HHS Public Access

Author manuscript

Phys Biol. Author manuscript; available in PMC 2015 February 02.

Published in final edited form as:

Phys Biol. 2013 August ; 10(4): 045001. doi:10.1088/1478-3975/10/4/045001.

Folding free energy surfaces of three small proteins under crowding: validation of the postprocessing method by direct simulation

Sanbo Qin[1], **Jeetain Mittal**[2], and **Huan-Xiang Zhou**[1,*]

[1]Department of Physics and Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida 32306, USA

[2]Department of Chemical Engineering, Lehigh University, Bethlehem, PA 18015, USA

Abstract

We have developed a "postprocessing" method for modeling biochemical processes such as protein folding under crowded conditions. In contrast to the direct simulation approach, in which the protein undergoing folding is simulated along with crowders, the postprocessing method requires only the folding simulation without crowders. The influence of the crowders is then obtained by taking conformations from the crowder-free simulation and calculating the free energies of transferring to the crowders. This postprocessing yields the folding free energy surface of protein under crowding. Here the postprocessing results for the folding of three small proteins under "repulsive" crowding are validated by those obtained previously by the direct simulation approach. This validation confirms the accuracy of the postprocessing approach and highlights its distinct advantages in modeling biochemical processes under cell-like crowded conditions, such as enabling an atomistic representation of the test proteins.

Keywords

protein folding; macromolecular crowding; cell-like conditions; postprocessing

1. Introduction

It is now widely recognized that the crowded conditions found in cellular environments can significantly affect the thermodynamic and kinetic properties of biochemical processes such as protein folding [1]. Experimental [2-6] and computational studies [7-9] have begun to move beyond a qualitative understanding [10-12] of crowding effects and are aiming to probe quantitative questions regarding the magnitudes of such effects and the nature of the interactions between test proteins and crowders. We have developed a computational approach called postprocessing that opens the door to making quantitative predictions on crowding effects [8]. In contrast to the direct simulation approach [7, 9], in which one simulates the folding process in the presence of crowders and has to resort to a coarse-grained representation of the test protein, the postprocessing approach requires only the

*Corresponding author. hzhou4@fsu.edu.

simulation in the absence of crowders and thereby enables an atomistic representation of the test protein. Here we validate the postprocessing approach by showing that the folding free energy surfaces of three small proteins under “repulsive” crowding, obtained previously by Mittal and Best (MB hereafter) [9] using direct simulation, are accurately predicted by postprocessing the crowder-free simulation.

In the postprocessing approach, one captures the effects of crowding by calculating how much the free energy surface of a biochemical process such as folding is changed by the crowders [1, 8, 13]. This change is simply the free energy of transferring the test protein in a given state (or microstate) from a dilute solution (modeled by a crowder-free simulation) to a crowded solution. The transfer free energy, Δμ, is given by

$$\exp\left(-\Delta\mu \,/\, k_{\mathrm{B}}T\right) = \left\langle \exp\left[-U_{\mathrm{int}}(\mathbf{X},\mathbf{R},\mathbf{\Omega}) \,/\, k_{\mathrm{B}}T\right]\right\rangle_{0;\,\mathrm{crowd}} \quad (1)$$

where $U_{\mathrm{int}}(\mathbf{X}, \mathbf{R}, \mathbf{\Omega})$ is the effective interaction energy between the test protein and the crowders when the former has conformation **X**, position **R**, and orientation **Ω**; k_{B} is Boltzmann's constant; T is the absolute temperature; $\langle\cdots\rangle_{0;\,\mathrm{crowd}}$ means averaging over the conformation, position, and orientation of the test protein and over the configuration of the crowders; and the subscript “0” signifies that the protein conformations are those sampled in the absence of crowders. For each protein conformation, the averaging over protein position and orientation and crowder configuration can be carried out first, leading to

$$\exp\left(-\Delta\mu \,/\, k_{\mathrm{B}}T\right) = \left\langle \exp\left[-G_{\mathrm{c}}(\mathbf{X}) \,/\, k_{\mathrm{B}}T\right]\right\rangle_{0} \quad (2)$$

$G_{\mathrm{c}}(\mathbf{X})$ can be viewed as the transfer free energy for the protein in conformation **X**. The postprocessing calculation is exact if the conformational space of the protein is exhaustively sampled so as to cover the important conformational region in the presence of crowders.

In a previous study [14], we have shown that the effects of crowding on the flap open-to-closed population ratio of the HIV-1 protease dimer predicted by the postprocessing approach from the crowder-free simulation agree well with those obtained from direct simulations in the presence of crowders [15]. Of course the conformational difference between the flap open and flap closed states is somewhat limited. Here we present a much more stringent validation of the postprocessing approach. We show that the entire folding free energy surfaces of three small proteins (figure 1), encompassing the fully unfolded state and the fully folded state, obtained by MB in direct simulations with crowders, are well predicted by postprocessing the crowder-free simulation. The validation also highlights the distinct advantages of the postprocessing approach, including superior statistics and the ability to more realistically model the test proteins, the crowders, and their interactions.

2. Computational details

2.1 Calculation of F(Q) from replica-exchange umbrella sampling

MB obtained the folding free energy function, $F(Q)$, along the fraction, Q, of native contacts from replica-exchange simulations with umbrella sampling. $F(Q)$ was calculated by the weighted histogram analysis method (WHAM) [16]. Below we extend WHAM to account for the effects of crowding. In preparation, we briefly summarize the original WHAM algorithm for a set of simulations (referred to by index i) applying biasing potentials $V_i(Q)$ that constrain the protein to different windows along Q. First consider the case where all the simulations are carried out at the same temperature T; for notational simplicity, we will now denote $1/k_B T$ as β. Suppose that n_i snapshots (referred to by index j) are sampled from simulation i, and from which a histogram $p_i(Q_l)$ in different bins, denoted by Q_l, are obtained:

$$p_i(Q_l) = \frac{\sum_{j=1}^{n_i} H(Q_{ij} \mid Q_l)}{n_i} \quad (3)$$

where Q_{ij} is the value of Q in snapshot j of simulation i, and $H(Q_{ij}|Q_l)$ is 1 of Q_{ij} is in bin Q_l and 0 otherwise. By removing the biasing potential $V_i(Q)$ and introducing an appropriate shift A_i, one obtains the equilibrium probability density:

$$P(Q_l) = p_i(Q_l)\exp\{\beta[V_i(Q_l) - A_i]\} \quad (4)$$

Better statistics are achieved when the data from different simulations are combined. To that end one rearranges equation (4) and then sums over i, leading to

$$P(Q_l) = \frac{\sum_i \sum_{j=1}^{n_i} H(Q_{ij} \mid Q_l)}{\sum_i n_i \exp\{-\beta[V_i(Q_l) - A_i]\}} \quad (5)$$

The shift A_i can be obtained from equation (4) by using the condition that the sum of $p_i(Q_l)$ over Q_l is 1 (see equation (3)):

$$\exp(-\beta A_i) = \sum_{Q_l} P(Q_l)\exp[-\beta V_i(Q_l)] \quad (6)$$

An implementation of the WHAM involves iterating equations (5) and (6) to convergence. Finally the free energy function is given by

$$\exp[-\beta F(Q)] = P(Q) \quad (7)$$

Now consider the case where each simulation is at a different temperature T_i (correspondingly $1/k_B T_i = \beta_i$). The goal again is to calculate the equilibrium probability density $P(Q_l)$ at temperature T. The difference of the simulation temperature T_i from the desired temperature T amounts to a biasing factor $\exp[-(\beta - \beta_i)U_{ij}]$, where U_{ij} is the energy of the protein in snapshot j of simulation i [17]. Defining

$$m_i = \sum_{j=1}^{n_i} \exp\left[-(\beta - \beta_i)U_{ij}\right] \quad (8)$$

and accounting for the biasing factor, one finds the counterpart of $p_i(Q_l)$ as

$$q_i(Q_l) = \frac{\sum_{j=1}^{n_i} H(Q_{ij} \mid Q_l)\exp\left[-(\beta - \beta_i)U_{ij}\right]}{m_i} \quad (9)$$

and now

$$P(Q_l) = q_i(Q_l)\exp\{\beta_i[V_i(Q_l) - A_i]\} \quad (10)$$

Equations (5) and (6) become

$$P(Q_l) = \frac{\sum_i \sum_{j=1}^{n_i} H(Q_{ij} \mid Q_l)\exp\left[-(\beta - \beta_i)U_{ij}\right]}{\sum_i m_i \exp\{-\beta_i[V_i(Q_l) - A_i]\}} \quad (11)$$

$$\exp(-\beta_i A_i) = \sum_{Q_l} P(Q_l)\exp[-\beta_i V_i(Q_l)] \quad (12)$$

Replica exchange simulations generate equilibrium distributions at multiple temperatures. The snapshots sampled from such simulations are equivalent to those in the case just considered. Iteration of equations (11) and (12) thus yields the equilibrium probability density and hence the free energy function at the desired temperature. MB carried out the simulations both in the absence and in the presence of crowders. The difference between the resulting free energy functions, $F(Q)$ and $F_c(Q)$, indicates the effects of crowding.

2.2 Calculation of $F_c(Q)$ by postprocessing

The goal of postprocessing is to generate the equilibrium probability density $P_c(Q)$ and hence the free energy function $F_c(Q)$ under crowding from the crowder-free simulations. For each crowder-free simulation, introducing an additional biasing factor $\exp[-\beta G_c(\mathbf{X}_{ij})]$ for each sampled conformation $\mathbf{X}_{ij}$ will result in the histogram under crowding. Defining

$$v_i = \sum_{j=1}^{n_i} \exp\left[-(\beta-\beta_i)U_{ij}\right]\exp\left[-\beta G_c(\mathbf{X}_{ij})\right] \quad (13)$$

we find the histogram under crowding as

$$\rho_i(Q_l) = \frac{\sum_{j=1}^{n_i} H(Q_{ij} \mid Q_l)\exp\left[-(\beta-\beta_i)U_{ij}\right]\exp\left[-\beta G_c(\mathbf{X}_{ij})\right]}{v_i} \quad (14)$$

Correspondingly the equilibrium probability density under crowding is

$$P_c(Q_l) = \rho_i(Q_l)\exp\{\beta_i[V_i(Q_l) - A_i]\} \quad (15)$$

In analogy to equations (11) and (12) we now have

$$P_c(Q_l)\frac{\sum_i \sum_{j=1}^{n_i} H(Q_{ij} \mid Q_l)\exp\left[-(\beta-\beta_i)U_{ij}\right]\exp\left[-\beta G_c(\mathbf{X}_{ij})\right]}{\sum_i v_i \exp\{-\beta_i[V_i(Q_l) - A_i]\}} \quad (16)$$

$$\exp(-\beta_i A_i) = \sum_{Q_l} P_c(Q_l)\exp\left[-\beta_i V_i(Q_l)\right] \quad (17)$$

2.3 Calculation of $\exp[-\beta G_c(\mathbf{X}_{ij})]$

In MB's simulations, the test protein had purely repulsive interactions with spherical crowders. If we further approximate this interaction energy as infinite when the test protein clashes with a crowder and 0 when no clash occurs with any crowder, then the Boltzmann factor $\exp[-U_{\text{int}}(\mathbf{X}, \mathbf{R}, \mathbf{\Omega})/k_B T]$ is either 0 or 1. Correspondingly $\exp[-G_c(\mathbf{X})k_B/k_B T]$ is the fraction of clash-free, or allowed attempts to randomly insert the test protein with conformation **X** into the crowders. We designed an algorithm for calculating the allowed fraction of insertion by mimicking the fictitious insertion process [8]. Subsequently we developed a semi-analytical method, referred to as the generalized fundamental measure theory (GFMT) [18], to substantially speed up the calculation. The GFMT was crucial for the present study since postprocessing calculations on a large number of protein conformations were required.

For a test protein in conformation **X**, analogous to the original fundamental measure theory [19, 20], the GFMT predicts the transfer free energy $G_c(\mathbf{X})$ as

$$G_c(\mathbf{X}) = \Pi_c v_p + \gamma_c s_p + \kappa_c l_p - k_B T \ln(1-\phi) \quad (18)$$

where ϕ is the volume fraction of the crowders; Π_c, γ_c, and κ_c are the osmotic pressure, surface tension, and bending rigidity, respectively, of the crowder suspension; and v_p, s_p, and l_p are the volume, surface area, and linear size, respectively, of the test protein. Equation (19) equates the transfer free energy to the work required to create a cavity in the crowder suspension to accommodate a test particle. The first term on the right-hand side is the work for creating the cavity volume, the second term is work for creating the interface area between the cavity and the crowder suspension, and the third term accounts for the curvature of this interface. The original fundamental measure theory [19, 20] was restricted to hard convex test particles. Our generalization allows atomistic test proteins to be treated.

The generalization involved appropriately defining the geometric quantities v_p, s_p, and l_p for an atomistic test protein. The calculation of v_p and s_p was based on the crowder-excluded surface, which depends on the crowder radius. This surface, when the probe radius is 1.4 Å (modeling a water molecule), is more commonly referred to as the molecular surface. The calculation of l_p involved a surface that is slightly inflated from the crowder-excluded surface. This inflated surface is composed of all points where rays emanating from the center of the protein first hit the closest approaching crowders around the protein; l_p was calculated as the radius of gyration for these hit points.

Π_c, γ_c, and κ_c are crowder-only quantities. Suppose that the crowder suspension consists of multiple species, with number density c_α, linear size l_α, surface area s_α, and volume v_α for species α. Note that

$$\phi = \sum_\alpha c_\alpha v_\alpha \quad (19)$$

Let the total number density, the number-averaged linear size and surface area of the crowders, be c, $\bar{l}_c$ and $\bar{s}_c$, respectively:

$$c = \sum_\alpha c_\alpha \quad (20)$$

$$\bar{l}_c = \sum_\alpha c_\alpha l_\alpha \Big/ c \quad (21)$$

$$\bar{s}_c = \sum_\alpha c_\alpha s_\alpha \Big/ c \quad (22)$$

Then

$$\beta\Pi_c = Y + \bar{l}_c \bar{s}_c Y^2 + \bar{s}_c^3 Y^3 \Big/ 12\pi \quad (23)$$

$$\beta\gamma_c = \bar{l}_c Y + \bar{s}_c^2 Y^2 \Big/ 8\pi \quad (24)$$

$$\beta\kappa_c = \bar{s}_c Y \quad (25)$$

where $Y = c/(1 - \phi)$.

The GFMT implemented for proteins represented at the all-atom level closely reproduced results obtained by the insertion algorithm for the effects of crowding on folding and binding free energies [18]. Here (as done previously [14]), for the purpose of validating the postprocessing approach against direct simulations, we kept MB's coarse-grained representation of one bead per residue for the three test proteins. In this representation, proposed by Karanicolas and Brooks [21], the beads are centered at the Cα atoms and have radii of 4.0 ± 0.8 Å; the interactions between residues in contact in the native structure are modeled as attractive and any other residue-residue contacts formed during conformational sampling are repulsive. MB found that this representation resulted in two-state folding behavior.

The postprocessing predictions involved calculating v_p, s_p, and l_p at a given crowder radius for each protein conformation sampled in a crowder-free simulation. Our computer code for generating the crowder-excluded surface (to calculate v_p and s_p) and inflated surface (to calculate l_p) was applied previously to proteins represented as all atoms with van der Waals radii [18]. This code was used here for the three proteins represented as Cα beads with enlarged radii.

3. Results

The three small proteins studied by MB are prb, protein G, and TNfn3, with 47, 56, and 90 residues respectively (figure 1). The 100 × 100 × 100 $Å^3$ simulation box contained a test protein without or with various numbers (N_c) of spherical crowders (radius: R_c). The interactions between the Cα beads of the protein and the crowders were purely repulsive. The nominal volume fraction, ϕ_0, of the crowders equals $4\pi N_c R_c^3 \Big/ \left(3 \times 10^6\right)$. Below we compare their simulation results for $F_c(Q)$ against those predicted by the postprocessing approach, which uses only the simulations without crowders. We also compare the resulting effects of crowding on the folding free energy and on the folding rate constant.

3.1 Folding free energy function under crowding

We first made the postprocessing calculations for $F_c(Q)$ by assuming ϕ_0 as the actual crowder volume fraction ϕ. The results generally show reasonable agreement with MB's direct simulation results, but as illustrated in figure 2a for protein G in the presence of 8 Å crowders, there is a systematic underestimation of crowding effects.

We then realized that the simulation system has a small, finite size. The test protein presents a volume that is inaccessible to the crowders; the actual volume fraction of the crowders around the test protein is thus higher than that (i.e., ϕ_0) calculated by assuming the absence of the test protein. We thus re-calculated ϕ by carving out from the total volume the region in and around the test protein inaccessible to each crowder (figure 3). The inaccessible volume v_{inac}, could be calculated by rolling around the protein a spherical probe with the crowder radius R_c, or could be estimated by Steiner's formula [22]

$$v_{inac} = v_p + R_c s_p + l_p s_c + v_c \quad (26)$$

where $s_c = 4\pi R_c^2$ and $v_c = 4\pi R_c^3 / 3$, respectively, are the surface area and volume of the spherical probe. We found that these two ways gave very similar results for v_{inac} and all the data presented below are by using equation (26). The correction in ϕ brings the postprocessing predictions into close agreement with MB's direct simulation results for $F_c(Q)$, as illustrated in figure 2b for protein G interacting with 8 Å crowders.

Note that calculating the transfer free energy $G_c(\mathbf{X})$ required for predicting $F_c(Q)$ amounts to inserting a fictitious test protein into the crowders around the real test protein (figure 3). The above correction for the finite size of the simulation system by carving out the region in and around the real test protein inaccessible to the crowders could result in overestimation of the crowding effect, since, when inserting the fictitious test protein, the inaccessible volume around the real test protein and the inaccessible volume around a crowder can overlap. This overlap volume is effectively counted twice in the above correction, once in calculating ϕ, and a second time when testing whether an insertion attempt is allowed. The overlap volume increases with increasing sizes of the test protein and the crowders. The resulting overestimation of the crowding effect was significant only for protein G interacting with 16 Å crowders and for TNfn3 interacting with 16 Å or 20 Å crowders.

A better correction for the finite size of the simulation system is to treat the insertion of a fictitious test protein into the crowders around the real test protein as insertion into the mixture of the real test protein and the crowders (figure 3). With this treatment, the postprocessing predictions for $F_c(Q)$ are in good agreement with MB's direct simulation results for all the three proteins interacting with all the different sized crowders. The results for protein G interacting with 8 Å crowders shown in figure 2c are representative of the extent of the agreement.

3.2 Effect of crowding on folding stability

As figure 2 shows, the free energy functions and $F(Q)$ and $F_c(Q)$ define two basins, one for the unfolded state (with free energy minimum at $Q \equiv Q_U \sim 0.1$), and the other for the native state (with free energy minimum at $Q \equiv Q_N \sim 0.9$). The folding stability is measured by the free energy difference, ΔF_{N-U}, which can be calculated from the free energy functions $F(Q)$ as

$$\Delta F_{N-U} = -k_B T \ln\left\{\int_0^{Q^\ddagger} \exp[-\beta F(Q)]dQ \Big/ \int_{Q^\ddagger}^1 \exp[-\beta F(Q)]dQ\right\} \quad (27)$$

where $Q^\ddagger$, with a value ~ 0.5, denotes the location of the free energy maximum separating the two stable states. An analogous free energy difference can be calculated for the situation under crowding. In figure 4 we compare the postprocessing predictions and MB's direct simulation results for the change in the free energy difference by crowding, $\Delta\Delta_{N-U}$, for the three proteins under various levels of crowding. It can be seen that, after correcting for the finite size of the simulation system, there is good agreement between postprocessing and direct simulation.

Some of the direct simulation results are apparently noisy, such as those for TNfn3 in the presence of 12 Å crowders at five ϕ_0 values. The noisy data can be attributed to the small number of crowders included in the simulations. For TNfn3 in the presence of 20 Å crowders, there were only 4 to 10 crowders in the simulation box. In contrast, in the postprocessing approach, the test protein fictitiously probes an essentially infinite suspension of crowders. The resulting better statistics yield the smooth dependences of $\Delta\Delta F_{N-U}$ on ϕ_0 shown in figure 4.

3.3 Effect of crowding on folding kinetics

The rate constant for the transition from the unfolded state to the native state can be calculated according to Kramers' theory [23]:

$$k_f = 1 \Big/ \int_0^{Q^\ddagger} \exp[-\beta F(Q)]dQ \int_{Q_U}^{Q_N} [D(Q)]^{-1} \exp[\beta F(Q)]dQ \quad (28)$$

where $D(Q)$ is the effective diffusion coefficient along Q. The folding rate constant under crowding, k_{fc}, can be similarly calculated. If the crowding does not affect $D(Q)$, the we have

$$\frac{k_{fc}}{k_f} = \frac{\int_0^{Q^\ddagger} \exp[-\beta F(Q)]dQ \int_{Q_U}^{Q_N} \exp[\beta F(Q)]dQ}{\int_0^{Q^\ddagger} \exp[-\beta F_c(Q)]dQ \int_{Q_U}^{Q_N} \exp[\beta F_c(Q)]dQ} \quad (29)$$

MB found that equation (29) predicts accurately the effect of crowding on the folding rate constant, supporting that crowding does not affect $D(Q)$ for the present systems.

In figure 5 we compare the k_{fc}/k_f results calculated from equation (29), using either the $F_c(Q)$ predicted by postprocessing or the $F_c(Q)$ from MB's direct simulation, for prb under various levels of crowding. This comparison again validates the postprocessing approach.

4. Discussion

We have shown that the postprocessing approach, using only the simulations in the absence of crowders, can predict well the entire free energy surfaces of three small proteins under repulsive crowding. Consequently the effects of crowding on the folding stability and the folding rate constant can be well predicted. Although we have previously validated the postprocessing approach against direct simulation with crowders for the flap open-to-closed population ratio [14], there the conformational space involved was relatively limited. In contrast, here the conformational space sampled encompasses the fully folded and fully folded state, and thus presents a much more stringent validation of the postprocessing approach.

To sample the entire folding free energy surface under crowding, MB used a relatively small $100 \times 100 \times 100$ Å^3 simulation box. For the largest sized (20 Å) crowders, only a few (4 to 10) of them were in the simulation box to represent the crowded environment. We found that the finite size of the simulation system has a significant effect. The crowded environment in the direct simulation is different from that if the single test protein were taken away. We found a way to correct for the finite size effect, by treating the insertion of a fictitious test protein into the crowded environment in the direct simulation as insertion into the mixture of the real test protein and the crowders. In our previous validation study [14], the simulation was much larger (at $510 \times 510 \times 510$ Å^3) and hence the finite size effect would have been negligible. We emphasize that the correction for the finite size effect was necessary only because we wanted to compare the postprocessing predictions against the direct simulation results. In all our other applications [8, 18, 24-28], we use an essentially infinite crowder suspension for the insertion to calculation the transfer free energy.

In addition to avoiding the finite size effect, in the postprocessing approach the crowder configurations are exhaustively sampled because the fictitious test protein is inserted everywhere in the essentially infinite crowder suspension. The postprocessing results thus have better statistics than the direct simulation counterparts.

The postprocessing calculation is formally exact if the conformational space of the protein is exhaustively sampled by crowder-free simulations, so as to cover the important conformational region in the presence of crowders. A potential concern is whether crowder-free simulations in practice can achieve this goal. The present validation study as well as the previous one [14] allays this concern. However, we do note that in both studies the test proteins had purely repulsive interactions with the crowders. It is possible that attractive protein-crowder interactions may alter the energy landscape of a protein such that new low-energy regions emerge in the presence of crowders. These regions could be under-sampled

in crowder-free simulations, leading to errors in the postprocessing calculation. Only further validation will tell the level of these errors.

The most important advantage of the postprocessing approach is the ability to represent the test protein at the atomistic level and its interaction with the crowders more realistically. As observed recently by Gruebele and co-workers [6], the atomistic level representation is required to capture the subtle effects of crowding on protein folding. With the present validation study, a major potential concern regarding the validity of the postprocessing approach is removed at least for repulsive crowders. We can now apply this powerful approach with confidence for realistic modeling of crowding effects.

Acknowledgements

This study was supported by grant GM88187 from the National Institutes of Health.

References

1. Zhou HX, Rivas G, Minton AP. Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. Annu Rev Biophys. 2008; 37:375–97. [PubMed: 18573087]
2. Ai X, Zhou Z, Bai Y, Choy WY. ^{15}N NMR spin relaxation dispersion study of the molecular crowding effects on protein folding under native conditions. J Am Chem Soc. 2006; 128:3916–7. [PubMed: 16551092]
3. Batra J, Xu K, Zhou H-X. Nonaddtive effects of mixed crowding on protein stability. Proteins. 2009; 77:133–8.
4. Mukherjee S, Waegele MM, Chowdhury P, Guo L, Gai F. Effect of macromolecular crowding on protein folding dynamics at the secondary structure level. J Mol Biol. 2009; 393:227–36. [PubMed: 19682997]
5. Miklos AC, Sarkar M, Wang Y, Pielak GJ. Protein crowding tunes protein stability. J Am Chem Soc. 2011; 133:7116–20. [PubMed: 21506571]
6. Guo M, Xu Y, Gruebele M. Temperature dependence of protein folding kinetics in living cells. Proc Natl Acad Sci U S A. 2012; 109:17863–7. [PubMed: 22665776]
7. Cheung MS, Klimov D, Thirumalai D. Molecular crowding enhances native state stability and refolding rates of globular proteins. Proc Natl Acad Sci U S A. 2005; 102:4753–8. [PubMed: 15781864]
8. Qin S, Zhou H-X. Atomistic modeling of macromolecular crowding predicts modest increases in protein folding and binding stability. Biophys J. 2009; 97:12–9. [PubMed: 19580740]
9. Mittal J, Best RB. Dependence of protein folding stability and dynamics on the density and composition of macromolecular crowders. Biophys J. 2010; 98:315–20. [PubMed: 20338853]
10. Minton AP. Effect of a concentrated “inert” macromolecular cosolute on the stability of a globular protein with respect to denaturation by heat and by chaotropes: a statistical-thermodynamic model. Biophys J. 2000; 78:101–9. [PubMed: 10620277]
11. Zhou H-X. Protein folding and binding in confined spaces and in crowded solutions. J Mol Recognit. 2004; 17:368–75. [PubMed: 15362094]
12. Minton AP. Models for excluded volume interaction between an unfolded protein and rigid macromolecular cosolutes: macromolecular crowding and protein stability revisited. Biophys J. 2005; 88:971–85. [PubMed: 15596487]
13. McGuffee SR, Elcock AH. Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm. PLoS Comput Biol. 2010; 6:e1000694. [PubMed: 20221255]
14. Qin S, Minh DD, McCammon JA, Zhou HX. Method to predict crowding effects by postprocessing molecular dynamics trajectories: application to the flap dynamics of HIV-1 protease. J Phys Chem Lett. 2010; 1:107–10. [PubMed: 20228897]

15. Minh DD, Chang CE, Trylska J, Tozzini V, McCammon JA. The influence of macromolecular crowding on HIV-1 protease internal dynamics. J Am Chem Soc. 2006; 128:6006–7. [PubMed: 16669648]
16. Kumar S, Bouzida D, Swendsen RH, Kollman PA, Rosenberg JM. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. J Comput Chem. 1992; 13:1011–21.
17. Gallicchio E, Andrec M, Felts AK, Levy RM. Temperature weighted histogram analysis method, replica exchange, and transition paths. Journal of Physical Chemistry B. 2005; 109:6722–31.
18. Qin S, Zhou HX. Generalized fundamental measure theory for atomistic modeling of macromolecular crowding. Phys Rev E. 2010; 81:031919.
19. Rosenfeld Y. Free-energy model for the inhomogeneous hard-sphere fluid mixture and density-functional theory of freezing. Phys Rev Lett. 1989; 63:980–3. [PubMed: 10041238]
20. Oversteegen SM, Roth R. General methods for free-volume theory. J Chem Phys. 2005; 122:214502. [PubMed: 15974749]
21. Karanicolas J, Brooks CL 3rd. The origins of asymmetry in the folding transition states of protein L and protein G. Protein Sci. 2002; 11:2351–61. [PubMed: 12237457]
22. Steiner J. Über parallele Flächen. Monatsber preuss Akad Wiss. 1840; 2:114–8.
23. Kramers HA. Brownian motion in a field of force and the diffusion model of chemical reactions. Physica. 1940; 7:284–304.
24. Batra J, Xu K, Qin S, Zhou H-X. Effect of macromolecular crowding on protein binding stability: modest stabilization and significant biological consequences. Biophys J. 2009; 97:906–11. [PubMed: 19651049]
25. Dong H, Qin S, Zhou HX. Effects of macromolecular crowding on protein conformational changes. PLoS Comput Biol. 2010; 6:e1000833. [PubMed: 20617196]
26. Tjong H, Zhou HX. The folding transition-state ensemble of a four-helix bundle protein: helix propensity as a determinant and macromolecular crowding as a probe. Biophys J. 2010; 98:2273–80. [PubMed: 20483336]
27. Qin S, Cai L, Zhou HX. A method for computing association rate constants of atomistically represented proteins under macromolecular crowding. Phys Biol. 2012; 9:066008. [PubMed: 23197255]
28. Phillip Y, Harel M, Khait R, Qin S, Zhou HX, Schreiber G. Contrasting factors on the kinetic path to protein complex formation diminish the effects of crowding agents. Biophys J. 2012; 103:1011–9. [PubMed: 23009850]

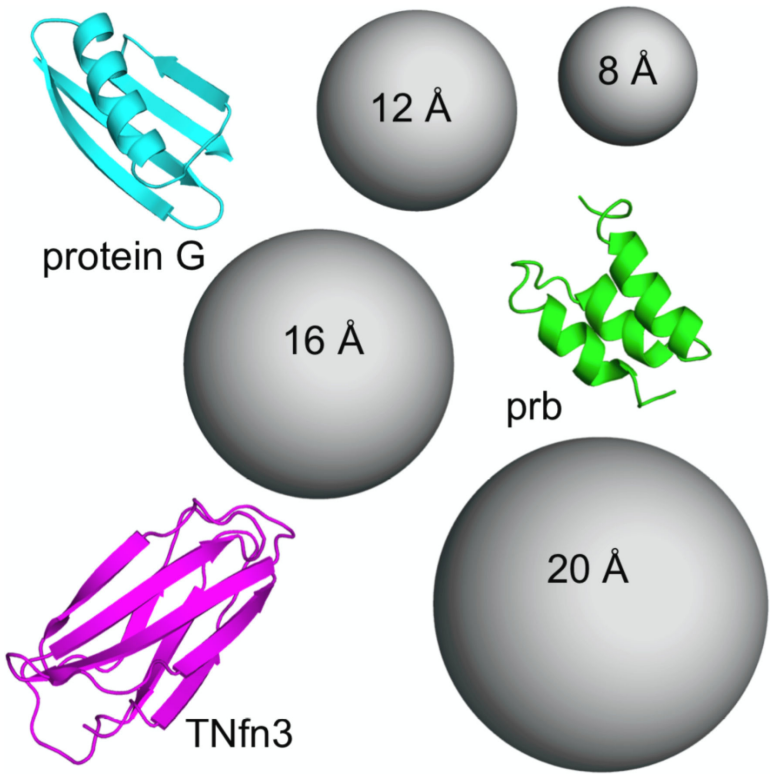

Figure 1.
The three proteins studied. Their structures and four different-sized crowders are shown to scale.

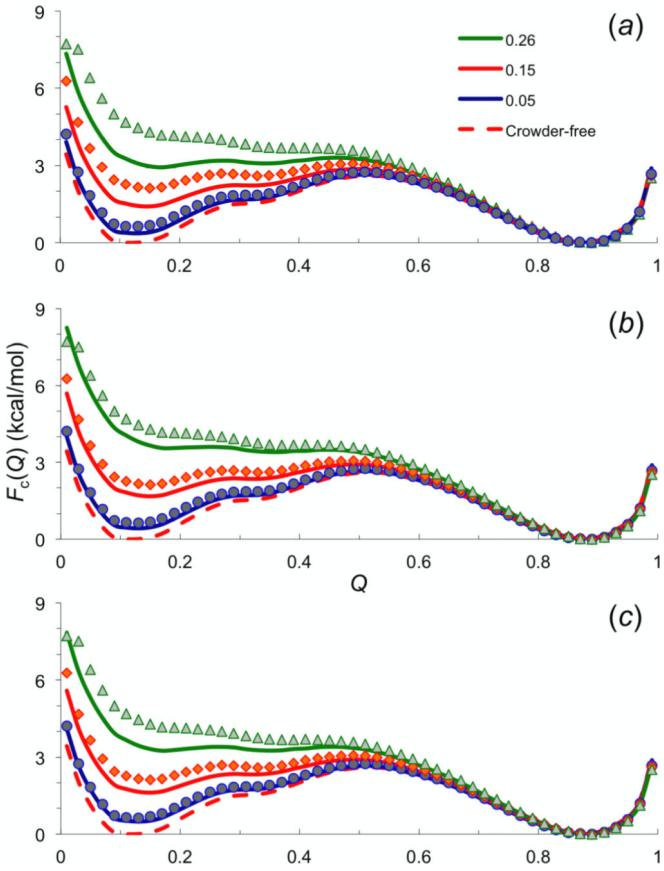

Figure 2.

Comparison of $F_c(Q)$ results from MB's direct simulations and predicted by the postprocessing approach. Circles, diamonds, and triangles show the simulation results for protein G in the presence of 8 Å crowders at ϕ_0 = 0.05, 0.15, and 0.26, respectively. The dashed curve shows the free energy function in the absence of crowders; the three solid curves show $F_c(Q)$ for the three ϕ_0 values by postprocessing the crowder-free simulations. The temperature is 310 K. The postprocessing results are obtained by either (a) assuming ϕ_0 as the actual crowder volume fraction ϕ; or correcting for the finite size of each simulation system, through (b) carving out the region in and around the real test protein inaccessible to a crowder in calculating ϕ or (c) treating the insertion of a fictitious test protein into the crowders around the real test protein as insertion into the mixture of the real test protein and the crowders.

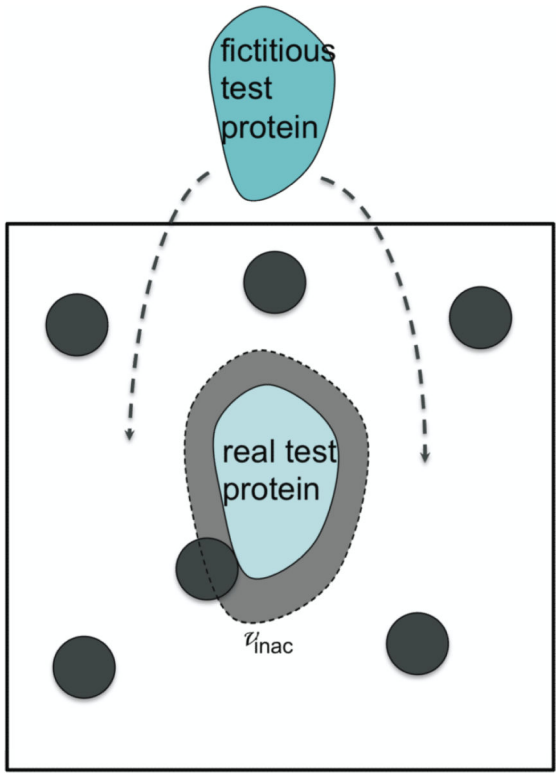

Figure 3.

Correcting for the finite size of MB's simulation system. Because the test protein in the direct simulation presents a volume (v_{inac}) inaccessible to the crowders, the actual volume fraction of the crowders is higher than the nominal volume fraction ϕ_0. Inserting a fictitious test protein into the crowders around the real test protein is equivalent to inserting into the mixture of the real test protein and the crowders.

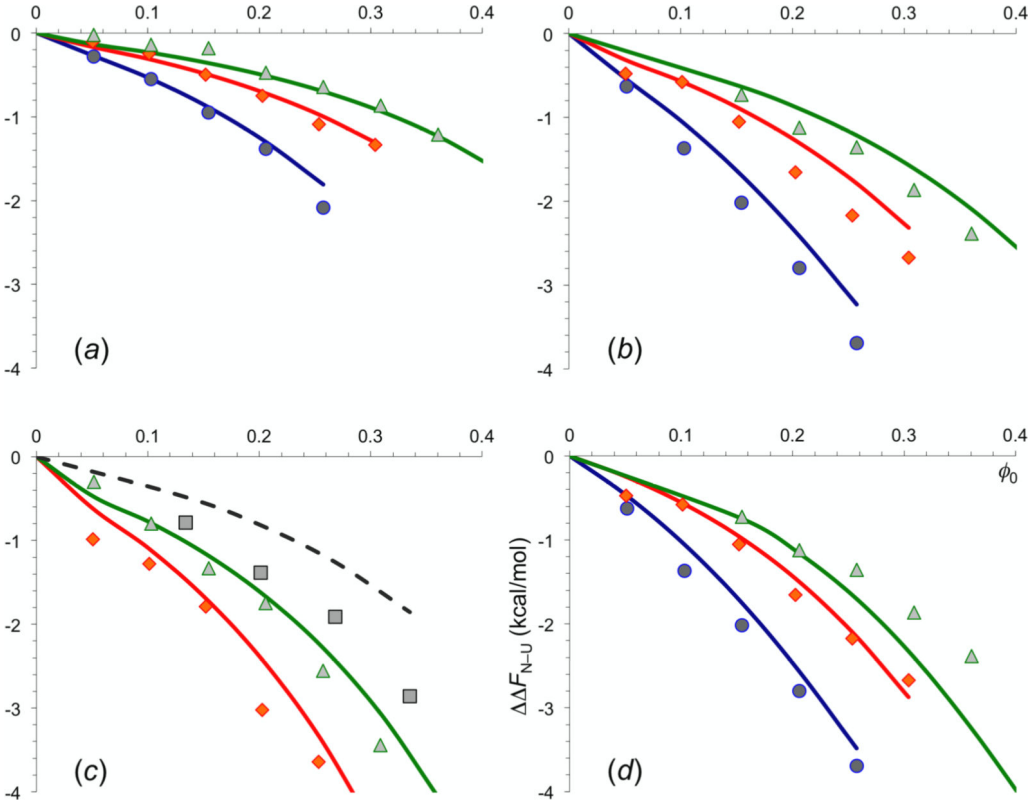

Figure 4.

Comparison of the postprocessing predictions and MB's direct simulation results for the effects of crowding on the folding stability of the three proteins. Circles, diamonds, triangles, and squares show the simulation results for crowders with radii at 8, 12, 16, and 20 Å, respectively. The corresponding curves show the postprocessing predictions after correcting for the finite size of each simulation system, by either (a-c) the mixture treatment or (d) the carving out treatment. (a) prb at 320 K; (b) and (d) protein G at 320 K; (c) TNfn3 at 300 K.

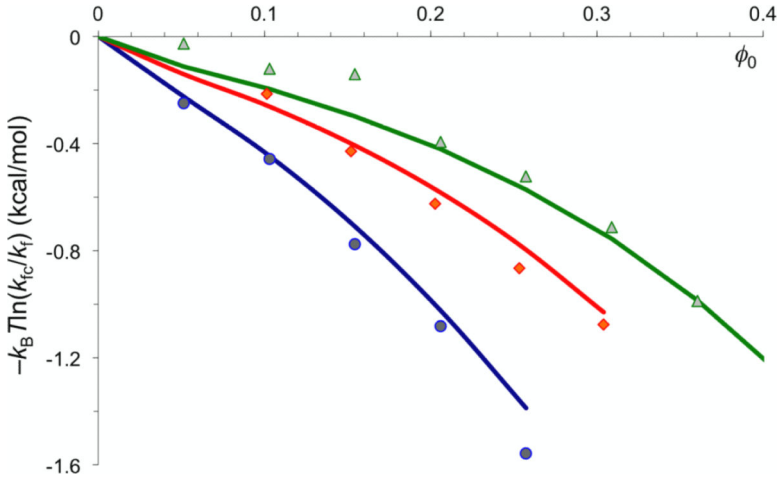

Figure 5.

Comparison of the postprocessing predictions and MB's direct simulation results for the effect of crowding on the folding rate constant of prb. Symbols have the same meaning as in figure 4a.