Direct Determination of the Nucleation Rates of Protein Crystals

Oleg Galkin[†] and Peter G. Vekilov*[,†,‡]

Center for Microgravity and Materials Research and Department of Chemistry, University of Alabama in Huntsville, Huntsville, Alabama 35899

Received: August 5, 1999

We have developed a novel method for direct determination of the steady-state rates of homogeneous nucleation. The method is applicable to studies of crystallization, aggregation, and similar first-order phase transitions in solutions of proteins or other soluble slow-growing materials with temperature-dependent solubility. Temperature *T* control of the solution supersaturation allows fast supersaturation changes from a level inductive of nucleation to a level where no nucleation occurs, but existing crystals grow to detectable dimensions. In this way, nucleation takes place only at the first *T* setting at a constant supersaturation before solution depletion due to crystal growth becomes significant. We use inert oil to cover nucleating solutions and suppress nucleation on the solution−air interface. To obtain reproducible statistical characteristics of the intrinsically random nucleation process, a large number of simultaneous trials take place under identical conditions. First data for the nucleation of the model protein lysozyme under typical crystallization conditions show that protein crystal nucleation may be occurring at or even beyond the boundary of applicability of classical, continuum nucleation models.

1. Introduction

Recently, much progress has been made in the understanding of the protein crystal growth mechanisms (see refs 1−5 and refs therein). Despite the structural complexity of the protein molecules, as well as the compositional complexity of protein solutions, there are pronounced similarities between the mechanisms and kinetics underlying the crystallization of proteins and inorganic materials.

By comparison, insight into protein nucleation processes is still very limited. Detailed studies of the nucleation and growth of bovine and porcine insulin described crystallization in terms of engineering mass-crystallization models, rather than crystal nucleation and growth theories.[6−8] In a first application of select nucleation concepts, particle size distributions were deduced from lysozyme light scattering data, implicitly assuming ideal solution behavior, i.e., noninteracting solute protein molecules.[9,10] In another series of works,[11−13] nucleation was postulated to proceed according to the protein self-assembly mechanism. The authors concluded that nucleation proceeds as a fourth-order chemical reaction, which would require the unlikely simultaneous collision of four monomers.

In a comprehensive study of nucleation of four macromolecular crystals (see refs 14 and 15 and refs therein), the time dependence of the average cluster radius was deduced from dynamic light scattering data. The results were interpreted in terms of the classical nucleation theories, with the critical nucleus ranging from 7 to 60 molecules. In another series of light scattering studies of lysozyme nucleation,[16,17] the authors argued that the nuclei possess fractal structure. This could significantly alter the nucleation mechanisms and kinetics. The authors' claim has not been corroborated as yet.

Traditionally, interpretations of light-scattering data to obtain cluster sizes assume ideal solution behavior,[9,10,14,15] equivalent to neglecting the interactions between the solute protein molecules. However, it is difficult to perceive how molecules aggregate to form nuclei without intermolecular interactions. Furthermore, recent investigations show that the light-scattering signal in both undersaturated and supersaturated solutions is sensitive to variations in the protein intermolecular interactions with concentration (see refs 18−20 and refs therein). In addition to the ambiguities arising from nonideal solution behavior, the derivation of narrow particle size distributions from light scattering signals is an intrinsically ill-posed problem.[21] Since quite different distributions can result in similar signatures, sophisticated, time-consuming angular dependence measurements of the scattering intensity are required for unambiguous conclusions. Thus, the separation of the contributions to the light-scattering signal due to molecular interactions from those due to particle-size evolutions in supersaturated solutions is still an unsolved problem.

From the above it is apparent that (a) protein nucleation is not well understood at this point, and (b) a direct method to measure nucleation rates, independent of assumptions about the molecular interactions in the solution, is needed. Note that the few techniques that have produced spectacular results for small-molecule nucleation are not applicable to studies of protein crystal nucleation. For instance, different variants of cloud chamber[22,23] and supersonic nozzle expansion technique[24] are specific for nucleation in the vapor−liquid phase transition. Techniques that use levitating droplets[25,26] are prone to evaporation of solution from the liquid−air interface. This may render the experimental conditions irreproducible or, given the relatively slow rate of transport of the large protein molecules, lead to higher protein concentration at the droplet boundary.

In this paper we describe a novel experiment setup for kinetics studies of protein crystal nucleation. This technique allows direct determinations of the nucleation rates from the number of

* To whom correspondence should be addressed. E-mail: peter@cmmr.uah.edu; fax: (256) 890-6944.
† Center for Microgravity and Materials Research.
‡ Department of Chemistry.

10.1021/jp992786x CCC: $18.00
Published on Web 11/18/1999

crystals appearing in a certain solution volume for a given time. The technique takes advantage of the slow growth rate of protein crystals and the temperature dependence of the solubility to separate nucleation from the ensuing growth by lowering supersaturation after nucleation has occurred. For experiment statistics, the technique provides for several hundred simultaneous experiments in the same device using a single solution sample. To suppress heterogeneous nucleation on the air−solution interface and to avoid solvent evaporation, the droplets of nucleating solution are suspended in inert oil. The first representative results for the model protein lysozyme are compared to predictions of classical and advanced nucleation models.

2. Experiment Setup and Procedures

2.1. Separation of Nucleation from Subsequent Growth. The main problem faced in nucleation kinetics experiments is the separation of nucleation from ensuing growth: at the supersaturations required for nucleation, crystals typically grow extremely rapidly. This results in supersaturation σ decreases sufficient to significantly reduce the nucleation rate. As a consequence, correlations between supersaturation and nucleation rate in bulk experiments become ambiguous. To overcome this problem, an approach proposed by Tammann[27] has been employed to numerous systems.[28] Its essence is that even before nuclei are detected, the supersaturation is lowered to levels where the nucleation rate is zero, but the crystals already formed can grow to detectable dimensions. To date, the most successful implementation of this approach is in nucleation on a substrate during electrocrystallization, where one can rapidly change σ through the electrode overpotential (see ref 29 and refs therein).

Proteins permit fast σ changes even in bulk solutions. This is due to the relatively slow growth kinetics even at the high supersaturations required for nucleation. The experiment protocol is illustrated by Figure 1a. At the beginning of the experiment the temperature of the protein solution is set to T_1 so that the nucleation process is initiated at a desired supersaturation level. After a time interval Δt_1, the temperature is changed to T_2 at which the supersaturation is sufficiently low to prevent further nucleation but allows existing nuclei to grow to visible size. Note that the positions of T_1 and T_2 with respect to one another are shown in Figure 1a for the case of normal temperature dependence of the protein solubility. If the protein exhibits a retrograde solubility dependence, i.e., higher solubility at lower temperature, $T_1 > T_2$ as shown in Figure 1b, and both temperatures are higher than the equilibrium T_{eq}.

At the end of the experiment the number of crystals N_Σ that have appeared in the solution is counted and nucleation rate can be determined as $J = N_\Sigma/(V\Delta t_1)$, where V is the volume of the solution. This protocol allows separation of nucleation from crystal growth, thus avoiding the problem of solution depletion during nucleation. The times required to change the temperature to T_1 and from T_1 to T_2 are determined by the thermal properties of the materials from which the crystallization container is built and in our case were about 1 min. To minimize the bias in the correlation between the number of crystals and Δt_1, which may be introduced by nucleation occurring during the T changes, the shortest Δt_1 was 12 min.

2.2. Experiment Statistics. The temperature regime discussed above ensures that all nuclei form at the same supersaturation. However, the presence in the container of crystals nucleated at earlier times may affect nucleation of new crystals at later experiment times, for instance by promoting secondary nucleation,[30,31] which is often found to occur for proteins.[32] Then,

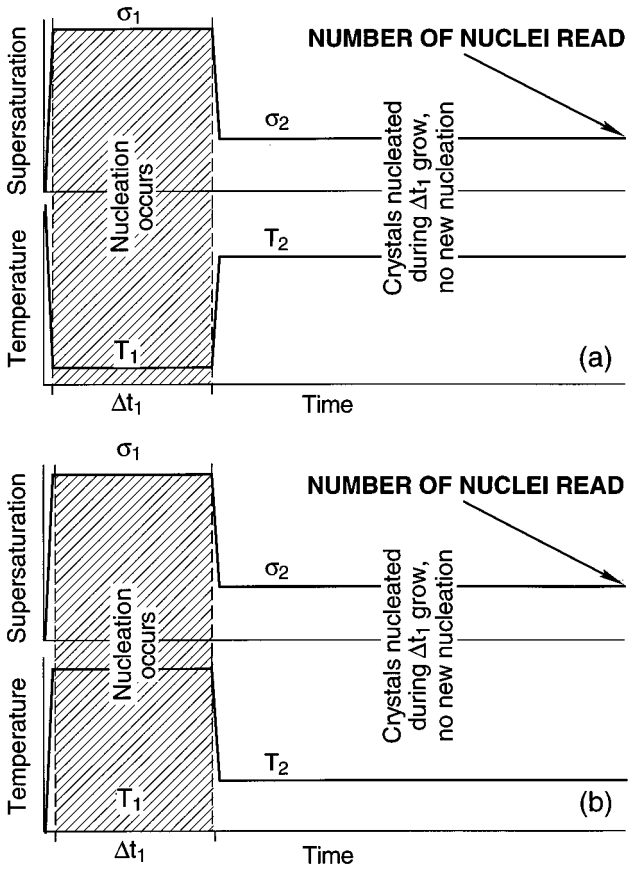

Figure 1. Temperature regimes and corresponding supersaturation levels used to induce nucleation during the time Δt_1 and to develop the crystals to detectable dimensions for a system with (a) normal temperature dependence of the solubility, such as lysozyme used here, and (b) retrograde temperature dependence of solubility, such as hemoglobin C or S.

the nucleation events in the container will not be statistically independent and a wrong value of the nucleation rate may obtain. Furthermore, if two or more nuclei are close, upon growth one of the crystals may engulf the others and lower the final crystal count. To minimize these unwanted effects, we used solution droplets with volume $V = 0.1-2.5\ \mu$L, usually $0.7\ \mu$L. Under the conditions tested in the experiments, one droplet contained at most 6−8, always well separated crystals.

These constraints on the number of crystals and the time constraints discussed in section 2.1 above, with the added limitation of a reasonable nucleation time of less than ~10 h, determine the dynamic range of nucleation rates accessible to our technique from ~0.7 to ~0.007 nuclei/cm^3 s.

The small volume of the nucleating solution elicits the stochastic nature of the individual nucleation events. Independent nucleation events occurring at a constant rate should result in a Poissonian distribution of the number of crystals in one droplet. Then, the requirement of a maximum number of crystals in a droplet < 8 leads to a mean number of crystals per droplet between 1 and 2. The standard deviation of this value is its square root, i.e., between 1 and 1.41, or 100% to 70% data scatter. Repeating the trial N_{trial} times reduces the scatter $\sqrt{N_{trial}}$ -fold. With 100 droplets, this would mean 10% to 7% scatter, while with 400 droplets the uncertainty is reduced to 5% to 3.5%.

We found that the best combination of the statistically required large number of trials and the limitations imposed by the material availability and results processing times is achieved

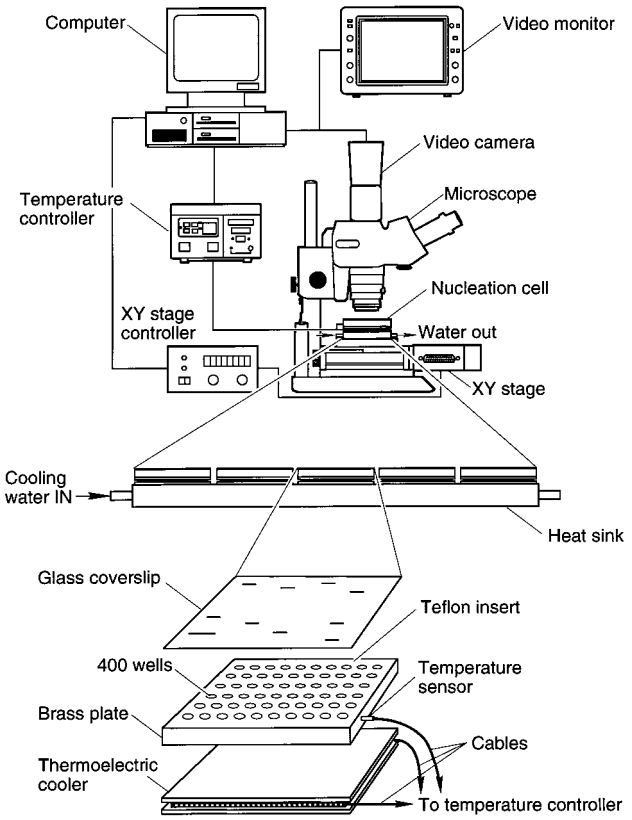

Figure 2. Schematic of experiment setup. For details, see text.

by using 400 identical droplets. To produce such a large number of droplets, a computer-controlled pump consisting of 0.5 mL syringe (Hamilton) and precision mechanical stage (Aerotech Inc.) was assembled. The reproducibility and accuracy (after calibration) of this assembly was determined by weighing droplets on an analytical balance and was better than 1%.

Using this pump assembly, the protein solution droplets are placed in round-bottom wells dug in a block of Teflon, see schematic of the experiment setup in Figure 2. For temperature uniformity, stability, and control, the Teflon block is pressed against a 6 mm thick brass plate. The temperature of the brass plate is measured by an embedded thermocouple and controlled by a 35×35 mm^2 thermoelectric (Peltier) cooler. The thermocouple and cooler are linked to temperature controllers (SE5010, Marlow Industries Inc.). Five identical blocks sit on a water-cooled plate that dissipates the heat from the thermoelectric coolers. The whole assembly is mounted on a computerized X−Y translation stage (Aerotech Inc.) under a microscope (Leica MZ8) with an attached 24-bit color video camera with 1280×1280 pixels resolution (SONY DKC-5000). A computer interfaced to the five temperature controllers regulates the temperatures in each block. This arrangement allows variations of T_1 and T_2 between the solution freezing point, for a typical aqueous salt solution of about −10 °C, and the limitation of the Peltier element of 65 °C. Keep in mind that temperature-sensitive proteins may not allow experiments throughout this whole range.

The use of independently controlled blocks allows five determinations of numbers of crystals nucleated for increasing time periods in samples of the same solution. This arrangement minimizes experiment error related to minor irreproducibility of solution preparation. The time dependence is used to determine the nucleation rate, see below. In this way, the duration of one nucleation rate determination is shortened to about 1 day.

2.3. Minimization of Nucleation on the Solution Surface. Preliminary experiments, in which the droplets deposited on the Teflon block were in contact with the air, were plagued by the fast nucleation of crystals on the solution−air interface. We minimized solvent evaporation from the droplet surface by loading the droplets in a closed glovebox. The partial vapor pressure of water in the glovebox was adjusted to that of the crystallizing protein solution by keeping in several Petri dishes in the glovebox fresh volumes of buffer-precipitant solutions with compositions identical to that in the crystallizing solution. Droplets of crystallizing solution of ~1 μL were stable in this environment for weeks. After deposition of the droplets on the Teflon−brass assembly, they were tightly covered with a glass lid and taken out of the glovebox for the experiment. Despite these precautions, abundant nucleation on the droplet surface was not suppressed. We speculate that nucleation on the droplet's surface is promoted by the protein adsorption. This conclusion is corroborated by a recent finding that full lysozyme monolayers form within a few milliseconds on this interface even at very low protein concentrations.[33,34] Note that nucleation on the air−solution interface is not limited to lysozyme. Trials with another protein of interest, hemoglobin C, also revealed extensive unwanted nucleation on the droplet surface.

Undesired nucleation on the solution surface was suppressed only by suspending the droplets in inert silicone oil (Poly-[dimethylsiloxane] 200 fluid, viscosity 5 cSt, density = 0.913 g/cm^3, from Aldrich). This approach has been used in optimizations of the crystallization conditions of a variety of proteins.[35,36] As an added advantage, the Archimedes force stemming from the closeness of the solution and oil densities provides for a very small area of contact between the droplet and the Teflon support. Although the Teflon surface does not induce heterogeneous nucleation of lysozyme, it may cause unwanted nucleation in other systems.

2.4. Solution Preparation. For validation and tests of the technique, initial experiments were performed with lysozyme. This enzyme hydrolyzes polysaccharides in bacterial cell walls and was one of the first proteins studied by X-ray diffraction.[37] Although it has lost its appeal to biochemists and structural biologists, it is particularly attractive for crystal growth research because its thermophysical properties are well characterized and because it was used in numerous prior studies (for a review see ref 2). Numerous recent crystallization mechanism studies using a wide range of other proteins have validated the results obtained with this material and justified lysozyme as a useful model system for growth studies.

We used lysozyme isolated from hen egg white (Seikagaku, 6× crystallized) without additional purification. Stock solution was prepared by dissolving the powdered protein in 0.05 M acetate buffer, pH = 4.5. It was then filtered through a 0.22 μm filter (Millipore Millex-GV) and stored at 4 °C for further experiments. Before each experiment a solution with desired composition was prepared by mixing the stock solution, buffer, and NaCl solution. The final concentration of protein was determined spectrophotometrically by measuring absorbance at 280 nm. Nucleation temperature used in our experiments T_1 = 12.6 °C, growth temperature T_2 = 20 °C; nucleation time Δt_1 ranged from 10 min to 8 h. With five arrays of 400 droplets with 0.7 μL of solution in each, the total volume of protein solution used in one run was about 1.4 mL.

3. Data Collection and Reproducibility

3.1. Reproducibility. Since nucleation is an intrinsically stochastic process, the number of crystals that appear in a certain

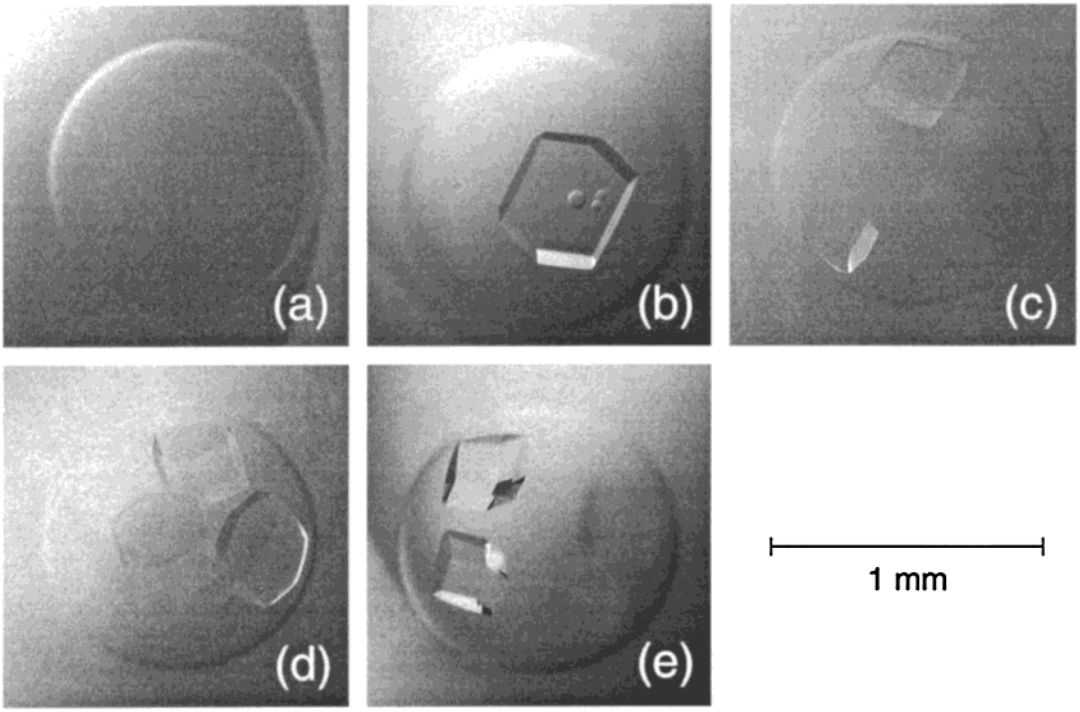

Figure 3. Variations in the number of crystals nucleated in a droplet under identical conditions.

volume is a random variable. Figure 3 illustrates that repetitions of an experiment under identical conditions could give, for instance, one, four, or no crystals. To extract a nucleation rate not one but a large number of statistically independent experiments is needed. Representative statistical distributions of the number of crystals in one droplet resulting from 400 simultaneous experiments under identical conditions are presented in Figure 4a. We compare the experimental distribution with the Poisson law

$$P(m) = \frac{N^m}{m!}\exp(-N) \tag{1}$$

where m is the number of crystals that appear in volume V during nucleation time Δt_1 and N is the mean value of the distribution. To evaluate the correspondence between the experimental data and the Poisson law we calculate the parameter χ^2 [38]

$$\chi^2 = \sum_m \frac{[F_m - P(m)N_{trial}]^2}{P(m)N_{trial}} \tag{2}$$

where F_m are the measured frequencies and $P(m)$ are Poissonian best-fit values. In most such determinations, for instance the data in Figure 4a, the χ^2 values correspond to confidence levels of 95% or better in the Poissonian character of the distributions. In a very few cases, this confidence level was only better than 85%. Hence, we conclude that all individual nucleation events observed in these studies are independent of each other.

The deviation of the measured N from its true value is characterized by the ratio $\sqrt{N}/N_{trial}$ and in this case is 6.5%.

A comparison of the distributions in Figure 4a illustrates that although the number of crystals in one droplet is random, the statistical distributions corresponding to experiments performed under the same conditions are reproducible. All five distributions have the same shape, follow the Poisson law, and have approximately the same mean values, Figure 4b.

3.2. Determination of the Steady-State Homogeneous Nucleation Rate. With increasing nucleation time Δt_1 the distribution of the number of crystals in one droplet shifts to the right, changing from near exponential at short times to close to Gaussian at long Δt_1. Figure 5 shows that the mean number of nucleated crystals changes linearly with time, indicating steady-state nucleation at the chosen experimental conditions. Several sources of unsteadiness can be discussed. There may be a competing process, such as denaturation or bacterial or enzymatic proteolyses, that consumes the protein in the solution and has characteristic times comparable to the crystal nucleation times. This would result in a nucleation rate lower than the steady state at longer nucleation time and a sublinear $N(t)$ dependence in Figure 5.

Furthermore, nucleation can be intrinsically unsteady in a closed system without sources of single molecules if the number of nuclei is so big that the monomer concentration is depleted. Then, the time-dependent nucleation flux $J_{ns}(t)$ is related to the steady-state flux J_{st} by[39]

$$J_{ns}(t) = J_{st}\left[1 + \sum_{j=1}^{\infty}(-1)^j \exp\left(-\frac{j^2 t}{\tau}\right)\right] \tag{3}$$

where

$$\tau = 4/\pi^3 Z^2 k^* n_1 \tag{4}$$

represents the nucleation induction period or time lag. In eq 4 Z is the so-called Zeldovich factor that accounts for a steady-state nuclei concentration lower than that at a hypothetical equilibrium, typically, $Z = 0.01-0.1$.[40] The constant k^* characterizes the kinetics of attachment of the single molecules to a nucleus, for the studied system it is $\sim 10^{-16}$ cm^3/s,[41] and n_1 is the number of molecules per cm^3, for lysozyme concentration of $\sim$50 mg/mL, $n_1 = 2 \times 10^{18}$ cm^{-3}. Physically, τ is a measure of the time needed for the transformation of the initial cluster size distribution to the steady-state distribution.[40] Thus, if the

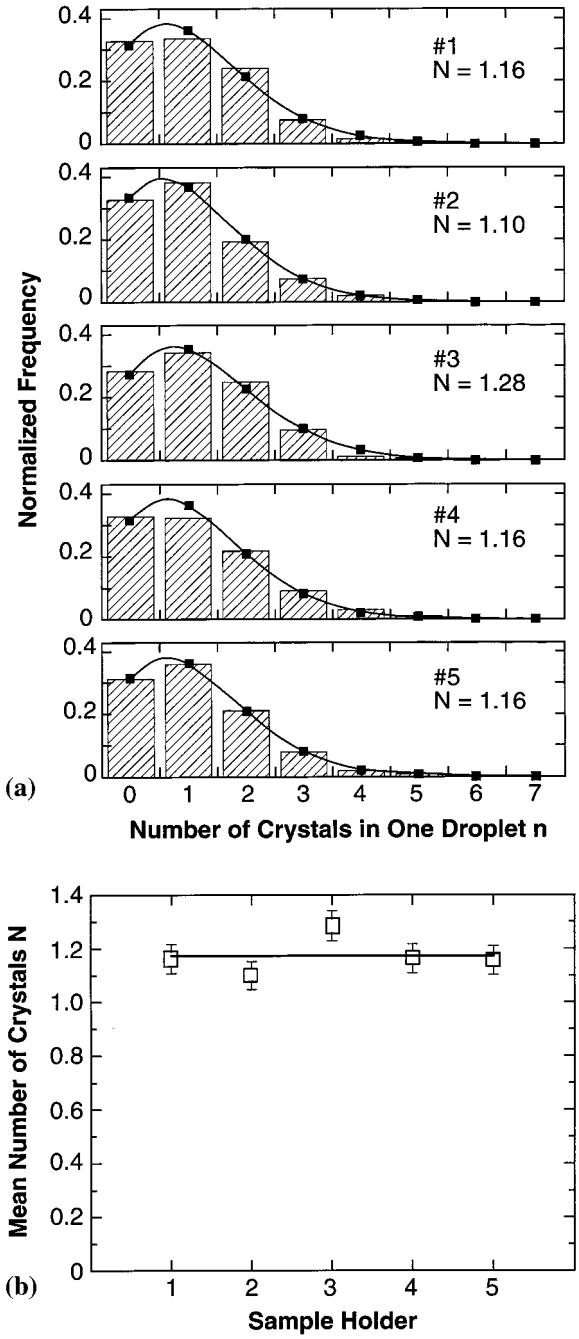

Figure 4. (a) Distributions of the number of lysozyme crystals appearing in one droplet at $T_1 = 12.6$ °C after $\Delta t_1 = 60$ min. Simultaneous experiments in five arrays of 400 droplets with volume $V = 0.7$ μL, lysozyme concentration $C = 69.5$ mg/mL, precipitant NaCl concentration $C_{NaCl} = 2.5\%$. Bars: experiment results; lines and symbols: fit with Poisson distribution. (b) Comparison between the mean number of crystals per droplet N from (a). Error bars correspond to the intrinsic error of such determinations, determined from the ratio of the mean number of crystals N to the number of independent trials N_{trial} as $\sqrt{N/N_{trial}}$.

characteristic time associated with significant concentration changes in the nutrient phase (due to nucleation and crystal growth) is several τ, nucleation will follow the steady-state equation. Using the above estimates, τ is of the order of 0.1−1 s in our system. This is significantly shorter than the nucleation times Δt_1, and we should get steady nucleation rates.

The data in Figure 5 illustrate another important feature of our experiment procedure. Although the use of oil to cover the solution droplets significantly reduces heterogeneous nucleation

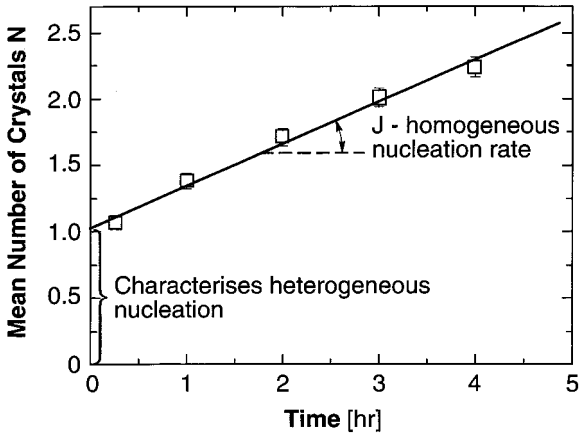

Figure 5. Time dependence of the mean number of crystals in one droplet N. Each N value determined from the Poisson distributions for simultaneous experiments in 400 droplets with volume $V = 0.7$ μL, lysozyme concentration $C = 55.5$ mg/mL, precipitant NaCl concentration $C_{NaCl} = 2.5\%$. Error bars correspond to $\sqrt{N/N_{trial}}$.

by preventing nucleation on the droplet surface, it still affects the data. However, since heterogeneous nucleation is characterized by lower energy barriers, it occurs much faster than homogeneous nucleation.[28] This leads to a constant additive to the number of nucleated crystals at various times. Thus, the intercept of the dependence with the y-axis in Figure 5 can be used to characterize the rate of heterogeneous nucleation. Linear regression analysis of the data yields the slope of the straight line and its uncertainty. From these, we determine the steady-state homogeneous nucleation rate and the error in its determination.

4. Representative Results and Discussion

In Figure 6a we present the dependence of the rate of homogeneous nucleation of lysozyme crystals on the protein concentration. As expected, the nucleation rate increases exponentially with protein concentration increase. For insight into the nucleation mechanisms, in Figure 6b we have recalculated these data in terms of a function of the supersaturation in terms of chemical potential difference in k_BT units $\sigma \equiv \Delta\mu/k_BT$. This variable can be calculated as $\sigma \cong \ln(C/C_{eq})$, C_{eq} being the equilibrium protein concentration (solubility) at the temperature of the determinations.[42,43]

We compare the experimental data with a formula derived for steady-state nucleation on the basis of continuum concepts and assuming spherical or other simple-shaped critical clusters. These approximations are often referred to as classical nucleation theory.[44] After adaptation to account for the heavier protein molecules,[41] the expected $J(\sigma)$ can be represented as

$$J = AC \exp(-B/\sigma^2) \tag{5}$$

where A and B are expression constants. The best fit to the data in Figure 6b yields $A = 18$ mg^{-1} sec^{-1}and $B = 65$. Using the expression for B [41]

$$B = \frac{16\pi}{3}\frac{\Omega^2\gamma^3}{(k_BT)^3} \tag{6}$$

with $\Omega = 3 \times 10^{-20}$ cm^3 being protein's molecular volume, we get for the surface free energy of the critical cluster $\gamma = 0.64$ mJ m^{-2}. Previous determinations from crystallization trials, in which nucleation kinetics data may have been contaminated by simultaneously occurring growth of the crystals, yielded

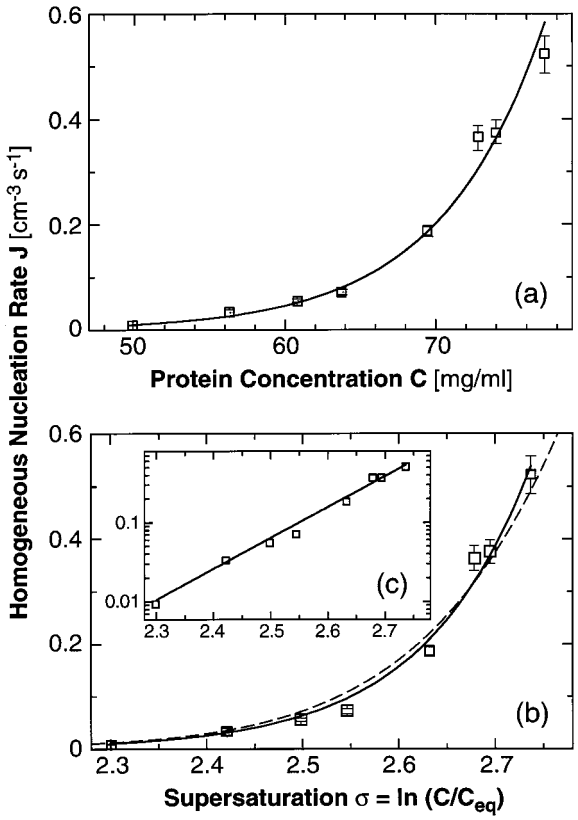

Figure 6. Dependence of lysozyme crystal homogeneous nucleation rate J at $T_1 = 12.6$ °C and with $C_{NaCl} = 2.5\%$ in 0.05 M acetate buffer, pH = 4.5 on (a) protein concentration C and (b) and (c) solution supersaturation σ. Solid lines are fits with an exponential function; dashed line is the fit with classical nucleation theory (see text for details).

somewhat higher values of 0.81−1.09 mJ m^{-2}.[41] The constant A is a complicated function of the molecular-level parameters of attachment kinetics. Its exact type strongly depends on features of the accepted model, which are difficult to verify experimentally. Hence, we do not attempt to extract any kinetic information from the value of A.

Although eq 5 provides a good quantitative fit to the data, arguably at intermediate supersaturations its predictions are somewhat high, while at higher supersaturations they are slightly lower than the actual nucleation rates. Quite surprisingly, a simple exponential dependence of the $J(\sigma)$ provides a better fit, see also the semilogarithmic plot in Figure 6c. To understand this exponential dependence of the nucleation rate on supersaturation, we use the nucleation theorem.[45,46] This is a universal, model-independent relationship between the work of nucleation, proportional to the logarithm of the nucleation rate, and the difference between the number of molecules in the nucleus and in an equal solution volume $n^* - n_0$

$$n^* - n_0 = \frac{\partial \ln J}{\partial \sigma} + \alpha \qquad (7)$$

The parameter α is a correction that may take values between 0 and 1.[46] From Figure 6c, $n^* - n_0 \cong 9.6 \pm 0.2$. We can roughly evaluate n_0 by comparing the molecular size, 30 Å, to the distance between the molecular centers $n_1^{-1/3} \approx 75$ Å. In the volume occupied by ∼10 crystal molecules there may be ∼1 solute molecule. Thus, $n^* \cong 10$ or 11 in the whole studied supersaturation range.

We can now go back to the correspondence between the classical nucleation theory and our data. This theory assumes that the nucleation rate depends on supersaturation for two reasons. (i) The number of molecules that overcome the energy barrier for nucleation increases, leading to an exponential dependence on this magnitude of the barrier. (ii) This barrier gets lower as the nucleus size decreases,[44] and this causes a decrease of the exponent. However, according to the nucleation theorem, in the studied supersaturation range the nucleus size is constant. Hence, $J(\sigma)$ increases more steeply than predicted by the classical theory.

To extend the supersaturation range of the nucleation kinetics studies, we determined the $J(\sigma)$ at $C_{NaCl} = 3\%$ and 4%. A detailed discussion of these results and of other phenomena occurring at higher supersaturations is provided elsewhere.[47] For the sake of completeness of the current presentation, we briefly summarize them below.

At 3% NaCl the dependence log $J(\sigma)$ is linear and the slope of the straight line corresponds to $n^* - n_0 = 4.2 \pm 0.2$ molecules. At 4% NaCl, this dependence breaks into two straight segments, corresponding to an abrupt change in $n^* - n_0$ from 4.7 ± 0.3 to 0.2 ± 0.3. At these precipitant concentrations, the $n^* - n_0$ and the volumes occupied by the nuclei are smaller, hence, n_0 is smaller and the correction it introduces in the nucleus size is comparable or smaller than α. Then, the n^* values corresponding to the $n^* - n_0$'s above are $n^* = 4$ or 5 at $C_{NaCl} = 3\%$, and $n^* = 4$ or 5 and then 1 or 2 at $C_{NaCl} = 4\%$. Note that critical clusters consisting of one molecule have been encountered before in investigations of electrochemically driven nucleation of new phases under high overvoltages/supersaturations.[29] Under those conditions, the nucleation rate is determined by the kinetics of attachment of molecules to this critical cluster.

These small numbers of molecules in the nuclei partially explain the constant n^* within certain σ intervals. One may still wonder why does the nucleus not cover all sizes between 10 and 1 to 2, but jumps from ∼10 to ∼ 4 to 1 or 2 molecules. Similar jumps were observed during crystallization of various metals and salts on crystalline and amorphous substrates.[29] Theoretical calculations of ΔG of clusters of various sizes assuming compact cluster shapes revealed that the molecular configurations consisting of different number of molecules have different stability.[48,49] As a result, the ΔG dependence on the cluster size is not monotonic and the locus of the maximum, i.e., the critical cluster size, jumps with increasing supersaturation over several cluster size units. Although the selection of the critical sizes depends on the structure of the cluster[50] and the symmetry of the bonds around the molecule, these calculations at least qualitatively explain the behavior of the nucleus size in the different growth regimes.

5. Conclusions

We have developed a novel experiment method for studies of the nucleation of protein crystals. The technique allows distinction between homogeneous and heterogeneous nucleation and allows determinations of the rates of homogeneous nucleation. The method is also applicable to studies of other phase transitions occurring in solutions or in low-temperature melts, provided that the transition kinetics is relatively slow and the equilibrium state is temperature dependent.

First experiments with the protein lysozyme show that the statistics and the kinetics of the nucleation for this model protein system follow the trends expected from theory and observed numerous times for small-molecule crystallization. The distribution of the number of crystals in one droplet follows the Poisson law. The nucleation rate depends exponentially on the protein concentration and thermodynamic supersaturation.

The size of critical nucleus, which can be extracted from our data, changes discretely with supersaturation and is constant within relatively wide supersaturation ranges. With increasing supersaturation, the found sequence of nuclei sizes is (10 or 11), (4 or 5), (1 or 2) molecules. These low numbers imply that the studied system, and likely other protein systems, may be beyond the boundary of applicability of classical nucleation theories that assume spherical nuclei and continuous changes of nucleus size.

Acknowledgment. We thank F. Rosenberger and D. Petsev for helpful discussions and L. Carver for expert graphics work. Support by the National Blood, Heart and Lung Institute, NIH (Grant # HL RO1 58038), the Life and Microgravity Sciences and Applications Division of NASA (Grant # NAG8 1354), and by the State of Alabama through the Center for Microgravity and Materials Research at the University of Alabama in Huntsville is gratefully acknowledged.

References and Notes

(1) Weber, P. In *Advances in protein chemistry*; Afinsen, C. B., Richards, F. M., Edsal, J. T., Eisenberg, D. S., Eds.; Academic Press: New York, 1991; Vol. 41.
(2) Rosenberger, F.; Vekilov, P. G.; Muschol, M.; Thomas, B. R. *J. Cryst. Growth* **1996**, *167*, 1−27.
(3) Durbin, S. D.; Feher, G. *Annu. Rev. Phys. Chem.* **1996**, *47*, 171−204.
(4) Vekilov, P. G.; Rosenberger, F.; Lin, H.; Thomas, B. R. *J. Cryst. Growth* **1999**, *196*, 261−275.
(5) Malkin, A. J.; Kuznetsov, Y. G.; McPherson, A. *J. Cryst. Growth* **1999**, *196*, 471−488.
(6) Schlichtkrull, J. *Acta Chem. Scand.* **1957**, *11*, 299−302.
(7) Schlichtkrull, J. *Acta Chem. Scand.* **1957**, *11*, 439−460.
(8) Schlichtkrull, J. *Acta Med. Scand.* **1965**, *177*, 103−113.
(9) Kam, Z.; Shore, H. B.; Feher, G. *J. Mol. Biol.* **1978**, *123*, 539−555.
(10) Feher, G.; Kam, Z. In *Methods in Enzymology*; Wyckoff, H. W., Hirs, C. H. W., Timasheff, S. N., Eds.; Academic Press: New York, 1985; Vol. 114; pp 77−112.
(11) Ataka, M.; Tanaka, T. *Biopolymers* **1986**, *58*, 337−350.
(12) Ataka, M.; Asai, M. *Biophys. J.* **1990**, *58*, 807−811.
(13) Besho, Y.; Ataka, M.; Asai, M.; Katsura, T. *Biophys. J.* **1994**, *66*, 310−313.
(14) Malkin, A. J.; McPherson, A. *Acta Crystallogr., Sect. D* **1994**, *50*, 385−395.
(15) Malkin, A. J.; McPherson, A. *J. Cryst. Growth* **1993**, *128*, 1232−1235.
(16) Georgalis, Y.; Umbach, P.; Raptis, J.; Saenger, W. *Acta Crystallogr., Sect. D* **1997**, *53*, 691−702.
(17) Georgalis, Y.; Umbach, P.; Raptis, J.; Saenger, W. *Acta Crystallogr., Sect. D* **1997**, *53*, 702−712.
(18) Muschol, M.; Rosenberger, F. *J. Chem. Phys.* **1995**, *103*, 10424−10432.
(19) Piazza, R. *J. Cryst. Growth* **1999**, *196*, 415−423.
(20) Tardieu, A.; Verge, A. L.; Malfois, M.; Bonnette, F.; Finet, S.; Ries-Kaut, M.; Belloni, L. *J. Cryst. Growth* **1999**, *196*, 193−203.
(21) Chu, B. *Laser Light Scattering*, 2nd ed.; Academic Press: New York, 1991.
(22) Hung, C.-H.; Krasnopoler, M. J.; Katz, J. L. *J. Chem. Phys.* **1989**, *90*, 1856−1865.
(23) Katz, J. L. *J. Chem. Phys.* **1970**, *52*, 4733−4748.
(24) Bartell, L. S.; Dibble, T. S. *J. Phys. Chem.* **1991**, *95*, 1159−1167.
(25) Arnold, S.; Goddard, N. L.; Wotherspoon, N. *Rev. Sci. Instrum.* **1999**, *70*, 1473−1477.
(26) Izmailov, A. F.; Myerson, A. S.; Arnold, S. *J. Cryst. Growth* **1999**, *196*, 234−242.
(27) Tammann, G. *Die Aggregatzustaende*, 2nd ed.; Voss: Leipsig, 1922.
(28) Chernov, A. A. *Modern Crystallography III: Growth of Crystals*; Springer: Berlin, 1984.
(29) Milchev, A. *Contemp. Phys.* **1991**, ***32***, 321−332.
(30) Denk, E. G.; Botsaris, G. D. *J. Cryst. Growth* **1972**, *13/14*, 493−499.
(31) Blem, K. E.; Ramanarayanan, K. A. *AIChE J.* **1987**, *33*, 677−680.
(32) Provost, K.; Robert, M. C. *J. Cryst. Growth* **1991**, *110*, 258−264.
(33) Stebe, K. J., Johns Hopkins University, Baltimore, MD 1999, unpublished.
(34) Lu, J. R.; Su, T. J.; Howlin, B. J. *J. Phys. Chem.* **1999**, *103*, 5903−5909.
(35) Chayen, N. E. *Protein Eng.* **1996**, *9*, 927−929.
(36) Chayen, N. E. *J. Cryst. Growth* **1999**, *196*, 434−441.
(37) Blake, C. C. F.; Johnson, L. N.; Mair, G. A.; North, A. C. T.; Phillips, D. C.; Sarma, V. R. *Proc. R. Soc. London, Ser. B* **1967**, *167*, 378.
(38) Young, H. D. *Statistical treatment of experimental data*; McGraw-Hill: New York, 1962.
(39) Kashchiev, D. *Surf. Sci* **1969**, *14*, 209−220.
(40) Kashchiev, D. In *Science and technology of crystal growth*; Eerden, J. P. v. d., Bruninsma, O. S. L., Eds.; Kluwer Academic Publishers: Dordrecht, 1995; pp 53−56.
(41) Vekilov, P. G.; Monaco, L. A.; Thomas, B. R.; Stojanoff, V.; Rosenberger, F. *Acta Crystallogr., Sect. D* **1996**, *52*, 785−798.
(42) Cacioppo, E.; Pusey, M. L. *J. Cryst. Growth* **1991**, *114*, 286−292.
(43) Rosenberger, F.; Howard, S. B.; Sowers, J. W.; Nyce, T. A. *J. Cryst. Growth* **1993**, *129*, 1−12.
(44) Mutaftschiev, B. In *Handbook of crystal growth*; Hurle, D. T. J., Ed.; Elsevier: Amsterdam, 1993; Vol. I., pp 189−247.
(45) Kashchiev, D. *J. Chem. Phys.* **1982**, *76*, 5098−5102.
(46) Oxtoby, D. W.; Kashchiev, D. *J. Chem. Phys.* **1994**, *100*, 7665−7671.
(47) Galkin, O.; Vekilov, P. G. *J. Am. Chem. Soc.*, in print.
(48) Stoyanov, S. In *Current Topics in Materials Science*; North-Holland: Amsterdam, 1974.
(49) Milchev, A.; Malinovski, J. *Surf. Sci.* **1985**, *156*.
(50) Yau, S.-T.; Vekilov, P. G. *Nature*, submitted.