From Surface Self-Assembly to Crystallization: Prediction of Protein Crystallization Conditions

2 AUTHORS:

Yanwei Jia
University of Macau
27 PUBLICATIONS   272 CITATIONS

SEE PROFILE

Xiang-Yang Liu
National University of Singapore
69 PUBLICATIONS   1,857 CITATIONS

SEE PROFILE

From Surface Self-Assembly to Crystallization: Prediction of Protein Crystallization Conditions

Yanwei Jia and Xiang-Yang Liu*

Department of Physics, National University of Singapore, 2 Science Drive 3, Singapore 117542

Received: July 1, 2005; In Final Form: January 20, 2006

A new criterion based on surface and volume diffusion kinetics was established to predict protein crystallization. Similar to the layer-by-layer crystal growth process of protein, the kinetics of the two-dimensional self-assembly of protein at the aqueous solution surface provides a convenient and reliable way to estimate the surface integration and the volume transport during protein crystallization. Both the surface and diffusion kinetics were estimated based on the protein self-assembly at the air/solution interface, which can be obtained by measuring the surface tension. A crystallization coefficient is found to provide an effective and reliable criterion to predict protein crystallization conditions. This criterion has been applied to lysozyme, concanavalin A and BSA crystallization, and it turns out to be very successful and more reliable than the second virial coefficient criterion.

Introduction

Protein crystallization has attracted much attention and has remained a hot topic in recent years because of its wide application in drug delivery and protein structure determination.[1−6] Protein structural data are necessary to explore biofunctionalities, drug design, disease treatment, and so forth.[7,8] Between the two main methods (NMR and X-ray crystallography) available for the determination of protein structure, X-ray crystallography has the advantage of a much higher resolution and is more cost-effective than NMR.[9] To take advantage of X-ray crystallography, sufficiently large and high-quality single crystals are essential. The key problem associated with this is that in many cases proteins will aggregate into an amorphous phase instead of the crystalline phase.[10,11] The above-mentioned issues have become a bottleneck for the determination of protein structure. To overcome this obstacle, a large scale of screening is conducted to search for the crystallization conditions, which can be extremely costly and time-consuming. Apart from this, the crystallization conditions may not be obtained because crystallization depends on a variety of factors in addition to external conditions.

To predict protein crystallization conditions without resorting to screening experiments, the second virial coefficient, B_{22}, has been used widely.[12−15] The value of the second virial coefficient is related to the effective interaction between a pair of macromolecules in solution, a positive value reflecting predominantly repulsive interactions and a negative value indicating attractive interactions. A necessary condition for protein crystallization is that B_{22} lies in a crystallization parameter window, $-8 \times 10^{-4} < B_{22} < -2 \times 10^{-4}$ mL mol/g^2. This criterion has its advantages because it requires only a small amount of protein and gives a discriminating response. However, the B_{22} criterion determines biomacromolecule crystallization only partially. Being macromolecules with complicated shapes and orientations, the crystallization of protein involves a nucleation and growth process, determined to a large extent by kinetics. Kinetics refers to the way protein molecules move in a solution, the rate at

which they are transported, and the way they are incorporated in protein crystals at the crystal surface. How protein molecules are incorporated into the "embryo" surface of protein crystal is of significance in controlling protein nucleation. Therefore, the kinetics of incorporating protein molecules into the kink site at the surface of protein crystal "embryos" should be taken into account. The crystallization window provided by the second virial coefficient disregards kinetic and other factors, which are unrelated to intermolecular interactions but nevertheless largely influence crystallization.

However, the study of the protein crystallization interfacial kinetics still remains very limited. In our recent research, it has been found that the protein self-assembly at the aqueous solution surface has the same structure as in the crystal structure.[16] This 2D assembly can act as a self-template for protein 3D crystallization. On the basis of this study, it should be possible to use the protein 2D assembly kinetics as a probe to study the kinetics of protein crystallization.

The aims of the investigations discussed here are to (i) study the kinetics of a surface assembly in a model protein system; (ii) apply the kinetics of a 2D assembly on the kinetics of protein 3D crystallization and obtain insight into the relationship between the kinetics and the protein crystallization conditions; and (iii) propose a new criterion for predicting protein crystallization based on kinetics and test the suitability of the criterion on other proteins.

Theory. *1. Protein Three-Dimensional Crystallization Kinetics.* Nucleation is a process of the creation of critical size of nuclei, which results from overcoming the nucleation barrier, in the context of the fluctuation and growth of embryos (Figure 1). (Embryos are metastable clusters of structural units with a broad distribution in size.) The molecular process of nucleation can be regarded as follows: the constituent atoms or molecules in the ambient phase may on collision join into groups of two or more particles forming dimers, trimers, tetramers, and so on. Before the embryos can reach a critical radius, r_c, they are unstable even when a positive thermodynamic driving force $\Delta\mu/k_BT$ is applied. ($\Delta\mu = \mu^f - \mu^s$, μ^f and μ^s are the chemical potential of the solute molecules in the fluid phase and in the

* Corresponding author: E-mail: phyliuxy@nus.edu.sg.
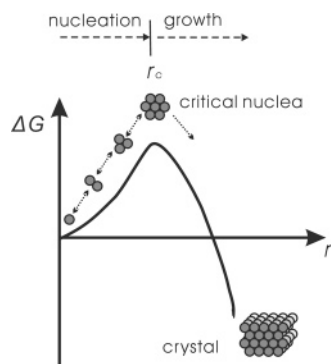
Figure 1. Illustration of the nucleation process. The monomers in solution collide and join together to form dimer, trimer, and higher order clusters. These clusters are defined as embryos. Only after the embryo is big enough to overcome the nucleation barrier can it grow up to a crystal. Otherwise, the embryo dissolves spontaneously.
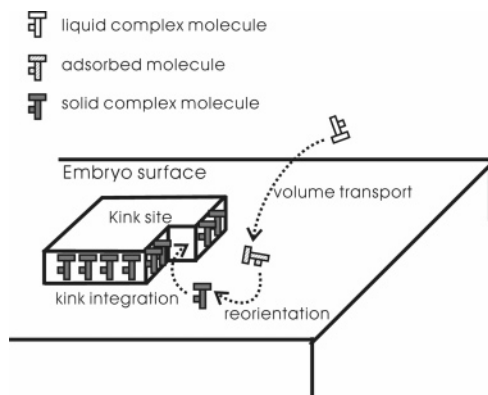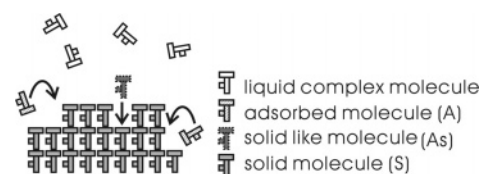

Figure 2. Illustration of the process of protein crystallization and aggregation. The molecules are transported from the bulk to the kink site of the embryo. Before they are incorporated into the kink site, the protein sheds small molecules adsorbed on it. At the same time, the macromolecules rearrange themselves to find their optimal orientations and conformations, as is required in the solid state.

solid phase, respectively; k_B is Boltzmann's constant and T is the temperature.) To reach r_c, an energy barrier, the so-called nucleation barrier, ΔG^*, needs to be overcome. When the nucleation barrier is overcome, the second stage of the phase transition begins: growth.[17,18]

As far as the interfacial process is concerned, growth units should be delivered from the bulk to the embryo surface and incorporated into kink sites (surface kinetics). This implies that the nucleation kinetics is determined mainly by overcoming the nucleation barrier, volume transport, and surface kinetics. In contrast to simple molecules, once complex macromolecules such as proteins become incorporated in the crystal structure, they acquire in addition a conformation and orientation at a certain site, as a consequence of the 3D symmetry. Statistically, liquid complex molecules have very different orientations and conformations from solid complex molecules. Only about 1% of the liquid molecules have the proper orientation and can be incorporated into kink sites.[19] Upon incorporation into the kink sites, liquid complex molecules should simultaneously acquire the same orientation and conformation as the solid molecules. In detail, intefacial kinetics includes the following parallel and cooperative steps: desolvation, the acquisition of the solid unitlike orientation and conformation, and surface integration (Figure 2).

At kink sites, because of the ordering of the fluid layers adjacent to the crystal surface, some adsorbed growth units have roughly the same orientations and conformations as required

Figure 3. Illustration of the protein molecular surface integration on the embryo surface. The solid unit-like molecules adsorbed at the interface can be incorporated directly, while in general the adsorbed molecules need to rearrange their conformations and orientations to be incorporated into the kink sites.

for the growth of the crystal surface (Figure 3).[20] Units of this type are defined as A_s units. However, the orientations or conformations of the majority of the adsorbed molecules are different from what is required for growth. Units of this type are defined A units. To transform an A_s unit into a solid unit S at a kink, the A_s unit needs to overcome a barrier ΔG^*_{kink} (>0, corresponding to desolvation free energy), defined as

$$A_s \xleftarrow{\Delta G^*_{kink}} S$$
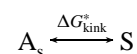
It follows that the relaxation time for such a transformation is

$$\tau_1 \approx \frac{1}{\nu_1} \exp(\Delta G^*_{kink}/k_B T)$$

where ν_1 denotes the frequency of thermal vibration of A_s units. However, the transformation from an A to an S unit is not so straightforward. Before an A unit can be incorporated into the crystal, it needs to be transformed into an A_s unit, for which it has to overcome an extra barrier $\Delta G^*_{rearrange}$ (≥ 0), as

$$A \xleftarrow{\Delta G^*_{rearrange}} A_s \xleftarrow{\Delta G^*_{kink}} S.$$
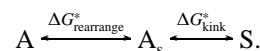
The relaxation time for this transformation is then given by

$$\tau_2 \approx \frac{1}{\nu_2} \exp(\Delta G^*_{rearrange}/k_B T)$$

where ν_2 denotes the frequency of thermal vibration of A units.

In the above analysis, τ_1 and τ_2 are related to the surface kinetics. This relaxation time is defined as surface integration time, $\tau_k = \tau_1 + \tau_2$. For a general adsorbed molecule $\tau_2 \neq 0$, and the molecular rearrangement time at the kink site cannot be neglected. The volume transport time for a fluid molecule to the embryo surface, τ_d, is inversely proportional to the self-diffusion coefficient of a protein.

Our recent study indicates that in the case of protein crystallization, when the volume transport is much faster than surface integration, amorphous aggregation is kinetically favorable above crystallization.[11] As illustrated in Figure 4a, if the surface integration is a rapid process, then the adsorbed molecules have enough time to rearrange before the arrival of additional molecules. Protein molecules aggregate into an ordered compact structure, which eventually grows into crystal. On the contrary, if the transport of growth units to the surface of a protein crystal is a much faster process, the first coming molecule does not have enough time to find its optimal orientation and conformation before the next molecule arrives. As a consequence, the formation of an amorphous phase becomes a kinetically more favorable process than the formation of a crystalline phase that is ordered in three dimensions. In this case, the adsorbed molecules will aggregate without

Prediction of Protein Crystallization Conditions

J. Phys. Chem. B, Vol. 110, No. 13, 2006 **6951**
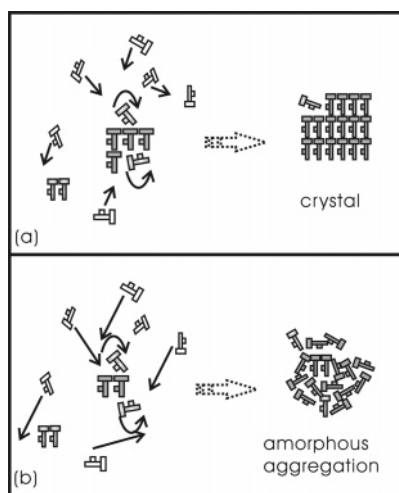
Figure 4. Kinetics of protein crystallization and amorphous aggregation. (a) Formation of a crystal when the rate constant of the molecular volume transport is comparable with that of a molecular rearrangement. (b) Formation of amorphous aggregation when the volume transport is much faster than the molecular rearrangement.

assuming an ordered arrangement. This type of nucleus will eventually give rise to amorphous aggregation (Figure 4b).

Essentially, this kinetically favorable formation of amorphous aggregation or crystallization is also an energetically favorable process of protein crystallization because

$$\tau \propto \exp\!\left(\frac{\Delta G}{k_{\mathrm{B}} T}\right)$$

where ΔG is the kinetic energy barrier from one state to another in the process of protein crystallization. A faster relaxation or shorter relaxation time τ corresponds to a lower energy barrier. If the diffusion is faster than the surface integration, then the energy barrier for the diffusion would be lower than that of kink integration. This makes the diffusion process easier than the subsequent surface integration process. As a consequence, amorphous aggregation as a semistable state exists before ordered crystal structure can occur.

To anticipate the probability of a nucleus to grow into a crystal or an amorphous aggregate, we introduce a crystallization coefficient, ξ_{c}, which is defined as the ratio between the diffusion rate and surface integration rate

$$\xi_{\mathrm{c}} = \frac{k_{\mathrm{d}}}{k_{\mathrm{k}}} = \frac{\tau_{\mathrm{k}}}{\tau_{\mathrm{d}}} \tag{1}$$

where $k_{\mathrm{d}} = 1/\tau_{\mathrm{d}}$ is the rate constant of diffusion and $k_{\mathrm{k}} = 1/\tau_{\mathrm{k}}$ is the rate constant of surface integration. Obviously, the larger the ξ_{c}, the faster the diffusion is as compared with surface integration and the more likely it becomes for a nucleus to form an amorphous aggregate. However, in an experiment on protein 3D crystallization it is hard to measure τ_{k} and τ_{d} directly. Because protein crystallization follows the 2D nucleation mechanism, the 2D assembly of protein at the air/solution interface can be employed to estimate τ_{k} and τ_{d}.

2. Protein Two-Dimensional Interface Assembly Kinetics. The growth of protein crystals is normally governed by a layer-by-layer mechanism.[18,21−25] This means that the growth of 3D protein crystals is accomplished by a sequence of 2D nucleation and spread to a 2D crystal layer on an existing crystal surface (Figure 5). Between the two steps, the formation of the "2D crystal" can be used as an analogy of the formation of a new
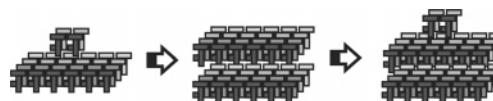
Figure 5. Layer-by-layer protein crystal growth mechanism. The growth of 3D protein crystals is accomplished by a sequence of 2D nucleation and spread to a 2D crystal layer on an existing crystal surface.
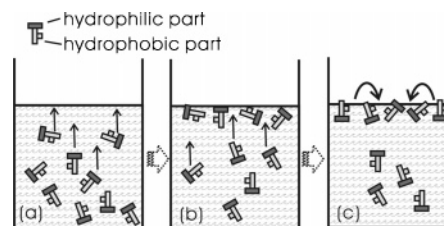
Figure 6. Protein 2D interface assembly kinetics. (a) Molecules diffuse to the air/solution interface immediately after mixing the solution. (b) When more molecules cover the surface, the subsequently impinging molecules need to penetrate a potential energy barrier to be adsorbed at the surface. (c) When the surface is almost fully covered, the molecules at the surface need to rearrange their orientations and conformations to make space for the arriving molecules.

layer of a crystal on an existing growing crystal surface. In our recent research, it has been found that the protein self-assembly at the surface of aqueous solution has the same structure as in the crystal structure.[16] This 2D self-assembly acts as a self-template for the 3D protein crystallization. This approach provides us with a new and simple method of studying the interfacial kinetics of protein crystallization.

Like all amphiphilic molecules, proteins tend to self-assemble at the air/solution interface and therefore lower the surface tension.[26] The dynamics of surface tension reflects the kinetics of the 2D self-assembly of proteins at the interface.

The rate of reduction of surface tension is determined by three consecutive or concurrent processes: (a) a diffusion of protein molecules to the interface; (b) the penetration of protein molecules through the interface film from subsurface to surface; and (c) molecular rearrangements of adsorbed molecules in the film (Figure 6).[26] Compared to the kinetics of 3D protein crystallization, the first step corresponds to volume transport, while the other two steps resemble the surface integration of the 2D assembly during nucleation and growth.

A useful analysis of the kinetics of protein surface self-assembly derives from the work of Ward and Tordai,[27] who considered the effects of diffusion from the bulk liquid to the surface and the energy barrier, which the molecule must overcome in order to be adsorbed. At the very beginning of the surface assembly, there is no energy barrier between the surface and the subsurface. In this step, diffusion is the rate-determining factor (Figure 6a). The surface tension is expressed as

$$\gamma(t) = \gamma_0 - 2C_{\mathrm{p}} k_{\mathrm{B}} T\!\left(\frac{D}{3.142}t\right)^{1/2} \tag{2}$$

where $\gamma(t)$ is the surface tension at time t, γ_0 is the solvent surface tension, C_{p} is the bulk protein concentration, and D is the diffusion coefficient of protein molecules. In this step, then the plot of γ against $t^{1/2}$ will be linear.[28,29] This happens only at the initial stage of the surface assembly.

At higher surface coverage, there are energy barriers to penetration of molecules into the surface. In this step, penetration is the rate-determining factor (Figure 6b). When the surface is almost fully covered, the reduction of surface tension will be determined by the molecular rearrangement in the film. For the situation where there is an energy barrier to assembly, eq 2 can

6952 *J. Phys. Chem. B, Vol. 110, No. 13, 2006*

Jia and Liu

be modified to give[28-30]

$$\ln \frac{d\Pi(t)}{dt} = \ln(\bar{k}C_p) - \Pi(t)\Delta A/k_B T \qquad (3)$$

where $\Pi(t) = \gamma_0 - \gamma(t)$ is the surface pressure at time t, $\bar{k}$ is a constant related to adsorption ability, and ΔA represents the mean area created in the film in order to adsorb a protein molecule. If ΔA is assumed to be a constant, then a plot of $\ln(d\Pi(t)/dt)$ versus Π should be linear with a slope giving ΔA. The values of ΔA, $\bar{k}$, and C_p can be considered to be constants within each rate-determining process, but these values will change when the rate-determining process is changed. Therefore, each successive linear part of $\ln(d\Pi(t)/dt)$ versus Π corresponds to each kinetic step.

When protein molecules penetrate into the interfacial region, they affect the surface free energy. Because in polymer adsorption the first layer effects are dominant in determining Π, changes in Π provide a convenient way of monitoring the penetration into the surface and configurational rearrangements of the adsorbed protein molecules. The rates of these processes can be analyzed by the first-order equation[29,30]

$$\ln\left(1 - \frac{\Pi(t)}{\Pi_e}\right) = -\frac{t}{\tau} \qquad (4)$$

where Π_e and $\Pi(t)$ are the surface pressure values in the equilibrium state and at any time, t, respectively, and τ is the relaxation time. Therefore, the rate constants, $k (= 1/\tau)$, can be estimated from the slopes of $\ln(1 - \Pi(t)/\Pi_e)$ versus t. Besides, the relaxation time of diffusion, τ_d, the penetration, τ_p, the self-rearrangement, τ_r, and therefore also the surface integration time, $\tau_k = \tau_p + \tau_r$, can be obtained.

Experiments

Materials. Hen egg white lysozyme of molecular weight, M_w, about 14.3 KDa (purifed six times by crystallization) was purchased from Seikagaku (MA). BSA with M_w about 66 KDa and concanavalin A (con A) with M_w about 23 KDa were purchased from Sigma. All of these proteins were used without further purification. Tris-Ac, sodium acetate, sodium chloride, and ammonium sulfate, all of analytical grade, were purchased from Merck (Darmstadt, Germany). High purity deionized water (18.2 MΩ) produced by a Millipore Milli-Q system (0.22 um) was used for preparing buffers, which were used as solvents for proteins and salts. The pH was adjusted by the corresponding acid and hydroxide within the accuracy of ±0.1. The stock solutions of proteins and salts were prepared and refrigerated at 4 °C for further use.

Methods. The kinetics of protein adsorption at the air−water interface was studied by the cylindrical Wilhelmy plate using a K14 Krüss tensiometer. The tensiometer trough (diameter 2 cm and height 1 cm) was placed in a circulating water bath maintained at 23 ± 0.5 °C. The solutions of 3 mL were mixed with protein and salt stock solutions right before the measurement. The change of the interfacial tension was recorded at 1-min intervals after mixing the solution for the maximum duration that the machine can sustain, which is 24 h. The experimental sample chamber was saturated with pure water vapor to keep consistently a humid environment. All of the containers used in this experiment were cleaned by chromic acid to eliminate potential organic contaminations.

In the data analysis, to obtain the equilibrium surface tension from the dynamic surface tension data, the "end points" or "last
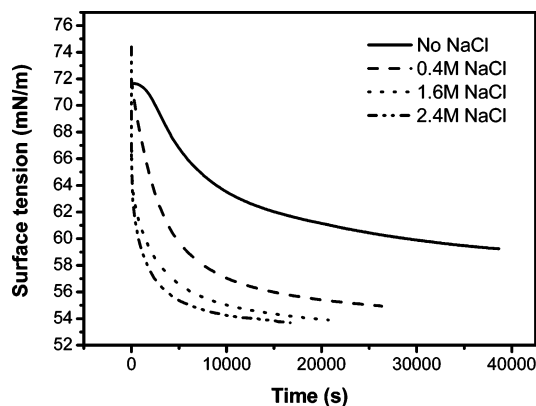
Figure 7. Time dependence of surface tension at the air/water interface for 1 mg/mL lysozyme solutions with different sodium chloride concentrations.

measurements" method was employed.[31] The equilibrium surface tension is calculated by averaging a certain number of data points collected at the end of each run until the standard deviation is less than 0.01 mN/m. In this study, 10 points were used.

To determine the limit of the three steps in the interface adsorption, we divided the curve of $\ln(d\Pi/dt)$ versus Π by a regression analysis into successive linear parts, each corresponding to a kinetic step. The limits of these steps were optimized and adjusted in order to be in agreement with those found by using the representation $\ln(1 - \Pi/\Pi_e)$ versus t. The resulting a diffusion step agreed with the corresponding step obtained from the plot of γ versus $t^{(1/2)}$. The limits were adjusted so that all of the fitting coefficients attained their largest values in all of the linear regression fittings.

Results and Discussion

The protein interface adsorption kinetics of 1 mg/mL lysozyme was studied by the dynamics of surface tension at a temperature of 23 °C. Figure 7 shows changes in the surface tension of lysozyme with different concentrations of sodium chloride. It appears that the presence of sodium chloride resulted in a decrease of the time (t_e) that is necessary to reach a steady state as well as the equilibrium surface pressure. The time t_e dropped from 11 h for pure lysozyme to 5 h for highly concentrated, 2.4 M, sodium chloride samples. This decrease corresponds to the increase of the diffusion rate of the protein molecules in the sodium chloride solutions, as indicated from Figure 8. As mentioned above in section II, the dynamics of surface tension indicates the kinetics of protein assembly at the air−water interface. It follows the three consecutive steps: diffusion, penetration, and rearrangement. As suggested by Tornberg and Graham,[28,29] the linear part in the commonly used plot $\gamma(t) = f(t^{(1/2)})$ (Figure 8) (refer to eq 2) illustrated the diffusion determined limit and $\ln d\Pi/dt = f(\Pi)$ (Figure 9) (refer to eq 3) revealed three steps for the interface assembly of each sample. The resulting slope of the curve $\ln(1 - \Pi/\Pi_e) = f(t)$ gave the rate constant for each step (Figure 10) (refer to eq 4).

Figure 11 shows the relationship between the concentration of sodium chloride and the rate constants of diffusion, k_d, penetration, k_p, and rearrangement, k_r. It appears that the rate constant of each step depends on the salt concentration. With the increase of sodium chloride, the diffusion rate increases dramatically. The changes in the rate constants of penetration and rearrangement were much less significant, although a little higher k_p value was obtained for the high salt sample. These changes can be explained by the electrical double layer theory.

Prediction of Protein Crystallization Conditions

J. Phys. Chem. B, Vol. 110, No. 13, 2006 **6953**
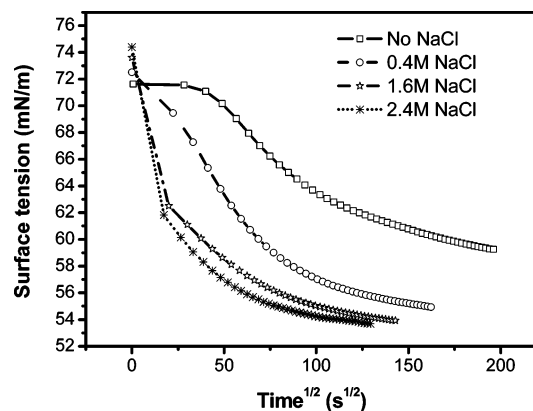
Figure 8. Surface tension as a function of the square root of time for 1 mg/mL lysozyme. The linear part shows the limit of the diffusion controlled step in the three-step process of the protein assembly at the aqueous solution surface.
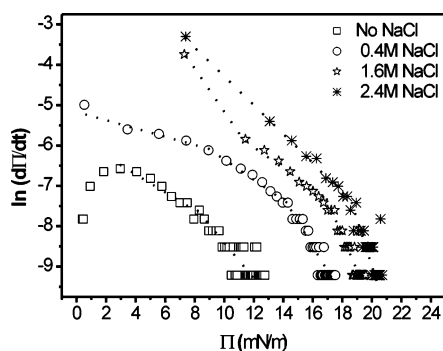

Figure 9. log($d\Pi/dt$) as a function of Π for 1 mg/mL lysozyme, which shows the three adsorption steps by the linear regression.
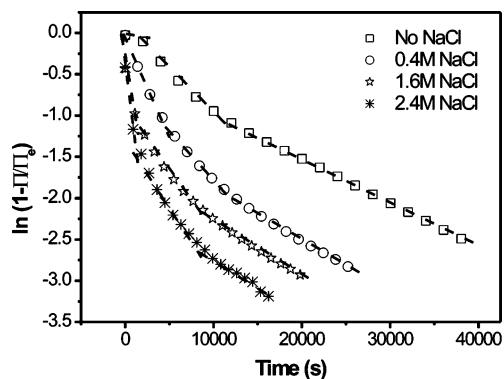

Figure 10. $\ln(1 - \Pi/\Pi_e)$ as a function of time for 1 mg/mL lysozyme. The rate constants for different steps can be obtained from the slope of the linear regression parts.

The isoelectric point, pI, of lysozyme is about pH 11.[32] At pH 4.5, lysozyme is positively charged. In the absence of salt, the electric double layer of the lysozyme molecule is thick and its influence is strong. The mobility of the protein molecule is low because as it moves, it drags the double layer surrounding it. In the presence of the salt, the surface charges of the protein molecule are screened by counterions in the solution, which decreases the thickness of the double layer around the protein.[33,34] Therefore, the effective size of the protein molecule also decreases, which results in an increase of the molecular mobility. The presence of salt increases the volume transport rate of the protein molecules in the bulk to the surface. However, the surface integration, penetration, and rearrangement is influenced by the intrinsic properties of protein molecules, such as their conformations. The presence of salt does not change
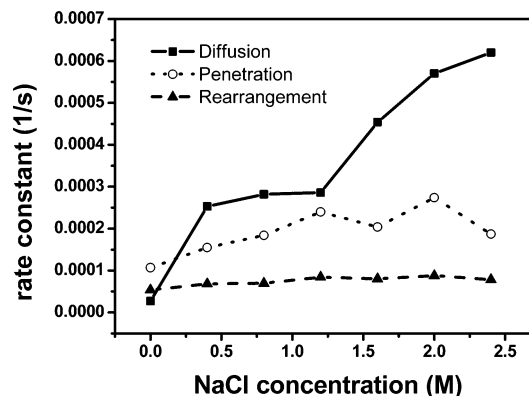
Figure 11. Rate constants of diffusion, k_d, penetration, k_p, and rearrangement, k_r, as a function of sodium chloride concentration.
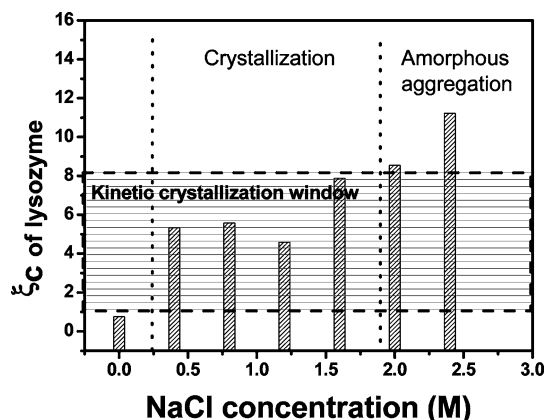
Figure 12. Crystallization coefficient, ξ_c, of lysozyme in a sodium chloride solution with different concentrations. The crystallization and amorphous aggregation in the salt conditions are also shown as a reference. A kinetic crystallization window is determined by comparison with the crystallization results.

the conformation and rigidity of a protein molecule. Therefore, the invariance of k_p and k_r with respect to salt concentration is understandable.

The energy barrier for each process could be obtained according to eq 5[35]

$$k = \frac{C_p}{\Gamma} \sqrt{\frac{RT}{2\pi M}} \exp\left(\frac{-\Delta G}{RT}\right) \qquad (5)$$

where k is the rate constant for each process, C_p is the protein bulk concentration, Γ is the surface coverage, R is the ideal gas constant, T is the temperature, and M is the molecular weight of the protein. At the initial stage, where the surface assembly is controlled by diffusion, there is no energy barrier for the first step, whereas in the following two steps, the energy barrier cannot be ignored. In our experiment, $C_p = 10^6$ mg/m^3, T is 296 K, M is 14 kg/mol, and Γ can be obtained from the literature[36] as 2.5 mg/m^2. This gives a potential energy barrier for penetration ΔG_p of about 23 RT and a rearrangement ΔG_r of about 24 RT. Because for the first stage there is no energy barrier, we then estimate ΔG to be about 47 RT. This energy barrier for surface assembly matches very well with the energy barrier of protein crystallization obtained by Ten Wolde[2] (50 RT) and by Grabe[37] (32.8−43.6 RT). This consistency confirms that the surface tension measurement gives a reasonable prediction of protein 3D crystallization.

The change of the crystallization coefficient, ξ_c, associated with the crystallization and amorphous aggregation conditions is shown in Figure 12 for lysozyme with sodium chloride. In

TABLE 1: Crystallization Coefficient for Various Proteins in Crystallization and Aggregation Conditions at 20 °C Compared with that of B$_{22}$

protein	salt concentration	pH	ξ_c	crystallization results	B$_{22}$ prediction	surface tension prediction
lysozyme	NaCl (<0.3 M)	4.5	0.8	solution	solution	solution
lysozyme	NaCl (0.3–2 M)	4.5	4–7	crystallization	crystallization	crystallization
lysozyme	NaCl (>2 M)	4.5	8–12	aggregation	aggregation	aggregation
lysozyme	(NH$_4$)$_2$SO$_4$ (2 M)	4.5	9.5	aggregation	crystallization (incorrect)	aggregation
con A	(NH$_4$)$_2$SO$_4$ (1 M)	7.0	7.6	crystallization		crystallization
BSA	(NH$_4$)$_2$SO$_4$ (52% saturated)	6.2	24	crystallization		aggregation (incorrect)

the crystallization experiments carried out by the batch method, good single crystals were obtained when the concentration of sodium chloride became larger than 0.2 M and less than 1.8 M.[11] When the concentration of sodium chloride was larger than 1.8 M, amorphous aggregation appeared immediately at supersaturation.

As discussed in section II, the possibility for a protein to crystallize under a certain condition can be reflected in the crystallization coefficient, ξ_c. Only when ξ_c is in an appropriate range where the diffusion rate and the surface integration rate are comparable, can a good single crystal appear. Higher values of ξ_c correspond to the case where the surface kinetics of protein crystallization is much slower than the volume diffusion. The solution is most likely to produce amorphous aggregation.

Comparing the crystallization conditions with the crystallization coefficient, it turns out that for crystallization, ξ_c lies in a certain range, $1 < \xi_c < 8$. When $\xi_c < 1$, the protein solution is in a steady state so that neither crystallization nor amorphous aggregation occur. The reason may be that when $\xi_c < 1$ the volume diffusion rate is less than the surface integration rate. The diffusion rate is so slow that it is hard for the molecules to collide. Therefore, the chance for embryos to grow to the stage of the critical radius and overcome the nucleation barrier ΔG^* is also low. However, when $\xi_c > 8$, the volume diffusion is faster by almost an order of magnitude than the surface integration. The molecules aggregate so fast that they grow up and overcome ΔG before they find their optimal orientation and conformation. In this case, amorphous aggregation is more favorable than crystallization. In the range where protein crystallization occurs, $1 < \xi_c < 8$, the diffusion and surface integration are comparable. This is called the "kinetic crystallization window".

To further test the applicability of this kinetic criterion on the prediction of protein crystallization, the crystallization and amorphous aggregation conditions obtained for lysozyme were applied to different solvents and different proteins, such as BSA and concanavalin A (con A). The results are shown in Table 1. It can be seen that in most cases the criterion is valid. The crystallization coefficient, ξ_c, falls in the kinetic crystallization window when crystallization takes place, and out of the window when amorphous aggregation, but no crystallization, occurs.

Compared with the B$_{22}$ criterion, the surface tension prediction offers many advantages. First, it is more reliable and physically more sound. This can be seen from the case of lysozyme in ammonium sulfate. Crystallizing lysozyme from ammonium sulfate is generally considered impossible.[32] In one questionable case the authors of ref 35 claimed that they obtained it from ammonium sulfate. It is, however, doubtful whether these authors used a genuine ammonium sulfate solution

because the ammonium sulfate concentration was very low; the additional compounds chloride acid and a relatively high concentration of sodium acetate used as buffer were introduced. As a consequence, sufficient amounts of sodium and chloride ions were presented to trigger the crystallization of lysozyme.[38] As shown in Table 1, our prediction of amorphous aggregation is correct, whereas the B$_{22}$ prediction is wrong. Second, the method of measuring the surface tension is easier to perform than the measurement of B$_{22}$, which is obtained by light scattering. Furthermore, the surface tension measurement requires only a small amount of protein because it is carried out at a single concentration, whereas a series of samples with different protein concentrations is needed to perform light scattering in order to determine B$_{22}$.

We notice that in some cases, such as BSA with 52% saturated ammonium sulfate, the prediction based on the surface tension also gives rise to an incorrect outcome (Table 1). In this case, our method predicts protein aggregation, whereas in practice, crystallization takes place. This may be attributed to the unfolding of BSA molecules at the air/water interface as shown in the literature.[39] The protein unfolding changes the surface assembly kinetics at the surface, which results in a deviation from the genuine crystallization condition relevant to the bulk phase. In this sense, if severe unfolding of the proteins occurs at the air/solution interface, then the prediction method based on the surface tension of the kinetic crystallization window becomes invalid.

In general, in cases where the dissociation and unfolding of the protein at the surface is not severe, the kinetic crystallization window offers a reliable and sound tool for predicting protein crystallization from solution. This criterion provides an easy and reliable way to crystallize new proteins based on a series of simple surface tension measurements.

Conclusions

We investigated the kinetics of protein 3D crystallization and 2D protein self-assembly at the interface using surface-tension measurements. The model protein lysozyme in the presence of the precipitant, sodium chloride, was used in this study. Protein crystallization is a kinetic process, which includes molecular volume transport and surface integration. The competition between volume transport and surface integration determines whether crystallization or amorphous aggregation will be the kinetically favorable process. The tendency to aggregate amorphously can be illustrated by the crystallization coefficient, ξ_c, which is the ratio between the volume diffusion rate and surface integration rate. From the study of protein 2D assembly at the air/solution interface, a kinetic crystallization window can be

determined, which lies in the range $1 < \xi_c < 8$. This kinetic crystallization window turns out to be more reliable and sound than the criterion based on the second virial coefficient. Patent pending.

References and Notes

(1) Galkin, O.; Vekilove, P. G. *J. Am. Chem. Soc.* **2000**, *122*, 156.

(2) Ten Wolde, P. R.; Frenkel, D. *Science* **1997**, *277*, 1975.

(3) Fiske, M. R. *Science* **1999**, *285*, 837.

(4) Georgalis, Y.; Umbach, P.; Saenger, W.; Ihmels, B.; Soumpasis, D. M. *J. Am. Chem. Soc.* **1999**, *121*, 1627.

(5) Zheng, B.; Roach, L. S.; Ismagilov, R. F. *J. Am. Chem. Soc.* **2003**, *125*, 11170.

(6) Caffrey, M. *J. Struct. Biol.* **2003**, *142*, 108.

(7) Matsuda, S.; Senda, T.; Itoh, S.; Kawano, G.; Mizuno, H.; Mitsui, Y. *J. Biol. Chem.* **1989**, *264*, 13381.

(8) Darby, J. J.; Creighton, T. E. *Protein Structure*; Oxford University Press: Oxford, 1993.

(9) Sybesma, C. *An Introduction to Biophysics*; Academic Press: New York, 1977.

(10) Chow, P. S.; Liu, X. Y.; Zhang, J.; Tan, R. B. H. *Appl. Phys. Lett.* **2002**, *81*, 1975.

(11) Zhang, J.; Liu, X. Y. *J. Chem. Phys.* **2003**, *119*, 10972.

(12) George, A.; Wilson, W. W. *Acta Crystallogr., Sect. D* **1994**, *50*, 361.

(13) Neal, B. L.; Asthagiri, D.; Velev, O. D.; Lenho, A. M.; Kaler, E. W. *J. Cryst. Growth* **1999**, *196*, 377.

(14) Bonnete, F.; Finet, S.; Tardieu, A. *J. Cryst. Growth* **1999**, *196*, 403.

(15) Narayanan, J.; Liu, X. Y. *Biophys. J.* **2003**, *84*, 523.

(16) Jia, Y. W.; Liu, X. Y. *Appl. Phys. Lett.* **2005**, *86*, 023903.

(17) Liu, X. Y. *J. Phys. Chem. B* **2001**, *105*, 11550.

(18) Chernov, A. A. *Modern Crystallography III - Crystal Growth*; Springer-Verlag: Berlin, 1984.

(19) Malkin, A. J.; McPherson, A. *J. Phys. Chem. B* **2002**, *106*, 6718.

(20) Liu, X. Y. *Phys. Rev. B* **1999**, *60*, 2810.

(21) Land, T. A.; Malkin, A. J.; Kuznetsov, Y. G.; McPherson, A.; Deyoreo, J. J. *Phys. Rev. Lett.* **1995**, *75*, 2774.

(22) Liu, X. Y. *J. Chem. Phys.* **2000**, *113*, 8807.

(23) Liu, X. Y.; Maiwa, K.; Tsukamoto, K. *J. Chem. Phys.* **1997**, *106*, 1870.

(24) Liu, X. Y.; Tsukamoto, K. *J. Chem. Phys.* **1997**, *107*, 10351.

(25) Durbin, S. D.; Feher, G. *Annu. Rev. Phys. Chem* **1996**, *47*, 171.

(26) Magdassi, S. *Surface Activity of Proteins*; Marcel Dekker: New York, 1996.

(27) Ward, A. F. H.; Tordai, L. *J. Chem. Phys.* **1946**, *14*, 453.

(28) Tornberg, E. *J. Colloid Interface Sci.* **1978**, *64*, 391.

(29) Graham, D. E.; Philips, M. C. *J. Colloid Interface Sci.* **1979**, *70*, 403.

(30) Subirade, M. *J. Colloid Interface Sci.* **1992**, *152*, 442.

(31) Chen, P.; Lanhooti, S.; Policova, Z.; Cabrerizo-Vilchez, M. A.; Neumann, A. W. *Colloids Surf., B* **1996**, *6*, 279.

(32) Ducruix, A.; Giege, R. *Crystallization of Nucleic Acids and Proteins*; Oxford University Press: New York, 1999.

(33) Adamson, A. W.; Gast, A. P. *Physical Chemistry of Surfaces*, 4th ed.; John Wiley: New York, 1997.

(34) Hunter, R. J. *Foundations of Colloid Science*; Oxford University Press: Oxford, 2001.

(35) Middelberg, A. P. J.; Radke, C. J.; Blanch, H. W. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5054.

(36) Lu, J. R.; Su, T. J.; Thomas, R. K.; Penfold, J.; Webster, J. *J. Chem. Soc., Faraday Trans.* **1998**, *94*, 3279.

(37) Grabe, M.; Neu, J.; Oyster, G.; Nollert, P. *Biophys. J.* **2003**, *84*, 854.

(38) Forythe, E. L.; Snell, E. H.; Pusey, M. L. *Acta Crystallogr., Sect. D* **1997**, *53*, 795.

(39) Van der Vegt, W.; Norde, W.; Van der Mei, H. C.; Busschre, H. J. *J. Colloid Interface Sci.* **1996**, *179*, 57.