

# Effects of macromolecular crowding on protein folding

Bui Phuong Thuy<sup>1,2</sup>, Hoang Thi Thu Huong<sup>1</sup> and Trinh X Hoang<sup>1</sup>

<sup>1</sup> Institute of Physics, Vietnam Academy of Science and Technology, 10 Dao Tan, Ba Dinh, Ha Noi, Vietnam

<sup>2</sup> Nam Dinh University of Technology Education, Phu Nghia, Loc Ha, Nam Dinh, Vietnam

E-mail: buiphuongthuyphys@gmail.com

**Abstract.** Folding of proteins *in vivo* is influenced by the presence of a high concentration of macromolecules in the cytosol. In this study we investigate the effects of macromolecular crowding on folding of proteins by using molecular dynamics simulation method with Langevin equations. The protein is considered in a Go-like model whereas crowders are modeled as soft spheres. In our simulations, protein and crowders are enclosed in a spherical container. It is shown that the crowders enhance the folding transition temperature as well as folding stability of protein. The effect of crowders on the folding free energy is found to agree with Minton's scaled particle theory.

## 1. Introduction

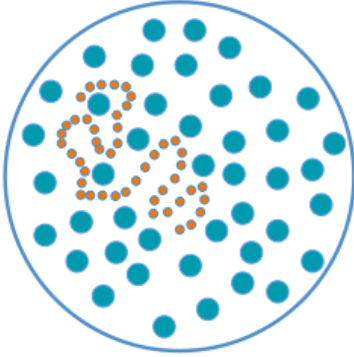
The cytosol of living cells contains a large number of macromolecules such as proteins, nucleic acids of a volume fraction up to about 40%. Such condition, known as 'macromolecular crowding' [1], significantly changes the properties of molecules in the solution and plays an important role in many cellular processes [2, 3, 4]. For example, macromolecular crowding is known to strengthen the binding affinity between DNA and DNA polymerase in DNA transcription and replication [5], to enhance protein-protein association [6], and to improve the functional efficiency of chaperonins [3]. Macromolecular crowding also has been shown to increase stability and folding rate of proteins [7] and at the same time promote their aggregation [8].

The primary effect of macromolecular crowding is due to the excluded volume occupied by the molecules of the crowd, or crowders, that becomes unavailable to other molecules. In protein folding, fluctuation of the void spaces left by the crowders disfavors the unfolded state more than the folded state. This is the basis of a number of theories based on scaled particle theory (SPT) [9, 10, 11] that allow one to estimate the change in the folding free energy due to a presence of the crowders. Molecular dynamics simulations have confirmed this picture and provided detailed assessments of these theories [12, 13]. The effects of crowding on protein folding have been shown to be essentially the same as those of confinement [11, 14].

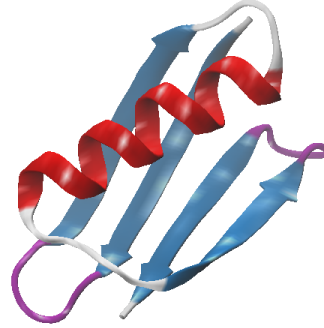
In this study, we investigate the effects of macromolecular crowding on folding of protein by molecular dynamics simulation. We consider simple models of protein and crowders that are essentially similar to those considered in Ref. [13] except that we use a close boundary condition instead of a periodic boundary condition.



## 2. Model and methods



**Figure 1.** Sketch of a macromolecular crowding model with a protein represented by a chain of amino acids (small dots) and other macromolecules (large dots) enclosed in a spherical container.



**Figure 2.** The native state of protein G (B1 domain) in a ribbon presentation. The structure has a  $\beta$ -sheet with 4 strands (blue) and an  $\alpha$ -helix (red).

### 2.1. Models of protein and macromolecular crowders

To mimic the cellular environment, we simulate folding of protein in presence of macromolecules modeled as soft spheres (Fig. 1). The protein is considered in a coarse-grained  $C_\alpha$ -based Go-like model [15, 16]. Each amino acid is considered as a single bead of radius  $R_a = 2.5 \text{ \AA}$  centered at the position of its  $C_\alpha$  atom. The potential energy of the protein in a given conformation is given by:

$$E_p(\{\mathbf{r}_i\}) = \sum_{i=1}^{N-1} K_b(r_{i,i+1} - b)^2 + \sum_{i=2}^{N-1} K_\theta(\theta_i - \theta_i^*)^2 + \sum_{n=1,3}^{N-2} \sum_{i=2}^{N-2} K_\phi^{(n)}[1 + \cos(n(\phi_i - \phi_i^*))] + \sum_{i+3 < j} 4\epsilon \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \Delta_{ij} + \sum_{i+3 < j} \epsilon \left( \frac{\sigma}{r_{ij}} \right)^{12} (1 - \Delta_{ij}), \quad (1)$$

where  $N$  is the number of amino acids;  $\mathbf{r}_i$  is the position of the bead  $i$  ( $i = 1, \dots, N$ );  $r_{ij}$  is the distance between beads  $i$  and  $j$ ;  $\theta$  and  $\phi$  are the bond angle and dihedral angle;  $n$  is equal to either 1 or 3; the superscript  $*$  corresponds to the native state;  $\Delta_{ij}$  is equal to 1 if there is a native contact between  $i$  and  $j$  and equal to 0 otherwise. A native contact is defined if the distance between two residues in the native state is less than  $7.5 \text{ \AA}$ . The first three terms of Eq. (1) correspond to the bonding potentials, bond angle potentials and dihedral angle potentials, respectively. The last two terms are Lennard-Jones (LJ) potentials for native contacts and repulsive potentials for non-native contacts. Energy is given in units of  $\epsilon$ , which corresponds to the depth of the LJ potential. The choice of LJ potential is such that the distance between residues in a native contact in the native state corresponds to its minimum, i.e.  $\sigma_{ij} = 2^{-1/6} r_{ij}^*$ . The parameters chosen for our model are  $b = 3.8 \text{ \AA}$ ,  $\sigma = 2R_a = 5 \text{ \AA}$ ,  $K_b = 100 \epsilon \text{ \AA}^{-2}$ ,  $K_\theta = 20 \epsilon (\text{rad})^{-2}$ ,  $K_\phi^{(1)} = -\epsilon$ ,  $K_\phi^{(3)} = -0.5\epsilon$ .

Macromolecular crowders are modeled as soft spheres of radius  $R_c$ . They repulse each other and the amino acids via a shifted and cut-off LJ potential:

$$U(r') = \begin{cases} 4\epsilon \left[ \left( \frac{\sigma}{r'} \right)^{12} - \left( \frac{\sigma}{r'} \right)^6 \right] + \epsilon & r' \leq r_c \\ 0 & r' > r_c. \end{cases} \quad (2)$$

with

$$r' = r - R - R_c + r_c, \quad (3)$$

where  $r$  is the center to center distance between the two molecules,  $R$  is either the radius of an amino acid ( $R_a$ ) or the radius of a macromolecule ( $R_c$ ), and  $r_c = 2^{1/6}\sigma$ .

We employ a boundary condition in which protein and crowders are enclosed in a spherical container of radius  $R_{\text{wall}}$  as shown in Fig. 1. The repulsive potential between amino acids and macromolecular crowders with the wall of the container has the same form as given in Eq. (2) but with:

$$r' = R_{\text{wall}} - r - R + r_c, \quad (4)$$

where  $r$  is the distance from the center of the amino acid or the crowder to the center of the spherical container. The volume fraction of  $M$  identical crowders in a volume  $V$  is given by:

$$\phi_c = \frac{(4\pi/3)R_c^3 M}{V}. \quad (5)$$

## 2.2. Molecular Dynamics method

The molecular dynamics (MD) simulation method employed in the present study is based on Langevin equation of motion [16]. For a given temperature  $T$ , Langevin equations for a protein of  $N$  amino acids and  $M$  crowders are given by:

$$m \ddot{\mathbf{r}}_i(t) = \mathbf{f}_i(t) - \zeta_a \dot{\mathbf{r}}_i(t) + \sqrt{2\zeta_a k_B T} \mathbf{g}(t) \quad (i = 1, \dots, N) \quad (6)$$

$$m_c \ddot{\mathbf{r}}_j(t) = \mathbf{f}_j(t) - \zeta_c \dot{\mathbf{r}}_j(t) + \sqrt{2\zeta_c k_B T} \mathbf{g}(t) \quad (j = 1, \dots, M) \quad (7)$$

where  $m$  and  $m_c$  are the masses of an amino acid and a crowder, respectively;  $\mathbf{f}_i$  and  $\mathbf{f}_j$  are molecular forces;  $\zeta_a$  and  $\zeta_c$  are friction coefficients for amino acid and crowder, respectively;  $\mathbf{g}$  denotes an “unit” random force whose components are given by a Gaussian white noise of zero mean and unit dispersion.

A Verlet algorithm is developed to numerically integrate the equations. Temperature is given in units of  $\epsilon/k_B$ , whereas time is measured in units of  $\tau = \sqrt{m\sigma^2/\epsilon}$ . The integration time step is chosen to be equal to  $\Delta t = 0.002\tau$ . We assume that the protein and the crowder have the same mass per volume, so that  $m_c = m(R_c/R_a)^3$ . The friction coefficients for amino acids and crowders are chosen to be  $\zeta_a = 5m/\tau$  and  $\zeta_c = \zeta_a(R_c/R_a)$ , so that they follow the Stokes law for viscous friction on spherical particles.

Equilibrium characteristics of the system can be calculated from data of MD simulations and the weighted histogram method [17]. The specific heat of protein is given by:

$$C_v(T) = \frac{\langle E^2 \rangle - \langle E \rangle^2}{k_B T^2}, \quad (8)$$

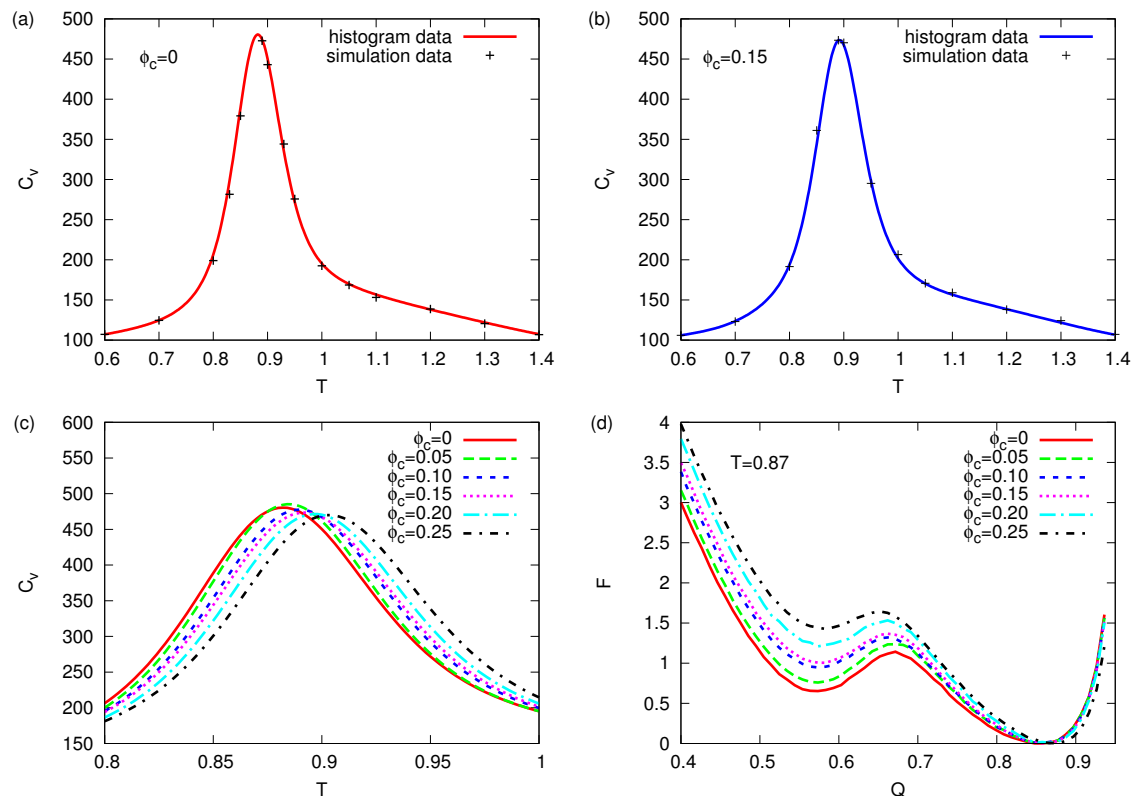
where  $E$  the total energy of the protein. The protein folding transition temperature or folding temperature  $T_f$  is defined as the temperature of the maximum of the specific heat. For a given conformation one defines the fraction of native contact  $Q$  as the number of native contacts in a given conformation divided by the total number of contacts in the native state. At a given temperature  $T$  an effective free energy as function of  $Q$  can be calculated as:

$$F(Q) = -k_B T \ln P(Q), \quad (9)$$

where  $P(Q)$  is the probability to find the protein having native contact fraction between  $Q$  and  $Q + dQ$ .

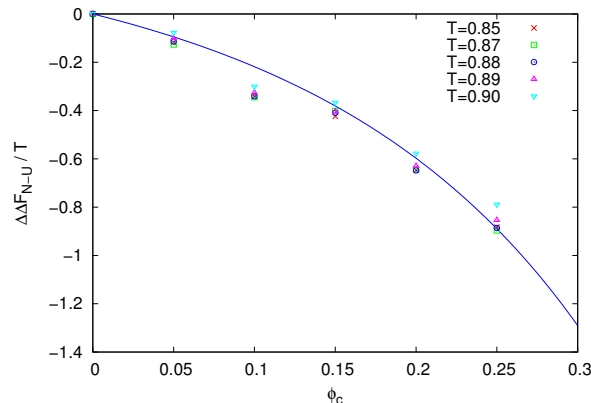
### 3. Results and discussion

We study the effects of crowding on folding of the B1 domain of protein G (GB1), a small protein of  $N = 56$  amino acids with native state shown in Fig. 2. The radius of the crowder is fixed to be  $R_c = 10.2\text{\AA}$ , that is equal to the radius of gyration of GB1, and the radius of the container to be  $R_{\text{wall}} = 106.4\text{\AA}$ . The size of the container was chosen so that it can accommodate a fully stretched protein. Extensive simulations are carried out to obtain equilibrium characteristics of the folding transition for different volume fractions of crowders.



**Figure 3.** (a-b) Temperature dependence of the specific heat,  $C_v$ , of protein G in absence (a) and presence (b) of crowders plotted using non-weighted data from MD trajectories (crosses) and those obtained by the weighted histogram method (solid lines). (c) Comparison of the specific heats for different crowders' volume fraction  $\phi_c = 0, 0.05, 0.1, 0.15, 0.2$  and  $0.25$  as indicated. (d) Dependence of effective free energy,  $F$ , on the fraction of native contacts,  $Q$ , at temperature  $T = 0.87\epsilon/k_B$  for different values of  $\phi_c$  as indicated. The data were obtained with crowder's radius  $R_c = 10.2\text{\AA}$  and container radius  $R_{\text{wall}} = 106.4\text{\AA}$ .

Fig. 3 shows that for both the absence and presence of crowders the folding transition of GB1 is cooperative as characterized by a prominent peak in the specific heat and is clearly two-state as displayed by the two minima separated by a barrier in the free energy profile (Fig. 3d). The position of the specific heat's peak (corresponding to  $T = T_f$ ) however is shifted towards higher temperature in presence of the crowders. As the crowders' volume fraction increases the folding temperature  $T_f$  also increases but the height of the specific heat peak somewhat decreases (Fig. 3c). These observations indicate that the crowders enhance the thermal stability of protein, but at the same time decreases folding cooperativity which usually is signified by the height of the specific heat peak. The crowders also enhance the thermodynamic stability of the native state or folding stability as can be seen from the free energy profile energy profiles (Fig. 3d).



**Figure 4.** Dependence of the change of the folding free energy on crowders' volume fraction. The points shown are obtained by simulations for different temperatures as indicated whereas the solid line presents an approximate fit by scaled particle theory (SPT) with  $R_c = 10.2\text{\AA}$ ,  $a_N = R_g^N = 10.2\text{\AA}$ , and  $a_U = 1.13R_g^N$ .

at a temperature below but close to the folding temperature. As  $\phi_c$  increases the free energy difference between the unfolded state and the folded state increases. The free energy barrier for folding decreases which indicates that folding is faster in presence of the crowders. One can also notice that the position of the minimum corresponding to the unfolded state is shifted towards higher  $Q$ .

Scaled particle theory (SPT) predicts that the free energy of inserting a hard sphere of radius  $R$  in a fluid of hard spheres of radius  $R_c$  is given by [9]:

$$\beta F = (3y + 3y^2 + y^3)\rho + (9y^2/2 + 3y^3)\rho^2 + 3y^3\rho^3 - \ln(1 - \phi_c), \quad (10)$$

where  $y = R/R_c$ ,  $\rho = \phi_c/(1 - \phi_c)$  and  $\phi_c$  is volume fraction of the fluid spheres. SPT has been used to estimate the effect of macromolecular crowding on the folding free energy of proteins, e.g. in Minton's theory [10], in which the folded and unfolded states are considered as effective spheres of radii  $a_N$  and  $a_U$ , respectively.

From the simulations, the folding free energy can be estimated as

$$\Delta F_{N-U} = F_N - F_U, \quad (11)$$

where  $F_N$  and  $F_U$  are the two free energy minima corresponding to the folded and unfolded states as found in the  $F(Q)$  plot. The change in folding free energy due to the effect of crowders can be calculated as:

$$\Delta\Delta F_{N-U} = \Delta F_{N-U}(\phi_c) - \Delta F_{N-U}(0). \quad (12)$$

Fig. 4 shows that  $\Delta\Delta F_{N-U}$  decreases monotonically with the volume fraction of crowders for a range of temperatures near the folding transition. The simulation data can be fitted quite well with SPT by using  $R_c = 10.2\text{\AA}$ ,  $a_N = R_g^N = 10.2\text{\AA}$ ,  $a_U = 1.13R_g^N$ .

#### 4. Conclusion

In this study, we have shown that macromolecular crowding leads to an increased folding temperature and increased folding stability as defined by the free energy difference between the folded and unfolded states. The change in folding stability depends monotonically with the volume fraction of the crowders in good agreement with scaled particle theory. The enhancement of folding stability found here are consistent with experiments and current understanding on the effects of crowding on protein folding.

## Acknowledgments

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number “103.01-2013.16”.

## References

- [1] Minton AP 1983 *Mol Cell Biochem.* **55** 119
- [2] Minton AP 2000 *Curr. Opin. Struct. Bio.* **10** 34
- [3] Ellis RJ 2001 *Curr. Opin. Struct. Bio.* **11** 114
- [4] Zhou HX, Rivas G, Minton AP 2008 *Ann. Rev. Biophys.* **37** 375
- [5] Zimmerman SB 1993 *Biochim. Biophys. Acta.* **1216** 175
- [6] Rivas G, Fernandez JA, Minton AP 1999 *Biochemistry* **38** 9379
- [7] Van den Berg B, Ellis RJ, Dobson CM 1999 *EMBO J.* **18** 6927
- [8] Van den Berg B, Wain R, Dobson CM, Ellis RJ 2000 *EMBO J.* **19** 3870
- [9] Lebowitz J L and Rowlinson J S 1964 *J. Chem. Phys.* **41** 133
- [10] Minton AP 2005 *Biophys. J.* **88** 971
- [11] Zhou HX 2008 *Arch. Biochem. Biophys.* **469** 76
- [12] Cheung MS, Klimov D and Thirumalai D 2005 *PNAS* **102** 4753
- [13] Mittal J, Best RB 2010 *Biophys J.* **98** 315
- [14] Cheung MS and Thirumalai D 2007 *J. Phys. Chem. B* **111** 8250
- [15] Go N 1983 *Annu. Rev. Biophys. Bioeng.* **12** 183
- [16] Hoang TX and Cieplak M 2000 *J. Chem. Phys.* **113** 8319
- [17] Ferrenberg A M and Swendsen R H 1989 *Phys. Rev. Lett.* **63** 1195