Effects of macromolecular crowding on protein folding and aggregation studied by density functional theory: Statics

Akira R. Kinjo[1,*] and Shoji Takada[1,2,†]
[1]*PRESTO, Japan Science and Technology Corporation, Kobe University, Kobe 675-8501, Japan*
[2]*Department of Chemistry, Faculty of Science, Kobe University, Kobe 657-8501, Japan*
(Received 4 April 2002; published 24 September 2002)

Proteins are neither purified nor diluted inside the living cell. Thus it is indispensable to take into account various interactions between the protein of interest and other macromolecules for understanding the properties of proteins in physiological conditions. Here we focus on excluded volume interactions which are omnipresent in dense or crowded solutions of proteins and macromolecules or ''crowding agents.'' A protein solution with macromolecular crowding agents is modeled by means of a density functional theory. Effects of macromolecular crowding on protein aggregation and stability are investigated in particular. Phase diagrams are obtained in various parameter spaces by solving the equation of state. Two generic features are found: the addition of the crowding agent (1) enhances the aggregation of the denatured proteins, and (2) stabilizes the native protein unless the aggregation occurs. The present theory is qualitatively in good agreement with experimental observations and unifies previous theories regarding the crowding effects on protein stability and aggregation.

DOI: 10.1103/PhysRevE.66.031911 PACS number(s): 87.15.Nn, 87.15.Aa

I. INTRODUCTION

Typically, physical and chemical properties of proteins have been studied under highly idealized solvent conditions in which proteins are purified and diluted so that the perturbations to them are minimized. However, inside, the living cell is inherently crowded with many kinds of proteins and other macromolecules which occupy in general 20–30 % of the space in the cell [1]. Such crowded conditions are expected to impose significant effects on a wide spectrum of properties including protein stability, association, enzymatic activity, diffusion, and so on [2–7]. There may be specific and/or nonspecific interactions between proteins and macromolecules. Among them, one particular class of nonspecific interactions, the excluded volume effect, is recently gaining attention [5–7] because it is always present as long as macromolecules exist around the protein. Such molecules that exert the excluded volume effect on the target protein are termed ''crowding agents'' [6]. In their prescient work, Asakura and Oosawa [8] have already investigated how the volume exclusion by macromolecules affects the interaction between colloidal particles and predicted that the osmotic pressure due to macromolecules can induce aggregation of the particles. More recently, much of theoretical studies of macromolecular crowding were put forward by Minton [9]. Corresponding experiments *in vitro* can be carried out by adding crowding agents such as dextran or bovine serum albumin to protein solutions. To date, crowding effects on, for example, protein folding [10,11], protein self-assembly [12], and amyloid formation [13], have been investigated to support theoretical predictions [5–7].

Although thorough, most previous theoretical works are based on the theories of uniform fluid so that they cannot treat inherently nonuniform phenomena such as the morphology of aggregation. Protein aggregation had been in many cases regarded as a nuisance factor in experimental studies. However, its fundamental importance has recently been recognized regarding its relevance to inclusion bodies, amyloidoses, and other protein deposition diseases [14]. Some simulation studies have been also performed to track protein aggregation [15–17]. Although these simulations can provide detailed pictures of protein conformational changes upon aggregation and/or folding, the number of protein molecules that can be handled in each simulation is limited. Since aggregation typically involves quite a large number of molecules, other theoretical frameworks are needed to complement detailed molecular simulations.

In the present paper, we present a theoretical framework to describe nonuniform system of proteins in crowded media. The theory presented in this paper is based on a static density functional theory which has been developed to treat nonuniform systems of infinite degree of freedom [18,19]. Density functional theories are especially suited to describe interfacial phenomena such as phase separations of colloidal dispersions and polymer blends, etc. [20]. Thus, the present theory provides a unified framework to treat protein stability and aggregation in inhomogeneous crowded solutions. Our model views the crowded system from a coarse-grained scale, and proteins and macromolecules are represented by the corresponding (number) density fields in space. The protein is characterized by the intrinsic free energies of its native and denatured states, and a native protein can transform to a denatured protein and vice versa. Using the model developed, phase diagrams are obtained and the crowding effects on protein stability and aggregation at equilibrium are examined.

*Electronic address: akinjo@theory.chem.sci.kobe-u.ac.jp
†Electronic address: stakada@kobe-u.ac.jp

The paper is organized as follows. In Sec. II, we formulate the free energy functional of the crowded system consisting of proteins and crowding agents, and derive equations for the equilibrium state. Some analytical results for the uniform phase are also presented. Then we apply these equations to obtain phase diagrams in Sec. III. In Sec. IV, we compare our results with experimental observations and other theories, and check the validity of the model. Sec. V concludes the present study.

II. THEORY AND MODELING

A. Free energy functional and equation of state

We consider a system which consists of one protein species (P) and one crowding agent species (C) . The protein can be in one of the two states: native (N) or denatured (D). In the limit of dilute solution, state of the protein is characterized by the intrinsic free energy of the native state (η_N) and of the denatured state (η_D). In principle, they can be further decomposed into energetic and entropic parts. However, since the temperature is fixed in the following calculations, decomposition of the intrinsic free energies is not necessary. The intrinsic free energy of the crowding agent does not play any role in the present theory, so we set $\eta_C=0$ for convenience. Next, we assume that the bare interaction ($u_{\alpha,\beta}$) between two molecules is the hard-core square-well potential, that is, with r being intermolecular distance,

$$u_{\alpha,\beta}(r)=\begin{cases}\infty & (r\leqslant R_\alpha+R_\beta),\\ \epsilon_{\alpha,\beta} & [R_\alpha+R_\beta<r\leqslant c(R_\alpha+R_\beta)],\\ 0 & [r>c(R_\alpha+R_\beta)],\end{cases} \tag{1}$$

where R_α and R_β are the radii of molecular species α and $\beta(=N,D,C)$, and c is a constant factor greater than unity. The excluded volume interaction is represented by the uppermost line in Eq. (1). Other longer ranged interactions are represented by $\epsilon_{\alpha,\beta}$. Since we are mainly concerned with the excluded volume interaction, $\epsilon_{\alpha,\beta}$ is set to 0 except for $\epsilon_{D,D}$. Because of exposed hydrophobic residues, isolated hydrogen bond donors and acceptors, denatured proteins are expected to attract each other. Therefore, $\epsilon_{D,D}$ is usually set negative in the following. The length scale in the present theory is set to $\sim$10 nm, which is taken as the unit length. Therefore the radii of the protein and crowding agent are of order of $\sim$0.1 unit length. We always set $c=3$ in this paper.

The system is represented by the number density fields of the native protein, denatured protein, and crowding agent: $\phi^N(\mathbf{r})$, $\phi^D(\mathbf{r})$ and $\phi^C(\mathbf{r})$, respectively. The density field of solvent, $\phi^S(\mathbf{r})$, is indirectly defined by $\phi^S(\mathbf{r})=\rho_0-\Sigma_\alpha\phi^\alpha(\mathbf{r})$, where the summation is over $\alpha=N,D,C$. Here, ρ_0 is the total bulk density of the system treated. The density functional (free energy functional) is given by

$$F[\{\phi^\alpha(\mathbf{r})\}]=F_i+F_n, \tag{2}$$

where F_i is the ideal part, and F_n is the nonideal part representing intermolecular interactions. The ideal part is

$$F_i[\{\phi^\alpha(\mathbf{r})\}]=\int d\mathbf{r}\Big[\sum_\alpha \eta_\alpha\phi^\alpha+T\sum_\alpha \phi^\alpha\ln\phi^\alpha + T\phi^S\ln\phi^S\Big]. \tag{3}$$

The first term (here and below, summation is always taken over α or $\beta=N,D$, and C) comes from the intrinsic free energy, and the second and third terms represent the entropy of mixing. If only up to two-body correlation is taken into account, the nonideal part of the free energy is given by

$$F_n[\{\phi^\alpha(\mathbf{r})\}]=\frac{1}{2}\sum_{\alpha,\beta}\int\int d\mathbf{r}_1 d\mathbf{r}_2\phi^\alpha(\mathbf{r}_1)\phi^\beta(\mathbf{r}_2)\times u^{\mathrm{eff}}_{\alpha,\beta}(|\mathbf{r}_1-\mathbf{r}_2|). \tag{4}$$

Here, $u^{\mathrm{eff}}_{\alpha,\beta}(|\mathbf{r}_1-\mathbf{r}_2|)$ is the effective interaction given by [21]

$$u^{\mathrm{eff}}_{\alpha,\beta}(|\mathbf{r}|)=T(1-e^{-u_{\alpha,\beta}(|\mathbf{r}|)/T}). \tag{5}$$

We assume that all the effective interactions are short ranged compared to the system size, the nonideal part of the free energy functional can be rearranged as follows. First, we transform the variables $\mathbf{r}_1$ and $\mathbf{r}_2$ using

$$\mathbf{r}=\frac{\mathbf{r}_1+\mathbf{r}_2}{2}, \tag{6}$$

$$\mathbf{s}=\mathbf{r}_1-\mathbf{r}_2. \tag{7}$$

Since $u^{\mathrm{eff}}_{\alpha,\beta}(|\mathbf{s}|)$ is of short range, it is meaningful for small $|\mathbf{s}|$. Thus $\phi^\alpha(\mathbf{r}_1)$ and $\phi^\beta(\mathbf{r}_2)$ in Eq. (4) can be approximated as follows:

$$\phi^\alpha(\mathbf{r}_1)\approx\phi^\alpha(\mathbf{r})+\frac{1}{2}\mathbf{s}\cdot\nabla\phi^\alpha(\mathbf{r}), \tag{8}$$

$$\phi^\beta(\mathbf{r}_2)\approx\phi^\beta(\mathbf{r})-\frac{1}{2}\mathbf{s}\cdot\nabla\phi^\beta(\mathbf{r}), \tag{9}$$

where the gradient operator (∇) is applied with respect to $\mathbf{r}$. With the above expansion, F_n becomes

$$F_n=\frac{1}{2}\int d\mathbf{r}[U_{\alpha,\beta}\phi^\alpha(\mathbf{r})\phi^\beta(\mathbf{r})-V_{\alpha,\beta}\nabla\phi^\alpha(\mathbf{r})\cdot\nabla\phi^\beta(\mathbf{r})]. \tag{10}$$

Here we have defined a set of effective interaction parameters:

$$U_{\alpha,\beta}=\int d\mathbf{s}u^{\mathrm{eff}}_{\alpha,\beta}(|\mathbf{s}|), \tag{11}$$

$$V_{\alpha,\beta}=\frac{1}{4d}\int d\mathbf{s}|\mathbf{s}|^2u^{\mathrm{eff}}_{\alpha,\beta}(|\mathbf{s}|), \tag{12}$$

where d is the spatial dimension (i.e., $d=3$). Thus defined $U_{\alpha,\beta}$ is (two times) the excluded volume parameter [22], and its positive or negative values correspond to repulsive or

attractive interaction, respectively. The parameter $V_{\alpha,\beta}$ in Eq. (12) corresponds to (negative) surface tension of the fluid. Mixing or separation of two species α and β is preferred when $V_{\alpha,\beta}$ is positive or negative, respectively.

Local chemical potential $\mu_\alpha(\mathbf{r})$ is defined as a functional derivative of the free energy F with respect to $\phi^\alpha(\mathbf{r})$:

$$\mu_\alpha(\mathbf{r}) = \frac{\delta F}{\delta \phi^\alpha(\mathbf{r})} = \eta_\alpha + T \ln\left(\frac{\phi^\alpha(\mathbf{r})}{\rho_0 - \sum_\beta \phi^\beta(\mathbf{r})} \right) + W_\alpha(\mathbf{r}), \tag{13}$$

where the last term $W_\alpha(\mathbf{r}) \equiv \Sigma_\beta (U_{\alpha,\beta} + V_{\alpha,\beta}\nabla^2)\phi^\beta(\mathbf{r})$ may be regarded as the potential of mean force for the species α. The chemical potential is spatially uniform at equilibrium. Let μ_α^0 be the equilibrium chemical potential. Then the equilibrium condition is given by $\mu_\alpha(\mathbf{r}) = \mu_\alpha^0$. This equation can be rearranged by the use of Eq. (13) to obtain a set of self-consistent equations of state:

$$\phi^\alpha(\mathbf{r}) = \frac{\rho_0 \exp[-\{\eta_\alpha + W_\alpha(\mathbf{r}) - \mu_\alpha^0\}/T]}{1 + \sum_\beta \exp[-\{\eta_\beta + W_\beta(\mathbf{r}) - \mu_\beta^0\}/T]}. \tag{14}$$

Note that each protein can be either native or denatured, requiring that $\mu_N^0 = \mu_D^0$ which we denote by μ_P^0. In the following, we solve this set of equations of state (14) to obtain phase diagrams.

B. Some analytical results for the uniform case

In order to grasp the qualitative behavior of the model presented above, we derive some analytical relations for the spatially uniform system. Namely, we describe the solution of the self-consistent equation (14) for ϕ^N with fixed values of $\phi^N + \phi^D (= \rho_P)$ and $\phi^C (= \rho_C)$. Setting $\mu_N = \mu_D$, substituting $\phi^D = \rho_P - \phi^N$ and defining $f_N = \phi^N/\rho_P$, we have

$$T \ln\left(\frac{f_N}{1-f_N}\right) - (2U_{N,D} - U_{N,N} - U_{D,D})\rho_P f_N + (U_{N,D} - U_{D,D})\rho_P + (U_{N,C} - U_{D,C})\rho_C + \eta_N - \eta_D = 0. \tag{15}$$

This equation can be graphically solved for f_N (Fig. 1). Graphical analysis with some algebra shows that Eq. (15) has one solution when

$$\rho_P(2U_{N,D} - U_{N,N} - U_{D,D}) \leqslant 4T, \tag{16}$$

and two solutions when

$$\rho_P(2U_{N,D} - U_{N,N} - U_{D,D}) > 4T \quad \text{and} \quad \rho_C^+ < \rho_C < \rho_C^-. \tag{17}$$

Here $\rho_C^\pm$ are defined by

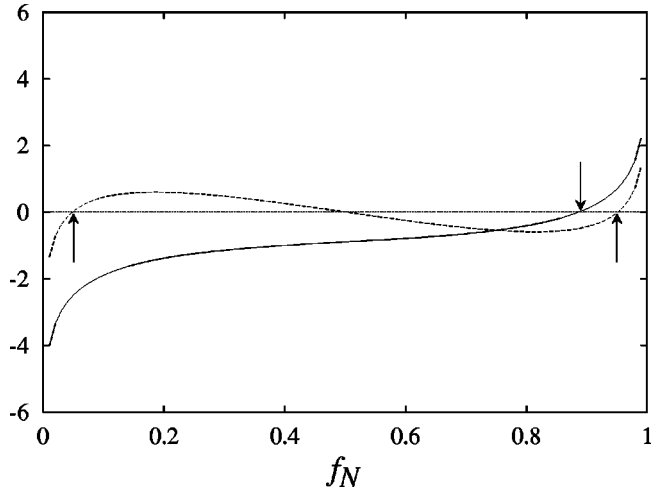

FIG. 1. Graphical solution of the self-consistent equation for the uniform case. The vertical axis indicates the value of the left-hand side of Eq. (15) in units of T (temperature). The horizontal axis is the fraction of the native protein (dimensionless). Stable solutions for Eq. (15) are indicated by arrows. There can be one (solid curve) or two (dashed curve) solutions depending on the strength of the interactions.

$$\rho_C^\pm = \frac{1}{U_{D,C} - U_{N,C}} \left[T \ln\left(\frac{f_N^\pm}{1-f_N^\pm}\right) - (2U_{N,D} - U_{N,N} - U_{D,D}) \times \rho_P f_N^\pm - \eta_D + \eta_N - (U_{D,D} - U_{N,D})\rho_P \right], \tag{18}$$

with

$$f_N^\pm = \frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{T}{(2U_{N,D} - U_{N,N} - U_{D,D})\rho_P}}. \tag{19}$$

Figure 1 illustrates these two cases. When the latter condition in Eq. (17) does not hold, there is again only one stable solution. From Eq. (15) and Fig. 1, we can see that the increase in ρ_C results in the increase in f_N provided that the denatured protein has larger excluded volume than the native one ($U_{D,C} > U_{N,C}$), which is the physically appropriate case. The increase in ρ_P leads to the decrease in ϕ^N when the denatured proteins attract each other strongly enough that $U_{D,D} < U_{N,D}$, which, again, is usually the case due to the exposed hydrophobic residues and isolated hydrogen-bonding groups in the denatured protein.

C. Numerics

Before going into the details of the calculation, we summarize the physical scales of the system treated. First, the unit of energy is defined as the folding temperature which is ~ 3 kJ/mol. Most of the calculations are carried out at the folding temperature in dilute solution, hence $\eta_N = \eta_D$ holds, and we set $\eta_N = \eta_D = 0$. These values are always used below otherwise stated. The unit length is taken as approximately 10 nm, so that the radii of the molecules are ~ 0.1 unit length. The experimentally measured radius of gyration of

the native state of the IgG binding domain of streptococcal protein *L* (62 amino acid residues) is 1.65 nm and that of the denatured state (denatured by GdmCl) is ~ 2.6 nm [23]. Hence $R_D \sim 1.5 R_N$. Depending on the protein, the radius of the denatured protein can be larger than that of the native protein by 100%.

In most of the calculations below, we set the radius of the native protein to 0.4 unit length and that of the denatured protin to 0.6. In order to keep the volume fraction of the denatured protein in a unit volume less than unity, we set $\rho_0 = 1$. In the following numerical calculations, the linear system size is set to 128 unit length.

1. Discretization scheme

The system is discretized into a simple cubic lattice system with $32 \times 32 \times 32$ lattice sites. Differential operators are discretized using central finite difference operators. For example, denoting $\mathbf{r} = (x,y,z)$, the partial derivative of an arbitrary function $f(x,y,z)$ with respect to x is defined as

$$\frac{\partial}{\partial x} f(\mathbf{r}) \rightarrow \frac{1}{h}[f(x+h/2,y,z) - f(x-h/2,y,z)], \qquad (20)$$

where h is the lattice constant $h = 128/32 = 4$. The discretized Laplacian operator includes up to next-next-nearest neighbors:

$$\nabla^2 f(\mathbf{r}_i) \rightarrow \frac{1}{h^2}\Bigg[w_1 \sum_{j \in n_1} f(\mathbf{r}_j) + w_2 \sum_{j \in n_2} f(\mathbf{r}_j) + w_3 \sum_{j \in n_3} f(\mathbf{r}_j) - z f(\mathbf{r}_i) \Bigg], \qquad (21)$$

where n_1, n_2, and n_3 stand for nearest neighbors, next-nearest neighbors, and next-next-nearest neighbors of $\mathbf{r}_i$, respectively. The weighting factors w_i, $(i = 1,2,3)$ are chosen to make the system isotropic. We use the values given by van Vlimmeren and Fraaije [24], that is, $w_1 = 0.294\,726$, $w_2 = 0.235\,425$, $w_3 = 0.175\,818$, and $z = 6$.

2. Calculation of equilibrium states

All the equilibrium properties are calculated from the density fields $\phi^\alpha(\mathbf{r})$ at equilibrium which are obtained by solving the self-consistent equations (14). Given the value of $\epsilon_{\alpha,\beta}$ and the radii of molecules, the interaction parameters $U_{\alpha,\beta}$ and $V_{\alpha,\beta}$ are calculated by numerical integration of $u^{\mathrm{eff}}_{\alpha,\beta}$ according to Eqs. (11) and (12), respectively. In all the cases below, all $\epsilon_{\alpha,\beta}$ but $\epsilon_{D,D}$ are set to 0. The periodic boundary condition is imposed in all directions throughout this study.

In order to solve Eq. (14) with predefined values of bulk densities of protein (ρ_P) and crowding agent (ρ_C), we use a nested functional iteration method. Note that the conserved bulk densities are those of protein, and crowding agent, not those of native protein, denatured protein and crowding agent. Thus we have to adjust two bulk chemical potentials, $\mu^0_P (= \mu^0_N = \mu^0_D)$ and μ^0_C, instead of three, so that the constraints $\rho_P = \langle \phi^N \rangle + \langle \phi^D \rangle$ and $\rho_C = \langle \phi^C \rangle$ are satisfied. The quantity $\langle \phi^\alpha \rangle$ is the spatial average of the density field $\phi^\alpha(\mathbf{r})$ defined by

$$\langle \phi^\alpha \rangle = V^{-1} \int \phi^\alpha(\mathbf{r}) d\mathbf{r}, \qquad (22)$$

where V is the volume of the system. Then the bulk chemical potentials are updated as

$$\mu^0_P \leftarrow \mu^0_P + T \ln[\rho_P / (\langle \phi^N \rangle + \langle \phi^D \rangle)], \qquad (23)$$

$$\mu^0_C \leftarrow \mu^0_C + T \ln[\rho_C / \langle \phi^C \rangle]. \qquad (24)$$

The overall nested functional iteration method proceeds as follows.

(1) Give the initial distribution of $\phi^\alpha(\mathbf{r})$ and initial bulk chemical potentials μ^0_P and μ^0_C. Calculate the potential of mean force $W_\alpha(\mathbf{r})$.

(2) Keeping $W_\alpha(\mathbf{r})$ fixed and using the current μ^0_P and μ^0_C, calculate (temporary) $\phi^\alpha(\mathbf{r})$ through Eq. (14), hence $\langle \phi^\alpha \rangle$. Update μ^0_P and μ^0_C using Eqs. (23) and (24). Neither $\phi^\alpha(\mathbf{r})$ nor $W_\alpha(\mathbf{r})$ are updated at this stage.

(3) Repeat step 2 until μ^0_P and μ^0_C converge.

(4) Using the current $W_\alpha(\mathbf{r})$, and μ^0_P and μ^0_C obtained in the previous step, update $\phi^\alpha(\mathbf{r})$ through Eq. (14). Then update $W_\alpha(\mathbf{r})$ using the new $\phi^\alpha(\mathbf{r})$.

(5) Repeat steps (2)–(4) until $\phi^\alpha(\mathbf{r})$ converge.

TABLE I. Parameter sets of the phase diagrams.

Plane	ρ_P	η_N	$\epsilon_{D,D}$	R_C	R_D
ρ_C-$\epsilon_{D,D}$	0.1	0		0.4	0.6
ρ_C-ρ_P		0	-0.058	0.4	0.6
ρ_C-η_N	0.1		-0.058	0.4	0.6
ρ_C-R_C	0.1	0	-0.058		0.6
ρ_C-R_D	0.1	0	-0.058	0.4	

III. NUMERICAL RESULTS

A. General remarks

Preliminary calculations of equilibrium states were conducted with various parameter values for ρ_P, ρ_C, $\epsilon_{D,D}$ as well as different initial conditions. It was found that there are two different phases. One is the uniform phase (U phase). The other is the phase in which there are aggregates of the denatured proteins and the remaining region is mostly uniform. In the latter phase, the lowest free energy state has one spherical aggregate. We call it the A_D phase. The phase boundary between the U and A_D phases can be determined by the relative spatial deviation, $\mathcal{D}_N [= \sqrt{\langle (\phi^N - \langle \phi^N \rangle)^2 \rangle} / \langle \phi^N \rangle]$, of the native protein density field $\phi^N(\mathbf{r})$. The system is in the uniform phase when $\mathcal{D}_N < 10^{-8}$ and in the A_D phase otherwise. The stability of the native protein can be monitored by the fraction of the native protein $f_N = \langle \phi^N \rangle / \rho_P$.

In the following phase diagrams, we always set $T = 1$, $\eta_D = 0$, and $R_N = 0.4$, and $\epsilon_{\alpha,\beta} = 0$ except for $\epsilon_{D,D}$. Other

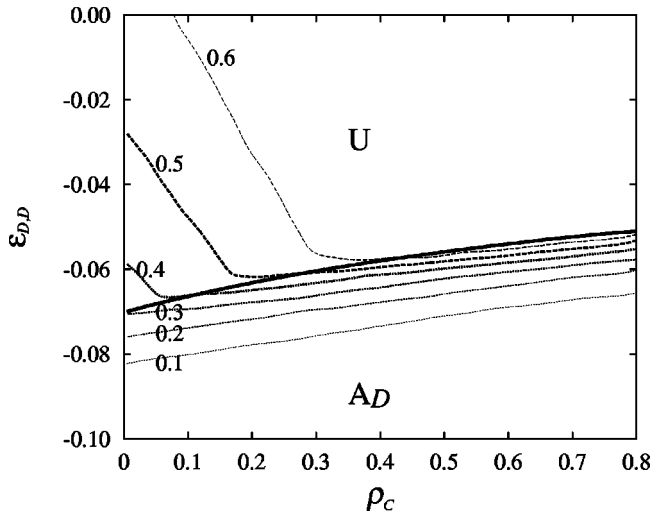

FIG. 2. Phase diagram on the ρ_C-$\epsilon_{D,D}$ plane with $\rho_P=0.1$. The thick solid line indicates the phase boundary. The contour of the fraction of the native protein, f_n, is shown as dashed and dotted lines with its values. U indicates the uniform phase and A_D the phase where the denatured proteins form an aggregate. Units: ρ_C and ρ_P in $[1/(R_N/0.4)^3]$; $\epsilon_{D,D}$ in [T].

parameters are summarized in Table I.

In most of the phase diagrams below, the bulk (number) density of the crowding agent (ρ_C) is varied from 0 to 0.8. When $R_C=0.4$ and $\rho_C=0.8$, the bulk volume fraction of the crowding agent is equal to 0.215. When $R_N=0.4$, $R_D=0.6$, and $\rho_P=0.1$, the bulk volume fraction of the protein resides in the range between 0.0268 and 0.0905. Thus the total volume fraction of molecules in the systems investigated roughly corresponds to that of the living cell, which is 0.2–0.3.

We first drew the phase diagram on the ρ_C-$\epsilon_{D,D}$ plane (Fig. 2) to determine the range of ρ_C and $\epsilon_{D,D}$ that can cause the aggregation of the denatured protein. From Fig. 2, we see that the system crosses the phase boundary when $\epsilon_{D,D}=-0.07$ to -0.05. This phase diagram exhibits two typical behaviors of the system. First, aggregation is enhanced as the bulk density of the crowding agent (ρ_C) increases. Second, the native protein is stabilized as ρ_C increases as long as the system is in the uniform (U) phase, which is the behavior predicted in the preceding section. In the following, we investigate these effects due to crowding in detail. The phase boundary in Fig. 2 suggests that, in order to study the aggregation of the denatured proteins, it is convenient to set $\epsilon_{D,D}\sim-0.06$ and we use the value $\epsilon_{D,D}=-0.058$ below.

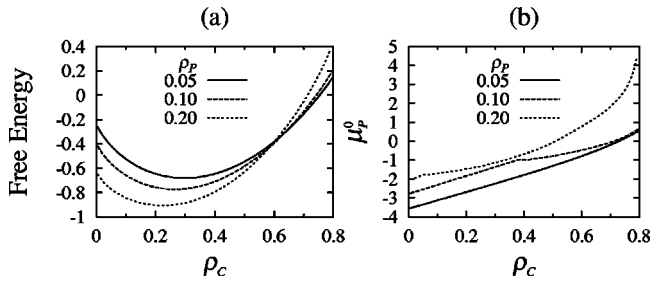

FIG. 4. Free energy and chemical potential of the protein as functions of ρ_C at different ρ_P. Units: free energy and chemical potential in $[T]$; ρ_C and ρ_P in $[1/(R_N/0.4)^3]$.

B. Crowding effects on aggregation and native stability

Although some essential features of the crowded protein solution can be grasped in the ρ_C-$\epsilon_{D,D}$ phase diagram (Fig. 2), it is experimentally difficult to manipulate the phenomenological parameter $\epsilon_{D,D}$. Therefore we choose other variables for phase diagrams. One natural choice is the bulk density of the protein, ρ_P. The ρ_C-ρ_P phase diagram is shown in Fig. 3. Here we again see the two features pointed out above: as ρ_C increases, the aggregation of the denatured protein is enhanced and the native protein is stabilized in the uniform phase. When the protein density ρ_P is sufficiently low, the aggregation is no longer observed. The contour of the relative fraction of the native protein (f_N) in the uniform

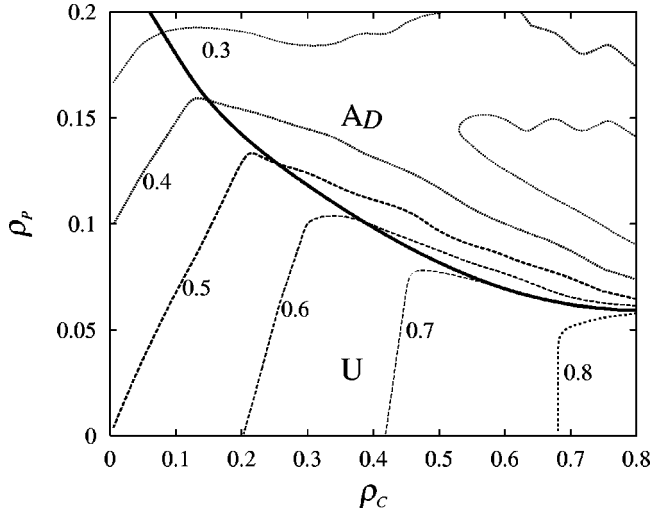

FIG. 3. Phase diagram on the ρ_C-ρ_P plane with the contour of f_N. Units of ρ_C and ρ_P are $[1/(R_N/0.4)^3]$.

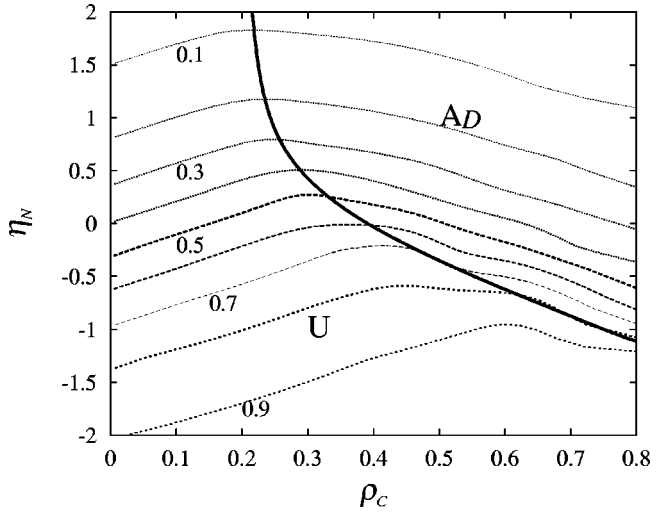

FIG. 5. Phase diagram on the ρ_C-η_N plane with the contour of f_N. Units: η_N in $[T]$; ρ_C in $[1/(R_N/0.4)^3]$.

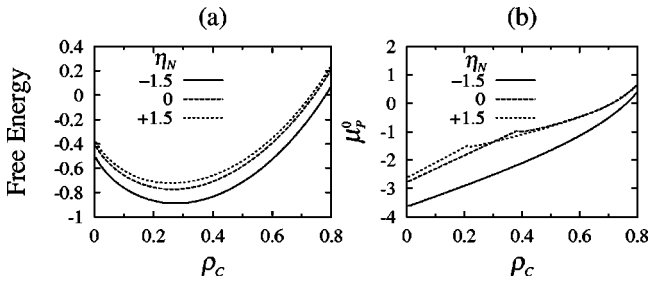

FIG. 6. Free energy and chemical potential of the protein as functions of ρ_C at different η_N. Units: free energy, chemical potential, and η_N in $[T]$; ρ_C in $[1/(R_N/0.4)^3]$.

phase are almost straight lines. This is a result derived in the preceding section. That is, Eq. (15) shows that, with f_N fixed, ρ_P is a linear function of ρ_C. The contour of the relative native fraction (f_N) is distorted in the upper right corner of the ρ_C-ρ_P phase diagram (Fig. 3). The reason for this distorted f_N contour seems to be an artifact due to the small size of the lattice system (i.e., the isotropy of the system is not completely preserved because of the large mesh size). Nevertheless, the free energy of the system as a function of ρ_C changes continuously and smoothly, and it is also concave up [Fig. 4(a)] for $\rho_P=0.05$, 0.1, and 0.2, indicating that the equation of state, Eq. (14), is successfully solved. When the system crosses the phase boundary, a cusp is observed in the chemical potential of the protein, μ_P^0, as a function of ρ_C [Fig. 4(b)], which is expected for a phase transition. The value of μ_P^0 increases as ρ_C increases, and the rate of its increase becomes more stringent for large ρ_C.

The intrinsic free energy η_N of the native protein is also chosen as a variable which can be experimentally manipulated by mutations. The ρ_C-η_N phase diagram (Fig. 5) again reveals the general tendency of the crowding effects: enhancing aggregation, and stabilizing the native protein in the uniform phase. When the native protein is highly destabilized intrinsically ($\eta_N>1$), the phase boundary on the ρ_C-η_N plane becomes nearly vertical. This indicates that the critical value of ρ_C for the aggregation becomes almost independent of η_N when the native protein is highly unstable. Note that when $\rho_C<0.2$ and $\eta_N>1$, the majority of the proteins is denatured ($f_N<0.2$). In this case, the system is already a solution of denatured proteins and crowding agents, but they are dispersed. The free energy changes continuously [Fig. 6(a)]. The chemical potential of the protein, μ_P^0, with different η_N [Fig. 6(b)] exhibits the same tendency as Fig. 4(b).

C. Dependence of protein stability and aggregation on molecular size

So far the radii of molecules have been fixed such that $R_N=R_C=0.4$ and $R_D=0.6$. Here we study how crowding effects vary for different sizes of the crowding agents and denatured proteins. Experimentally, the size of the crowding agent can be readily manipulated by changing the crowding agent to a different one. In the case when the crowding agent is a polysaccharide, the average degree of polymerization may also be manipulated. In the calculation, the radius of the crowding agent R_C was varied from 0.1 to 0.45. The ρ_C-R_C phase diagram [Fig. 7(a)] shows that, with the number density ρ_C fixed, larger crowding agent enhances the aggregation of the denatured proteins and stabilizes the native proteins in the uniform phase more significantly. That is, the increase in R_C imposes effects similar to the increase in the number density ρ_C, which is as expected since a larger crowding agent occupies more space at a constant number density ρ_C. In order to exclude this apparent dependence, we redraw the phase diagram with the volume fraction of the crowding agent, $\tilde{\rho}_C=(4\pi/3)R_C^3\rho_C$, as a variable [Fig. 7(b)]. If the phase boundary is dependent solely on the volume excluded by the crowding agent, it would appear as a vertical line on the $\tilde{\rho}_C$-R_C plane, but this was not found to be the case. Within the range of R_C studied, we see that, at a constant volume fraction $\tilde{\rho}_C$, smaller crowding agents have stronger effects both on the native protein stabilization in the uniform phase and on the aggregation. The increase in R_C at fixed $\tilde{\rho}_C$ reduces the native protein stability in the uniform phase, and tends to prevent aggregation. Thus, the crowding effects on protein stabilization and aggregation are dependent not only on the volume excluded by the crowding agent, but also on the size of the crowding agent. For $R_C>0.45$, the nested functional iterations did not converge to yield reliable solutions of the self-consistent equations of state.

Next we investigate how the size of the denatured protein, R_D, affects the crowding effects. For this purpose, we depict the ρ_C-R_D phase diagram (Fig. 8). We can see the increase in ρ_C enhances the aggregation when R_D is large ($R_D>0.5$). For $R_D>R_N=0.4$, the increase in ρ_C stabilizes the native

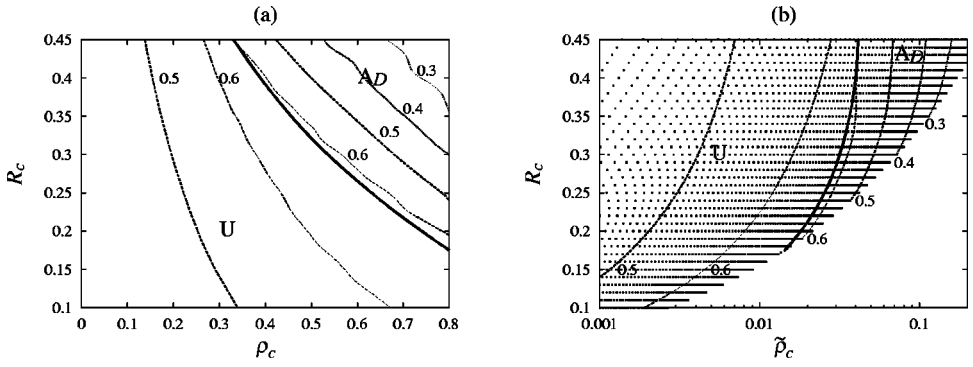

FIG. 7. Phase diagrams on the ρ_C-R_C plane (a) and on the $\tilde{\rho}_C$-R_C plane (b). $\tilde{\rho}_C$ is the volume fraction of the crowding agent. The dotted region in (b) corresponds to a subspace of (a). The contour of f_N is also shown. Units: R_C in $[R_N/0.4]$; ρ_C in $[1/(R_N/0.4)^3]$; $\tilde{\rho}_C$ (dimensionless).

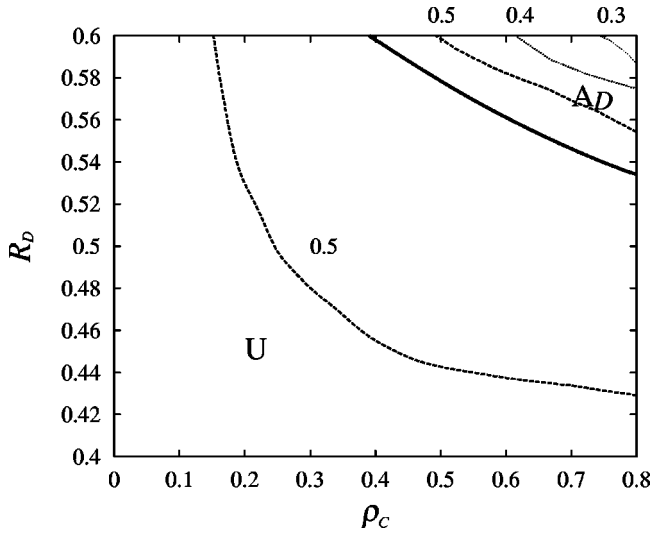

FIG. 8. Phase diagram on the ρ_C-R_D plane with the contour of f_N. Units: R_D in $[R_N/0.4]$; ρ_C in $[1/(R_N/0.4)^3]$.

protein in the uniform phase as in the previous examples. These crowding effects on aggregation and native protein stability are stronger for larger R_D. Although unfolded proteins are always more expanded than the native ones, we may ask what would happen theoretically if it were not the case. We calculated the phase diagram for $R_D < R_N$ in which the increase in the crowding agent density destabilized the native protein in the uniform phase (data not shown). This result clarifies that the effects of macromolecular crowding observed here are caused by the fact that the denatured proteins are larger than the native ones.

IV. DISCUSSION

Regarding the crowding effect on protein stability, Minton [25] proposed a statistical-thermodynamic model based on the McMillan-Mayer theory of multicomponent solutions [26]. He modeled the native protein as a hard sphere and the denatured protein as an ensemble of (convex hulls of) random coils of varying compactness. The main result of Minton's theory is that increasing the concentration of crowding agent leads to the enhanced thermostability of native proteins. The system treated in the Minton theory corresponds to the uniform phase in our case. The ρ_C-$\epsilon_{D,D}$ (Fig. 2), ρ_C-ρ_P (Fig. 3), and ρ_C-η_N (Fig. 5) phase diagrams all show that, in the uniform phase and with sufficiently low ρ_P, the increase of ρ_C is accompanied by the increase of the fraction of native protein, thus our results are in agreement with the Minton theory. A qualitatively similar result was also obtained by Zhou and Hall [27] that the addition of crowding agent whose size is comparable with the protein enhances the stability of the native protein. They also modeled the native protein as a hard sphere, and the denatured protein was modeled as a chain-molecule by a thermodynamic perturbation theory. However, as pointed out by Minton [25], the perturbation theory of Zhou and Hall does not actually treat the chain connectivity of the denatured protein, and the degree of native stabilization by the crowding agent measured in terms of the equilibrium constant $f_N/(1-f_N)$ was predicted to be unrealistically high ($\sim 10^9$) [27]. In the present theory, the value of $f_N/(1-f_N)$ in the uniform phase can be calculated from Eq. (15). Using the same parameters as in the ρ_C-ρ_P phase diagram (Table I) with $\rho_P = 0.05$; as the bulk density of the crowding agent ρ_C changes from 0 to 0.9, the equilibrium constant $f_N/(1-f_N)$ monotonically increases from 0.8 to 6.5. These values are of the order similar to that obtained by Minton [25] who predicted the maximum increase to be $\sim$10–100. This partial agreement may be attributed to the treatment of the conformation of the denatured protein: The convex hull of random coils as a model of the denatured state introduced by Minton somewhat resembles the present hard-sphere model. Difference between the results in the present theory and those of the Minton theory may be due to the different size of the denatured protein relative to the native protein, and due to the attraction between the denatured proteins in the present theory.

There is an experimental observation that the native stability of *G*-actin was increased in the presence of polyethylene glycol 6000 [28]. However, experimental investigations to test the theoretical predictions regarding stability in general have been difficult because actual systems exhibit quasi-irreversible aggregation of denatured proteins [25]. This difficulty was partly overcome by Eggers and Valentine [29] who investigated thermal stability of four kinds of proteins (lysozyme, α-lactalbumin, metmyoglobin, and apomyoglobin) by the sol-gel method in which proteins are encapsulated in a silica glass matrix. The silica glass matrix confines proteins inside its pores and thus exerts the molecular confinement effect similar to macromolecular crowding effect [5]. The sol-gel method is useful, in that encapsulating proteins in different pores of the glass prevents aggregation of proteins. Eggers and Valentine [29] reported that three out of four proteins they studied retained nativelike structures and exhibited increased thermostability. One exception was apomyoglobin which was found to be totally unfolded inside the pore. This was later shown to be due to the perturbed water structure inside the pores, hence weakened hydrophobic effect [30]. Recent molecular simulations of a β-hairpin forming peptide confined in a pore also exhibited enhanced stability of compact native states [31]. Thus it appears that the theoretical prediction is validated by experiment as well as by simulations in the majority of the cases, although the above experimental results suggest the limitation of the present theoretical model which does not include the hydration effect explicitly.

The theory presented in this paper can treat not only the uniform phase but also the aggregation phase. Van den Berg *et al.* [10] studied the effect of macromolecular crowding on the refolding of oxidized as well as reduced lysozyme using four kinds of crowding agents, namely, ficoll 70, dextran 70, bovine serum albumin, and ovalbumin. They reported that the aggregation of the reduced lysozyme was enhanced with increased concentration of any crowding agent while refolding of oxidized lysozyme was hardly affected. In the reduced lysozyme, the native disulphide bonds are broken and therefore the protein structures at stages of refolding are expected to be more expanded compared to the oxidized lysozyme. The ρ_C-R_D phase diagram (Fig. 8) shows that the denatured

protein with larger size is more liable to aggregation. Thus, disregarding other complications, the present model with the above steric consideration alone can qualitatively describe the difference between the refolding of oxidized and reduced lysozyme.

Recently, molecular simulation studies on the competition between protein folding and aggregation have been performed [15–17]. These studies with detailed representations of protein molecules revealed the sensitivity of the native and aggregate conformations [15,17]. Dima and Thirumalai [17] also obtained a phase diagram of a system containing two protein molecules. Simulations of a single peptide molecule confined in a pore have also been performed to examine molecular confinement effects on protein folding [31], in which native as well as denatured conformations of the peptide were analyzed in detail. However, with these detailed models, systems consisting of only two to eight [16] proteins were treated, possibly because of computational limitation. The model presented in this paper, on the other hand, can treat systems containing infinite number of proteins together with crowding agents at the cost of microscopic details. Therefore, the present model serves to study large scale properties of protein aggregation to complement molecular simulation studies. We note that it is possible, at least in principle, to incorporate amino acid sequence information to the density functional formulation by combining the path integral and self-consistent field descriptions as is done in simulations of polymer solutions and melts [22,32]. However, the uniqueness of the native protein conformation complicates the use of such formulation.

In the present paper, we have considered only the excluded volume interaction between proteins and crowding agents, except for the attraction between the denatured proteins. It should be stressed that the term ''excluded volume'' is defined in this paper as the volume excluded by a macromolecule (protein or crowding agent) to other macromolecules. The excluded volume of a protein in the present meaning should be distinguished from the partial molar volume. Although partial molar volume of proteins may decrease upon unfolding due to the change of hydration structure, the excluded volume always increases upon unfolding because unfolded structures of protein cannot be more compact than its native structure. As discussed above, the present highly simplified picture of crowded protein solutions is at least qualitatively in good agreement with some experimental observations, and thus can elucidate the importance of the excluded volume effects in crowded environments. Further, elaboration of the theory to include more realistic interactions such as electrostatic ones is straightforward if the interactions are isotropic. Since macromolecules (proteins and crowding agents) and solvent (water) molecules largely differ in their sizes, explicit treatment of hydration structures may be difficult and is left for future studies.

V. CONCLUSION

We have proposed a density functional theory that describes crowded protein solutions. Neglecting solvent water molecules, the present theory can handle systems consisting of proteins and macromolecules of comparable sizes. By solving the self-consistent equations of state, it was shown that uniform and nonuniform equilibrium states can be successfully obtained for wide ranges of parameter sets. Various phase diagrams suggest that macromolecular crowding (1) enhances aggregation of the denatured proteins, (2) stabilizes the native protein as long as the system is uniform. The present theoretical results agree qualitatively with some experimental results. An extension of the present theory to treat dynamical phenomena will be given in a forthcoming paper [33].

ACKNOWLEDGMENT

The authors thank Professor Yuji Goto for helpful discussion and for providing some references.

[1] A.B. Fulton, Cell **30**, 345 (1982).
[2] S.B. Zimmerman and A.P. Minton, Annu. Rev. Biophys. Biomol. Struct. **22**, 27 (1993).
[3] A.P. Minton, Curr. Opin. Biotechnol. **8**, 65 (1997).
[4] A.P. Minton, Curr. Opin. Struct. Biol. **10**, 34 (2000).
[5] A.P. Minton, J. Biol. Chem. **276**, 10577 (2001).
[6] R.J. Ellis, Curr. Opin. Struct. Biol. **11**, 114 (2001).
[7] R.J. Ellis, Trends Biochem. Sci. **26**, 597 (2001).
[8] S. Asakura and F. Oosawa, J. Polym. Sci. **33**, 183 (1958).
[9] A.P. Minton, Methods Enzymol. **295**, 127 (1998).
[10] B. van den Berg, R.J. Ellis, and C.M. Dobson, EMBO J. **18**, 6927 (1999).
[11] B. van den Berg, R. Wain, C.M. Dobson, and R.J. Ellis, EMBO J. **19**, 3870 (2000).
[12] G. Rivas, J.A. Fernández, and A.P. Minton, Proc. Natl. Acad. Sci. U.S.A. **98**, 3150 (2001).
[13] D.M. Hatters, A.P. Minton, and G.J. Howlett, J. Biol. Chem. **277**, 7824 (2002).
[14] A.L. Fink, Fold Des **3**, R9 (1998).
[15] P.M. Harrison, H.S. Chan, S.B. Prusiner, and F.E. Cohen, J. Mol. Biol. **286**, 593 (1999).
[16] A.V. Smith and C.K. Hall, J. Mol. Biol. **312**, 187 (2001).
[17] R.I. Dima and D. Thirumalai, Protein Sci. **11**, 1036 (2002).
[18] J.P. Hansen and I.R. McDonald, *Theory of Simple Liquids*, 2nd ed. (Academic Press, London, 1986).
[19] K. Kawasaki, *Non-equilibrium and Phase Transition* (Asakura Shoten, Tokyo, 2000), in Japanese.
[20] S.A. Safran, *Statistical Thermodynamics of Surfaces, Interfaces, and Membranes*, Frontiers in Physics Vol. 90 (Addison-Wesley, Reading, MA, 1994).
[21] L.D. Landau and I.M .Lifshitz, *Statistical Physics*, 3rd ed. (Iwanami Shoten, Tokyo, 1980), Pt. I, Japanese translation.
[22] T. Kawakatsu, *Foundations of Polymer Physics* (Saiensu Sha, Tokyo, 2001), in Japanese.
[23] A. Fersht, *Structure and Mechanism in Protein Science* (Freeman, New York, 1999).

[24] B.A.C. van Vlimmeren and J.G.E.M. Fraaije, Comput. Phys. Commun. **99**, 21 (1996).
[25] A.P. Minton, Biophys. J. **78**, 101 (2000).
[26] W.G. McMillan, Jr. and J.E. Mayer, J. Chem. Phys. **13**, 276 (1945).
[27] Y. Zhou and C.K. Hall, Biopolymers **38**, 273 (1996).
[28] R.L. Tellam, M.J. Sculley, and L.W. Nichol, Biochem. J. **213**, 651 (1983).
[29] D.K. Eggers and J.S. Valentine, Protein Sci. **10**, 250 (2001).
[30] D.K. Eggers and J.S. Valentine, J. Mol. Biol. **314**, 911 (2001).
[31] D.K. Klimov, D. Newfield, and D. Thirumalai, Proc. Natl. Acad. Sci. U.S.A. **99**, 8019 (2002).
[32] J.G.E.M. Fraaije, B.A.C. van Vlimmeren, N.M. Maurits, M. Postma, O.A. Evers, C. Hoffmann, P. Altevogt, and G. Goldbeck-Wood, J. Chem. Phys. **106**, 4260 (1997).
[33] A.R. Kinjo and S. Takada (unpublished).