Available online at www.sciencedirect.com
ScienceDirect
Archives of Biochemistry and Biophysics 469 (2008) 76–82
ABB
www.elsevier.com/locate/yabbi

Review

Protein folding in confined and crowded environments

Huan-Xiang Zhou *

Department of Physics and Institute of Biophysics and School of Computational Science, Florida State University, Tallahassee, FL 32306, USA

Received 10 May 2007, and in revised form 12 July 2007
Available online 1 August 2007

Abstract

Confinement and crowding are two major factors that can potentially impact protein folding in cellular environments. Theories based on considerations of excluded volumes predict disparate effects on protein folding stability for confinement and crowding: confinement can stabilize proteins by over $10k_BT$ but crowding has a very modest effect on stability. On the other hand, confinement and crowding are both predicted to favor conformations of the unfolded state which are compact, and consequently may increase the folding rate. These predictions are largely borne out by experimental studies of protein folding under confined and crowded conditions in the test tube. Protein folding in cellular environments is further complicated by interactions with surrounding surfaces and other factors. Concerted theoretical modeling and test-tube and *in vivo* experiments promise to elucidate the complexity of protein folding in cellular environments.

Keywords: Protein folding; Confinement; Crowding; Excluded volume

The complexity of the cellular environments in which proteins function is increasingly been appreciated [1]. Potential impacts of two related, but distinct factors, confinement and macromolecular crowding, on protein folding are receiving extensive attention [2–5]. Confinement arises from encapsulation of proteins in compartments with dimensions on the scales of 10s to 100s Å, which are only moderately larger than the sizes of the proteins themselves. Examples of such compartments include the Anfinsen cage of chaperonins and the exit tunnel of ribosomes. Crowding refers to the exclusion of volumes to proteins of interest by the presence of other macromolecules. The aims of this paper are to discuss potential impacts of confinement and crowding on protein folding and to review the progress made by recent studies on protein folding in confined and crowded conditions and in cellular environments.

The review is focused on the literature that has appeared from 2004 onward. In 2003, EMBO sponsored a workshop on "Biological Implications of Macromolecular Crowding." Organized by John Ellis, Allen Minton, and Germán Rivas, that workshop presented an overview of studies on confinement and crowding up to that point. Papers presented there are published in a special issue of Journal of Molecular Recognition (2004 September–October issue). Readers interested in the earlier literature will find that special issue useful. The contributions of confinement and crowding to protein folding stability, in the context of entropy-based factors, are also reviewed in a 2004 paper [6].

In what follows, first the basic principles for understanding confinement and crowding are outlined. Given the complexity of modeling confinement and crowding and the desire to give the general reader a useful guide, the focus will be on basic ideas and qualitative predictions. Therefore the many studies using molecular dynamics simulations to elucidate details will not be extensively covered. Then an overview into experimental studies on confinement and crowding will be presented. While a major effect of confinement and crowding is excluded volume, it has become clear that interactions of proteins with confining surfaces and with macromolecular crowders can also exert significant influence. This and other complicating factors are discussed next. The review concludes with the positive outlook that much of the complexity of *in vivo* protein fold-

* Fax: +1 850 644 7244.
E-mail address: zhou@sb.fsu.edu

0003-9861/$ - see front matter
doi:10.1016/j.abb.2007.07.013

ing can be elucidated through theoretical modeling and parallel experiments under simulated and actual cellular conditions.

Theoretical background

In the simplest case, protein folding involves the transition between the unfolded state and the folded state. Folding stability is measured by the free-energy difference between the folded and unfolded states. The folding rate is determined by the dynamic processes that bring a protein molecule from the unfolded state to the folded state. An essential dynamic step is the formation of native contacts [7–10]. Here a theoretical framework for understanding the effects of confinement and crowding on folding free energy and contact formation rate is outlined.

Confinement and folding free energy

As noted in the Introduction, a major effect of confinement and crowding is excluded volume. In a confined space, conformations or positions in which parts of the protein chain lie outside the walls cannot be sampled. The unfolded state is more adversely affected, since in this state the conformations are more open and hence are more likely to cross the walls of the confined space. Thus in general confinement is expected to favor the folded state and increase folding stability.

This basic idea can be made concrete by modeling the folded protein as a solid sphere (with radius a_F) and the unfolded protein as a polymer chain (see Fig. 1) [11,12]. In a spherical cage with radius R_c, a shell with thickness a_F is excluded to the center of the folded protein, reducing the accessible volume from $4\pi R_c^3/3$ to $(4\pi R_c - a_F)^3/3$. This

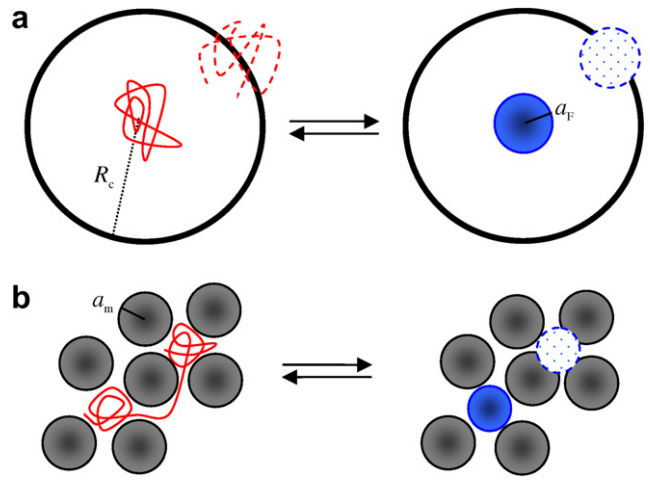

Fig. 1. Illustration of protein folding (a) inside a spherical cage (with radius R_c) and (b) in a solution of macromolecular crowders (with radius a_m). The unfolded protein is shown as a chain in red and the folded protein is shown as a sphere in blue. Conformations/positions in which the protein (unfolded or folded) overlaps with the confining surface or with crowders are eliminated.

volume exclusion reduces the statistical weight of the folded protein by a factor $f_F = (1 - a_F/R_c)^3$. A spherical cage is most suited for a spherical protein. Any other shape, such as an ellipsoid, will be more difficult to be accommodated in the spherical cage, and will thus have a lower f_F value. For other shapes, f_F can be obtained by placing the protein inside the cage with uniform distributions for the location and orientation; f_F is the fraction of placements in which the protein is fully within the confining cage.

The simplest polymer chain, for modeling the unfolded protein, has Gaussian statistics for the displacement vector, $\mathbf{r}$, between two residues separated by n peptide bonds:

$$p_s(\mathbf{r}, n) = (3/2\pi b^2 n)^{3/2} \exp(-3\mathbf{r}^2/2b^2 n) \tag{1}$$

where b is the effective bond length and the subscript "s" signifies that the polymer chain is in a dilute solution. If the sequence separation n is approximated as a continuous variable, then the probability density satisfies the diffusion equation

$$\frac{\partial p_s(\mathbf{r}, n)}{\partial n} = \frac{b^2}{6} \nabla^2 p_s(\mathbf{r}, n) \tag{2}$$

with n as the analog of time and $b^2/6$ as an apparent diffusion constant. Eq. (2) shows that the conformations of the polymer chain correspond to Brownian trajectories launched from a fixed starting position. Suppose that the polymer chain has $N+1$ residues, corresponding to $n = 0$ to $n = N$. The root mean square end-to-end distance of the chain is

$$\bar{R}_s = bN^{1/2} \tag{3}$$

The walls of the confining cage serve as absorbing boundaries since any conformation of the polymer chain that crosses them is eliminated. Let $p(\mathbf{r},n;\mathbf{r}_0)$ be the probability density of residue n in conformations that started from a fixed position $\mathbf{r}_0$ and have not yet crossed the walls of the confining cage. It satisfies Eq. (2) with a zero value on the walls of the cage and an "initial" value $p(\mathbf{r}, 0; \mathbf{r}_0) = \delta(\mathbf{r} - \mathbf{r}_0)$. The "survival" fraction of all conformations started from $\mathbf{r}_0$ is

$$S(\mathbf{r}_0) = \int \mathrm{d}^3\mathbf{r} p(\mathbf{r}, N; \mathbf{r}_0) \tag{4}$$

where the volume integration is restricted to within the cage. Averaging $S(\mathbf{r}_0)$ over starting positions uniformly distributed in the confining cage then gives the factor f_U by which the statistical weight of the unfolded protein is reduced by the confinement. For a spherical cavity, one finds [13]

$$f_U = \frac{6}{\pi^2} \sum_{k=1}^{\infty} \frac{\exp(-\pi^2 y_c^2 k^2)}{k^2} \tag{5}$$

where

$$y_c = \bar{R}_s / 6^{1/2} R_c \tag{6}$$

The first term of the summation dominates except when $y_c \ll 1$. Hence

$$\ln f_U \approx -\pi^2 y_c^2 - \ln(\pi^2/6) \quad (7)$$

The negative $\ln f_U$ corresponds to an entropy reduction for the unfolded protein; confinement thus can be viewed as destabilizing the unfolded state entropically.

The polymer chain considered so far does not have internal excluded-volume interactions. Realistic modeling of unfolded proteins requires that self-avoiding be taken into consideration [14]. Self-avoiding will swell the polymer chain, and the root mean square of the end-to-end distance scales with the number of residues like

$$\bar{R}_s = b' N^{\nu} \quad (8)$$

The exponent ν is close to 0.6 according to different theories [15]. For unfolded proteins, it was found that $\bar{R}_s = 5.31N^{0.6}$ [14]. The swelling increases the probability that the chain crosses the walls of the confining cage. As a result, f_U will be further reduced. Arguments based on scaling lead to the following result for a self-avoiding polymer confined in a spherical cage [16,17]:

$$\ln f_U \approx -B y_c^{\gamma} \quad (9)$$

where the exponent, $\gamma = 3/(3\nu - 1)$, is 3.75 for $\nu = 0.6$, and B is a constant. [Note that simply setting ν to 1/2 in the expression for γ does not recover the exponent of 2 of Eq. (7).] Monte Carlo simulations [18] show that Eq. (9) works for a self-avoiding chain with 256 residues inside a confining sphere, with $\gamma = 3.61$ and $B = 7.36$, for y_c between 1.3 and 3.

The effect of the confinement on the folding free energy is given by

$$\Delta\Delta G_F = -k_B T \ln(f_F/f_U) \quad (10)$$

where $k_B T$ is thermal energy. Fig. 2 displays $\Delta\Delta G_F$ for proteins with 101 and 201 residues confined in a spherical cage. When the folded protein is modeled as a sphere, $\Delta\Delta G_F$ can be over $20k_B T$ when R_c, the radius of the cage, decreases toward the size of the folded protein. For such small confining cages, modeling the unfolded protein as either Gaussian or self-avoiding chain leads to a significant difference in $\Delta\Delta G_F$. However, as R_c increases, the two models appear to merge. Thus the Gaussian-chain model will be useful when R_c is beyond twice the size of the folded protein. Its predictions at large R_c have indeed been found to agree well with results obtained from molecular dynamics simulations of protein folding inside a cage [19].

As noted earlier, a folded protein with a non-spherical shape will have a smaller f_F, leading to less stabilization by confinement. Fig. 2 also shows the $\Delta\Delta G_F$ result when the folded protein is modeled as a prolate ellipsoid with an axis ratio of 2. Relative to the situation when the folded protein is modeled as a sphere, $\Delta\Delta G_F$ now has a smaller magnitude and starts to turn over when the size of the cage shrinks down to the length of the longer semi-axis. A similar turnover also occurs if the folded protein is modeled as a sphere, when R_c approaches its radius. That turnover requires a smaller confining cage since the radius of a sphere is smaller than the longer semi-axis of a corresponding ellipsoid with the same volume. Ziv et al. [20] has used a theory much like what is outlined here to rationalize the stabilization of helix formation by confinement in a cylinder mimicking the ribosomal exit tunnel.

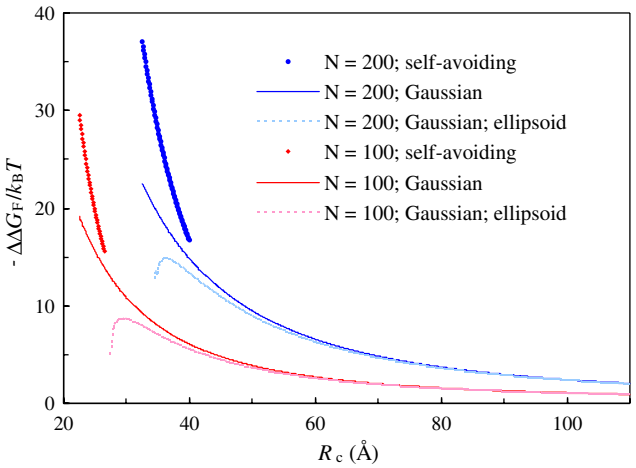

Fig. 2. Effect of confinement on the folding free energy. A protein with either 101 or 201 residues is confined in a sphere with radius R_c. The unfolded protein is modeled as either a Gaussian chain or a self-avoiding chain. The folded protein is modeled as a sphere (with radius a_F) or as a prolate ellipsoid with an axis ratio of 2. The root mean square end-to-end distance of the unfolded protein is given by $5.31N^{0.6}$ Å [14]. In the case of a spherical model, the radius of the folded protein is given by $3.73N^{1/3}$ Å [58]; the ellipsoidal model has the same volume as the corresponding spherical model. For the ellipsoidal model, the reduction factor f_F is found by uniformly placing the ellipsoid center within the confining sphere, with the long axis uniformly oriented, and counting the fraction of placements that are fully enclosed in the confining cage.

Crowding and folding free energy

Macromolecular crowders also serve to eliminate chain conformations that cross the formers' boundaries. Continuing the analogy between Gaussian chain and Brownian trajectory, the fraction of surviving trajectories is equivalent to the survival probability of a point ligand in the sea of the crowders as static traps [11]. If the macromolecular crowders are spheres with radius a_m and are present at concentration c, the survival probability, to lowest order in c, is the Smoluchowski result [21–23]

$$S_0(N) = \exp\{-4\pi(Nb^2/6)a_m c[1 + 2a_m(\pi N b^2/6)^{-1/2}]\} \quad (11)$$

where replacements of time by N and diffusion constant by $b^2/6$ have taken place. The crowders also restrict positions where the Brownian trajectories can be launched to regions outside the excluded volumes. If the volume fraction, $4\pi a_m^3 c/3$, is denoted as ϕ, this restriction contributes a reduction factor $1 - \phi$. Together, the factor f_U by which the statistical weight of the unfolded protein is reduced by the crowders is

$$f_U = S_0(N)(1-\phi)$$
$$= \exp\{-3\phi(\bar{R}_s^2/6a_m^2)[1+2(\pi\bar{R}_s^2/6a_m^2)^{-1/2}]\}(1-\phi)$$

Thus

$$-\ln f_U = 3\phi y_m^2 + 6\pi^{-1/2}\phi y_m - \ln(1-\phi) \quad (12)$$

where

$$y_m = \bar{R}_s/6^{1/2}a_m \quad (13)$$

Eq. (12) is the same as a result obtained by Eisenriegler et al. [24] for the reduction factor for inserting a sphere in a sea of Gaussian-chain polymers when $-\ln(1-\phi)$ is approximated as ϕ. As already noted, inserting a Gaussian-chain polymer into a solution of spherical obstacles is equivalent to the problem of a single point Brownian particle in the presence of a concentration of static traps. By the same token, inserting a sphere into a solution of Gaussian-chain polymers is equivalent to the problem of a single static trap in the presence of a concentration of Brownian particles. Eq. (12) gives the exact result for the latter pair of problems, but is valid for the former pair of problems only to the lowest order in concentration and overestimates the survival probability at higher concentrations [22]. An approximate correction, applicable when $y_m > 1$, is [23]

$$S_1(N) = S_0(N)\exp(9\phi_m^2 y_m^2 \ln y_m) \quad (14)$$

This leads to

$$-\ln f_U = 3\phi y_m^2 + 6\pi^{-1/2}\phi y_m^{1/2} - \ln(1-\phi) - 9\phi^2 y_m^2 \ln y_m \quad (15)$$

The crowders also reduce the statistical weight of the folded protein by eliminating positions at which the protein can be placed. If the folded protein is modeled as a sphere with radius a_F, the reduction factor f_F can be obtained from scaled-particle theory [25], which gives

$$-\ln f_F = (3z+3z^2+z^3)\rho + (9z^2/2+3z^3)\rho^2 + 3z^3\rho^3 - \ln(1-\phi) \quad (16)$$

where $z = a_F/a_m$ and $\rho = \phi/(1-\phi)$.

The effect of macromolecular crowding on the folding free energy is again given by Eq. (10). Fig. 3 displays $\Delta\Delta G_F$ for placing proteins with 101 and 201 residues in solutions of crowders with radii at 15 and 30 Å and at concentrations up to 300 g/L. In contrast to the situation in confinement, the effect of crowding is found to be quite small, reaching maximal stabilization of only $0.5k_BT$ to $1.6k_BT$. An alternative treatment of the unfolded protein leads to significantly stronger stabilization effect [26].

For long chain lengths (i.e., $y_m \gg 1$), the difference in f_U between a Gaussian chain and a self-avoiding polymer chain becomes significant. The leading term in Eq. (12), $3\phi y_m^2$, is replaced by [27,28]

$$A\phi y_m^{1/\nu}$$

for a self-avoiding chain. The self-avoiding chain has an exponent ~1.7, lower than the corresponding value of 2 for the Gaussian chain. Thus the self-avoiding chain will have a smaller reduction factor; consequently the small stabilizing effect from crowding will be further reduced. The reason for the lower reduction factor of the self-avoiding chain relative to the counterpart for the Gaussian chain is that the former is more open, leaving interior spaces for the crowders to occupy. A related argument can also be made to explain the modest stabilization effect of crowding [11]. While the unfolded chain may be able to explore all the interstitial voids between the crowders, the folded protein can only be accommodated in the fraction of voids with sizes greater than itself (see Fig. 1b). That penalty against the folded protein nearly balances the penalty experienced by the unfolded protein when chain conformations are eliminated by the crowders, leading to a modest overall effect. At higher concentrations of the crowders, the former penalty increases more rapidly than the latter; as a result, $\Delta\Delta G_F$ turns over and eventually becomes positive, corresponding to a destabilizing effect.

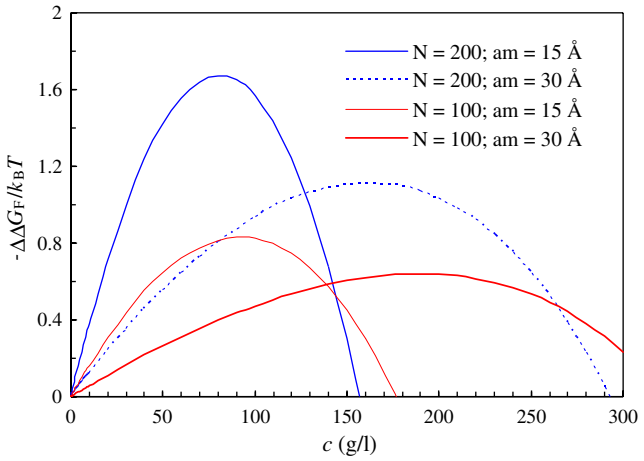

Fig. 3. Effect of crowding on the folding free energy. The protein, with 101 or 201 residues, is modeled as a Gaussian chain in the unfolded state and as a sphere in the folded state. The radius, a_m, of the crowders is shown. For calculating the volume fraction, the molecular weight of the crowder is set to $2.125a_m^3$ (with a_m in units of Å).

Speed up of contact formation and protein folding by confinement and crowding

The probability density for the end-to-end distance of a polymer chain is maximal around the root mean square $\bar{R}_s$. Contact formation entails bringing the ends to a separation much less than $\bar{R}_s$. The much lower probability at such a separation amounts to an entropy barrier. For a Gaussian chain, the diffusion-controlled rate of contact formation at a separation σ is [29]

$$k_f = 3(6/\pi)^{1/2}D\sigma/\bar{R}_s^3 \quad (17)$$

where D is the relative diffusion constant of the chain ends.

Confinement and crowding will eliminate chain conformations that cross surrounding boundaries. The chance of elimination is higher for the more open conformations.

Therefore the surviving conformations will be more compact, with a root mean square end-to-end distance smaller than $\bar{R}_s$. Eq. (17) will then predict a higher contact formation rate. However, confinement and crowding will also likely lead to a reduction in the diffusion constant [11]. Hence the increase in k_f due to chain compaction will be moderated (and possibly even reversed) by the accompanying decrease due to reduced diffusivity. This qualitative analysis is borne out by simulations of polymer chain dynamics under crowded conditions [30].

The folding rates of proteins correlate well with the contact order, which is a measure of chain compactness [31]. To the extent that confinement and crowding favor conformations that are more compact, they are expected to increase the folding rate.

Experimental studies in the test tube

Enhancement of folding stability by encapsulation

The interest to mimic the effects of cellular encapsulation, coupled with theoretical predictions of significant stabilization, has spurred experimental studies of protein folding stability in artificially confined environments. Two kinds of encapsulation are now widely used. The first is formed by nanoporous silica gels or glasses [32–35] or polyacrylamide gels [36]; the second is formed by sodium bis(2-ethylhexyl) sulfosuccinate (AOT) reverse micelles [37–41]. In both cases, the cage sizes can be controlled. The inner diameters of AOT reverse micelles are easily varied by changing W_0, the molar ratio of water to AOT, also known as water loading. With isooctane as cosolvent, the cage size follows a linear relation with water loading: $R_c = 1.5W_0 + 4.5$ Å [42].

These studies have confirmed that significant stabilization can arise from protein encapsulation. For example, Ravindra et al. [32] found that, upon being confined within the pores, with sizes averaging ~25 Å, of a silica glass, the melting temperature of ribonuclease A is raised by ~30 °C. The measured increase in melting temperature for the 124-residue protein is in quantitative agreement with predicted stabilization by pores of such sizes (see Fig. 2). Mukherjee et al. [41] observed increased helix formation of alanine-rich peptides in AOT reverse micelles, but attributed the increased stability to a different reason (see below).

Perhaps the most dramatic demonstration of the stabilizing effect of confinement is provided by the study of Peterson et al. [38]. These authors extensively mutated a designed three helix-bundle protein, α_3W, so that it becomes unfolded in a dilute solution, as indicated by an unresolved ^{15}N heteronuclear single quantum correlation (HSQC)[1] spectrum. When the α_3W variant is encapsulated in AOT reverse micelles, the ^{15}N HSQC peaks are sharpened. For a low water loading, corresponding to a small cage, the ^{15}N HSQC spectrum resembles that of the original α_3W, indicating the formation of a three-helix bundle structure.

Effect of macromolecular crowding on folding stability

In contrast to the significant stabilization effects seen for confinement, experimental studies on macromolecular crowding from different laboratories have consistently found a modest effect on protein stability [43–48]. The theoretical prediction of modest stabilizing effect by crowding, presented earlier, thus appears to have experimental support. The disparate effects of confinement and crowding on protein folding stability emerging from the earlier literature were already noted [6,11,12].

In an earlier study, Qu and Bolen [43] attempted to stabilize, by adding the crowding agent dextran 70, a purposely destabilized variant of ribonuclease T1. The design of this experiment is much like the one used by Peterson et al. [38] for confinement. However, even with 400 g/L of dextran 70, Qu and Bolen found only a small stabilizing effect $\sim 2k_BT$. Tokuriki et al. [44] observed that the CD spectrum of ribonuclease A at 2.4 M urea reverts to that at 0.2 M urea when 350 g/L PEG 20000 or Ficoll 70 is added. Unfortunately a full urea denaturation curve was not reported; hence a quantitative assessment of the stabilizing effect is not possible. Spencer et al. [45] also used Ficoll 70 as a crowding agent and, through urea denaturation, found that it has a stabilizing effect of $\sim k_BT$ on the FK506-binding protein. McPhie et al. [46] studied the effect of dextran 30 on the heat and cold denaturation of the molten globule form of apomyoglobin; a stabilization effect of $\sim 0.5k_BT$ was observed at 270 g/L of the crowding agent. Based on NMR relaxation dispersion data obtained under native conditions, Ai et al. [47] found that the folding stability of a redesigned four helix-bundle protein is increased by $\sim 04k_BT$ in the presence of 85 g/L PEG 20000. Ignatova et al. [48] even reported preliminary data indicating a small destabilizing effect by Ficoll 70 on cellular retinoic acid-binding protein I.

Contact formation and protein folding rates under macromolecular crowding

Neuweiler et al. [49] measured the rate of end-to-end contact formation of peptides in the presence of Ficoll 70 by fluorescence correlation spectroscopy. At 300 g/L Ficoll 70, they found that the diffusion constants of the peptides are reduced by a factor of 7. A slowing down in contact formation was observed in the presence of 300 g/L Ficoll 70. However, the reduction in the contact formation rate is less than the reduction in diffusivity; the former ranges from 5 for a 10-residue peptide to 3 for a 30-residue peptide. The discrepancy was attributed to chain compaction by the crowding agent.

The experimental results of Ai et al. [47], referred to earlier, for the effect of crowding on folding stability were derived from measurements on the folding and unfolding

[1] *Abbreviation used:* HSQC, heteronuclear single quantum correlation.

rates. The small stabilizing effect of crowding arose mostly from a speeding up of the folding process, in agreement with theoretical predictions under the assumption of compaction of the unfolded protein by crowding. Ai et al. found that the ^{15}N chemical shift differences between the folded and unfolded states in the presence of the crowding agent show a strong correlation with those in the absence of the crowding agent. The former are 5-10% lower than the latter, perhaps a sign that the unfolded protein in the presence of crowding is slightly more compact than the counterpart in the absence of crowding.

Beyond volume exclusion

So far the focus has been on the excluded volumes presented by confining boundaries or by crowders. In reality, the surrounding surfaces will also interact with the proteins under study. Depending on whether the interactions are more favorable in the folded state or in the unfolded state, the folding equilibrium can be further shifted to the former state or latter state. The wall of the Anfinsen cage of chaperonins can even change from nonpolar, favoring the unfolded protein when it is initially encapsulated, to polar, favoring the folded protein when it transitions toward the folded conformation. Thus the wall here plays an active role during the folding cycle (also see below). In the case of confinement in reverse micelles, the role of surface interaction can be investigated by modifying the headgroups of the surfactant [39].

Mukherjee et al. [41] attributed the increased helix formation of alanine-rich peptides in AOT reverse micelles to decreased hydration of backbone amide and carbonyl groups. The decreased hydration is plausible perhaps because water molecules interact more strongly with the AOT headgroups. It can lead to helix stabilization because backbone hydrogen bonding will then be favored. Regardless of whether dehydration is the dominant stabilization mechanism, it certainly is an idea worth further exploring. For example, one may introduce ion pairs at i and $i+4$ positions. The dehydration hypothesis would lead to the easily testable consequence that the stabilization arising from ion-pair formation in the reverse micelle will be stronger than found in a dilute solution.

In recent molecular dynamics simulations, Pande and co-workers have treated the solvent explicitly [19,50]. Contrary to previous studies, their simulations found that helix formation and protein folding are disfavored in confined cages relative to in dilute solutions. To rationalize the decreased helix stability of an alanine peptide inside a nanotube observed in their simulations, Sorin and Pande [50] argued that water released upon peptide backbone–backbone hydrogen bonding has less translational entropy inside the nanotube than in a dilute solution and the reduced gain in translational entropy upon helix formation is the cause for the destabilization. However, their theoretical analysis along this line of argument contain several crucial sign errors. The simulation results of Sorin and Pande have been explained based on the idea that the inner surface of the nanotube, by its nonpolar nature, increases the activity of the confined water [51]. The increased water activity favors the coil state since the backbone amides and carbonyls are then free to make hydrogen bonds with the water molecules. The implementation of this idea within the Lifson–Roig theory [52] for helix-coil transition was able to quantitatively reproduce the decrease in helical content with decreasing nanotube diameter observed in Sorin and Pande's simulations. The idea of an increase in water activity by the nonpolar inner surface of the nanotube can be viewed as the flip side of Mukherjee et al.'s [41] proposal of decreased hydration of backbone amide and carbonyl groups inside an AOT reverse micelle, which features charged headgroups on the inner surface.

Protein folding in cellular environments

The exit tunnel of the ribosome is ~100 Å long with a diameter ~10 Å in the narrowest central part and ~20 Å toward the ends. Lu and Deutsch [53] found that helix formation is promoted at the wider portions of the exit tunnel. Experimental and theoretical studies aimed at untangling possible contributions from entropy reduction by confinement, surface interaction, and modulation of water activity will shed light on the functional roles of the ribosomal exit tunnel in the folding of nascent proteins.

Tang et al. [54] designed mutations of the *Escherichia coli* chaperonin GroEL to test the theoretical prediction that the Anfinsen cage facilitates protein folding by favoring compact intermediates. They reduced or enlarged the cage size by adding or removing sequences to the C-terminals of the GroEL subunits, which protrude into the central cavity. Tang et al. found that the folding rate can be modulated by the cage size. There is an optimal cage size for each substrate protein, and the optimum shifts to larger sizes for larger proteins (cf. Fig. 2). These results thus "are remarkably consistent with prediction." Indeed, Hayer-Hartl and Minton [55] showed that the results can be quantitatively rationalized by a theoretical model similar to one used to generate the plots shown in Fig. 2, with the transition state (instead of the folded state) modeled as a sphere. However, Tang et al. also presented data demonstrating that more than just entropy reduction from confinement is at play in GroEL/ES-assisted folding. The C-terminal sequences and a number of conserved negative charges lining the cavity wall are critical in facilitating the folding of some proteins. Importantly, Tang et al. found that the ability of overexpressed GroEL/ES mutants to suppress aggregation of overexpressed substrate proteins in *E. coli* correlates with the folding rates of the substrate proteins within the respective GroEL/ES mutants.

α-Synuclein is a naturally disordered protein, which, in fibrillar form, is the primary component of Lewy bodies found in Parkinson's disease patients. Overexpressed α-synuclein is found in the periplasm of *E. coli* [56].

Using in-cell NMR spectroscopy, McNulty et al. [57] investigated the conformations of α-synuclein in the periplasm of *E. coli*. They found that α-synuclein adopts a more compact form than is observed in a dilute solution. Interestingly, the more compact form can be induced by adding 300 g/L of bovine serum albumin as a crowding agent, in line with theoretical predictions for conformational compaction by crowding.

These studies demonstrate that, in concert with theoretical modeling and parallel experiments in simulated conditions, much of the complexity of protein folding in cellular environments can be elucidated.

Acknowledgments

I thank Robert L. Baldwin, Yawen Bai, and Feng Gai for reading the manuscript. This work was supported in part from Grant GM058187 from the National Institutes of Health.

References

[1] B.J. Marsh, D.N. Mastronarde, K.F. Buttle, K.E. Howell, J.R. McIntosh, Proc. Natl. Acad. Sci. USA 98 (2001) 2399–2406.
[2] A.P. Minton, Curr. Biol. 16 (2006) R269–R271.
[3] R.J. Ellis, A.P. Minton, Biol. Chem. 387 (2006) 485–497.
[4] A.P. Minton, J. Cell Sci. 119 (2006) 2863–2869.
[5] R.J. Ellis, Adv. Exp. Med. Biol. 594 (2007) 1–13.
[6] H.-X. Zhou, Acc. Chem. Res. 37 (2004) 123–130.
[7] S.J. Hagen, J. Hofrichter, A. Szabo, W.A. Eaton, Proc. Natl. Acad. Sci. USA 93 (1996) 11615–11617.
[8] D.E. Makarov, C.A. Keller, K.W. Plaxco, H. Metiu, Proc. Natl. Acad. Sci. USA 99 (2002) 3535–3539.
[9] H.-X. Zhou, J. Chem. Phys. 118 (2003) 2010–2015.
[10] T.R. Weikl, M. Palassini, K.A. Dill, Protein Sci. 13 (2004) 822–829.
[11] H.-X. Zhou, J. Mol. Recog. 17 (2004) 368–375.
[12] H.-X. Zhou, Biochemistry 43 (2004) 2141–2154.
[13] H.-X. Zhou, K.A. Dill, Biochemistry 40 (2001) 11289–11293.
[14] H.-X. Zhou, J. Phys. Chem. B 106 (2002) 5769–5775.
[15] M. Doi, S.F. Edwards, The Theory of Polymer Dynamics, Crarendon Press, Oxford, 1986.
[16] A.Y. Grosberg, A.R. Khokholov, Statistical Physics of Macromolecules, American Institute of Physics, New York, 1994.
[17] T. Sakaue, E. Raphael, Macromolecules 39 (2006) 2621–2628.
[18] A. Cacciuto, E. Luijten, Nano Lett. 6 (2006) 901–905.
[19] D. Lucent, V. Vishal, V.S. Pande, Proc. Natl. Acad. Sci. USA 104 (2007) 10430–10434.
[20] G. Ziv, G. Haran, D. Thirumalai, Proc. Natl. Acad. Sci. USA 102 (2005) 16961–18956.
[21] M.V. Smoluchowski, Z. Phys. Chem. 92 (1917) 129–168.
[22] A. Szabo, R. Zwanzig, N. Agmon, Phys. Rev. Lett. 61 (1988) 2496–2499.
[23] A.M. Berezhkovskii, Y.A. Makhnovskii, G.H. Weiss, J. Chem. Phys. 115 (2001) 5376–5380.
[24] E. Eisenriegler, A. Hanke, S. Dietrich, Phys. Rev. E 54 (1996) 1134–1152.
[25] J.L. Lebowitz, J.S. Rowlinson, J. Chem. Phys. 41 (1964) 133–138.
[26] A.P. Minton, Biophys. J. 88 (2005) 971–985.
[27] P.G. de Gennes, C.R. Acad. Sci. B 288 (1979) 359–361.
[28] A. Hanke, E. Eisenriegler, S. Dietrich, Phys. Rev. E 59 (1999) 6853–6878.
[29] A. Szabo, K. Schulten, Z. Schulten, J. Chem. Phys. 72 (1980) 4350–4357.
[30] N.M. Toan, D. Marenduzzo, P.R. Cook, C. Micheletti, Phys. Rev. Lett. 97 (2006) 178302.
[31] K.W. Plaxco, K.T. Simons, D. Baker, J. Mol. Biol. 277 (1998) 985–994.
[32] R. Ravindra, S. Zhao, H. Gies, R. Winter, J. Am. Chem. Soc. 126 (2004) 12224–12225.
[33] B. Campanini, S. Bologna, F. Cannone, G. Chirico, A. Mozzarelli, S. Bettati, Protein Sci. 14 (2005) 1125–1133.
[34] K.E. Wheeler, J.M. Nocek, B.M. Hoffman, J. Am. Chem. Soc. 128 (2006) 14782–14783.
[35] S. Zhao, H. Gies, R. Winter, Z. Phys. Chem. Int. J. Res. Phys. Chem. Chem. Phys. 221 (2007) 139–154.
[36] D. Bolis, A.S. Politou, G. Kelly, A. Pastore, P. Andrea Temussi, J. Mol. Biol. 336 (2004) 203–212.
[37] C.R. Babu, V.J. Hilser, A.J. Wand, Nat. Struct. Mol. Biol. 11 (2004) 352–357.
[38] R.W. Peterson, K. Anbalagan, C. Tommos, A.J. Wand, J. Am. Chem. Soc. 126 (2004) 9498–9499.
[39] Z.S. Shi, R.W. Peterson, A.J. Wand, Langmuir 21 (2005) 10632–10637.
[40] A.K. Simorellis, P.F. Flynn, J. Am. Chem. Soc. 128 (2006) 9580–9581.
[41] S. Mukherjee, P. Chowdhury, F. Gai, J. Phys. Chem. B 110 (2006) 11615–11619.
[42] A. Maitra, J. Phys. Chem. 88 (1984) 5122–5125.
[43] Y. Qu, D.W. Bolen, Biophys. Chem. 101–102 (2002) 155–165.
[44] N. Tokuriki, M. Kinjo, S. Negi, M. Hoshino, Y. Goto, I. Urabe, T. Yomo, Protein Sci. 13 (2004) 125–133.
[45] D.S. Spencer, K. Xu, T.M. Logan, H.-X. Zhou, J. Mol. Biol. 351 (2005) 219–232.
[46] P. McPhie, Y.S. Ni, A.P. Minton, J. Mol. Biol. 361 (2006) 7–10.
[47] X. Ai, Z. Zhou, Y. Bai, W.Y. Choy, J. Am. Chem. Soc. 128 (2006) 3916–3917.
[48] Z. Ignatova, B. Krishnan, J.P. Bombardier, A.M. Marcelino, J. Hong, L.M. Gierasch, Biopolymers 88 (2007) 157–163.
[49] H. Neuweiler, M. Lollmann, S. Doose, M. Sauer, J. Mol. Biol. 365 (2007) 856–869.
[50] E.J. Sorin, V.S. Pande, J. Am. Chem. Soc. 128 (2006) 6316–6317.
[51] H.-X. Zhou, J. Chem. Phys. (in press).
[52] S. Lifson, A. Roig, J. Chem. Phys. 34 (1961) 1963–1974.
[53] J.L. Lu, C. Deutsch, Nat. Struct. Mol. Biol. 12 (2005) 1123–1129.
[54] Y.C. Tang, H.C. Chang, A. Roeben, D. Wischnewski, N. Wischnewski, M.J. Kerner, F.U. Hartl, M. Hayer-Hartl, Cell 125 (2006) 903–914.
[55] M. Hayer-Hartl, A.P. Minton, Biochemistry 45 (2006) 13356–13360.
[56] C. Huang, G. Ren, H. Zhou, C.-c. Wang, Protein Expr. Purif. 42 (2005) 173–177.
[57] B.C. McNulty, G.B. Young, G.J. Pielak, J. Mol. Biol. 355 (2006) 893–897.
[58] M. Vijayakumar, H.-X. Zhou, J. Phys. Chem. B 104 (2000) 9755–9764.