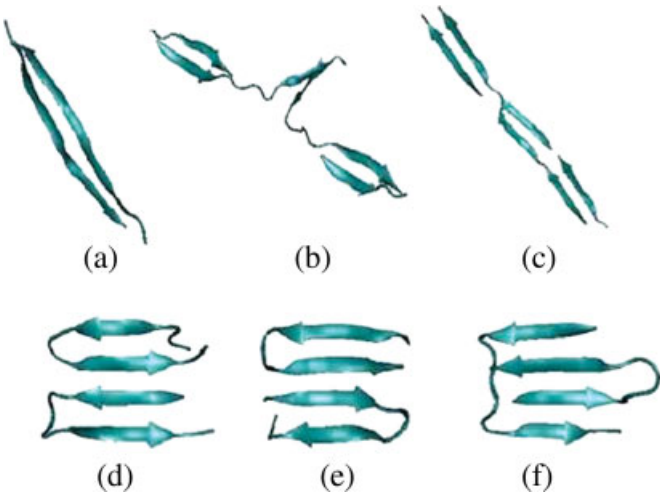

Figure 14. The typical dimeric structures for EAK16-IV peptide. (a) antiparallel β-sheet dimer, (b) mixed β-sheet dimer, (c) parallel β-sheet dimer, (d) parallel β-hairpin dimer, (e) antiparallel β-hairpin dimer, (f) mixed β-hairpin dimer. (a), (b), and (c) are classified into sheet-like dimer, and (d), (e), and (f) are hairpin-like dimer. The sheet-like dimers are kinetically preferable to hairpin-like dimers as the concentration increases. The figure is taken from Ref. (*161*), with permission.

helix-β structure conformational conversion and self-assembling. We developed a five-bead discrete model in which five beads, including the α-carbon (Cα), the carbonyl carbon (C) and oxygen (O), the amide nitrogen (N), and the first atom of side chain (Cβ), and the hydrogen bonds between C, N, and O in the peptide plane were explicitly modeled. The structure dynamics of monomers was studied and it was found that there is a competition between two types of structural motifs, α-helix and β-hairpin (Fig. 13); different strength of electrostatic interaction favors different structures. The hairpin-like structures would be dominant for strong electrostatic interactions both thermodynamically and kinetically, showing the importance of this interaction on the formation of hairpin-like structures. Another interesting observation was that the molecular concentration contributes essentially to the shape of dimers; the sheet-like dimers are kinetically preferable to hairpin-like dimers as the concentration increases (Fig. 14). This work demonstrated that the electrostatic interactions and kinetic factors might be important for the ordered amyloid aggregates.

METAL-COUPLED PROTEIN FOLDING

Many proteins need the help of cofactors for their proper functions. Metal ions, *e.g.*, Zn(II), Ca(II), Cu(II), etc., are a kind of important cofactors assisting the folding and stabilization of a large number of proteins (*163*). Because of the involvement of metal ions, the protein folding exhibits quite dif-

ferent behaviors compared with the protein folding without these cofactors. Understanding the metal coupled folding can significantly extend the current knowledge of the protein folding, which is mostly obtained based on the studies of spontaneous protein folding.

The zinc coupled folding of Cys_2His_2 type Zinc-finger is small protein motif in which the zinc ion is essential for its proper folding and stabilization. The native structure of the classical zinc-finger contains an N-terminal β-hairpin and a C-terminal α-helix which are stabilized by zinc binding and hydrophobic core (*164*). It represents a wide class of proteins for which the folding or other functional motions depends on the binding of metal-cofactors. Although many features have been revealed experimentally, a detailed atomistic picture of the coupling between metal binding and conformational motions is still lacking due to the limited temporal and spatial resolutions in experiments.

The MD method has proven to be a powerful tool in revealing the detailed mechanism of protein folding. However, when applied to the metal coupled folding, the conventional MD models need to be modified due to the following new features introduced by the involvement of metal ions. First, the formation and breakup of the coordination bonds between metal ion and its ligands occur frequently and the charge can transfer between metal ion and ligand atoms during the folding/unfolding process. Such quantum mechanical process was usually not modeled in the force field of conventional classical MD simulation explicitly. Second, the titratable ligands can be deprotonated when coordination bond forms, due to the strong oxidative activity of involved metal ions. This process contributes to the experimentally observed enthalpy change significantly. Again, it was not modeled in the conventional classical MD.

In ref. (*165*), we developed a theoretical model for zinc coupled protein folding by modifying the conventional MD method. In this model, the coordination bonds between the Zn(II) and liganding atoms were described by a nonbonded model (van de Waals and electrostatic interactions) (*166*). The charge transfer effect was included by appropriately integrating the quantum-mechanically derived quantities into the classical MD model following the idea of Sakharov and Lim (*167*). The metal ion induced protonation/deprotonation of ligand residues and the metal-induced polarization effect of the imidazole ring in the histidine residues were implemented by an empirical way (See ref. (*165*) for the the details). With these modifications, the metal coupled protein folding and other functional motions can be appropriately described by classical MD method.

On the basis of this model, we performed long time scale all-atom MD simulations for a Cys2His2 type zinc-finger with explicit solvent using the replica exchange MD method. The simulation results showed that the Zn(II) binds to the peptide at the very beginning of the folding, indicating that the Zn(II) participates in the whole folding process, not just stabilizing the native structure. The Zn(II) binds to the two conserved cystein residues earlier than the two conserved histidine residues, which is consistent with the experimental observations based on Raman spectra and Uv–vis absorption spectra (*168, 169*). Before all the four native coordinate bonds are formed, other non-native ligands, *i.e.*, water molecules or/and atoms of other residues, can be coordinated to the zinc ion. Therefore, during the zinc-coupled folding process, the ligand exchange is necessary for the protein to achieve correct coordination structure. Such misligation and ligand exchange may be a unique feature in metal-coupled protein folding and other functional motions.

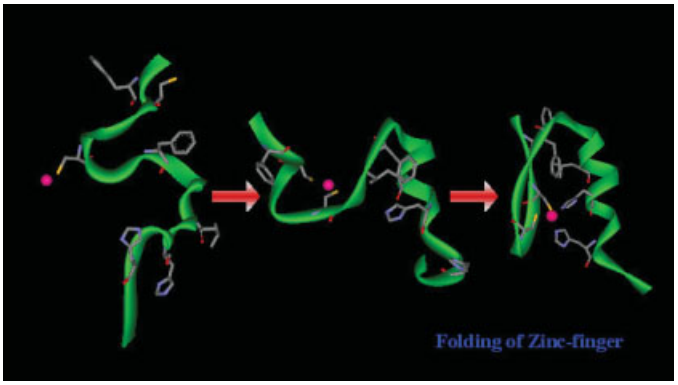

Figure 15. A schematic representation of the Zn(II) coupled folding of the Cys_2His_2 Zinc-finger. From left to right, the figures are the denatured state, an intermediate state and the native structure, respectively.

By analyzing the folding free energy landscape, a folding pathway was deduced. It was found that the folding initiates with a hydrophobic collapse, during which the α-helix may be partially formed although the stability is low. Then the folding proceeds with the full formation of the β-hairpin and partial formation and stabilization of the α-helix. This is a rate-limiting step since a high barrier needs to be overcome. Finally, the folding finishes with the full formation of the α-helix. Note that such a folding scenario is strongly coupled with the binding of the Zn(II) (Fig. 15). Comparisons between the folding free energy landscapes with and without zinc binding indicated that the zinc binding has dramatic contributions to the folding free energy landscape. Instead of just changing the relative populations of each major state in the conformational space, binding of the Zn(II) completely changes the pattern of the free energy landscape by eliminating some major states and inducing several new basins of attractions. These results again suggested that the Zn(II) is actively involved in the whole folding process, and its binding can direct and modulate the folding and stabilization of the secondary structures. Also, this crucial role of zinc binding is mediated by the packing of the conserved hydrophobic residues (*165*).

In ref. (*170*), Dahivat et al. designed a new protein FSD-1 based on the classical zinc-finger by replacing the binding zinc with a much larger hydrophobic core. This new protein is capable of folding to the same structure as the original zinc-finger.

In ref. (*171*), we performed intensive all atom MD simulations for protein FSD-1. The results showed that the designed protein has quite different folding pathway compared with the original zinc-finger. This difference also supports that the zinc binding plays a crucial role in folding of zinc-finger.

Recently, accumulating evidence suggests that the aggregation of the amyloid-β peptide (Aβ) is the main pathogeny of Alzheimer's disease (AD), and the metal binding can promote the aggregation and the development of the AD (*172–174*). It is helpful for treating the related diseases to investigate the molecular mechanism of the metal induced aggregation and fibrillation of Aβ peptide and to detect possible factors which can inhibit the pathogenic binding of metal ions. It is believed that two possible mechanisms can result in the metal induced aggregation of Aβ (*175–177*): (1) the metal ion binds to two Aβ monomers simultaneously and promotes the aggregation by bridging the peptides; (2) each metal ion binds to one monomer and results in some specific conformational transition, which contributes to the Aβ aggregation. In ref. (*177*), by comparing the conformational distributions of Aβ (Aβ_{40}) peptides with and without zinc binding based on all-atom MD simulations, we showed that the zinc binding can affect the conformational distribution of Aβ monomer dramatically. Particularly, we found that the zinc binding can increase the formation probabilities of the β-strand in the central hydrophobic cluster, the salt bridge Asp23-Lys28, and the turn comprising the residues 23–28. These local structures play important roles in the Aβ aggregation (*179–181*). Therefore, the peptide with zinc binding samples more conformations which are prone to aggregate, suggesting that the metal induced conformational variation is one of the possible mechanisms for metal promoted aggregation of Aβ peptide.

Although the earlier results were obtained mostly based on the study of a zinc-finger peptide, knowledge of the metal-coupled folding for this motif could provide insight into the general mechanism of metal-cofactor-dependent protein folding because the processes such as metal binding, ligand exchange, and metal-induced secondary structure conversion exhibited by zinc-finger may also occur during the folding and functional motions of other metalloproteins.

CONCLUSION

The MD simulation is a paramount tool for studying protein folding. It has achieved great success in the past two decades. Besides the recent developments described in the earlier context, there are more new applications of computer simulations, such as the study of protein folding in *in vivo* cell environment (the crowdness effect), the cotranslational folding, the chaperone assisted folding, and the simulation of amyloid fibril aggregation and oligomer formation. The territory of computer simulations has even been expanded to large protein complex or molecular machines. For example, it has been used to study the stability of nucleosomes, the dynamics of ion channels and virus capsides, the motion of molecular motors, and the synthesis process of proteins in ribosome (*182–190*). With the continuing improvement in computational power as well as in methodology, both temporal and spatial region accessible to simulations will be broadened further; it can be expected that computer simulations will prove invaluable in an even wider field and provide more exciting insights into structural biology.

ACKNOWLEDGEMENTS

This work was supported by National Natural Science Foundation of China (Nos. 10834002, 10504012 and 10704033) and National Basic Research Program of China (Nos. 2006CB910302 and 2007CB814800). The authors acknowledge Shanghai Supercomputer Center for providing computational resources.

REFERENCES

1. Anfinsen, C. B., Haber, E., Sela, M., and White, F. H., Jr. (1961) The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc. Natl. Acad. Sci. USA* **47**, 1309–1314.
2. Anfisen, C. B. (1973) Principles that govern the folding of protein chains. *Science* **181**, 223–230.
3. Levinthal, C. (1968) Are there pathways for protein folding? *J. Chim. Phys.* **65**, 44–45.
4. Shakhnovich, E. I., Farztdinov, G., Gutin, A. M., and Karplus, M. (1991) Protein folding bottlenecks: a lattice Monte Carlo simulation. *Phys. Rev. Lett.* **67**, 1665–1668.
5. Chan, H. S. and Dill, K. A. (1991) Polymer principles in protein structure and stability. *Anuu. Rev. Biophys. Biophys. Chem.* **20**, 447–490.
6. Chan, H. S. and Dill, K. A. (1994) Transition states and folding dynamics of proteins and heteropolymers. *J. Chem. Phys.* **100**, 9238–9257.
7. Chan, H. S. and Dill, K. A. (1996) Comparing folding codes for proteins and polymers. *Proteins: Struct. Funct. Genet.* **24**, 335–344.
8. Li, H., Helling, R., Tang, C., and Wingreen, N. (1996) Emergence of preferred structures in a simple model of protein folding. *Science* **273**, 666.
9. Bryngelson, J. D., Onuchic, J. N., Socci, N. D., and Wolynes, P. G. (1995) Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins: Struct. Funct. Genet.* **21**, 167–195.
10. Camacho, C. J. and Thirumalai, D. (1993) Kinetics and thermodynamics of folding in model proteins. *Proc. Natl. Acad. Sci. USA* **90**, 6369–6372.
11. Sali, A., Shakhnovich, E., and Karplus, M. (1994) How does a protein fold? *Nature* **369**, 248–251.
12. Dill, K. A., Bromberg, S., Yue, K., Fiebig, K. M., Yee, D. P., Thomas, P. D., Chan, H. S. (1995) Principles of protein folding—a perspective from simple exact models. *Protein Sci.* **4**, 561–602.
13. Socci, N. D., Onuchic, J. N., and Wolynes, P. G. (1996) Diffusive dynamics of the reaction coordinate for protein folding funnels. *J. Chem. Phys.* **104**, 5860–5868.
14. Kolinski, A., Milik, M., and Skolnick, S. (1991) Static and dynamic properties of a new lattice model of polypeptide chains. *J. Chem. Phys.* **94**, 3978.
15. Kolinski, A. and Skolnick, J. (1992) Discretized model of proteins. I. Monte Carlo study of cooperativity in homopolypeptides. *J. Chem. Phys.* **97**, 9412.
16. Kolinski, A. (2004) Protein modeling and structure prediction with a reduced representation. *Acta Biochim. Polonica* **51**, 349–371.
17. Wolynes, P. G., Onuchic, J. N., and Thirumalai, D. (1995) Navigating the folding routes. *Science* **267**, 1619–1620.

18. Onuchic, J. N. Wolynes, P. G., Luthey-Schulten, Z., and Socci, N. D. (1995) Towards an outline of the topography of a realistic protein folding funnel. *Proc. Natl. Acad. Sci. USA* **92**, 3626–3630.
19. Dill, K. A. and Chen, H. S. (1997) From Levinthal to pathways to funnels. *Nat. Struct. Biol.* **4**, 10–19.
20. Colombo, C. and Micheletti, C. (2006) Protein folding simulations: combining coarse-grained models and all-atom molecular dynamics. *Theor. Chem. Acc.* **116**, 75–86.
21. Wang, Z. X. (1994) Assessing the accuracy of protein secondary structure. *Nat. Struct. Biol.* **1**, 145–146.
22. Wang, Z. X. (1996) How many fold types of protein are there in nature? *Proteins: Struct. Funct. Genet* **26**, 186–191.
23. Liu, X. S., Fan, K., and Wang, W. (2004) The number of protein folds and their distribution over families in Nature. *Proteins* **54**, 491–499.
24. Veitshans, T., Klimov, D., and Thirumalai, D. (1996) Protein folding kinetics: timescales, pathways and energy landscapes in terms of sequence-dependent properties. *Fold. Des.* **2**, 1–22.
25. Clementi, C., Vendruscolo, M., Maritan, A., and Domany, E. (1996) Folding Lennard-Jones proteins by a contact potential. *Proteins: Struct. Funct. Genet.* **37**, 544–553
26. Clementi, C., Jennings, P. A., and Onuchic, J. N. (2000) How native-state topology affects the folding of dihydrofolate reductase and interleukin-1beta. *Proc. Natl. Acad. Sci. USA* **97**, 5871–5876.
27. Koga, N. and Takada, S. (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.
28. Alm, E., Morozov, A. V., Kortemme, T., and Baker, D. (2002) Simple physical models connect theory and experiment in protein folding kinetics. *J. Mol. Biol.* **322**, 463–476.
29. Paci, E., Vendruscolo, M., and Karplus, M. (2002) Native and non-native interactions along protein folding and unfolding pathways. *Proteins: Struct. Funct. Genet.* **47**, 379–392.
30. Moghaddam, M. S., Shimizu, S., and Chan, H. S. (2005) Temperature dependence of three-body hydrophobic interactions: potential of mean force, enthalpy, entropy, heat capacity and nonadditivity. *J. Am. Chem. Soc.* **127**, 303–316.
31. Li, A. and Daggett, V. (1994) Characterization of the transition state of protein unfolding by use of molecular dynamics: chymotrypsin inhibitor 2. *Proc. Natl. Acad. Sci. USA* **91**, 10430–10434.
32. Guo, Z., Brooks, C. L., III, and Boczko, E. M. (1997) Exploring the folding free energy surface of a three-helix bundle protein. *Proc. Natl. Acad. Sci. USA* **94**, 10161–10166.
33. Duan, Y. and Kollman, P. A. (1998) Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science* **282**, 740–744.
34. Schaefer, M., Bartels, C., and Karplus, M. (1998) Solution conformations and thermodynamics of structured peptides: molecular dynamics simulation with an implicit salvation model. *J. Mol. Biol.* **284**, 835–848.
35. Ferrara, P. and Caflisch, A. (2000) Folding simulations of three-stranded antiparallel beta-sheet peptide. *Proc. Natl. Acad. Sci. USA* **97**, 10780–10785.
36. Garcia, A. E. and Sanbonmatsu, K. Y. (2001) Exploring the energy landscape of a beta hairpin in explicit solvent. *Proteins: Struct. Funct. Genet.* **42**, 345–354.
37. Zhou, R. H., Berne, B. J., and Germain, R. (2001) The free energy landsacpe for beta hairpin folding in explicit water. *Proc. Natl. Acad. Sci. USA* **98**, 14931–14936.
38. Snow, C. D., Nguyen, H., Pande, V. S., and Gruebele, M. (2002) Absolute comparison of simulated and experimental protein-folding dynamics. *Nature* **420**, 102–106.
39. Shea, J. E., Onuchic, J. N., and Brooks, C. L., III. (2002) Probing the folding free energy landscape of the src-SH3 protein domain. *Proc. Natl. Acad. Sci. USA* **99**, 16064–16068.
40. Shen, M. Y. and Freed, K. F. (2002) All-atom fast protein folding simulations: the villin headpiece. (2002) *Proteins: Struct. Funct. Genet.* **49**, 439–445.
41. Bolhuis P. G. (2003) Transition-path sampling of beta-hairpin folding. *Proc. Natl. Acad. Sci. USA* **100**, 12129–12134.
42. Pitera, J. W. and Swope, W. (2003) Understanding folding and design: replica-excahnge simulations of trp-cage miniproteins. *Proc. Natl. Acad. Sci. USA* **100**, 7587–7592.
43. Zhou, R. H. (2003) Trp-cage: folding free energy landscape in explicit water. *Proc. Natl. Acad. Sci. USA* **100**, 13280–13285.
44. Yang, W. Y., Pitera, J. W., Swope, W. C., and Gruebele, M. (2004) Heterogeneous folding of trpzip hairpin: full atom simulation and experiment. *J. Mol. Biol.* **336**, 241–251.
45. Zhang, J. Qin, M., and Wang, W. (2006) Folding mechanism of beta-hairpin studied by replica exchange molecule simulations. *Proteins* **62**, 672–685.
46. Okazaki, K., Koga, N., Takada, S., Onuchic, J. N., and Wolynes, P. G. (2006) Multiple-basin energy landscapes for large-amplitude conformational motions of proteins: structure-based molecular dynamics simulations. *Proc. Natl. Acad. Sci. USA* **103**, 11844–11849.
47. Wu, L., Zhang, J., Qin, M., Liu, F., and Wang, W. (2008) Folding of proteins with an all-atom Go-model. *J. Chem. Phys.* **128**:235103, 1–8.
48. Liu, H., Luo, Q., Duan, Z., and Shi, Y. (1999) Structure-based ligand design by dynamically assembling molecular building blocks at active site, *Proteins* **36**, 462–470.
49. Liu, H., Zhang, Y., and Yao, W. (2000) How is the active site of enolase organized to catalyze two different reaction steps? *J. Am. Chem. Soc.* **122**, 6560–6570.
50. Ding, H., Yang, Y., Zhang, J., Wu, J., Liu, H., and Shi, Y. (2005) Structural basis for SUMO-E2 interaction revealed by a complex model using docking approach in combination with NMR data. *Proteins* **61**, 1050–1058.
51. Service, R. F. (2008) Problem solved (sort of). *Science* **321**, 784–786.
52. Bradley, P., Misura, K. M. S., and Baker, D. (2005) Toward high-resolution de novo structure prediction for small proteins. *Science* **309**, 1868–1871.
53. Shea, J. E. and Brooks, C. L., III. (2001) From folding theories to folding proteins: a review and assessment of simulation studies of protein folding and unfolding. *Annu. Rev. Phys. Chem.* **52**, 499.
54. Norberg, J. and Nilsson, L. (2003) Advances in biomolecular simulations: methodology and recent applications. *Qtly. Rev. Biophys.* **36**, 257.
55. Scheraga, H. A., Khalili, M., and Liwo, A. (2007) Protein-folding dynamics: overview of molecular simulation techniques. *Annu. Rev. Phys. Chem.* **58**, 57.
56. Taketomi, H., Ueda, Y., and Go, N. (1975) Studies on protein folding, unfolding and fluctuations by computer simulation. I. The effect of specific amino acid sequence represented by specific inter-unit interactions. *Int. J. Pept. Protein Res.* **7**, 445.
57. Clementi, C., Nymeyer, H., and Onuchic, J. N. (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.
58. Plotkin, S. S. and Onuchic, J. N. (2000) Investigation of routes and funnels in protein folding by free energy functional method. *Proc. Natl. Acad. Sci. USA* **97**, 6509–6514.
59. Zhang, J., Qin, M., and Wang, W. (2005) Multiple folding mechanisms of protein ubiquitin. *Proteins* **59**, 565–579.
60. Xu, W. X., Wang, J., and Wang, W. (2005) Folding behavior of chaperonin-mediated substrate protein. *Proteins.* **61**, 777–794.
61. Lindberg, M. O., Tångrot, J., Otzen, D. E., Dolgikh, D. A., Finkelstein, A. V., and Oliveberg, M. (2001) Folding of circular permutants with decreased contact order: general trend balanced by protein stability. *J. Mol. Biol.* **314**, 891–900.

62. Lingberg, M., Tångrot, J., and Oliveberg, M. (2002) Complete change of the protein folding transition state upon circular permuation. *Nat. Struct. Biol.* **9**, 818–822.
63. Otzen, D. E. and Oliveberg, M. (2002) Conformational plasticity in folding of the split beta-alpha-beta protein S6: evidence for burst-phase disruption of the native state. *J. Mol. Biol.* **317**, 613–627.
64. Hubner, I. A., Oliveberg, M., and Shakhnovich, E. I. (2004) Simulation, experiment, and evolution: Understanding nucleation in protein S6 folding. *Proc. Natl. Acad. Sci. USA* **101**, 8354–8359.
65. Matysiak, S., Clementi, C. (2004) Optimal combination of theory and experiment for the characterization of the protein folding landscape of S6: How far can a minimalist model go? *J. Mol. Biol.* **343**, 235–248.
66. Wu, L., Zhang, J., Wang, J., Li, W. F., and Wang, W. (2007) Folding behavior of ribosomal protein S6 studied by modified Gō-like model. *Phys. Rev. E* **75**, 1–8.
67. Fersht, A. R. (1999) *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*. Freeman, New York, NY.
68. Daggett, V. and Fersht, A. R. (2003) The present view of the mechanism of protein folding. *Nat. Rev. Mol. Cell Biol.* **4**, 497–502.
69. Sevier, C. S. and Kaiser, C. A. (2002) Formation and transfer of disulfide bonds in living cells. *Nat. Rev. Mol. Cell Biol.* **3**, 836–847.
70. Goldenberg, D. P. (1992) Native and non-native intermediates in the BPTI folding pathway. *Trends Biochem. Sci.* **17**, 257–261.
71. Robinson, A. S. and King, J. (1997) Disulphide-bonded intermediate on the folding and assembly pathway of a non-disulphide bonded protein. *Nat. Struct. Biol.* **4**, 450–455.
72. Flory, P. J. (1956) Theory of elastic mechanisms in fibrous proteins. *J. Am. Chem. Soc.* **78**, 5222–5235.
73. Camacho, C. J. and Thirumalai, D. (1995) Modeling the role of disulfide bonds in protein folding: entropic barriers and pathways. *Proteins: Struct., Funct. Genet.* **22**, 27–40.
74. Chaffotte, A. F., Guillou, Y., and Goldberg, M. E. (1992) Kinetic resolution of peptide bond and side chain far-UV circular dichroism during the folding of hen egg white lysozyme. *Biochemistry* **31**, 9694–9702.
75. Schonbrunner, N., Koller, K. P., and Kiefhaber, T. (1997) Folding of the disulfide-bonded-sheet protein tendamistat: rapid two-state folding without hydrophobic collapse. *J. Mol. Biol.* **268**: 526–538.
76. Schonbrunner, N., Pappenberger, G., Scharf, M. Engels, J., and Kiefhaer, T. (1997) Effect of preformed correct tertiary interactions on rapid two-state tendamistat folding: evidence for hairpins as initiation sites for b-sheet formation. *Biochemistry* **36**: 9057–9065.
77. Bachmann, A. and Kiefhaber, T. (2001) Apparent two-state tendamistat folding is sequential process along a defined route. *J. Mol. Biol.* **306**: 375–386.
78. Pappenberger, G., Aygun, H., Engels, J. W., Reimer, U., Fischer, G., and Kiefhaber, T. (2001) Nonprolyl cis peptide bonds in unfolded proteins cause complex folding kinetics. *Nat. Struct. Biol.* **8**: 452–458.
79. Volgl, T., Brengelmann, R., Hinz, H. J., Scharf, M., Lotzbeger, M., and Engels, J. W. (1995) Mechanism of protein stabilization by disulfide bridges: calorimetric unfolding studies on disulfide-deficient mutants of the a-amylase inhibitor tendamistat. *J. Mol. Biol.* **254**: 481–496.
80. Qin, M., Zhang, J., and Wang, W. (2006) Effects of disulfide bonds on folding behavior and mechanism of the beta-sheet protein Tendamistat. *Biophysical J.* **90**: 272–286.
81. Fersht, A. (1997) Nucleation mechanisms in protein folding. *Curr. Opin. Struct. Biol.* **7**, 3–9.
82. Sugita, Y. and Okamoto, Y. (1999) Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **314**, 141–151.
83. Xu, C., Wang, J., and Liu, H. (2008) A Hamiltonian replica exchange approach and its application to the study of side chain type and neighbor effects on peptide backbone conformations. *J. Chem. Theor. Comput.* **4**, 1348–1359.
84. Munoz, V., Henry, E. R., Hofrichter, J., and Eaton, W. A. (1998) A statistical model for beta-hairpin kinetics. *Proc. Natl. Acad. Sci. USA* **95**, 5872–5879.
85. Dinner, A. R., Lazaridis, T., and Karplus, M. (1999) Understanding beta-hairpin formation. *Proc. Natl. Acad. Sci. USA* **96**, 9068–9037.
86. Pande, V. S. and Rokhsar, D. S. (1999) Molecular dynamics simulations of unfolding and refolding of a beta-hairpin fragment of protein G. *Proc. Natl. Acad. Sci. USA* **96**, 9062–9067.
87. Klimov, K. and Thirumalai, D. (2000) Mechanisms and kinetics of beta-hairpin formation. *Proc. Natl. Acad. Sci. USA* **97**, 2544–2549.
88. Zhou, Y. Q. and Linhananta, A. (2002) Role of hydrophilic and hydrophobic contacts in folding of the second beta-hairpin fragment of protein G: molecular dynamics simulation studies of an all-atom model. *Proteins* **47**, 154–162.
89. Munoz, V., Thompson, P. A., Hofrichter, J., and Eaton, W. A. (2003) Folding dynamics and mechanism of beta-hairpin formation. *Nature* **390**, 196–199.
90. Wei, G. H., Mousseau, N., and Derreumaux, P. (2004) Complex folding pathways in a simple beta-hairpin. *Proteins* **56**, 464–474.
91. Du, D. G., Zhou, Y. J., Huang, C. Y., and Gai, F. (2004) Understanding the key factors that control the rate of beta-hairpin folding. *Proc. Natl. Acad. Sci. USA* **101**, 15915–15920.
92. Baker, D. (2000) A surprising simplicity to protein folding. *Nature* **405**, 39–42.
93. Onuchic, J. N., Luthey-Schelten, Z., and Wolynes, P. G. (1997) Theory of protein folding: the energy landscape perspective. *Annu. Rev. Phys. Chem.* **48**, 545–600.
94. Shakhnovich, E. I. and Finkelstein, A. V. (1989) Theory of cooperative transitions in protein molecules. I. Why denaturation of globular protein is a first-order phase transition. *Biopolymers* **28**, 1667–1680.
95. Kaya, H. and Chan, H. S. (2003) Solvation effects and driving forces for protein thermodynamic and kinetic cooperativity: how adequate is native-centric topological modeling? *J. Mol. Biol.* **326**, 911–930.
96. Chan, H. S., Shimizu, S., and Kaya, H. (2004) Cooperativity principles in protein folding. *Method Enzym.* **380**, 350–379.
97. Wang, J. and Wang, W. (2003) Folding transition of model protein chains characterized by partition function zeros. *J. Chem. Phys.* **118**, 2952–2963.
98. Pande, V. S., Grosberg, A., Tanaka, T., and Rokhsar, D.S. (1998) Pahtways for protein folding: is a new view needed? *Curr. Opin. Struct. Biol.* **8**, 68–79
99. Zhou, Y. Q. and Karplus, M. (1999) Interpreting the folding kinetics of helical proteins. *Nature* **401**, 400–403.
100. Brockwell, D. J., Simith, D. A., and Radford, S. E. (2000) Protein folding mechanisms: new methods and emerging ideas. *Curr. Opin. Struct. Biol.* **10**, 16–25.
101. Fersht, A. R. (1994) Characterizing transition states in protein folding an essential step in the puzzle. *Curr. Opin. Struct. Biol.* **5**, 79–84.
102. Chen, J., Wang, J., and Wang, W. (2004) Transition states for folding of circular – permuted proteins. *Proteins* **57**, 153–171.
103. Mirny, L. and Shakhnovich, E. I. (2001) Protein folding theory: from lattice to all-atom models. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 361–396.
104. Shakhnovich, E. I. (2006) Protein folding thermodynamics and dynamics: where physics, chemistry, and biology meet. *Chem. Rev.* **106**, 1559–1588.
105. Dill, K. A., Ozkan, S.B., Shell, M.S., and Weikl, T.R. (2008) The protein folding problem. *Annu. Rev. Biophys.* **37**, 289–316.
106. Plotkin, S. S. and Wolynes, P. G. (2003) Buffed energy landscapes: another solution to the kinetic paradoxes of protein folding. *Proc. Natl. Acad. Sci. USA* **100**, 4417–4422.
107. Yang, W. Y. and Gruebele, M. (2004) Folding lamda-repressor at its speed limit. *Biophys. J.* **87**, 596–608.
108. Liu, F. and Gruebele, M. (2007) Tuning lambda (6–85) towards downhill folding at its melting temperature. *J. Mol. Biol.* **370**, 574–584.

109. Sabelko, J., Ervin, J., and Gruebele, M. (1999) Folding lamda-repressor at its speed limit. *Proc. Natl. Acad. Sci. USA* **96**, 6031–6036.
110. Garcia-Mira, M. M., Sadqi, M., Fischer, N., Sanchez-Ruiz, J. M., and Munoz, V. (2002) Experimental Identification of Downhill Protein Folding. *Science* **298**, 2191–2195.
111. Ferguson, M., Schartau, P. J., Sharpe, T. D., Sato, S., and Fersht, A. R. (2004) One-state downhill versus conventional protein folding. *J. Mol. Biol.* **344**, 295–301.
112. Ferguson, M., Sharpe, T. D., Schartau, P. J., Sato, S., Allen, M. D., Johnson, C. M., Rutherford, T. J., and Fersht, A. R. (2005) Ultra-fast barrier-limited folding in the peripheral subunit-binding domain family. *J. Mol. Biol.* **353**, 427–446.
113. Yu, W., Chung, K., Cheon, M., Heo, M., Han, K.H., Ham, S., and Chang, I. (2008) Cooperative folding kinetics of BBL protein and peripheral subunit-binding domain homologues. *Proc. Natl. Acad. Sci. USA* **105**, 2397–2402.
114. Kubelka, J., Hofrichter, J., and Eaton, W. A. (2004) The protein folding "speed limit". *Curr. Opin. Struct. Biol.* **14**, 76–88.
115. Sadqi, M., Fushman, D., and Munoz, V. (2006) Atom-by-atom analysis of global downhill protein folding. *Nature* **442**, 317–321.
116. Liu, F., Du, D., Fuller, A. A., Davoren, J. E., Wipf, P., Kelly, J. W., and Gruebele, M. (2008) An experimental survey of the transition between two-state and downhill protein folding scenarios. *Proc. Natl. Acad. Sci. USA* **105**, 2369–2374.
117. Li, P., Oliva, F. Y., Naganathan, A. N., and Munoz, V. (2009) Dynamics of one-state downhill protein folding. *Proc. Natl. Acad. Sci. USA* **106**, 103–108.
118. Cho, S. S., Weinkam, P., and Wolynes, P. G. (2008) Origins of barriers and barrierless folding in BBL. *Proc. Natl. Acad. Sci. USA* **105**, 118–123.
119. Knott, M. and Chan, H. S. (2006) Criteria for downhill protein folding: calorimetry, Chevron plot, kinetic relaxation, and single-molecule radius of gyration in chain models with subdued degrees of cooperativity. *Proteins* **65**, 373–391.
120. Zhang, J., Li, W. F., Wang, J., Qin, M., and Wang, W. (2008) All-atom replica exchange molecular simulation of protein BBL. *Proteins* **72**, 1038–1047.
121. Zuo, G. H. Wang, J., and Wang, W. (2006) Folding with downhill behavior and low cooperativity of proteins. *Proteins* **63**, 165–173.
122. Zuo, G. H., Zhang, J. Wang, J., and Wang, W.2005. Folding behavior for protein BBL and E3BD with Go-like model. *Chin. Phys. Lett.* **22**: 1809–1812.
123. Zhao, Z., Peng, Y., Hao, S. F., Zeng, Z. H., and Wang, C. C. (2003) Dimerization by domain hybridization bestows chaperone and isomerase activities. *J. Biol. Chem.* **278**, 43292–43298.
124. Martin, J. and Hartl, F. U. (1997) The effect of macromolecular crowding on chaperonin-mediated protein folding. *Proc. Natl. Acad. Sci. USA* **94**, 1107–1112.
125. Chen, J., Song, J. L., Zhang, S., Wang, Y., Cui, D. F., and Wang, C. C. (1999) Chaperone activity of DsbC. *J. Biol. Chem.* **274**, 19601–19604.
126. Yao, Y., Zhou, Y. C., and Wang, C. C. (1997). Both the isomerase and chaperone activities of protein disulfide isomerase are required for the reactivation of reduced and denatured acidic phospholipase A2. *EMBO J.* **16**, 651–658.
127. Abulimiti, A., Fu, X., Gu., L., Feng, X., and Chang, Z. (2003) Mycobacterium tuberculosis Hsp16.3 nonamers are assembled and reassembled via trimer and hexamer intermediates. *J. Mol. Biol.* **326**, 1013–1023.
128. Hong, W., Jiao, W., Hu, J., Zhang, J., Liu, C., Fu, X., Shen, D., Xia, B., and Chang, Z. (2005) Periplasmic protein HdeA exhibits chaperone-like activity exclusively within stomach pH range by transforming into disordered conformation. *J. Biol. Chem.* **280**, 27029–27034.
129. Ellis, R. J. (2001) Macromolecular crowding: obvious but underappreciated. *Trends Biochem. Sci.* **26**, 597–604.
130. Minton, A. P. (2006) Macromolecular crowding. *Curr. Biol.* **16**, R269–R271.
131. Hardesty, B. and Kramer, G. (2001) Folding of a nascent peptide on the ribosome. *Prog. Nucleic Acid Res. Mol. Biol.* **66**, 41–46.
132. Lu, J. L. and Deutsch, C. (2005) Folding zones inside the ribosomal exit tunnel. *Nature Struct. Biol.* **12**, 1123–1129.
133. Das, D., Das, A., Samanta, D., Ghosh, J., Dasgupta, S., Bhattacharya, A., Basu, A., Sanyal, S., and Gupta, C. D. (2008) Role of the ribosome in protein folding. *Biotech. J.* **3**, 999–1009.
134. Weibezahn, J., Schlieker, C., Tessarz, P., Mogk, A., and Bukau, B. (2005) Novel insights into the mechanism of chaperone-assisted protein disaggregation. *Biol. Chem.* **386**, 739–744.
135. Hatter, D. M., Minton, A. P., and Howlett, G. J. (2002) Macromolecular crowding accelerates amyloid formation by human apolipoprotein C-II. *J. Biol. Chem.* **277**, 7824–7830.
136. Munishikana, L. A., Cooper, E. M., Uversky, V. N., and Fink, A. L. (2004) The effect of macromolecular crowding on protein aggregation and amyloid fibril formation. *J. Mol. Recog.* **17**, 456–464 (2004).
137. Thirumalai, D., Klimov, D. K., and Lorimer, G. H. (2003) Caging helps proteins fold. *Proc. Natl. Acad. Sci. USA* **100**, 11195–11197.
138. Takagi, F., Koga, N., and Takada, S. (2003) How protein thermodynamics and folding mechanisms are altered by the chaperonin cage: molecular simulations. *Proc. Natl. Acad. Sci. USA* **100**, 11367–11372.
139. Mittaland, J. and Best, R. B. (2008) Thermodynamics and kinetics of protein folding under confinement. *Proc. Natl. Acad. Sci. USA* **105**, 20233–20238.
140. Li, J., Zhang, S., and Wang, C. C. (2001) Effects of macromolecular crowding on the refolding of glucose-6-phosphate dehydrogenase and protein disulfide isomerase. *J. Biol. Chem.* **276**, 4396–34401.
141. Chapman, E., Farr, G. W., Usaite, R., Furtak, K., Fenton, W. A., Chaudhuri, T. K., Hondorp, E. R., Matthews, R. G., Wolf, A. G., Yates, J. R., Pypaert, M., and Horwich, A. L. (2006) Global aggregation of newly translated proteins in an Escherichia coli strain defcient of the chaperonin GroEL. *Proc. Natl. Acad. Sci. USA* **103**, 15800–15805.
142. Lin, Z., Madan, D., and Rye, H. S. (2008) GroEL stimulates protein folding through forced unfolding. *Nat. Struct. Biol.* **15**, 303–311.
143. Burt, M. C. and Dave, B. C. (2006) Externally tunable dynamic confinement effect in organosilica sol—gels. *J. Am. Chem. Soc.* **128**, 11750–11751.
144. Sorin, E. J. and Pande, V.S. (2006) Nanotube Confinement Denatures Protein Helices. *J. Am. Chem. Soc.* **128**, 6316–6317.
145. Horovitz, A. and Willison, K. R. (2005) Allosteric regulation of chaperonins. *Curr. Opin. Struct. Biol.* **15**, 646–651.
146. Jiao, W., Qian, M., Li, P., Zhao, L., and Chang, Z. (2005) The essential role of the flexible termini in the temperature-responsiveness of the oligomeric state and chaperone-like activity for the polydisperse small heat shock protein IbpB from *Escherichia coli*. *J. Mol. Biol.* **347**, 871–884.
147. Gu, L., Abulimiti, A., Li, W., and Chang, Z. (2002) Monodisperse Hsp16.3 nonamer exhibits dynamic dissociation and reassociation, with the nonamer dissociation prerequisit for chaperone-like activity. *J. Mol. Biol.* **319**, 517–526.
148. Chan, H.S. and Dill, K. A. (1996) A simple model of chaperonin-mediated protein folding. *Proteins* **24**, 345–351.
149. Lin, Z. and Rye, H. S. (2006) GroEL-mediated protein folding: making the impossible, possible. *Critical Rev. Biochem. Mol. Biol.* **41**, 211–239.
150. Baumketner, A., Jewett, A., and Shea, J. E. (2003) Effects of confinement in chaperonin assisted protein folding: rate enhancement by decreasing the roughness of the folding energy landscape. *J. Mol. Biol.* **332**, 701–313.
151. Brinker, A., Pfeifer, G., Kerner, M. J., Naylor, D. J., Hartl, F. U., and Hayer-Hartl, M. (2001) Dual function of protein confinement in chaperonin-assisted protein folding. *Cell* **107**, 223–233.

152. Stenberg, G. and Fersht, A. R. (1997) Folding of barnase in the presence of the molecular chaperone SecB. *J. Mol. Biol.* **274**, 268–275.
153. Itzhaki, L. S., Otzen, D. E., and Fersht, A. R. (1995) Nature and consequences of GroEL-protein interactions. *Biochemistry* **34**, 14581–14587.
154. Urbanc, B., Cruz, L., Yun, S., Buldyrev, S. V., Bitan, G., Teplow, D. B., and Standley, H. E. (2004) In silico study of amyloid beta-protein folding and oligomerizatin. *Proc. Natl. Acad. Sci. USA* **101**, 17345–17350.
155. Urbanc, B., Cruz, L., Ding, F., Sammond, D., Khare, S., Buldyrev, S. V., Sandley, H. E., and Dokholyan, N. V. (2004) Molecular dynamics simulation of amyloid beta-dimer formation. *Biophys. J.* **87**, 2310–2321.
156. Nguyen, H. D. and Hall, C. K. (2004) Molecular dynamics simulations of spontaneous fibril formation by random-coil peptides. *Proc. Natl. Acad. Sci. USA* **101**, 16180–16185.
157. Dokholyan, N. V. (2006) Studies of folding and misfolding using smplicied models. *Curr. Opin. Struct. Biol.* **16**, 79–85.
158. Chen, Y., Ding, F., Nie, H., Serohijos, A. W., Sharma, S., Wilcox, K. C., Yin, S., and Dokholyan, N. V. (2008) Protein folding: Then and now. *Arch. Biochem. Biophys.* **469**, 4–19.
159. Zhou, Y. Q. and Karplus, M. (1997) Folding thermodynamics of a model three-helix-bundle protein. *Proc. Natl. Acad. Sci. USA* **94**, 14429–14432.
160. Smith S. W., Hall, C. K., and Freeman, B. D. (1997) Molecular dynamics for polymeric fluids using discontinuous potentials. *J. Comp. Phys.* **134**, 16–30.
161. Yan, Z. Q., Wang, J., and Wang, W. (2008) Folding and dimerization of the ionic peptide EAK16-IV. *Proteins: Strct. Funct. Genet.* **72**, 150–162.
162. Jun, S., Hong, Y., Imamura, H., Ha, B. Y., Bechhoefer, J., and Chen, P. (2004) Self-assembly of the ionic peptide EAK16: the effect of charge distributions on self-assembly. *Biophys. J.* **87**, 1249–1259.
163. Wittung-Stafshede, P. (2002) Role of cofactors in protein folding. *Acc. Chem. Res.* **35**, 201–208.
164. Frankel, A. D., Berg, J. M., and Pabo, C. O. (1987) Metal-dependent folding of a single zinc finger from transcription factor IIIA. *Proc. Natl. Acad. Sci. USA* **84**, 4841–4845.
165. Li, W. F., Zhang, J., Wang, J., and Wang, W. (2008) Metal-coupled folding of Cys2His2 zinc-finger. *J. Am. Chem. Soc.* **130**, 892–900.
166. Stote, R. H. and Karplus, M. (1995) Zinc-binding in proteins and solution: a simple but accurate non-bonded representation. *Proteins* **23**, 12–31.
167. Sakharov, D. V. and Lim, C. (2005) Zn protein simulations including charge transfer and local polarization effects. *J. Am. Chem. Soc.* **127**, 4921–4929.
168. Shi, Y., Beger, R. D., and Berg, J. M. (1993) Metal binding properties of single amino acid deletion mutants of zinc finger peptides: studies using cobalt(II) as a spectroscopic probe. *Biophys. J.* **64**, 749–753.
169. Miura, T., Satoh, T., and Takeuchi, H. (1998) Role of metal–ligand coordination in the folding pathway of zinc finger peptides. *Biochim. Biophys. Acta* **1384**, 171–179.
170. Dahiyat, B. I. and Mayo, S. L. (1997) De novo protein design: fully automated sequence selection. *Science* **278**, 82–87.
171. Li, W. F., Zhang, J., and Wang, W. (2007) Understanding the folding and stability of a zinc finger-based full sequence design protein with replica exchange molecular dynamics simulations. *Proteins* **67**, 338–349.
172. Selkoc, D. J. (2003) Folding proteins in fatal ways. *Nature* **426**, 900–904.
173. Chiti, F. and Dobson, C. M. (2006) Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **75**, 333–366.
174. Adlard, P. A. and Bush, A. I. (2006) Metals and Alzheimer's disease. *J. Alzheimer's Disease* **10**, 145–163.
175. Miura, T., Suzuki, K., Kohata, N., and Takeuchi, H. (2000) Metal binding modes of Alzheimer's amyloid β-peptide in insoluble aggregates and soluble complexes. *Biochemistry* **39**, 7024–7031.
176. Karr, J. W., Akintoye, H., Kaupp, L. J., and Szalai, V. A. (2005) N-terminal deletions modify the Cu2+ binding site in Amyloid-β. *Biochemistry* **44**, 5478–5487.
177. Talmard, C., Guilloreau, L., Coppel, Y., Mazarguil, H., and Faller, P. (2007) Amyloid-beta peptide forms monomeric complexes with Cu(II) and Zn(II) prior to aggregation, *Chem. Bio. Chem.* **8**, 163–165.
178. Li, W., Zhang, J., Su, Y., Wang, J., Qin, M., and Wang, W. (2007) Effects of zinc binding on the conformational distribution of the Amyloid-β peptide based on molecular dynamics simulations. *J. Phys. Chem. B* **111**, 13814–13821.
179. Miravalle, L., Tokuda, T., Chiarle, R., Giaccone, G., Bugiani, O., Tagliavini, F., Frangione, B., and Ghiso, J. (2000) Substitutions at codon 22 of Alzheimer's Aβ peptide induce diverse conformational changes and apoptotic effects in human cerebral endothelial cells. *J. Biol. Chem.* **275**, 27110–27116.
180. Morimoto, A., Irie, K., Murakami, K., Masuda, Y., Ohigashi, H., Nagao, M., Fukuda, H., Shimizu, T., and Shirasawa, T. (2004) Analysis of the secondary structure of β-Amyloid (Aβ42) fibrils by systematic proline replacement. *J. Biol. Chem.* **279**, 52781–52788.
181. Sciarretta, K. L., Gordon, D. J., Petkova, A. T., Tycko, R., and Meredith, S. C. (2005) Aβ40-Lactam(D23/K28) models a conformation highly favorable for nucleation of amyloid. *Biochemistry* **44**, 6003–6014.
182. Aqvist, J. and Luzhkov, V. (2000) Ion permeation mechanism of the potassium channel. *Nature* **404**, 881–884.
183. Berneche, S. and Roux, B. (2001) Energetics of ion conduction through the K+ channel. *Nature* **414**, 73–77.
184. Ma, B. Y. and Nussinov, R. (2002) Stabilities and conformaitons of Alzheimer's beta-amyloid peptide oligomers ($Abeta_{16–22}$, $Abeta_{16–35}$, and $Abeta_{10–35}$): Sequence effects. *Proc. Natl. Acad. Sci. USA* **99**, 14126–14131.
185. Arkhipov, A., Freddolino, P. L., and Schulten, K. (2006) Stability and dynamics of virus capsides described by coarse-grained modeling. *Structure* **14**, 1767–1777.
186. Koga, N. and Takada, S. (2006) Folding-based molecular simulations reveal mechanisms of the rotary motor F-1-ATPase. *Proc. Natl. Acad. Sci. USA* **103**, 5367–5372.
187. Ding, F., LaRocque, J. J., and Dokholyan, N. (2005) Direction observation of protein folding, aggregation, and a prion-like conformational conversion. *J. Biol. Chem.* **280**, 40235–40240.
188. Elcock, A. H. (2006) Molecular simulations of cotranslational protein folding: Fragment stabilities, folding cooperativity, and trapping in the ribosome. *PLOS Comput. Biol.* **2**, e98, 0824–0841.
189. Ziv, G., Haran, G., and Thirumalai, D. (2006) Ribosome exit tunnel can entropically stabilize alpha-helices. *Proc. Natl. Acad. Sci. USA* **103**, 18956–18961.
190. Sharma, S., Ding, F., and Dokholyan, N. V. (2007) Multiscale modeling of nucleosome dynamics. *Biophys. J.* **92**, 1457–1470.