ELSEVIER

Biophysical Chemistry 133 (2008) 71 – 80

Biophysical Chemistry

http://www.elsevier.com/locate/biophyschem

How native proteins aggregate in solution: A dynamic Monte Carlo simulation

Lin Zhang, Diannan Lu, Zheng Liu *

Department of Chemical Engineering, Tsinghua University, Beijing, 100084, China

Received 29 October 2007; received in revised form 16 December 2007; accepted 16 December 2007
Available online 27 December 2007

Abstract

Aggregation of native proteins in solution is of fundamental importance with regard to both the processing and the utilization of proteins. In the present work, a dynamic Monte Carlo simulation has been performed to give a molecular insight into the way in which native proteins aggregate in solution and to explore means of suppressing aggregation, using two proteins of different compositions and conformations represented by a two-dimensional (2D) lattice model (HP model). It is shown that the native HP protein with accessible hydrophobic beads on its surface is prone to aggregation. The aggregation of this protein is intensified when the solution conditions favor the partially unfolded conformation as opposed to either the native or fully unfolded conformations. In this case, the partially unfolded proteins form the cores of aggregates, which may also encapsulate the native protein. One way to inhibit protein aggregation is to introduce polymers of appropriate hydrophobicity and chain length into the solution, such that these polymer molecules wrap around the hydrophobic regions of both the unfolded and folded proteins, thereby segregating the protein molecules. Our simulation is consistent with experimental observations reported elsewhere and provides a molecular basis for the behavior of proteins in liquid environments.

Keywords: Protein aggregation; Hydrophobic interaction; Polymer; Dynamic Monte Carlo simulation; HP model

1. Introduction

Protein aggregation in solution is a fundamental issue in relation to both the downstream processing of proteins, such as protein refolding [1], membrane separations [2–3], and preparative chromatography [4], as well as the utilization of proteins as biocatalysts [5] and pharmaceuticals [6]. While the amino acid composition and structure intrinsically determine the aggregation behavior of a protein, the effects of solution properties, including temperature, pH, salt type and concentration, co-solutes, preservatives, and surfactants, are also significant [7]. It has been shown [8–16] both experimentally and through simulations that aggregation starts from a partially unfolded conformation with more exposed hydrophobic regions as compared to its native counterpart. Current studies on protein aggregation are mainly concerned with protein folding where the protein is initially in the random coil conformation and efforts are directed towards delineating the competition between folding, misfolding, and aggregation [10,15,17–23]. Molecular descriptions of the aggregation of native protein in solution are not yet adequate, though the addition of water-soluble polymers is extensively applied to stabilize proteins in solution, selected examples of which are listed in Table 1.

Molecular simulation, as a powerful tool for exploring micro conformational transitions, has been widely used to study protein structural transitions. Dill et al. [31] proposed a two-dimensional lattice protein model (HP model), which highlights the hydrophobic interaction as the driving force for protein folding and aggregation, and which has since been widely used as a model protein. Istrail et al. [32] proposed a computer model for protein aggregation with competing productive folding, and were the first to present some background into the nature and significance of protein aggregation and the use of lattice Monte Carlo simulations in understanding other aspects of protein

* Corresponding author.
E-mail address: liuzheng@mail.tsinghua.edu (Z. Liu).

0301-4622/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2007.12.008

Table 1
Inhibition of protein aggregation during refolding by the addition of polymers

Polymer	Protein	Mechanism	References
PEG	Carbonic anhydrase B	Interaction of PEG with the first intermediate	[24–26]
	Lactate dehydrogenase	Preservation of α-helix with PEG	[27]
PNIPAAm	Carbonic anhydrase B	Formation of complexes with folded intermediates through hydrophobic interactions	[28]
	Lysozyme	Formation of PNIPAAm–lysozyme complex through hydrophobic interactions	[1,29,30]
Dextran	Aviscumine	Hydrogen-bonding interaction between protein and dextran molecules	[6]

folding. Leonhard et al. [33] proposed an interaction energy scale based on the Miyazawa–Jernigan forcefield to develop amino acid residue–solvent interactions for lattice Monte Carlo simulations of model proteins in water, and performed evaluations on a 27-mer 3D lattice model. Gupta and Hall [10,17,18,34,35] have subjected a 20-bead HP (HP20) protein model to a dynamic Monte Carlo simulation to study the folding pathway, the transition states of a single protein, and protein aggregation during the refolding process, starting from the fully denatured conformation. Bratko et al. [19], Costello et al. [20], Broglia et al. [15] and Cellmer et al. [36] have each used a 3D lattice model with Monte Carlo simulation to explore the competition between folding and aggregation. More computer simulation results on protein aggregation have been listed in a recent review by Cellmer et al. [37]. In a previous work, Lu et al. [38,39] performed dynamic Monte Carlo simulation of the refolding of HP20 protein assisted by surfactant and polymer molecules, and showed that the formation and dissociation of protein–surfactant or protein–polymer complexes facilitate the evolution from the partially folded conformation to the native conformation. In continuation of this work, Lu et al. [29] have proposed and demonstrated the idea of using a thermally responsive polymer to establish a transient environment that is in tune with the kinetics of folding during the collapse and rearrangement stage. All of these endeavors have confirmed the validity of the HP model in capturing the essential physical nature of protein folding versus aggregation, i.e., the formation of a hydrophobic core driven by intramolecular hydrophobic interactions as opposed to the formation of a molecular assembly driven by intermolecular hydrophobic interactions.

Describing the physical nature of protein aggregation in the context of protein stabilization, however, requires a different approach to those described above. Here, the starting state is the correctly folded native conformation rather than a fully denatured random coil conformation with fully exposed hydrophobic beads, which can thus interact with their neighboring counterparts leading to the formation of aggregates. The aims of the present study were to provide molecular insights into: 1) how the native protein aggregates in solution, and 2) how and to what extent the aggregation may be suppressed by adjusting the composition of the solution or introducing a weakly hydrophobic polymer. Simulations of protein stabilization have been focused on the interaction of the polymer with the native protein, the resulting polymer–protein complex, as well as how this impacts on: 1) protecting the native structure of the protein, and 2) inhibiting the formation of protein aggregate. However, these issues have not yet been adequately addressed.

For the present study, we used two different HP model proteins, namely 13-bead HP protein (HP13) and 20-bead HP protein (HP20), with a view to establishing a more comprehensive understanding of the aggregation of native protein as a function not only of the protein composition and structure, but also of its concentration, the solution composition, and the presence of different kinds of polymers. To mimic the aggregation that occurs

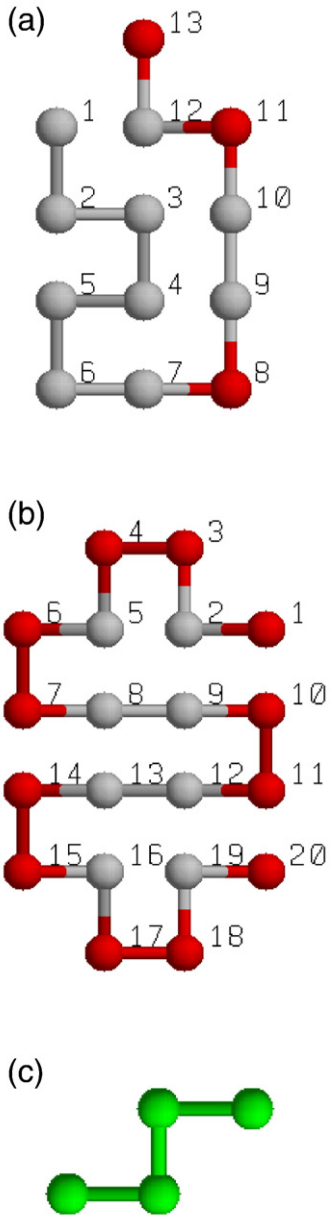

Fig. 1. Protein and polymer models: structures of native HP13 (a), native HP20 (b), and model polymer (c).

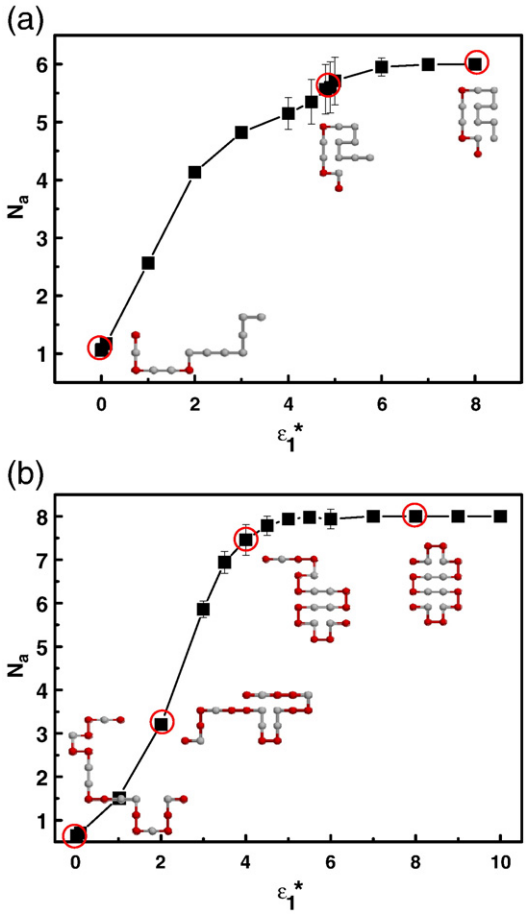

Fig. 2. Unfolding transition curves of the model proteins HP13 (a) and HP20 (b).

in protein liquid products during storage, unfolded protein molecules were also added to the native protein solution as seeds to initiate aggregation. Finally, polymers of different hydrophobicities and chain lengths were introduced into the protein solution at different concentrations and their effects on stabilizing the native conformation of the protein as opposed to the formation of aggregate were examined. The results provide a molecular basis for the behavior of proteins in solution.

2. Models and methods

2.1. Model and potential function

2.1.1. Protein model

Two HP model proteins [40], HP13 and HP20, composed of two types of "bead", polar (P) and hydrophobic (H), have been applied in the present study. The sequence of HP13 is HHHHHHHPHHPHP. As shown in Fig. 1(a), HP13 is aggregation-prone because it has seven accessible hydrophobic beads. In contrast, HP20, PHPPHPPHHPPHHPPHPPHP, does not aggregate so readily because all of the hydrophobic beads are embedded in the interior region of the protein, as shown in Fig. 1(b). Both HP13 and HP20 have been widely used to investigate protein refolding and aggregation behavior, both in solution and in confinement [10,17,18,34,35,38,39].

The simple square-well potential function has been applied to describe the hydrophobic interaction between the protein beads. The Hamiltonian of the model protein chain is:

$$u(r) = \begin{cases} \infty, r = 0 \\ -\varepsilon_1^* \delta_{ij}, 0 < r \le \sigma \\ 0, r > \sigma \end{cases} \quad (1)$$

where $\varepsilon_2^* = \varepsilon_1 / k_B T$ is the dimensionless hydrophobic interaction among the protein beads. A low ε_1^* indicates a weak hydrophobic interaction due to the presence of a high concentration of denaturant. On the contrary, a high ε_1^* indicates a strong intramolecular hydrophobic interaction as a result of a low denaturant concentration. δ_{ij} is the Dirac function; δ_{ij}= 1 when both i and j are hydrophobic beads, otherwise δ_{ij}= 0; σ is the lattice unit.

2.1.2. Polymer model

The polymer [39] is modeled as a connected chain made up of homogeneous hydrophobic beads, as shown in Fig. 1(c).

The Hamiltonian of the polymer is:

$$u(r) = \begin{cases} \infty, r = 0 \\ -\varepsilon_2^*, 0 < r \le \sigma \\ 0, r > \sigma \end{cases} \quad (2)$$

where $\varepsilon_2^* = \varepsilon_2 / k_B T$ is the dimensionless hydrophobic interaction among the polymer beads.

The Hamiltonian between the protein and polymer beads is obtained by using the Lorentz–Berthelot mixing rule, as given in Eq. (3):

$$u(r) = \begin{cases} \infty, r = 0 \\ -\sqrt{\varepsilon_1^* \varepsilon_2^*}\, \delta_{ij}, 0 < r \le \sigma \\ 0, r > \sigma \end{cases} \quad (3)$$

where ε_1^* and ε_2^* are the same as above. δ_{ij} is the hydrophobic interaction of bead i in the protein and bead j in the polymer and *vice versa*; δ_{ij} = 1 when both i and j are hydrophobic beads, otherwise δ_{ij} = 0; σ is the lattice unit.

2.2. Simulation details

NVT ensemble and dynamic Monte Carlo methods have been applied in the simulations. The set of moves for both the proteins and polymers consists of chain-end one-bead rearrangement, chain-end two-bead rearrangement, two-bond move, three-bond move, four-bond move, and translational move. One Monte Carlo simulation step (MC) is finished when all beads in the system have moved once, i.e.:

$$1\mathrm{MC} = N_1 L_1 + N_2 L_2 \quad (4)$$

where L_1 is the chain length of the model protein, N_1 is the number of protein chains, L_2 is the chain length of the model polymer, and N_2 is the number of polymer molecules.

In each simulation, an initial configuration containing N_1 protein chains and N_2 polymer molecules of chain length L_2 was randomly generated in the simulation box. The values of ε_1^*

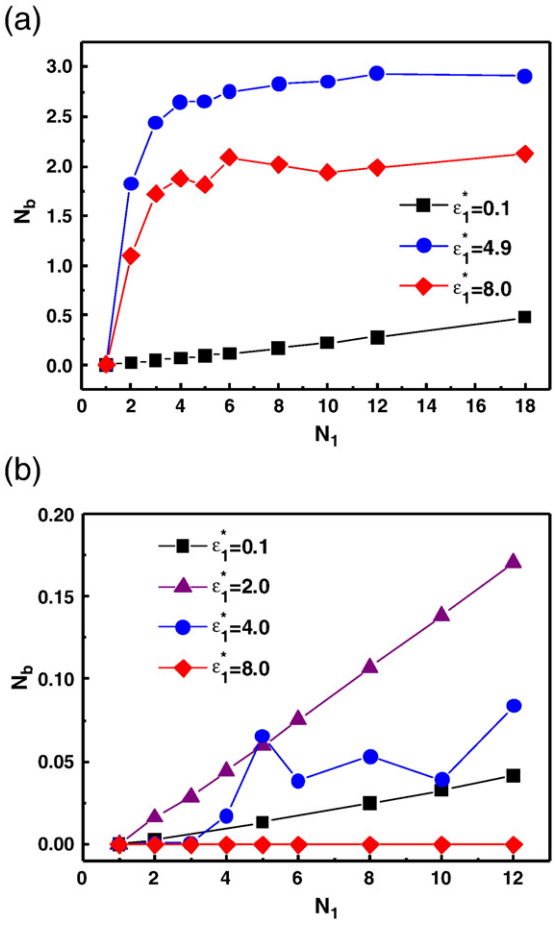

Fig. 3. Aggregation behavior of the model proteins in the absence of polymer: (a) HP13, (b) HP20.

and ε_2^* were adjusted as required and the system was allowed to reach equilibrium. The side length of the simulation box was 28σ for HP13 and 42σ for HP20. All simulations were run for 7.5×10^5 MC steps to reach equilibrium, and thereafter for a further 2.5×10^5 steps to give the average values. Twenty independent simulations were performed for each given set of conditions. All snapshots including the protein and polymer were prepared with the program RASMOL [41].

2.3. Analysis

Hydrophobic contact number (NC), which gives a measure of the total number of hydrophobic interaction pairs of a given protein, is defined by Eq. (5):

$$NC = N_a + N_b$$
$$N_a = \sum_{\substack{i \le j-3 \\ \text{intra-protein}}} \delta_{ij},\ N_b = \sum_{\substack{i,j \\ \text{inter-protein}}} \delta_{ij} \qquad (5)$$

where δ_{ij} is the same as in Eq. (1). N_a is the intramolecular hydrophobic contact number per protein, which is 6 for the native HP13 and 8 for the native HP20, as shown in Fig. 1. N_b is the intermolecular hydrophobic contact number per protein and is thus an index of the aggregation.

Fig. 2 shows the structure transition curves of the model proteins with respect to ε_1^* as well as representative conformations of the proteins at different values of ε_1^*; the transition curve for HP20 is consistent with the reported results of Nguyen et al. [10]. It can be seen from Fig. 2 that $\varepsilon_1^* = 0.1$, 4.9, and 8.0 can be used to represent the solution conditions for the unfolded, partially unfolded, and native conformation for HP13, while the corresponding representative values for HP20 are $\varepsilon_1^* = 0.1$, 2.0, 4.0, and 8.0.

3. Results and discussions

3.1. An overview on the aggregation of native protein in solution

Starting from their respective native conformations, the amounts of aggregates of HP13 (Fig. 3(a)) and HP20 (Fig. 3(b)),

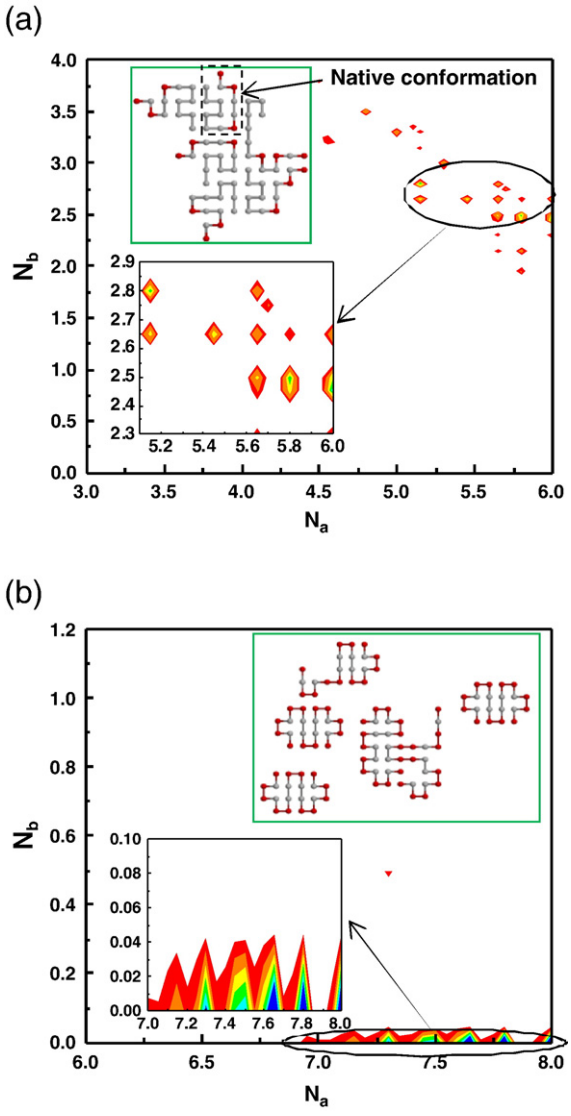

Fig. 4. Protein aggregates obtained in the absence of polymer. (a) HP13 at $\varepsilon_1^* = 4.9$; (b) HP20 at $\varepsilon_1^* = 4.0$. Legend: 0.00 0.01 0.02 0.03 0.04 0.05 0.06

as indicated by N_b, were obtained as a function of the solution condition, ε_1^*, and the protein concentration, N_1. The protein conformation distributions for HP13 and HP20 under the above conditions are plotted in Fig. 4(a) and (b), respectively.

As shown in Fig. 3, HP13 presents a significantly higher N_b than HP20 under all conditions. This is intrinsically due to its native structure, which has more exposed and accessible hydrophobic beads than HP20, as shown in Fig. 1. For HP13, N_b at $\varepsilon_1^* = 4.9$ is higher than that at $\varepsilon_1^* = 0.1$ or $\varepsilon_1^* = 8.0$, at which this protein exists in completely extended and compact conformations, respectively (Fig. 2(a)). A similar response of N_b to ε_1^* is also seen for HP20, for which a significantly higher N_b is obtained at $\varepsilon_1^* = 2.0$, corresponding to a partially unfolded condition (Fig. 2(b)), as compared to the values obtained at $\varepsilon_1^* = 0.1$ and $\varepsilon_1^* = 8.0$. Thus, for both HP13 and HP20, aggregation is more extensive under the solution conditions that favor the partially unfolded protein as compared to that under the solution conditions favoring either the fully denatured or the native protein. A similar conclusion has also been drawn by Nguyen et al. [10] following their study on the refolding of HP20 starting from the fully denatured conformation. The increase in N_b in response to N_1 is due to the increased frequency of protein−protein interactions as a result of the higher protein concentration, which is in line with experimental observations[16] on both protein stabilization and protein refolding.

To highlight the protein aggregation, Fig. 4 shows the conformational distributions obtained at $\varepsilon_1^* = 4.9$ for HP13 and $\varepsilon_1^* = 4.0$ for HP20. The number of protein molecules in the simulation box (N_1) is 6.

Fig. 4(a) shows that most of the HP13 protein molecules are located in the region of N_a=5.1−6.0 and N_b=2.3−2.9, indicating that both native (N_a=6, denoted by "Native conformation" in Fig. 4(a)) and partially unfolded (N_a< 6) states are present in the aggregates. The snapshot shows that the partially unfolded protein, which has a more extended structure and more accessible hydrophobic regions, forms the "core" of the aggregate, in which the native conformation is also found. For HP20, as shown in Fig. 4(b), most of protein molecules are located in the region N_a= 7.0−8.0 and N_b= 0−0.05, i.e., the HP20 protein is much less prone to aggregation than HP13, as a result of having fewer exposed hydrophobic regions.

In practice, Zhang et al. [11] have monitored the temperature-induced denaturation of adenylate kinase by fluorescence spectroscopy and found accelerated aggregation in response to exposure of the hydrophobic surfaces of the partially denatured

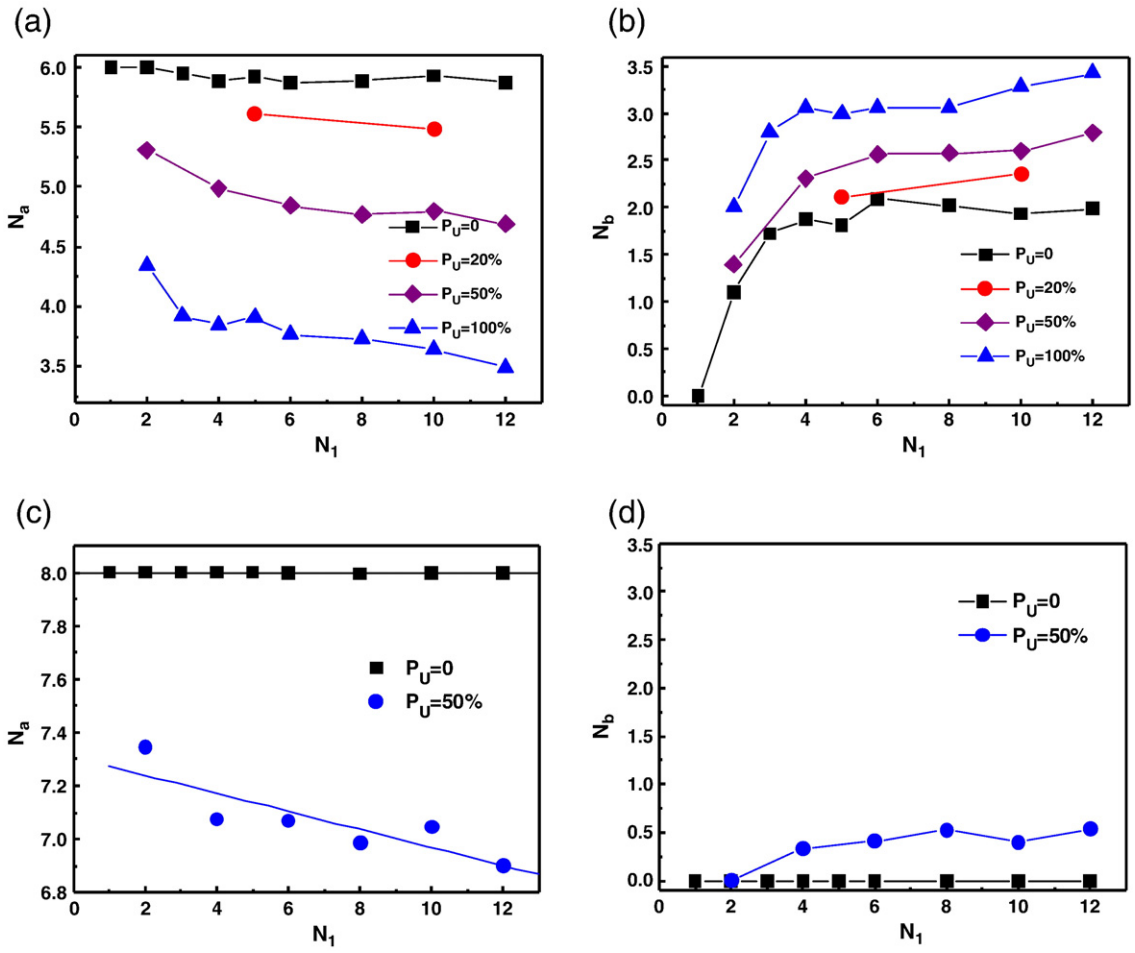

Fig. 5. Effect of adding unfolded protein on the aggregation. (a) and (b) for HP13; (c) and (d) for HP20, all at $\varepsilon_1^* = 8.0$.

enzyme. Raman et al. [13] have shown that lysozyme of intermediate status has more hydrophobic regions than either its native or fully unfolded counterparts and is thus more prone to aggregation. The simulations in Figs. 3 and 4 not only reproduce such experimental results, but also suggest that strengthening the "hydrophobic core" may offer a means of suppressing the protein aggregation that occurs for extended proteins. In practice, this can be realized by adjusting the solution conditions to give a suitably high value of ε_1^*.

3.2. How partially unfolded protein "induces" protein aggregation

To mimic the aggregation that occurs in separation process, e.g. membrane separation or chromatography of different kinds, we introduced a certain number of unfolded protein molecules into the simulation box at the beginning of the simulation and monitored the aggregation process. The results are shown in Fig. 5, in which N_a and N_b are plotted against both the overall protein concentration and the content of unfolded protein, P_U. Here, ε_1^* was set at 8.0, i.e., the solution conditions favoring the native conformation for both HP13 and HP20.

It can be seen in Fig. 5(a) and (c) that, for both HP13 and HP20, N_a decreases with increasing N_1 at all values of P_U. In Fig. 5(b) and (d), it can be seen that N_b increases with both N_1 and P_U. For HP20, the maximum N_b is below 0.5 when 50% of the total protein is in the unfolded conformation (Fig. 5(d)). Under the same conditions, N_b for HP13 reaches 2.7 (Fig. 5(b)). Here again, HP20 is seen to be much less prone to aggregation than HP13. In other words, HP20 is much more stable in solution.

Fig. 6 shows the time course of the aggregation and the structures of the protein aggregates at different values of P_U. It can be seen in Fig. 6(a) and (b) that the apparent aggregation rate increases with P_U in the initial stage, suggesting that partially unfolded proteins "induce" protein aggregation.

Fig. 6(c) and (d) shows the conformational distributions of the model proteins and typical snapshots at P_U = 83.3% for HP13 and P_U = 50% for HP20, respectively. Compared with

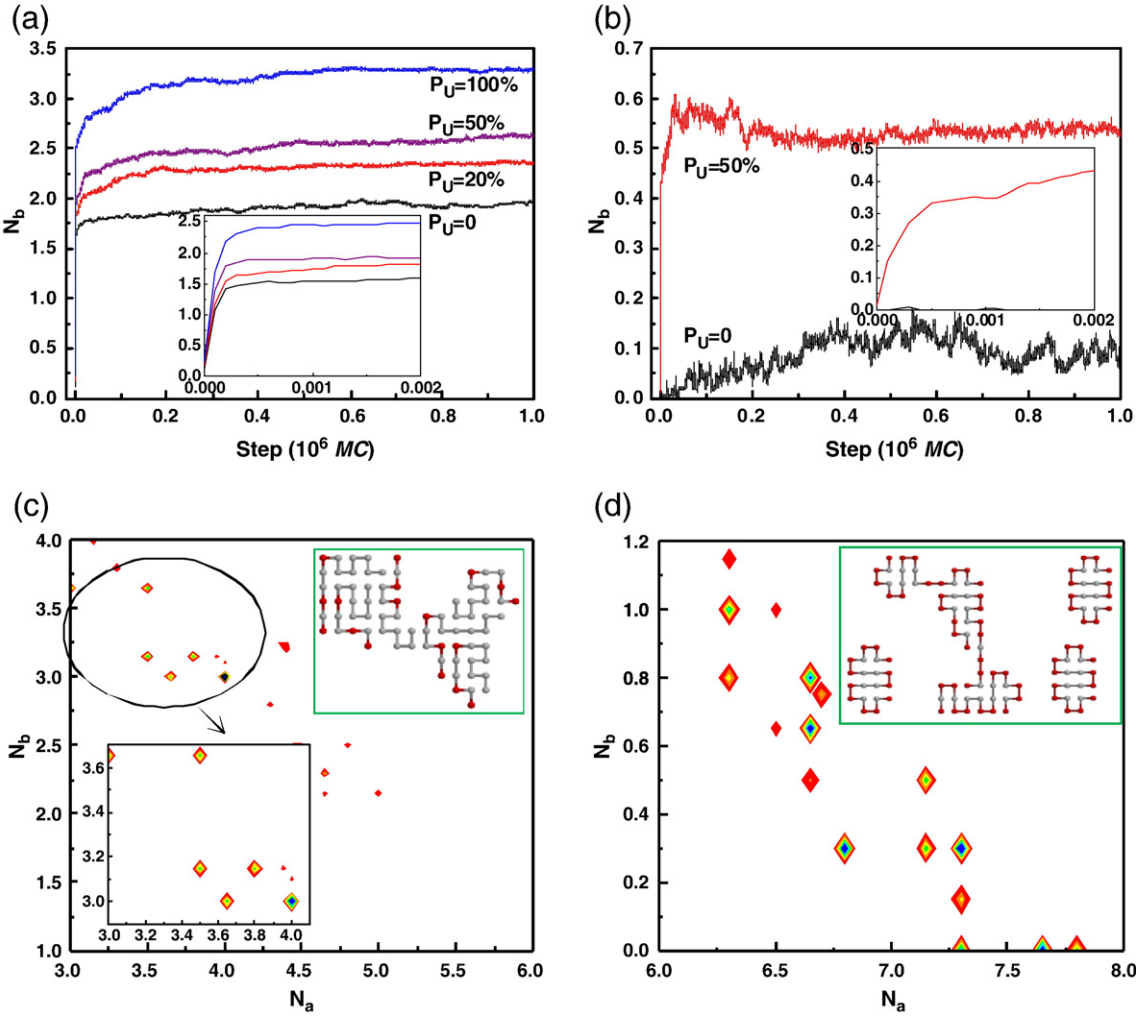

Fig. 6. Effect of adding unfolded protein on the aggregation process and products. (a) HP13 at N_1=10, (b) HP20 at N_1=12, (c) protein products of HP13 at P_U=83.3%, N_1=6, and (d) protein products of HP20 at P_U=50%, N_1=6.

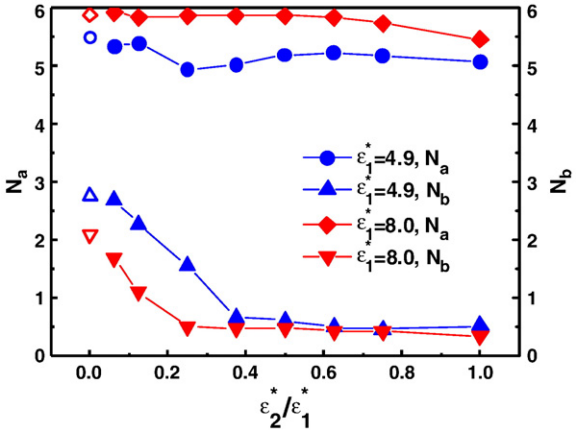

Fig. 7. Protein aggregation in the presence of polymers of different hydrophobicities, where the hollow symbols denote the corresponding values of N_a and N_b in the absence of polymers.

Fig. 4, more conformations are distributed in the area of low N_a and high N_b, indicating enhanced aggregation. The snapshots show that the conformations with more exposed hydrophobic sites represent "nuclei" for aggregation, even under conditions that favor the native conformation. Again, the partially unfolded protein "induces" aggregation.

3.3. Inhibition of protein aggregation by weakly hydrophobic polymers

The addition of polymers such as PEG [24–27] and PNIPAAm [1,28] is a well-established means of inhibiting protein aggregation in solution during protein refolding processes that start from the unfolded conformation. PNIPAAm and dextran-grafted-PNIPAAm have also been used to promote protein refolding [1,29,30], whereby the hydrophobic interaction between the polymer and protein plays the key role in inhibiting protein aggregation during protein refolding. In the present work, we added polymer to aqueous native protein solution and then investigated how it interacts with the native protein and suppresses protein aggregation. For this study, HP13 was used as the model protein because its native structure is more prone to aggregation.

3.3.1. Protein aggregation in the presence of polymers with different backbone hydrophobicities

The backbone hydrophobicity of a polymer intrinsically determines the strength of its interaction with protein molecules. Fig. 7 shows the N_a and N_b of HP13 as a function of the relative hydrophobicities of the polymer and protein ($\varepsilon_2^*/\varepsilon_1^*$). The concentrations of protein (N_1) and polymer (N_2) were 6 and 30, respectively.

At $\varepsilon_1^* = 8.0$, the conditions favoring the native conformation of HP13, a slight decrease in N_a is observed when $\varepsilon_2^*/\varepsilon_1^*$ exceeds 0.75, suggesting that a strong hydrophobic interaction between the polymer and protein may also cause denaturation, i.e., loss of the hydrophobic core. On the other hand, at $\varepsilon_1^* = 4.9$ and 8.0, N_b decreases monotonously in response to polymer hydrophobicity when $\varepsilon_2^*/\varepsilon_1^* \leq 0.375$ and then reaches a minimum value, indicating an effective inhibition of protein aggregation. These findings suggest that the addition of a polymer of appropriate hydrophobicity can effectively inhibit protein aggregation in solution without sacrifice of the protein native conformation.

3.3.2. Protein aggregation in the presence of polymers of different chain lengths

Polymer length is another important parameter determining the intensity of polymer–protein interactions. Fig. 8 shows N_a and N_b as a function of the chain length of the polymer (L_2) when 6 protein and 30 polymer molecules were introduced into the simulation box.

It can be seen that the polymer chain length (L_2) has less effect on N_a and N_b at both $\varepsilon_1^* = 4.9$ and $\varepsilon_1^* = 8.0$. At $\varepsilon_1^* = 4.9$, a decrease of N_b to 0.5 is observed when L_2 is increased from 2 to 4. This suggests that a long-chain polymer has a better stabilizing effect as compared to a shorter one. A similar conclusion has been drawn from experiments on PEG

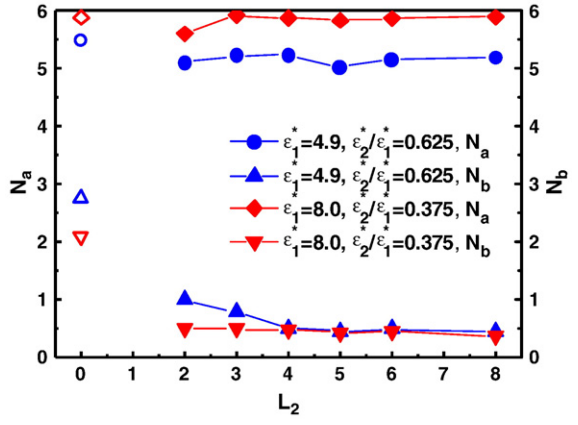

Fig. 8. Protein aggregation in the presence of polymers of different chain lengths, where the hollow symbols denote the corresponding values of N_a and N_b in the absence of polymers.

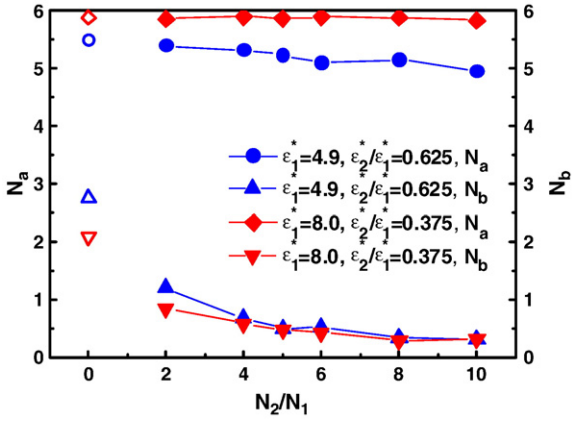

Fig. 9. Protein aggregation in the presence of different concentrations of polymers, where the hollow symbols denote the corresponding values of N_a and N_b in the absence of polymers.

stabilizing proteins, in which PEG8000 proved to be more effective than PEG400 in suppressing the aggregation of lactate dehydrogenase [27]. At $\varepsilon_1^* = 8.0$, the solution conditions favoring the native conformation of HP13, N_b is maintained at 0.5 with increasing L_2. Considering the N_b obtained at $\varepsilon_1^* = 4.9$, it is concluded that the stabilizing effect of the polymer is more significant under the solution conditions that favor partially folded conformations (Fig. 2(a)).

3.3.3. Protein aggregation in the presence of different concentrations of polymer

Fig. 9 shows N_a and N_b as a function of polymer concentration (N_2). Here, $N_1 = 6$ and $L_2 = 4$. The ratio $\varepsilon_2^*/\varepsilon_1^*$ was chosen for a given value of ε_1^* according to Fig. 7.

It can be seen that at $\varepsilon_1^* = 4.9$, the solution conditions favoring the partially unfolded state of HP13 (Fig. 2(a)), N_a decreases with respect to N_2/N_1, i.e., the presence of the polymer destroys the native structure of HP13. At $\varepsilon_1^* = 8.0$, which favors the native state, N_a is much less sensitive to the polymer concentration. On the other hand, N_b decreases with respect to N_2/N_1, i.e., a higher concentration of polymer inhibits protein aggregation more effectively. Mi et al. [27] have reported a higher residual activity of lactate dehydrogenase obtained in the presence of a high concentration of PEG. A similar stabilizing effect has also been reported by Chen et al. [28], whereby an increase in the concentration of PEG4000 led to an increased refolding yield.

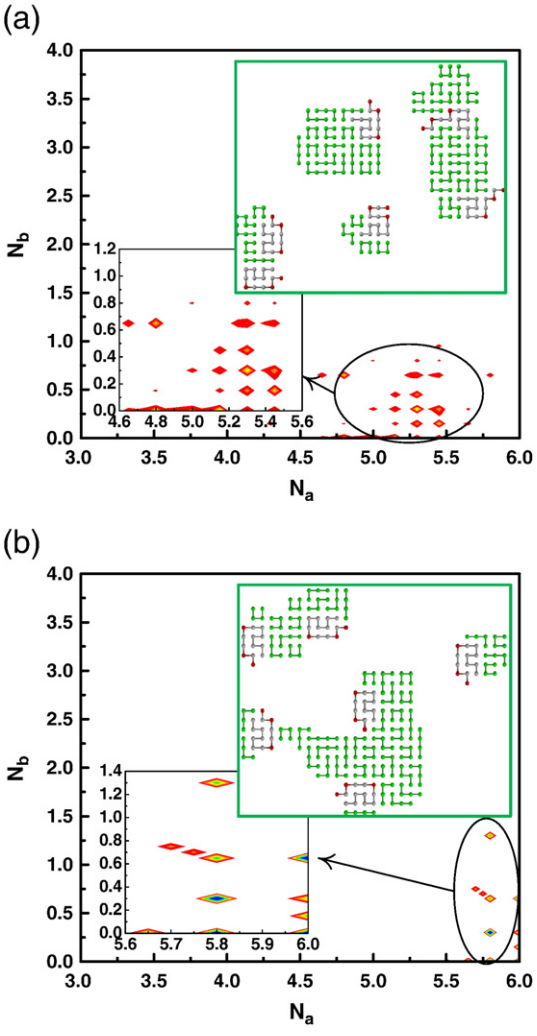

Fig. 10. Distribution of HP13 in the presence of polymer at $\varepsilon_1^* = 4.9$, $\varepsilon_2^*/\varepsilon_1^* = 0.625$ (a), and $\varepsilon_1^* = 8.0$, $\varepsilon_2^*/\varepsilon_1^* = 0.375$ (b), where N_2=48 and L_2=4.

Fig. 10 shows both the conformational distribution of the model proteins in the presence of polymers and some typical snapshots. The total number of protein molecules in the simulation box (N_1) is 6.

It can be seen in Fig. 10(a) that at $\varepsilon_1^* = 4.9$, the solution conditions favoring the partially unfolded protein, most of protein molecules are located in the region of N_a=4.6−5.6 and N_b=0.0−1.2 in the presence of the polymer. Compared to Fig. 4 (a), the addition of polymer of appropriate backbone hydrophobicity significantly inhibits protein aggregation. On the other hand, the interaction with the polymer may also induce a small degree of unfolding of the protein, as indicated by the shift of the conformation distribution to lower N_a.

At $\varepsilon_1^* = 8.0$, the solution conditions favoring the native conformation of the protein, as shown in Fig. 10(b), most of protein molecules are located in the region of N_a=5.6−6.0 and N_b=0.0−1.4 in the presence of polymer. The protein chains are well-dispersed and thus aggregation is inhibited. This again suggests that an optimal stabilization effect is obtained once the hydrophobicity of the polymer backbone matches that of the target protein.

The snapshots in Fig. 10 also show that the partially unfolded protein constitutes the “nucleus” of the aggregate. Polymers of appropriate hydrophobicity wrap around both the native and partially unfolded protein through hydrophobic interactions, forming a protein–polymer complex. This shields the hydrophobic interactions between protein molecules in different folded states and inhibits the formation of protein aggregate. In practice, PEG, PNIPAAm, and dextran have long been applied to assist protein refolding, whereby they form complexes with folded intermediates and thus suppress the formation of protein aggregate [1,24]. Here, the results shown in Fig. 10 illustrate how these polymers stabilize the native protein in solution. The ability to form complexes with both the native and partially unfolded protein, as shown in Fig. 10, is key to the polymer’s capacity to segregate single protein molecules and thus inhibit the formation of the protein aggregate.

4. Conclusion

A dynamic Monte Carlo simulation on 2D lattice HP protein models of different compositions and conformations has been performed in order to provide a molecular insight into the aggregation of native protein in solution. The partially unfolded protein has more exposed hydrophobic regions as compared to its native counterpart, and thus functions as the “nucleus” of the aggregate. Strengthening intramolecular hydrophobic interactions is one way of enhancing the stability of the native protein and suppressing the formation of protein aggregates.

On the other hand, a polymer of appropriate hydrophobicity and suitable chain length can wrap around the protein to form a protein–polymer complex. This shields the hydrophobic regions of the protein and thereby inhibits protein aggregation. Being consistent with experimental results, our simulation provides a molecular view of polymer–protein interactions and the way in which a polymer can both stabilize the native conformation of the protein and suppress protein aggregation. This is essential for the design and application of polymers as stabilizers for proteins in solution.

Acknowledgments

This work was supported by the National Natural Science Foundation under Grant No. 20576061, Grant No. 20636040 and by the Ministry of Science and Technology through 973 Project under Grant No. 2003CB716004.

References

[1] D.N. Lu, Z.X. Liu, M.L. Zhang, Z. Liu, H.M. Zhou, The mechanism of PNIPAAm-assisted refolding of lysozyme denatured by urea, Biochem. Eng. J. 24 (2005) 55–64.
[2] J. Zhan, Z. Liu, B.G. Wang, F.X. Ding, Modification of a membrane surface charge by a low temperature plasma induced grafting reaction and its application to reduce membrane fouling, Sep. Sci. Technol. 39 (2004) 2977–2995.
[3] R. Chan, V. Chen, The effects of electrolyte concentration and pH on protein aggregation and deposition: critical flux and constant flux membrane filtration, J. Membr. Sci. 185 (2001) 177–192.
[4] C. Machold, R. Schlegl, W. Buchinger, A. Jungbauer, Continuous matrix assisted refolding of alpha-lactalbumin by ion exchange chromatography with recycling of aggregates combined with ultradiafiltration, J. Chromatogr., A 1080 (2005) 29–42.
[5] M. Yan, Z.X. Liu, D.N. Lu, Z. Liu, Fabrication of single carbonic anhydrase nanogel against denaturation and aggregation at high temperature, Biomacromolecules 8 (2007) 560–565.
[6] O. Gloger, K. Witthohn, B.W. Muller, Lyoprotection of aviscumine with low molecular weight dextrans, Int. J. Pharm. 260 (2003) 59–68.
[7] E.Y. Chi, S. Krishnan, T.W. Randolph, J.F. Carpenter, Physical stability of proteins in aqueous solution: mechanism and driving forces in nonnative protein aggregation, Pharm. Res. 20 (2003) 1325–1336.
[8] Z.J. Qin, D.M. Hu, M. Zhu, A.L. Fink, Structural characterization of the partially folded intermediates of an immunoglobulin light chain leading to amyloid fibrillation and amorphous aggregation, Biochemistry-Us 46 (2007) 3521–3531.
[9] A. Tanksale, M. Ghatge, V. Deshpande, Alpha-crystallin binds to the aggregation-prone molten-globule state of alkaline protease: implications for preventing irreversible thermal denaturation, Protein Sci. 11 (2002) 1720–1728.
[10] H.D. Nguyen, C.K. Hall, Effect of rate of chemical or thermal renaturation on refolding and aggregation of a simple lattice protein, Biotechnol. Bioeng. 80 (2002) 823–834.
[11] Y.L. Zhang, X.M. Pan, J.M. Zhou, Surface hydrophobicity and thermal aggregation of adenylate kinase, Biochem. Mol. Biol. Int. 44 (1998) 949–960.
[12] M.A. Speed, D.I. Wang, J. King, Specific aggregation of partially folded polypeptide chains: the molecular basis of inclusion body composition, Nat. Biotechnol. 14 (1996) 1283–1287.
[13] B. Raman, T. Ramakrishna, C.M. Rao, Refolding of denatured and denatured/reduced lysozyme at high concentrations, J. Biol. Chem. 271 (1996) 17067–17072.
[14] G. Georgiou, P. Valax, M. Ostermeier, P.M. Horowitz, Folding and aggregation of tem beta-lactamase–analogies with the formation of inclusion-bodies in *Escherichia coli*, Protein Sci. 3 (1994) 1953–1960.
[15] R.A. Broglia, G. Tiana, S. Pasquali, H.E. Roman, E. Vigezzi, Folding and aggregation of designed proteins, Proc. Natl. Acad. Sci. U. S. A 95 (1998) 12930–12933.
[16] E.D. Clark, D. Hevehan, S. Szela, J. Maachupalli-Reddy, Oxidative renaturation of hen egg-white lysozyme. Folding vs aggregation, Biotechnol. Prog. 14 (1998) 47–54.
[17] P. Gupta, C.K. Hall, A. Voegler, Computer simulation of the competition between protein folding and aggregation, Fluid Phase Equilib. 160 (1999) 87–93.
[18] P. Gupta, C.K. Hall, A.C. Voegler, Effect of denaturant and protein concentrations upon protein refolding and aggregation: a simple lattice model, Protein Sci. 7 (1998) 2642–2652.
[19] D. Bratko, H.W. Blanch, Competition between protein folding and aggregation: a three-dimensional lattice-model simulation, J. Chem. Phys. 114 (2001) 561–569.
[20] G. Costello, S.R. Euston, A Monte Carlo simulation of the aggregation, phase-separation, and gelation of model globular molecules, J. Phys. Chem., B 110 (2006) 10151–10164.
[21] L.A. Clark, Protein aggregation determinants from a simplified model: cooperative folders resist aggregation, Protein Sci. 14 (2005) 653–662.
[22] D. Bratko, T. Cellmer, J.M. Prausnitz, H.W. Blanch, Molecular simulation of protein aggregation, Biotechnol. Bioeng. 96 (2007) 1–8.
[23] L. Toma, S. Toma, A lattice study of multimolecular ensembles of protein models. Effect of sequence on the final state: globules, aggregates, dimers, fibrillae, Biomacromolecules 1 (2000) 232–238.
[24] J.L. Cleland, C. Hedgepeth, D.I. Wang, Polyethylene glycol enhanced refolding of bovine carbonic anhydrase B. Reaction stoichiometry and refolding model, J. Biol. Chem. 267 (1992) 13327–13334.
[25] J.L. Cleland, S.E. Builder, J.R. Swartz, M. Winkler, J.Y. Chang, D.I. Wang, Polyethylene glycol enhanced protein refolding, Nat. Biotechnol. 10 (1992) 1013–1019.
[26] J.L. Cleland, D.I. Wang, Cosolvent assisted protein refolding, Nat. Biotechnol. 8 (1990) 1274–1278.
[27] Y.L. Mi, G. Wood, L. Thoma, S. Rashed, Effects of polyethylene glycol molecular weight and concentration on lactate dehydrogenase activity in solution and after freeze–thawing, PDA J. Pharm. Sci. Tech. 56 (2002) 115–123.
[28] Y.J. Chen, L.W. Huang, H.C. Chiu, S.C. Lin, Temperature-responsive polymer-assisted protein refolding, Enzyme Microb. Technol. 32 (2003) 120–130.
[29] D.N. Lu, Z.X. Liu, M.L. Zhang, X.G. Wang, Z. Liu, Dextran-grafted-PNIPAAm as an artificial chaperone for protein refolding, Biochem. Eng. J. 27 (2006) 336–343.
[30] D.N. Lu, K. Zhang, Z. Liu, Protein refolding assisted by an artificial chaperone using temperature stimuli responsive polymer as the stripper, Biochem. Eng. J. 25 (2005) 141–149.
[31] K.F. Lau, K.A. Dill, A lattice statistical mechanics model of the conformational and sequence spaces of proteins, Macromolecules 22 (1989) 3986–3997.
[32] S. Istrail, R. Schwartz, J. King, Lattice simulations of aggregation funnels for protein folding, J. Comput. Biol. 6 (1999) 143–162.
[33] K. Leonhard, J.M. Prausnitz, C.J. Radke, Solvent-amino acid interaction energies in three-dimensional-lattice Monte Carlo simulations of a model 27-mer protein: folding thermodynamics and kinetics, Protein Sci. 13 (2004) 358–369.
[34] P. Gupta, C.K. Hall, Effect of solvent conditions upon refolding pathways and intermediates for a simple lattice protein, Biopolymers 42 (1997) 399–409.
[35] P. Gupta, C.K. Hall, Computer-simulation of protein refolding pathways and intermediates, AIChE J. 41 (1995) 985–990.
[36] T. Cellmer, D. Bratko, J.M. Prausnitz, H. Blanch, Protein-folding landscapes in multichain systems, Proc. Natl. Acad. Sci. U. S. A 102 (2005) 11692–11697.

[37] T. Cellmer, D. Bratko, J.M. Prausnitz, H.W. Blanch, Protein aggregation in silico, Trends Biotechnol. 25 (2007) 254–261.
[38] D.N. Lu, Z. Liu, Z.X. Liu, M.L. Zhang, P. Ouyang, Molecular simulation of surfactant-assisted protein refolding, J. Chem. Phys. 122 (2005) 134902–134911.
[39] D.N. Lu, Z. Liu, Molecular simulation of polymer assisted protein refolding, J. Chem. Phys. 123 (2005) 134903–134908.
[40] H.S. Chan, K.A. Dill, Transition-states and folding dynamics of proteins and heteropolymers, J. Chem. Phys. 100 (1994) 9238–9257.
[41] R. Sayle, E. Milnerwhite, RASMOL−−biomolecular graphics for all, Trends Biochem. Sci. 20 (1995) 374–376.

Glossary

Symbol

L_1, L_2: Chain length of protein and polymer respectively
N_1, N_2: Number of protein and polymer respectively
NC, N_a, N_b: Total, intra-protein and inter-protein hydrophobic contact number respectively
ε_1^*,ε_2^*: Dimensionless hydrophobicity of protein beads and polymer beads respectively
σ: Lattice length unit