JOURNAL OF MOLECULAR RECOGNITION
J. Mol. Recognit. 2004; **17**: 368–375
DOI:10.1002/jmr.711

Protein folding and binding in confined spaces and in crowded solutions†

Huan-Xiang Zhou*

Department of Physics and Institute of Molecular Biophysics, Florida State University, Tallahassee, FL 32306, USA

Simple theoretical models are presented to illustrate the effects of spatial confinement and macromolecular crowding on the equilibria and rates of protein folding and binding. Confinement is expected to significantly stabilize the folded state, but for crowding only a marginal effect on protein stability is expected. In confinement the unfolded chain is restricted to a cage but in crowding the unfolded chain may explore different interstitial voids. Because confinement and crowding eliminate the more expanded conformations of the unfolded state, folding from the compact unfolded state is expected to speed up. Crowding will shift the binding equilibrium of proteins toward the bound state. The significant slowing down in protein diffusion by crowding, perhaps beneficial for chaperonin action, could result in a decrease in protein binding rates.

Keywords: **spatial confinement; macromolecular crowding; excluded-volume effect; protein folding; protein binding; protein aggregation; chaperonin action**

Received 3 November 2003; revised 6 March 2004; accepted 6 May 2004

INTRODUCTION

The crowded environment inside the cytoplasm is very different from the dilute solutions typically used in *in vitro* studies of proteins and may significantly affect the behaviors of protein molecules (Minton, 2000a; Ellis, 2001). In early *in vitro* studies of DNA replication (Fuller *et al.*, 1981; Zimmerman, 1993), the addition of macromolecular crowding agents was found to be a requirement for replication. Recently, macromolecular crowding was implicated in promoting self-association of FtsZ and accelerating the rate of amyloid formation (Rivas *et al.*, 2000; Hatters *et al.*, 2002). Here simple theories are presented to provide qualitative estimates for the magnitudes of the effects of macromolecular crowding and spatial confinement on both the equilibria and the rates of protein folding and protein–protein binding.

The theoretical models presented here are intentionally simple. They are devised to capture the essence of the process (folding or binding) under study and the essential effects of crowding or confinement, but lack realistic detail. The simplicity of the models hopefully will give rise to an intuitive picture for the effect of crowding and confinement. Whenever appropriate, connections will be made between the theoretical models and experimental data; however, the simplicity of the models should be kept in mind.

*Correspondence to: Huan-Xiang Zhou, Department of Physics and Institute of Molecular Biophysics, Florida State University, Tallahassee, FL 32306, USA.
E-mail: zhou@sb.fsu.edu

†This paper is published as part of a special issue entitled 'EMBO Workshop on Biological Implications of Macromolecular Crowding held on 14–18 June 2003, Las Navas del Marqués, Avila, Spain'.

Contract/grant sponsor: NIH; *contract/grant number:* GM58187.

Abbreviations used: BSA, bovine serum albumin.

Spatial confinement is related to but distinct from macromolecular crowding (Fig. 1). Both will reduce volumes accessible to the protein molecule under study. However, in the former case the accessible volume is confined within a single 'cage', whereas in the latter case the accessible volumes are dispersed throughout space. This distinction brings out interesting differences in the effects on the equilibrium and rate of protein folding.

MODELS OF FOLDING AND BINDING

The simplest model for a folded protein is a sphere, and the simplest model for an unfolded protein is a Gaussian chain. The folded state will be assumed to be separated from the unfolded state by a free-energy difference ΔG. The effects of confinement and crowding on the folding equilibrium will be measured by the change in ΔG, $\Delta\Delta G$.

Folding rate

The process of protein folding may be viewed as the accumulation of native contacts. Here I will use the formation of a single native contact (Zhou, 2003) to illustrate the folding process. The treatment can be extended to the formation of additional native contacts, which lead to the folded state (Makarov *et al.*, 2002; Makarov and Metiu, 2002).

Consider the contact formation between two residues from the unfolded chain (see Fig. 2). The unfolded protein is modeled as a Gaussian chain, so the probability density for finding the two residues at a distance r is

$$P(r) = \left(3/2\pi < r^2 >\right)^{3/2} \exp\left(-3r^2/2 < r^2 >\right) \qquad (1a)$$

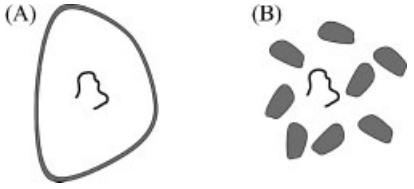

Figure 1. The contrast between (A) confinement and (B) crowding. In confinement a protein chain is restricted to a single cage, but in crowding the protein's accessible volumes are dispersed throughout space.

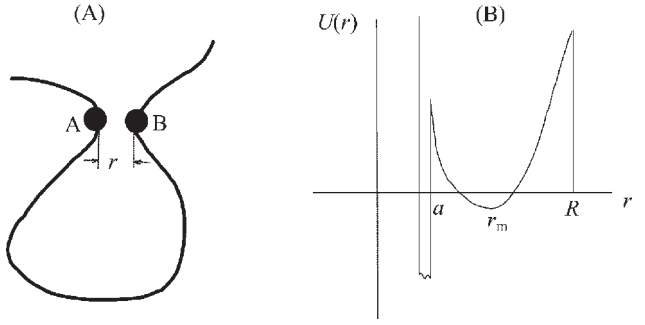

Figure 2. (A) The formation of a contact between two residues (represented as two small spheres) along a protein chain. (B) The equivalent potential $U(r)$ giving rise to the equilibrium distribution $P(r)$. The labeled distances are: a, the contact distance (at which the contact starts to be stabilized by short-range interactions); r_m, distance at which $4\pi r^2 U(r)$ is minimal; and R, the maximum distance between the two residues allowed by the protein chain.

where $<r^2>$ is the mean square distance. $<r^2>$ has a linear dependence on the number of peptide bonds separating the residues,

$$<r^2> = nb^2 \tag{1b}$$

where b is an effective bond length. An equivalent potential giving rise to the same equilibrium distribution as $P(r)$ is given by

$$\exp[-\beta U(r)] = P(r) \tag{2}$$

where $\beta = (k_B T)^{-1}$.

Let us first assume that every encounter between the residues leads to productive contact formation. The problem of contact formation is then identical to the problem of a Brownian particle moving in an potential $U(r)$ and being absorbed at the contact distance $r = a$, provided the relative motion of the residues can be described by diffusion. According to Kramers (1940), the rate constant is

$$k_{f0} = D \Big/ \int_a^R \exp(-\beta U) 4\pi r^2 \, dr \int_a^{r_m} \exp(\beta U) \left(4\pi r^2\right)^{-1} dr$$

where D is the relative diffusion constant, R is the maximum inter-residue distance, and r_m is the distance at which $4\pi r^2 U(r)$ is minimal. Specialization to a Gaussian chain leads to (Szabo *et al.*, 1980)

$$k_{f0} = 3(6/\pi)^{1/2} Da / <r^2>^{3/2} \tag{3}$$

The formation of a native contact requires that the two residues are in their correct local conformations. Assuming that the transitions into and out of the correct local conformation of each residue are rate processes and the transitions of the two residues are independent, then the rate constant of native contact formation is smaller than k_f^0 by

$$\frac{k_{f0}}{k_f} = \frac{\omega_{A-}\omega_{B-}}{\omega_{A+}\omega_{B+}\left[aD^{-1/2}(\omega_{A+} + \omega_{A-} + \omega_{B+} + \omega_{B-})^{1/2} + 1\right]} \tag{4}$$

where ω_+ and ω_- are the transition rates into and out of the correct local conformation of a residue (A or B).

Binding equilibrium constant

Now consider the binding of two rigid proteins modeled as spheres [Fig. 3(A)]. If the complex of the two proteins is held together by a strong short-range inter-protein potential $U(r)$, then the binding constant is given by (Shoup and Szabo, 1982)

$$K_s = \int_a^{r^*} \exp[-\beta U(r)] 4\pi r^2 dr \tag{5a}$$

where a is the contact distance and r^* is the upper boundary defining the bound state. The subscript 's' signifies that the result holds for the spherical geometry.

A protein complex is stereospecific [Fig. 3(B)]. That is, the complex is considered formed only if strict constraints on the relative translation and rotation of the two proteins are fulfilled. Let the relative displacement vector be $\mathbf{r}$ and the rotational angles of the two proteins be collectively denoted as $\mathbf{\Omega}_A$ and $\mathbf{\Omega}_B$, then the binding constant is

$$K = \int_\Gamma \exp[-\beta U(\mathbf{r}, \mathbf{\Omega}_A, \mathbf{\Omega}_B)] \, d^3\mathbf{r} \, d^3\mathbf{\Omega}_A d^3\mathbf{\Omega}_B / \left(8\pi^2\right)^2 \tag{5b}$$

where Γ represents the configurational space of the bound state.

Binding rate

The classical result of Smoluchowski (1917), $k_s = 4\pi Da$, gives the rate constant for the instantaneous coagulation of two spherical particles at a contact distance a. The particles freely diffuse with a relative diffusion constant D (without the influence of a long-range potential). For future reference, it is noted that the above Smoluchowski result is for

Figure 3. Models for the bound state of two binding proteins. (A) Spherical model. (B) Stereospecific model.

the steady state, when the supply of reactant pairs by diffusion is exactly depleted by coagulation. Suppose initially there is an equilibrium distribution (which is uniform in the present case). Then at time t the reactive flux, which becomes the Smoluchowski rate constant at steady state, is

$$k_s(t) = 4\pi Da\left[1 + a/(\pi Dt)^{1/2}\right] \quad (6)$$

Under the influence of a potential $U(r)$, Debye (1942) derived the following result for the steady-state binding rate constant:

$$k_s = D\Big/\int_a^\infty \exp[\beta U(r)]\left(4\pi r^2\right)^{-1} \mathrm{d}r$$

In general, predicting the effect of a potential on the rate constant for stereospecific binding is difficult. However, when the range of the potential is long relative to the size of the configurational space of the bound state, a simple approximate formula has been obtained for the diffusion-limited binding rate constant (Zhou *et al.*, 1997; Zhou, 1997):

$$k = k_0 < \exp(-\beta U) >^* \quad (7)$$

where k_0 is the rate constant in the absence of the potential and $< \cdots >^*$ signifies averaging over the outer boundary of the bound state.

SHIFT OF FOLDING EQUILIBRIUM BY CONFINEMENT AND CROWDING

Confinement and crowding restrict the motional freedom of a protein molecule. If the unfolded state and folded states are affected to different extents, then the folding equilibrium will be shifted. For an unfolded chain, the probability density $G(\mathbf{x}, \mathbf{x}_0, n)$ that the chain starting at position $\mathbf{x}_0$ will end at position $\mathbf{x}$ after n peptide bonds satisfies a diffusion equation (Casassa, 1967; Doi and Edwards, 1986)

$$\frac{\partial G(\mathbf{x}, \mathbf{x}_0, n)}{\partial n} = \frac{b^2}{6}\nabla^2 G(\mathbf{x}, \mathbf{x}_0, n) \quad (8)$$

The discrete variable n, treated as continuous (a good approximation for chains with residue number $N \gg 1$), plays the same role as time t in the diffusion of a Brownian particle with a diffusion constant $D = b^2/6$. The equivalence of an unfolded chain and a Brownian particle stems from the fact the probability densities for both problems are Gaussian [eq. (1a)], with the mean square distance proportional to the chain separation in the former [eq. (1b)] and the lapsed time in the latter. Since the chain cannot cross any obstacle (e.g. a crowding macromolecule) or the walls of a confining cage, an absorbing boundary condition $G(\mathbf{x}, \mathbf{x}_0, n) = 0$ applies. The confining cage or crowding macromolecules serve to eliminate conformations otherwise available to the unfolded chain. The fraction f_u of chain conformations that do not cross any boundaries is given by

$$f_u = \int \mathrm{d}^3\mathbf{x} \int \mathrm{d}^3\mathbf{x}_0 G(\mathbf{x}, \mathbf{x}_0, N)/V \quad (9)$$

where the integration is over the whole volume V of the confining cage or the solution (in the case of crowding).

Confinement

The diffusion equation [eq. (8)] subject to an absorbing boundary condition can be solved for a number of symmetric confining cages (Casassa, 1967; Zhou and Dill, 2001). Results for the fraction of accessible conformations are

$$f_u = \frac{8}{\pi^2}\sum_{k=1,3,5,\ldots}\frac{1}{k^2}\exp\left(-\pi^2 R_g^2 k^2/s^2\right)$$

between two parallel walls at separation s

$$= \frac{6}{\pi^2}\sum_{k=1,2,3,\ldots}\frac{1}{k^2}\exp\left(-4\pi^2 R_g^2 k^2/d^2\right)$$

in a sphere with diameter d

$$= \frac{32}{\pi^2}\sum_{k=1,3,5,\ldots}\frac{1}{k^2}\exp\left(-\pi^2 R_g^2 k^2/h^2\right) \times \sum_{k=1,2,3,\ldots}\frac{1}{x_k^2}\exp\left(-4R_g^2 x_k^2/d^2\right)$$

in a cylinder with diameter d and height h

where $R_g = (Nb^2/6)^{1/2}$ is the radius of gyration of the unfolded chain and x_k are the roots of the Bessel function of the first kind of order zero. In the case of confinement, the fraction f_u of accessible conformations is the partition coefficient of the unfolded chain.

The folded state, here modeled as a rigid sphere, is also restricted by confinement. Specifically, V_{acc}, the volume accessible to the center of the folded protein is reduced from the volume V of the confining space. Let $f_f = V_{acc}/V$. In particular,

$$f_f = 1 - d_f/s, \text{between two parallel walls at separation } s$$
$$= (1 - d_f/d)^3, \text{ in a sphere with diameter } d$$
$$= (1 - d_f/h)(1 - d_N/d)^2, \text{in a cylinder with diameter } d \text{ and height } h$$

where d_f is the diameter of the folded protein. The net effect of confinement on the folding free energy is

$$\Delta\Delta G = -k_B T \ln(f_f/f_u) \quad (10)$$

The results for $\Delta\Delta G$ by confinement in cubic and spherical cages are shown in Fig. 4(A). Stabilization in excess of $15\,k_B T$ ($\sim$9 kcal/mol) is predicted.

Experimental evidence for such dramatic stabilization has indeed been presented. Eggers and Valentine (2001) encapsulated α-lactalbubin in the pores of silica glass and observed an increase of over 30 °C in the melting temperature.

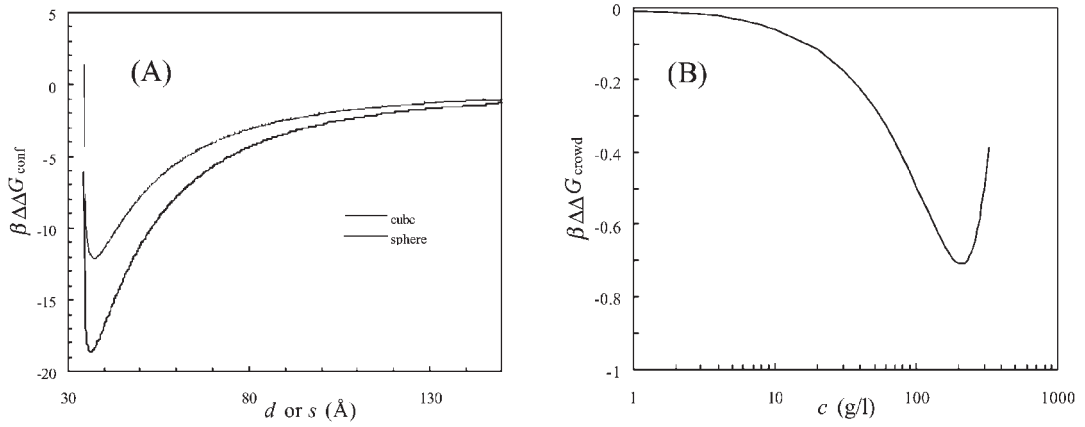

Figure 4. Effects of (A) confinement and (B) crowding on the folding free energy. The dimensions of the folded and unfolded protein are: $d_f/2 = 17.2$ Å and $R_g = 30$ Å. (B) Effect of crowding on the folding stability. In (A) black and red curves represent results in a cubic and a spherical cage, respectively. The value of $\Delta\Delta G$ for a cubic cage is three times that for two parallel walls. In (B) the crowding macromolecules have a radius $a_c = 15.4$ Å and a molecular weight $M_c = 17\,000$.

Crowding

The equivalence between an unfolded chain and a Brownian particle has already been noted. The problem of an unfolded chain crowded by macromolecules also has a well-known counterpart, i.e., the problem of a Brownian particle being trapped by static absorbing traps. In the trapping problem, $\int d^3\mathbf{x}G(\mathbf{x}, \mathbf{x}_0, N)$ maps to the survival probability of the Brownian particle. The additional integration over $\mathbf{x}_0$ to obtain f_u [eq. (9)] is equivalent to averaging over different realizations of trap distributions. Thus f_u is the same as the survival probability in the trapping problem at an equivalent time $t = N$.

Let the concentration of the crowding macromolecules, or, equivalently, traps, be c. When the traps are modeled as spheres, then the Smoluchowski results apply. For short to moderate times, the survival probability $S(t)$ of a Brownian particle is given by [eq. (6); Szabo *et al.*, 1988]

$$-\ln\ S(t) = c\int_0^t \mathrm{d}t' k_s(t') = 4\pi D a_c c t[1 + 2a_c(\pi D t)]^{-1/2} \tag{11a}$$

where a_c is the radius of traps or crowding macromolecules. Theory (Nieuwenhuizen, 1989) and simulations (Barkema *et al.*, 2001) suggest that eq. (11a) is accurate up to extremely small values of $S(t)$. Departure from the Smoluchowski theory occurs at around $S(t) = \exp\left(-6.68/\phi^{1/2}\right)$, where $\phi = 4\pi a_c^3 c/3$ is the volume fraction of traps. For example, at $\phi = 0.25$, eq. (11a) is expected to hold up to $-\ln\ S(t) = 13.4$.

Mapping eq. (11a) to the problem of a Gaussian chain crowded by spherical macromolecules ($D = b^2/6$ and $t = N$), one finds

$$-\ln f_u = 3\phi y^2\left(1 + 2/\pi^{1/2}y\right) \tag{11b}$$

where $y = R_g/a_c$. At $y = 1$, $-\ln f_u$ never exceeds the upper limit of 13.4 for the validity of eq. (11b). At $y = 2$, 3, and 4, the range of validity of eq. (11b) extends to ϕ = 0.50, 0.32, and 0.23, respectively. These results suggest that eq. (11b) will be reliable under conditions of biological interest.

The effect of macromolecular crowding on the unfolded chain was studied by Minton (2000b) in a heuristic treatment, based on modeling the unfolded protein as an expanded hard sphere [Kinjo and Takada (2002a, b, 2003) have also used an expanded-sphere model for the unfolded state]. Minton's treatment for the effect of crowding on the folded protein, based on the scaled-particle theory, appears far more realistic. This treatment for the folded state is adopted here. Crowding affects the folded protein because many of the placements of the protein will not be allowed due to overlap with the crowding molecules. In physical chemistry, the inverse of the fraction, f_f, of allowed placements is called the activity coefficient of the folded protein. When both the folded protein and the crowding macromolecules are modeled as spheres, the scaled-particle theory predicts (Lebowitz and Rowlinson, 1964)

$$-\ln f_f = -\ln(1-\phi) + \left(3z + 3z^2 + z^3\right)\phi/(1-\phi) + \left(9z^2/2 + 3z^3\right)[\phi/(1-\phi)]^2 + 3z^3[\phi/(1-\phi)]^3$$

where $z = d_f/2a_c$. If the concentration and radius of the macromolecules are in g/l and Å, respectively, then $\phi = 2.52 \times 10^{-3} \times a_c^3 c/M_c$ where M_c is the molecular weight of the macromolecules.

As illustration, Fig. 4(B) displays the effect of crowding by ribonuclease A on the folding free energy of α-lactalbumin. The relevant parameters for this case are: $R_g = 30$ Å, $d_f/2 = 17.2$ Å, $a_c = 15.4$ Å, and $M_c = 17\,000$. When the crowder concentration is increased, the difference between $-\ln f_u$ and $-\ln f_f$ increases to a maximum (to $\sim$0.7, corresponding to $\Delta\Delta G \approx -0.4$ kcal/mol at room temperature) at $c \approx 200$ g/l.

The effect of crowding on protein stability predicted by the current model is rather modest. This finding appears to rationalize experimental observations on macromolecular crowding. van den Berg *et al.* (1999) studied the refolding of oxidized hen lysozyme from 8 M urea or 6 M guanidine in the presence of crowding agents. The yields of correctly folded protein were hardly affected by the crowding agents. Only at crowder concentrations greater than 250 g/l did the refolding yield decreased by $\sim$10%. Qu and Bolen (2002)

measured an increase of 0.9 kcal/mol in the unfolding free energy of a ribonuclease T_1 variant by 300 g/l of dextran. Sasahara *et al.* (2003) recently studied the thermal unfolding of hen lysozyme in the presence of dextran. Even at 300 g/l of dextran, the melting temperature increased by just 2.5 °C. Recent *in vivo* experiments by Ghaemmaghami and Oas (2001) and by Ignatova and Gierasch (2004) showed that stabilities of monomeric lambda repressor and cellular retinoic acid-binding protein I within the cell are the same as those measured in simple buffers *in vitro*.

In macromolecular crowding, the folded protein and the unfolded protein apparently are restricted to similar extents. On the other hand, confinement has a modest effect on the folded protein but a significant restrictive effect on the unfolded protein. The disparate consequences of confinement and crowding on protein stability can be understood in the following way. In the former situation, walls of the confining space fully enclose the unfolded chain, hence all conformations that cross any part of the walls are eliminated. However, in the latter situation, there are always interstitial spaces that allow the unfolded chain to escape. At high concentrations of crowding macromolecules, the interstitial voids may become too small in serving as routes of translocation for the compact folded protein but will continue to allow the unfolded chain to leak. In this sense crowding is like confinement in a cage with holes, which compensate for the excess restriction of the cage on the unfolded chain.

As the concentration of crowding macromolecules increase even further, most of the unoccupied spaces become too small to accommodate the folded protein as modeled by a hard sphere. On the other hand, the unfolded protein modeled as a Gaussian chain can always changes its conformation to have it accommodated. Then $-\ln f_f$ may become even greater than $-\ln f_u$. This situation is akin to what happens in spatial confinement when the confining space is shrunk to the extent of not being able to accommodate the folded hard sphere; nonetheless there will be a small fraction of extremely compact chain conformations of the unfolded protein that can be accommodated within the confining space. In that situation the unfolded state also becomes favored [Fig. 4(A)].

EFFECT ON FOLDING RATE BY CONFINEMENT AND CROWDING

Both confinement and crowding will likely affect the rate of contact formation in the folding process in two opposing ways. They will eliminate the more expanded conformations of the unfolded chain, thus residues will have higher probabilities of being near contact. On the other hand, intrachain diffusion in a confined space and a crowded solution will be slowed down.

Confinement

To illustrate the effect of confinement on the rate of contact formation, consider a Gaussian chain confined to a spherical cage (with diameter d). For concreteness, it is assumed one residue forming the contact is rigidly held at the center of the confining sphere while the other residue freely moves within the confines of the sphere. The probability density of the second residue, separated by n peptide bonds, at a radial distance r is (Park and Sung, 1998)

$$G(r) = \frac{2}{d^2 r} \sum_{k=1,2,3,\ldots} k \sin(2\pi k r/d) \exp\left(-4\pi^2 r_g^2 k^2/d^2\right)$$

where $r_g = (nb^2/6)^{1/2}$. The mean square distance is

$$\langle r^2 \rangle = (d/2)^2 \sum_{k=1,2,3,\ldots} (-1)^{k-1}\left(1 - 6/\pi^2 k^2\right) \exp\left(-4\pi^2 r_g^2 k^2/d^2\right) \Big/ \sum_{k=1,2,3,\ldots} (-1)^{k-1} \exp\left(-4\pi^2 r_g^2 k^2/d^2\right) \quad (12)$$

At $d = 2r_g$, the mean square distance is predicted by eq. (12) to be 15.3-fold smaller than the value without confinement. The reduction in $\langle r^2 \rangle$ will contribute to a higher rate of contact formation [see eq. (3)].

An indirect consequence of the elimination of the more expanded conformations by the confinement might be that the propensities for native conformations, which tend to be compact, will be enhanced. As a result, the equilibrium constant ω_+/ω_- for forming correct local conformations might be increased. This increase also contributes to a higher rate of contact formation [see eq. (4)].

On the other hand, confinement also slows down intrachain diffusion. For a small sphere with diameter a at a distance b from the wall of a container with dimension d, the diffusion constant of the small sphere becomes $D_{\text{conf}} = D[1 - (a/d)f(b/d)]$, where $f(x)$ is a function determined by the shape of the container (Happel and Brenner, 1983). When the small sphere is not located in the immediate vicinity of the container wall, $f(x)$ has a value close to 2. For $a = 5$ Å, $d = 30$ Å, and $f(x) = 2$, one finds $D_{\text{conf}} \approx 0.67\,D$. This calculation suggests that confinement normally will reduce the diffusion constant by no more than a factor of 2.

Given the small effect of confinement on the diffusion constant and that this small effect is likely offset by enhanced propensities of local correct conformations, it is then expected that the main consequence of confinement is the reduction in $\langle r^2 \rangle$, which serves to increase the contact formation rate. The resulting increase in the rate constant, calculated according to eq. (3), is shown in Fig. 5. It is seen that confinement may speed up contact formation by as much as two orders of magnitudes.

Klimov *et al.* (2002) have carried out detailed simulations of the β-hairpin formation of a 16-residue peptide inside a confining sphere. Their results were in qualitative and perhaps even semi-quantitative agreement with those predicted by the simple model presented above. In particular, for their peptide, $R_g \approx 16$ Å and $d_f \approx 30$ Å, Klimov *et al.* found a stabilization of 1.2 kcal/mol by confinement in a spherical cage with a 35 Å diameter, whereas eq. (10) predicts a stabilization of 1.5 kcal/mol. Klimov *et al.* observed a doubling of folding rate by confinement in a spherical cage with a 40 Å diameter, a result predicted by eqs (12) and (3) if forming a contact separated by seven peptide bonds (e.g. the hydrophobic contact between Tyr5 and Phe12) is the rate-limiting step. However, other details,

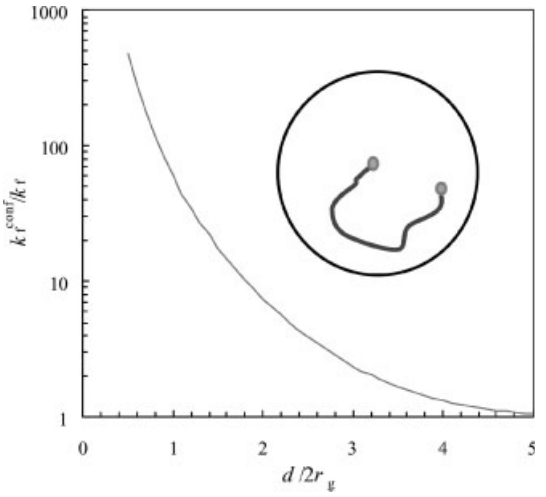

Figure 5. Speed up of protein folding by confinement. Effect of confinement in a spherical cage on the rate of contact formation is displayed. k_f^{conf} is calculated from eq. (3) with $<r^2>$ given by eq. (12).

such as kinetic traps, are outside the scope of the simple model. Qualitatively similar results were obtained in other simulation studies on the effects of confinement (Friedel *et al.*, 2003; Baumketner *et al.*, 2003; Takagi *et al.*, 2003; Thirumalai *et al.*, 2003). Takagi *et al.* (2003) found that folding rate enhancement by confinement is greater for proteins with long-range tertiary contacts. This finding is entirely consistent with eqs (12) and (3), which predict larger acceleration for the formation of contacts with a longer chain separation.

The central cavity (also called the Anfinsen cage) of chaperonins provides an important example of confinement. Recently Brinker *et al.* (2001) have indeed obtained experimental evidence that folding inside the Anfinsen cage is faster than in bulk solution.

Crowding

The size of a polymer chain in the presence of obstacles will contract. Indeed, Sasahara *et al.* (2003) and Tokuriki *et al.* (2004) have obtained direct experimental evidence of the compaction of an unfolded protein by crowding. However, there are conflicting theories for the size contraction. A number of Monte Carlo simulations have been carried out, in part motivated by the theoretical conflicts (Baumgartner and Muthukumar, 1987; Honeycutt and Thirumalai, 1989; Dayantis *et al.*, 1998). These simulations show that, up to an obstacle volume fraction of 0.5 and up to a chain length of a hundred freely jointed units, the mean square distance is at most 5-fold smaller than the value in the absence of the obstacles. Crowding thus has a much smaller potential for speeding up protein folding than confinement.

CHANGE OF BINDING EQUILBRIUM AND RATE BY CROWDING

In contrast to the unimolecular process of protein folding, the bimolecular process of protein binding has an equilibrium that is concentration dependent. Since in a confined cage concentration cannot be defined in the sense of the thermodynamic limit (whereby both the number of molecules and the volume go to infinity but their ratio remains constant), here I will focus attention on the effect of crowding. The binding of two folded proteins will be modeled as the approach of two spherical particles to closest contact, but stereospecificity will be partly accounted for.

Binding equilibrium

Consider two identical tracer hard spheres (with diameter d_f) in the presence of other hard spheres (with radius a_c) occupying a fraction ϕ of space. In the absence of the crowding hard spheres, the tracer particles will have a uniform relative distribution. The crowding particles increase the probability that the two tracer particles are near each other. The crowding-induced radial distribution function, $g_c(r)$, is equivalent to an effective attractive potential $U_c(r)$ [cf. eq. (2)]. At contact, the scaled-particle theory predicts (Lebowitz and Rowlinson, 1964)

$$g_c(d_f) = \exp[-\beta U_c(d_f)] = (1 - \phi + 3z\phi/2)/(1 - \phi)^2 \quad (13)$$

The effective potential $U_c(r)$ adds to the interaction potential between the two binding proteins, resulting in an increase in the equilibrium constant [eq. (5a) or (5b)]:

$$K^{crowd}/K = \exp[-\beta U_c(d_f)] \quad (14)$$

This account of the effect of crowding is framed in languages different from those used by Minton (1998). In his account, both reactant and product species are treated as rigid particles, with crowding contributing to their activity coefficients. In the above account, it is recognized that, even when the reactant species are treated as rigid particles, the product species will have inter-molecular freedom (as represented by the inter-particle distance r). However, after the specialization is made that the product species has a unique inter-particle distance (i.e. the contact distance), then the result for K^{crowd}/K should be the same as that obtained from calculating the activity coefficients of the reactant (a hard sphere) and product (two contacting hard spheres) species. Previously the activity coefficient for a species represented by two contacting spheres has not been obtained. [Hall and Minton (2003) recently considered the situation where the interaction between a protein and a crowding macromolecule is 'soft.']

Figure 6 displays the effects of crowding on the binding constant predicted by eqs (13) and (14). For comparison, results obtained by modeling the reactant species as a sphere with twice the volume of the reactant species are also displayed. Significantly smaller effects are seen for the contacting-spheres model, a result anticipated by Minton (1998) and rationalized by the compactness of the enlarged-sphere model.

Binding rate

While crowding increases the equilibrium probability for the reactant species to be near each other, it also slows down

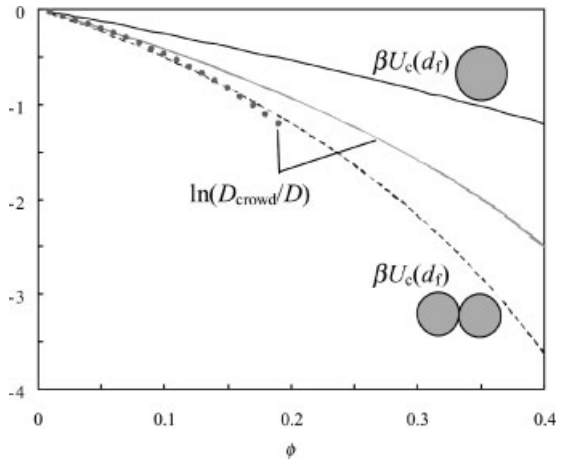

Figure 6. Effects of crowding on the binding equilibrium and the diffusion constant of proteins. Solid and dashed curves represent the crowding-induced effective potential $U_c(d_f)$ at contact distance for two models of the bound state (the contacting-spheres model and the enlarged-sphere model, shown next to the curves). A gray line represents the decrease in diffusion constant by crowding. Solid circles are from fitting to the experimental data of Muramatsu and Minton (1988) for the diffusion constant of labeled BSA in concentrated BSA solutions. Crowding macromolecules and the protein monomer are assumed to have the same radius.

the diffusional approach of the reactants. Here I will use eq. (7) to estimate the overall effect of crowding on the binding rate when the binding is diffusion-limited. It should be recalled that eq. (7) is a good approximation if the interaction potential is long-ranged. While the crowding-induced effective potential $U_c(r)$ is not a contact potential, it is expected to have strong distance dependence. Though this dependence will affect the applicability of eq. (7), this equation should suffice for the present purpose of obtaining qualitative estimates.

It is instructive to mention a simulation of a diffusion-limited reaction in a very crowded solution carried out some years ago (Zhou and Szabo, 1991). In a box of hard spheres occupying a volume fraction $\phi = 0.412$, the reaction of labeled hard spheres was found to be described well by the Smoluchowski theory after the crowding-induced effective potential and the effect of crowding on the diffusion constant were incorporated.

The decrease in the diffusion constant of a hard sphere by crowding hard spheres was studied by Tokuyama and Lebowitz (1994). Their result is

$$D_{\text{crowd}}/D = \frac{1 - 9\phi/32}{1 + H(\phi) + (\phi/0.5718)(1 - \phi/0.5718)^2} \tag{15}$$

where

$$H(\phi) = \frac{2b_1^2}{1 - b_1} - \frac{b_2}{1 + 2b_2} - \frac{2b_1 b_2(2 + b_2)}{(1 + b_2)(1 - b_1 + b_2)}$$

with $b_1 = (9\phi/8)^{1/2}$ and $b_2 = 11\phi/16$. Equation (15) appears to be in reasonable agreement with the experimental data of Muramatsu and Minton (1988) for the diffusion constant of labeled bovine serum albumin (BSA) in concentrated BSA solutions (Fig. 6).

If the binding is reaction-limited, then the slowed-down diffusion does not have any effect on the binding rate. However, in the diffusion-limited situation, the binding rate will be reduced in proportion to the slowing down of the diffusion [eq. (7)]. When the contribution of the crowding-induced effective potential and the effect of the slowed-down diffusion are combined, it can be seen from Fig. 6 that cancellation of the two opposing effects leaves a moderate overall effect of crowding on the binding rate.

The slowing down of protein diffusion by crowding may play an important role in the action of chaperonins. There is evidence from experiments and simulations (Martin and Hartl, 1997; Elcock, 2003) that this slow down has a beneficial effect on the yield of proteins undergoing chaperonin-assisted folding, because it increases the chance that a partially folded chain will be recaptured by the chaperonin for new rounds of folding.

CONCLUSIONS

Based on the simple models, the qualitative effects of confinement and crowding on the equilibria and rates of protein folding and binding can be summarized as follows:

(1) confinement stabilizes proteins and may accelerate their folding significantly;
(2) crowding is expected to have only a marginal effect on protein stability, but may accelerate folding;
(3) crowding may significantly shift the binding equilibrium of proteins toward the bound state; as such it may contribute to protein aggregation and amyloid formation;
(4) crowding significantly slows down protein diffusion; for diffusion-limited protein binding, this slowing down will moderate the positive effect of crowding in enhancing the equilibrium probability near contact, and may even result in an overall decrease in the binding rate.

Acknowledgements

I thank Allen Minton and Robert Baldwin for discussion. This work was supported in part by NIH grant GM58187.

REFERENCES

Barkema GT, Biswas P, van Beijeren H. 2001. Diffusion with random distribution of traps. *Phys. Rev. Lett.* **87**(17): 170601.

Baumgartner A, Muthukumar M. 1987. A Trapped polymer chain in random porous media. *J. Chem. Phys.* **87**(5): 3082–3088.

Baumketner A, Jewett A, Shea JE. 2003. Effects of confinement in chaperonin assisted protein folding: rate enhancement by decreasing the roughness of the folding energy landscape. *J. Mol. Biol.* **332**(3): 701–713.

Brinker A, Pfeifer G, Kerner MJ, Naylor DJ, Hartl FU, Hayer-Hartl M. 2001. Dual function of protein confinement in chaperonin-assisted protein folding. *Cell* **107**(2): 223–233.

Casassa EF. 1967. Equilibrium Distribution of flexible polymer chains between a macroscopic solution phase and small voids. *J. Polym. Sci. B (Polym. Lett.)* **5**: 773–778.

Dayantis J, Abadie MJM, Abadie MRL. 1998. Statstics of chains confined inside cubes with fixed random obstacles. *Comp. Theor. Polym. Sci.* **8**: 273–282.

Debye P. 1942. Reaction rates in ionic solutions. *Trans. Am. Electrochem. Soc.* **82**: 265–272.

Doi M, Edwards SF. 1986. *The Theory of Polymer Dynamics*. Clarendon Press: Oxford.

Eggers DK, Valentine JS. 2001. Molecular confinement influences protein structure and enhances thermal protein stability. *Protein Sci.* **10**: 250–261.

Elcock AH. 2003. Atomic-level observation of macromolecular crowding effects: escape of a protein from the GroEL cage. *Proc. Natl Acad. Sci. USA* **100**: 2340–2344.

Ellis RJ. 2001. Macromolecular crowding: obvious but underappreciated. *Trends Biochem. Sci.* **26**: 597–604.

Friedel M, Sheeler DJ, Shea JE. 2003. Effects of confinement and crowding on the thermodynamics and kinetics of folding of a minimalist beta-barrel protein. *J. Chem. Phys.* **118**(17): 8106–8113.

Fuller RS, Kaguni JM, Kornberg A. 1981. Enzymatic replication of the origin of the *Escherichia coli* chromosome. *Proc. Natl Acad. Sci. USA* **78**: 7370–7374.

Ghaemmaghami S, Oas TG. 2001. Quantitative protein stability measurement *in vivo*. *Nat. Struct. Biol.* **8**(10): 879–882.

Hall D, Minton AP. 2003. Macromolecular crowding: qualitative and semiquantitative successes, quantitative challenges. *Biochim. Biophys. Acta* **1649**(2): 127–139.

Happel J, Brenner H. 1983. *Low Reynolds Number Hydrodynamics*. Martinus Nijhoff: Dordrecht.

Hatters DM, Minton AP, Howlett GJ. 2002. Macromolecular crowding accelerates amyloid formation by human apolipoprotein C-II. *J. Biol. Chem.* **277**(10): 7824–7830.

Honeycutt JD, Thirumalai D. 1989. Static properties of polymer chains in porous media. *J. Chem. Phys.* **90**(8): 4542–4559.

Ignatova Z, Gierasch LM. 2004. Monitoring protein stability and aggregation *in vivo* by real-time fluorescent labeling. *Proc. Natl Acad. Sci. USA* **101**(2): 523–528.

Kinjo AR, Takada S. 2002a. Effects of macromolecular crowding on protein folding and aggregation studied by density functional theory: statics. *Phys. Rev. E* **66**(3): 031911.

Kinjo AR, Takada S. 2002b. Effects of macromolecular crowding on protein folding and aggregation studied by density functional theory: dynamics. *Phys. Rev. E* **66**(5): 051902.

Kinjo AR, Takada S. 2003. Competition between protein folding and aggregation with molecular chaperones in crowded solutions: insight from mesoscopic simulations. *Biophys. J.* **85**(6): 3521–3531.

Klimov DK, Newfield D, Thirumalai D. 2002. Simulations of β-hairpin folding confined to spherical pores using distributed computing. *Proc. Natl Acad. Sci. USA* **99**: 8019–8024.

Kramers HA. 1940. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7**: 284–304.

Lebowitz JL, Rowlinson JS. 1964. Thermodynamic properties of mixtures of hard spheres. *J. Chem. Phys.* **41**: 133–138.

Makarov DE, Metiu H. 2002. A model for the kinetics of protein folding: Kinetic Monte Carlo simulations and analytical results. *J. Chem. Phys.* **116**(12): 5205–5216.

Makarov DE, Keller CA, Plaxco KW, Metiu H. 2002. How the folding rate constant of simple, single-domain proteins depends on the number of native contacts. *Porc. Natl Acad. Sci. USA* **99**(6): 3535–3539.

Martin J, Hartl FU. 1997. The effect of macromolecular crowding on chaperonin-mediated protein folding. *Porc. Natl Acad. Sci. USA* **94**: 1107–1112.

Minton AP. 1998. Molecular crowding: analysis of effects of high concentrations of inert cosolutes on biochemical equilibria and rates in terms of volume exclusion. *Meth. Enzym.* **25**: 127–149.

Minton AP. 2000a. Implications of macromolecular crowding for protein assembly. *Curr. Opin. Struct. Biol.* **10**: 34–39.

Minton AP. 2000b. 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.* **78**: 101–109.

Muramatsu N, Minton AP. 1988. Tracer diffusion of globular proteins in concentrated protein solutions. *Proc. Natl Acad. Sci. USA* **85**: 2984–2988.

Nieuwenhuizen TM. 1989. Trapping and Lifshitz in random media, self-attracting polymers, and the number of distinct sites visited: a renormalization approach in three dimensions. *Phys. Rev. Lett.* **62**(4): 357–360.

Park PJ, Sung W. 1998. Polymer release out of a spherical vesicle through a pore. *Phys. Rev. E* **57**(1): 730–734.

Qu Y, Bolen DW. 2002. Efficacy of macromolecular crowding in forcing proteins to fold. *Biophys. Chem.* **101–102**: 155–165.

Rivas G, Lopez A, Mingorance J, Ferrandiz MJ, Zorrilla S, Minton AP, Vicente M, Andreu JM. 2000. Magnesium-induced linear self-association of the FtsZ bacterial cell division protein monomer: the primary steps for FtsZ assembly. *J. Biol. Chem.* **275**(16): 11740–11749.

Sasahara K, McPhie P, Minton AP. 2003. Effect of dextran on protein stability and conformation attributed to macromolecular crowding. *J. Mol. Biol.* **326**: 1227–1237.

Shoup D, Szabo A. 1982. Role of diffusion in ligand-binding to macromolecules and cell-bound receptors. *Biophys. J.* **40**(1): 33–39.

Smoluchowski MV. 1917. Versuch einer mathematischen Theorie der Koagulationskinetik kolloider Loeschungen. *Z. Phys. Chem.* **92**: 129–168.

Szabo A, Schulten K, Schulten Z. 1980. First passage time approach to diffusion controlled reactions. *J. Chem. Phys.* **72**: 4350–4357.

Szabo A, Zwanzig R, Agmon N. 1988. Diffusion-controlled reactions with mobile traps. *Phys. Rev Lett.* **61**(21): 2496–2499.

Takagi F, Koga N, Takada S. 2003. How protein thermodynamics and folding mechanisms are altered by the chaperonin cage: molecular simulations. *Proc. Natl Acad. Sci. USA* **100**: 11367–11372.

Thirumalai D, Klimov DK, Lorimer GH. 2003. Caging helps proteins fold. *Proc. Natl Acad. Sci. USA* **100**: 11195–11197.

Tokuriki N, Kinjo M, Negi S, Hoshino M, Goto Y, Urabe I, Yomo T. 2004. Protein folding by the effects of macromolecular crowding. *Protein Sci.* **13**(1): 125–133.

Tokuyama M, Lebowitz I. 1994. Dynamics of hard-sphere suspensions. *Phys. Rev. E* **50**(1): R16–R19.

van den Berg B, Ellis RJ, Dobson CM. 1999. Effects of macromolecular crowding on protein folding and aggregation. *EMBO J.* **18**: 6927–6933.

Zhou HX. 1997. Enhancement of protein-protein association rate by interaction potential: accuracy of prediction based on local Boltzmann factor. *Biophys. J.* **73**: 2441–2445.

Zhou HX. 2003. Theory for the rate of contact formation in a polymer chain with local conformational transitions. *J. Chem. Phys.* **118**: 2010–2015.

Zhou HX, Dill KA. 2001. Stabilization of proteins in confined spaces. *Biochemistry* **40**: 11289–11293.

Zhou HX, Szabo A. 1991. Comparison between molecular dynamics simulations and the Smoluckowski theory of reactions in a hard sphere liquid. *J. Chem. Phys.* **95**: 5948–5952.

Zhou HX, Wong KY, Vijayakumar M. 1997. Design of fast enzymes by optimizing interaction potential in active site. *Proc. Natl Acad. Sci. USA* **94**: 12372–12377.

Zimmerman SB. 1993. Macromolecular crowding effects on macromolecular interactions: some implications for genome structure and function. *Biochim. Biophys. Acta* **1216**: 175–185.

