Are Nucleation Kinetics of Protein Crystals Similar to Those of Liquid Droplets?

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · DECEMBER 1999

Impact Factor: 12.11 · DOI: 10.1021/ja9930869

2 AUTHORS, INCLUDING:

Peter G Vekilov
University of Houston
167 PUBLICATIONS **5,516** CITATIONS

Are Nucleation Kinetics of Protein Crystals Similar to Those of Liquid Droplets?

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

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

Received August 24, 1999. Revised Manuscript Received November 4, 1999

Abstract: We have studied the nucleation of crystals of a model protein from aqueous solutions using a novel technique that allows direct determinations of homogeneous nucleation rates. At a constant temperature of 12.6 °C we varied the thermodynamic supersaturation by changing the concentrations of protein and precipitant. We found a broken dependence of the homogeneous nucleation rate on supersaturation that is beyond the predictions of the classical nucleation theory. The nucleation theorem allows us to relate this to discrete changes of the size of the crystal nuclei with increasing supersaturation as (10 or 11) → (4 or 5) → (1 or 2). Furthermore, we observe that the existence of a second liquid phase at high protein concentrations strongly affects crystal nucleation kinetics: (i) Crystal nucleation rates are lower than expected in the phase region of liquid−liquid demixing. (ii) In the immediate proximity of this region, nucleation rates vary by factors of up to 2 in identical experiments. Since for this region theory predicts a sharp rate maximum, we attribute this kinetic instability to minor shifts of the experimental conditions toward or away from the phase boundary.

Introduction

The main impetus for protein crystallization studies has been the realization that the preparation of sufficiently large single crystals is the main bottleneck[1,2] of X-ray diffraction studies of protein structure. Protein structural data are needed, for instance, for rational drug design, or for the understanding of genome structure−function correlations.[3] Recently increased interest in protein crystallization stems from the fact that protein crystallization and aggregation occur in the human body and are responsible for severe pathological conditions. For instance, the aggregation of Hemoglobin S and crystallization of Hemoglobin C in human blood cells underlies the sickle cell disease and related anemia[4−7] and aggregation of eye-lens proteins is the cause of cataract.[8,9] Understanding the physics of these processes could provide a means to influence them in the human body. Furthermore, the slow dissolution rate of protein crystals is utilized to achieve sustained release of medications, such as insulin and interferon-α.[10−13] If the administered dose consists of a few, larger, equidimensional crystallites, steady medication release rates can be maintained for longer periods than for doses with broad crystal size distributions. To achieve narrow size distributions, nucleation should be constrained to very short time spans, i.e., crystals should nucleate almost simultaneously, so that they can grow at the same decreasing supersaturation.

In view of these factors, much progress has recently been made in the understanding of the protein crystal growth mechanisms (see refs 14−18 and references therein). Despite the structural complexity of the protein molecules, and the compositional complexity of protein solutions, there are pronounced similarities between the mechanisms and kinetics underlying the growth of protein and inorganic crystals**.** Protein crystal growth has even been viewed as a particularly convenient model for phase transitions that occur in a variety of systems: water freezing in clouds and oceans, metal and alloy solidification, semiconductor single-crystal production, etc.[19,20]

* Address correspondence to this author.
‡ Center for Microgravity and Materials Research.
† Department of Chemistry.

(1) DeLucas, L. J.; Bugg, C. E. *Trends Biotechnol.* **1987**, *5*, 188−193.
(2) Chayen, N. E.; Boggon, T. J.; Casseta, A.; Deacon, A.; Gleichmann, T.; Habash, J.; Harrop, S. J.; Helliwell, J. R.; Neih, Y. P.; Peterson, M. R.; Raftery, J.; Snell, E. H.; Hädener, A.; Niemann, A. C.; Siddons, D. P.; Stojanoff, V.; Thompson, A. W.; Ursby, T.; Wulff, M. *Q. Rev. Biophys.* **1996**, *29*, 227−278.
(3) Darby, N. J.; Creighton, T. E. *Protein Structure*; Oxford University Press: Oxford, 1993.
(4) Hirsch, R. E.; Raventos-Suarez, C.; Olson, J. A.; Nagel, R. L. *Blood* **1985**, *66*, 775−778.
(5) Eaton, W. A.; Hofrichter, J. In *Advances in protein chemistry*; Anfinsen, C. B., Edsal, J. T., Richards, F. M., Eisenberg, D. S., Eds.; Academic Press: San Diego, 1990; Vol. 40; pp 63−279.
(6) Lin, M. J.; Nagel, R. L.; Hirsch, R. E. *Blood* **1989**, *74*, 1823−1825.
(7) Feeling-Taylor, A. R.; Banish, R. M.; Hirsch, R. E.; Vekilov, P. G. *Rev. Sci. Instrum.* **1999**, *70*, 2845−2849.
(8) Broide, M. L.; Berland, C. R.; Pande, J.; Ogun, O. O.; Benedek, G. B. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*.
(9) Berland, C. R.; Thurston, G. M.; Kondo, M.; Broide, M. L.; Pande, J.; Ogun, O.; Benedek, G. B. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 1214−1218.
(10) Reichert, P.; McNemar, C.; Nagabhushan, N.; Nagabhushan, T. L.; Tindal, S.; Hruza, A. In US Patent, 5,441,734,1995.
(11) Matsuda, S.; Senda, T.; Itoh, S.; Kawano, G.; Mizuno, H.; Mitsui, Y. *J. Biol. Chem* **1989**, *264*, 13381−13382.
(12) Peseta, S.; Langer, J. A.; Zoon, K. C.; Samuel, C. E. In *Annual Review of Biochemistry*; Richardson, C. C., Boyer, P. D., Dawid, I. B., Meister, A., Eds.; Annual Reviews: Palo Alto, 1989; Vol. 56, pp 727−778.
(13) Brange, J. *Galenics of Insulin*; Springer: Berlin, 1987.
(14) 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.
(15) Rosenberger, F.; Vekilov, P. G.; Muschol, M.; Thomas, B. R. *J. Crystal Growth* **1996**, *167*, 1−27.
(16) Durbin, S. D.; Feher, G. *Annu. Rev. Phys. Chem.* **1996**, *47*, 171−204.
(17) Vekilov, P. G.; Rosenberger, F.; Lin, H.; Thomas, B. R. *J. Crystal Growth* **1999**, *196*, 261−275.
(18) Malkin, A. J.; Kuznetsov, Y. G.; McPherson, A. *J. Crystal Growth* **1999**, *196*, 471−488.

10.1021/ja9930869 CCC: $19.00
Published on Web 12/16/1999

By comparison, insight into the first stage of protein crystallization, crystal nucleation, is still very limited. In some cases, abundant experimental data have been fitted to engineering or other preconceived models.[21−28] In other studies (see refs 29−32 and references therein), select nucleation concepts were applied and the conclusion was drawn that crystal nucleation of the studied proteins follows the classical nucleation theories, i.e., protein crystal nucleation, similarly to crystal growth, follows the mechanisms established for the nucleation of simple liquids or inorganic salts from vapor or solution. A conflicting point of view is that mechanisms entirely different from those found for inorganic molecules should apply to protein crystals.[33]

Furthermore, recent results on the molecular interactions and phase behavior in protein solutions (that typically also contain buffer, another electrolyte and often a smaller organic additive[34]) have shown a rich variety of phenomena that do not occur in molecular solutions of inorganic substances (see refs 8, 9, and 35−39 and references therein). Particularly intriguing is the existence of another condensed phase, a high-concentration liquid, at low temperatures.[8,35] The thermodynamics of these phenomena are more akin to the thermodynamics of phase transitions in colloid solutions[37,40−43] than to those of small-molecule inorganic solutions. Molecular dynamics simulations and phase-field modeling predict that the processes of phase separation in the solution will affect the kinetics and the mechanisms of protein crystal nucleation.[42,44,45]

Hence, the aims of the investigations discussed here are to (i) study the kinetics of crystal nucleation in a model protein system, (ii) obtain insight into relevant mechanisms with particular attention to possible concurrent processes, e.g., the appearance of the second liquid phase that may affect the nucleation pathways and kinetics, and (iii) test the applicability to protein crystal nucleation of the classical nucleation theory,[46,47] which adequately describes nucleation kinetics of many substances, including small molecule crystallization from solutions,[48] or condensation of vapors into liquid droplets.[49−51]

For these studies, we used lysozyme isolated from hen egg white. This enzyme has a molecular mass of ∼14500 Da, hydrolyzes polysaccharides in bacterial cell walls, and was one of the first proteins studied by X-ray diffraction.[52] It is still broadly used in, for instance, studies of protein folding dynamics,[53] and it is a particularly attractive crystallization model because its thermophysical properties are well-known and it has been used in numerous prior investigations (for a review see ref 15). Numerous recent crystallization mechanism investigations using a wide range of other proteins have validated the results obtained with this material and justified lysozyme as a useful model system for protein crystallization studies.

Experimental Section

Solution Preparation. Hen egg white lysozyme, from Seikagaku, 6× crystallized, was used without additional purification. A stock solution was prepared by dissolving the protein powder 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. A 20% solution of NaCl in the same buffer was used to add precipitant in the chosen concentration. Before each experiment a solution with the desired composition was prepared by mixing the protein stock, buffer, and NaCl solutions. The final protein concentration was determined spectrophotometrically by measuring absorbance at 280 nm using absorbance $\alpha^{280\text{nm}} = 2.64\ \text{mL}\cdot\text{mg}^{-1}\cdot\text{cm}^{-1}$.[54] The protein solution volume used in one run was about 1.5 mL.

Technique for Nucleation Rates Determinations. The existing experimental methods for determinations of homogeneous nucleation rates would be inapplicable or produce ambiguous results if applied to protein systems. For instance, different variants of cloud chamber[55] and supersonic nozzle expansion technique[56] are specific for vapor−liquid nucleation. Techniques that use levitating droplets[57,58] are prone to evaporation of solution from the liquid−air interface. Light scattering,[29,31] although a powerful technique, is heavily dependent on assumptions about the interactions between the molecules for data interpretation.

Hence, we developed a novel technique that allows direct determinations of the steady-state rate of homogeneous nucleation. In the beginning of a run, the protein solution is loaded at a temperature chosen to prevent nucleation or liquid−liquid demixing. Then the temperature is lowered to a selected T_1 at which nucleation occurs. In analogy to a

(19) Vekilov, P. G.; Alexander, J. I. D. *Chem. Rev.* Submitted for publication.
(20) Rosenberger, F. *Cryst. Res. Technol.* **1999**, *34*, 163−165.
(21) Schlichtkrull, J. *Acta Chem. Scand.* **1957**, *11*, 299−302.
(22) Schlichtkrull, J. *Acta Chem. Scand.* **1957**, *11*, 439−460.
(23) Schlichtkrull, J. *Acta Med. Scand.* **1965**, *177*, 103−113.
(24) Ataka, M.; Tanaka, T. *Biopolymers* **1986**, *58*, 337−350.
(25) Ataka, M.; Asai, M. *Biophys. J.* **1990**, *58*, 807−811.
(26) Besho, Y.; Ataka, M.; Asai, M.; Katsura, T. *Biophys. J.* **1994**, *66*, 310−313.
(27) Georgalis, Y.; Umbach, P.; Raptis, J.; Saenger, W. *Acta Crystallogr. Sect. D* **1997**, *53*, 691−702.
(28) Georgalis, Y.; Umbach, P.; Raptis, J.; Saenger, W. *Acta Crystallogr. Sect. D* **1997**, *53*, 702−712.
(29) Kam, Z.; Shore, H. B.; Feher, G. *J. Mol. Biol.* **1978**, *123*, 539−555.
(30) 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.
(31) Malkin, A. J.; McPherson, A. *Acta Crystallogr. Sect. D* **1994**, *50*, 385−395.
(32) Malkin, A. J.; McPherson, A. *J. Cryst. Growth* **1993**, *128*, 1232−1235.
(33) Littlechild, J. A. *J. Phys. D: Appl. Phys.* **1991**, *24*, 111−118.
(34) McPherson, A. In *Crystallization of Membrane Proteins*; Michel., H., Ed.; CRC Press.: Boca Raton, 1990; pp 1−51.
(35) Muschol, M.; Rosenberger, F. *J. Chem. Phys.* **1995**, *103*, 10424−10432.
(36) Velev, O. D.; Kaler, E. W.; Lenhoff, A. M. *Biophys. J.* **1998**, *75*.
(37) Adams, M.; Dogic, Z.; Kellery, S. L.; Fraden, S. *Nature* **1998**, *393*, 349−352.
(38) Piazza, R. *J. Cryst. Growth* **1999**, *196*, 415−423.
(39) Tardieu, A.; Verge, A. L.; Malfois, M.; Bonnette, F.; Finet, S.; Ries-Kaut, M.; Belloni, L. *J. Cryst. Growth* **1999**, *196*, 193−203.
(40) Pusey, P. N.; Van Megen, W. *Nature* **1986**, *320*, 340−343.
(41) Van Megen, W.; Underwood, S. M. *Nature* **1993**, *362*, 616−619.
(42) Hagen, M. H. J.; Frenkel, D. *J. Chem. Phys.* **1994**, *101*, 4093−4097.
(43) Asherie, N.; Lomakin, A.; Benedek, G. B. *Phys. Rev. Lett.* **1996**, *77*, 4832−4835.
(44) Wolde, P. R. t.; Frenkel, D. *Science* **1997**, *277*, 1975−1978.
(45) Talanquer, V.; Oxtoby, D. W. *J. Chem. Phys.* **1998**, *109*, 223−227.
(46) Mutaftschiev, B. In *Handbook of crystal growth*; Hurle, D. T. J., Ed.; Elsevier: Amsterdam, 1993; Vol. I, pp 189−247.
(47) Kashchiev, D. In *Science and technology of crystal growth*; Eerden, J. P. v. d., Bruinsma, O. S. L., Eds.; Kluwer Academic Publishers: Hingham, MA, 1995; pp 53−56.
(48) Leeden, M. C. V. D.; Kashchiev, D.; Rosmalen, G. M. V. *J. Colloid. Interface Sci.* **1992**, *152*, 338−3509.
(49) Katz, J. L.; Ostermier, B. J. *J. Chem. Phys.* **1967**, *47*, 478−487.
(50) Katz, J. L. *J. Chem. Phys.* **1970**, *52*, 4733−4748.
(51) Stoyanova, V.; Kashchiev, D.; Kupenova, T. *J. Aerosol. Sci.* **1994**, *25*, 867−877.
(52) 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.
(53) Rariy, R. V.; Klibanov, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 13520−13523.
(54) Sophianopoulos, A. J.; Rhodes, C. K.; Holcomb, D. N.; VanHolde, K. E. *J. Biol. Chem.* **1962**, *237*, 1107−1112.
(55) Hung, C.-H.; Krasnopoler, M. J.; Katz, J. L. *J. Chem. Phys.* **1989**, *90*, 1856−1865.
(56) Bartell, L. S.; Dibble, T. S. *J. Phys. Chem.* **1991**, *95*, 1159−1167.
(57) Arnold, S.; Goddard, N. L.; Wotherspoon, N. *Rev. Sci. Instrum.* **1999**, *70*, 1473−1477.
(58) Izmailov, A. F.; Myerson, A. S.; Arnold, S. *J. Cryst. Growth* **1999**, *196*, 234−242.

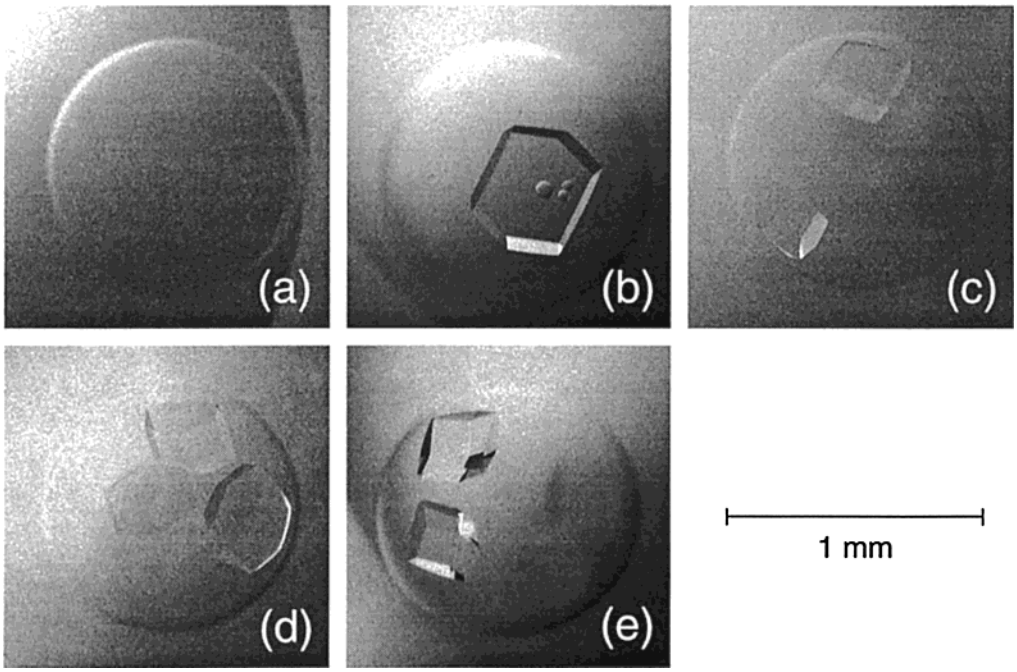

Figure 1. Variations in the numbers of lysozyme crystals in a droplet nucleating under identical conditions.

technique developed for studies of electrochemical nucleation on a substrate,[59] after a time period of Δt_1 the temperature is raised from the nucleation temperature T_1 to the growth temperature T_2 (in all studies reported here, $T_1 = 12.6$ °C and $T_2 = 20$ °C). At T_2, supersaturation is at levels where nucleation rate is practically zero, but the crystals already formed can grow to detectable dimensions.[60] This allows separation of the nucleation from the ensuing growth. After this growth stage, the crystals nucleated at T_1 during Δt_1 are counted.

To obtain reproducible statistical characteristics of the random nucleation process, 400 simultaneous trials take place under identical conditions. Each of these trials took place in a solution droplet of volume 0.7 μL. To suppress the undesired nucleation at the solution air interface, the droplets were suspended in inert silicone oil, used[61,62] in optimizations of the crystallization conditions of a variety of proteins. To extract the nucleation rate from the time dependence of the number of nucleated crystals, five arrays of 400 droplets are subjected to the nucleation supersaturation at increasing time intervals Δt_1. These Δt_1 values ranged from 12 min to 8 h. In all, the determination of one nucleation rate data point is based upon statistics over 2000 protein solution droplets. The experiment setup and procedures are described in detail elsewhere.[63]

Data Collection and Reproducibility. Since the number of crystals that appear in a certain volume is a random variable, successive repetitions of an experiment under identical conditions could give, for instance, one, four, or no crystals, Figure 1. Statistical distributions of the number of crystals in one droplet resulting from 400 simultaneous experiments under identical conditions are presented in Figure 2. We compared the experimental distributions similar to those in Figure 2 with the Poisson law

$$P(n) = \frac{N^n}{n!}\exp(-N) \qquad (1)$$

Here n is the number of crystals that appear in one droplet during Δt_1 and N is the mean value of the distribution. We evaluated the goodness of the fits using the χ^2 criterion.[64] The found χ^2 values show correspondence between the data and the Poisson law with a confidence level of 0.92−0.99. This indicates that the individual nucleation events are independent of each other. The determination error, i.e., the deviation of the measured N value from its true value, was evaluated from the ratio $\sqrt{N/N_{\text{trial}}}$ and was typically ~5%.

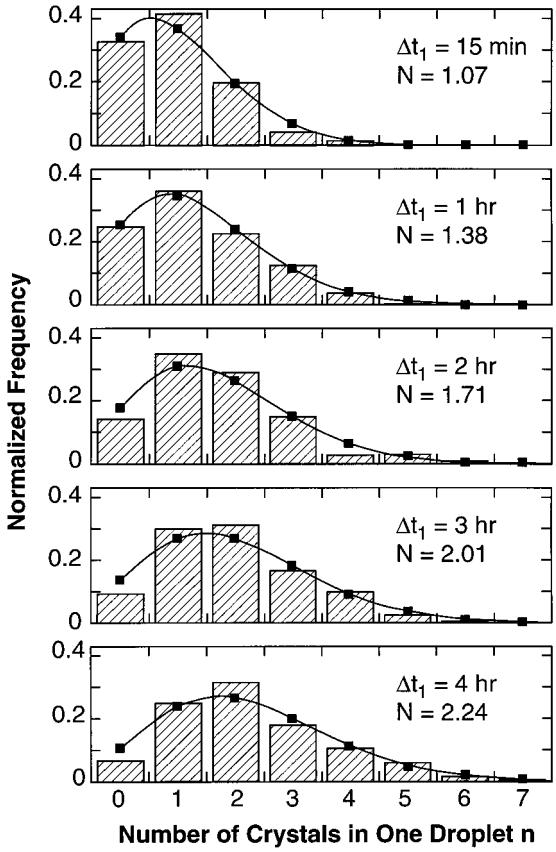

Figure 2. Distributions of the number of lysozyme crystals n appearing in one droplet as a result of increasing nucleation times Δt_1 indicated in the plots. Each distribution presents the result of simultaneous experiments in 400 droplets with volume $V = 0.7$ μL each, lysozyme concentration $C = 55.5$ mg/mL, precipitant NaCl concentration $C_{\text{NaCl}} = 3\%$. Bars: experiment results. Lines and symbols: fits with Poisson distribution.

(59) Milchev, A. *Contemp. Phys.* **1991**, *32*, 321−332.
(60) Tammann, G. *Die Aggregatzustaende*, 2nd ed.; Voss: Leipsig, 1922.
(61) Chayen, N. E. *Protein Eng.* **1996**, *9*, 927−929.
(62) Chayen, N. E. *J. Cryst. Growth* **1999**, *196*, 434−441.
(63) Galkin, O.; Vekilov, P. G. *J. Phys. Chem.* **1999**, *103*, 10965−10971.
(64) Young, H. D. *Statistical treatment of experimental data*; McGraw-Hill: New York, 1962.

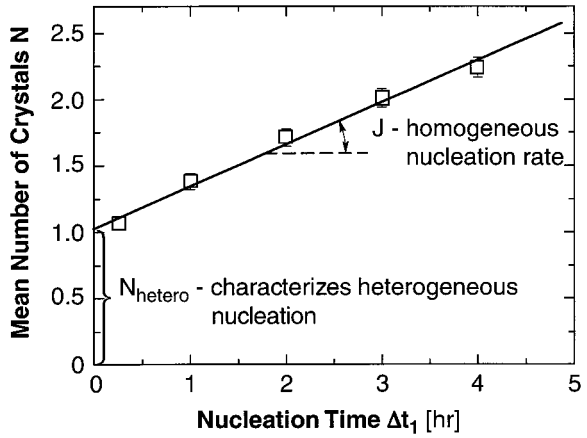

Figure 3. Time Δt_1 dependence of the mean number of crystals in one droplet N. Each N value is determined from a distribution in Figure 2. Error bars correspond to $\sqrt{N/N_{\text{trial}}}$. The slope of the straight line is used to calculate the nucleation rate J in $\text{cm}^{-3}\ \text{s}^{-1}$, the intercept N_{hetero} characterizes heterogeneous nucleation.

The reproducibility of this determination of the mean number of crystals in a droplet N was estimated by repeating an experiment five times using identical conditions and procedures. The five N values were within 5−6%, as expected from the above estimate of the determination error.[63]

Determination of the Homogeneous Nucleation Rate. Figure 2 presents the changes of the distribution of number of crystals in one droplet with nucleation time Δt_1. Figure 3 shows that the mean number of nucleated crystals increases linearly with time, indicating steady-state nucleation at the chosen experimental conditions. Several sources of unsteadiness may appear. 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 nucleation rates lower than the steady state at longer nucleation time and a sublinear $N(t)$ dependence. 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. However, evaluation of the nucleation time lag τ,[65] i.e., the characteristic time needed for the transformation of the initial cluster size distribution to the steady-state distribution,[47] yields for this system 0.1−1 s.[66] This is significantly shorter than the nucleation times Δt_1 in Figure 3, and we should get steady nucleation rates. The straight line in Figure 3 indicates that none of these factors affect our results.

Although the use of oil to cover the droplets significantly reduces heterogeneous nucleation by preventing nucleation on the droplet surface, apparently it still occurs (note the nonzero intercept of the data in Figure 3 at $t = 0$). Likely centers for this process may be the small ($<0.22\ \mu$m) particles remaining in the solution after filtration. The good linearity of the $N(t)$ at $t > 0$ indicates that the heterogeneous nucleation is limited to times shorter than those accessible to our technique. This mode of fast crystal nucleation leads to a constant additive to the number of nucleated crystals at all times. Thus, the intercept of the dependence with Y-axis in Figure 3 can be used to characterize the rate of *heterogeneous* nucleation and the slope of this dependence yields the *homogeneous* nucleation rate.

Results

Homogeneous Nucleation Rates. The variations of the homogeneous nucleation rate with protein concentration at three different concentrations of the precipitant, NaCl, are presented in Figure 4. In agreement with general expectations, the nucleation rate increases exponentially with protein concentration at constant precipitant concentration and, overall, is higher

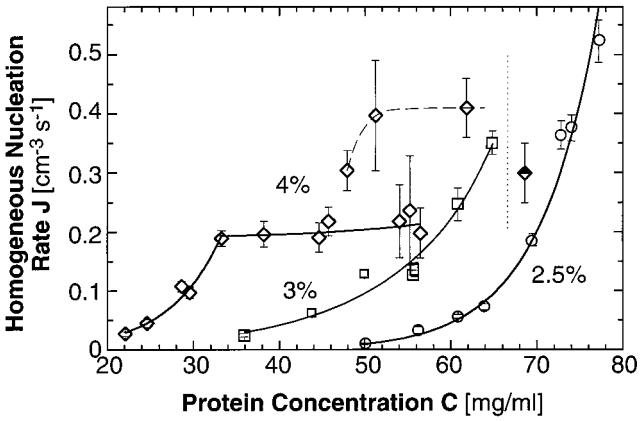

Figure 4. Dependencies of the homogeneous nucleation rates J of lysozyme crystals on protein concentration C at $T = 12.6$ °C and the three precipitant concentrations C_{NaCl} indicated in the plots. Solid lines: fits with exponential functions. Dashed line for data points at $C_{\text{NaCl}} = 4\%$ is just a guide for the eye. Datum point at $C_{\text{NaCl}} = 4\%$ and lysozyme $C = 68$ mg/mL was obtained in a cloudy solution and was not used in fitting procedures. Vertical dotted line at $C = 66$ mg/mL indicates the liquid−liquid demixing boundary at this T and $C_{\text{NaCl}} = 4\%$.

at higher precipitant concentrations. However, at the highest precipitant concentration, $C_{\text{NaCl}} = 4\%$, the $J(C)$ dependence contains three peculiarities.

(i) The dependence breaks at $C^* = 33.5$ mg/mL, with the sections at $C < C^*$ and $C > C^*$ following different exponents.

(ii) The data point at the highest lysozyme concentration, $C = 68$ mg/mL, although not apparently deviating from the exponent through the lower concentration data, is lower than the data point determined at a lower $C = 64$ mg/mL and a lower $C_{\text{NaCl}} = 3\%$. We noticed that during the determination of the nucleation rate at $C_{\text{NaCl}} = 4\%$ and $C = 68$ mg/mL the crystallization solutions in all 2000 droplets became cloudy immediately after temperature was lowered to the nucleation temperature of $T_1 = 12.6$ °C. The solutions were cloudy throughout the nucleation period Δt_1 and became clear again when T was raised at the end of the period Δt_1 to T_2 (no crystals or any other formations were detected in the solution droplets at that time). This indicates that this set of conditions (C, C_{NaCl}, T) is below the liquid−liquid coexistence boundary for the lysozyme−water−NaCl system. After the growth stage, most of the crystals found in the droplets appeared like regular tetragonal lysozyme crystals. However, in a few (∼20 out of 2000) droplets spherulitic crystals with thin needles growing radially outward from a center ("sea urchin" morphology) were detected, see Figure 5. No correlation between the nucleation times and the number of droplets with "sea urchin" morphology crystals was found. The presence of crystals with such unusual shape has been related to crystallization starting below the liquid−liquid coexistence curve.[67]

(iii) At $C > 48$ mg/mL the dependence bifurcates with the data points belonging to either of two branches. We noticed that in the runs leading to J values from the upper branch, in a few droplets (∼2−3 out of 2000) spherulitic crystals, similar to those in Figure 5, appeared. This suggests that the closeness of the liquid−liquid coexistence boundary may be affecting crystallization in those runs.

Heterogeneous Nucleation. The mean numbers of heterogeneously nucleated crystals N_{hetero} are plotted for all studied conditions in Figure 6. These N_{hetero} values are extracted from

(65) Kashchiev, D. *Surf. Sci* **1969**, *14*, 209−220.

(66) Vekilov, P. G.; Monaco, L. A.; Thomas, B. R.; Stojanoff, V.; Rosenberger, F. *Acta Crystallogr. Sect. D* **1996**, *52*, 785−798.

(67) Muschol, M.; Rosenberger, F. *J. Chem. Phys.* **1997**, *107*, 1953−1962.

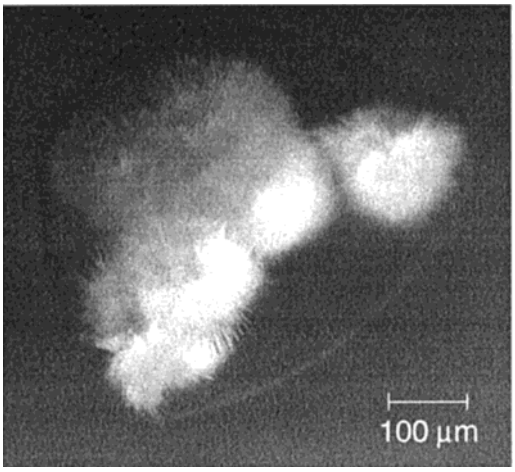

Figure 5. "Sea urchin" morphology of crystals observed in a few of the droplets during runs at $C_{NaCl} = 4\%$ and $T = 12.6$ °C yielding higher J values, as well as the run at $C_{NaCl} = 4\%$ and lysozyme $C = 68$ mg/mL in which the crystals nucleated in a cloudy solution.

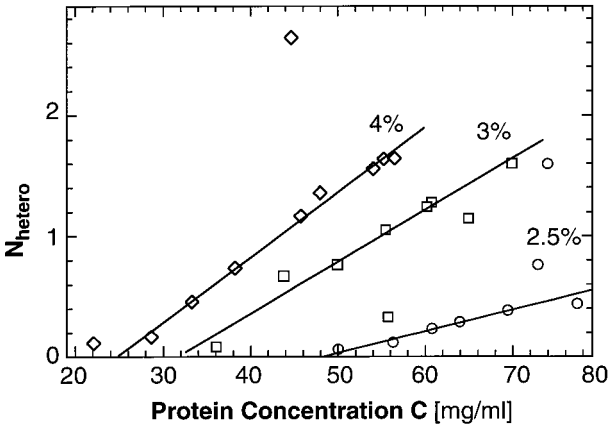

Figure 6. Dependencies of the number of heterogeneously nucleated crystals N_{hetero} determined from plots similar to Figure 3, on protein concentration at $T = 12.6$ °C and the three precipitant concentrations C_{NaCl} indicated on the plots. Solid straight lines are just guides for the eye.

the intercepts of the straight lines of the time dependence of the mean number of nucleated crystals per droplet, similar to the one in Figure 3. The number N_{hetero} increases with both protein and precipitant concentration. Note that although the homogeneous nucleation rates in Figure 4 are exponential functions of the protein concentration, the increase in N_{hetero} is close to linear.

Discussion

The Nucleus Size. The nucleus or critical cluster of the new phase is a cluster that has equal probability to grow or to dissolve and is in a labile equilibrium with the supersaturated solution.[68−71] The number of molecules in the nucleus is the most important characteristic of the nucleation process. The nucleus size n^* largely determines the height of the free energy barrier for nucleation (i.e., the reversible work for nucleation) ΔG and hence the nucleation rate J. The relation between ΔG and the number difference between the molecules in the nucleus and in an equal solution volume $n^* - n_0$ is treated by the nucleation

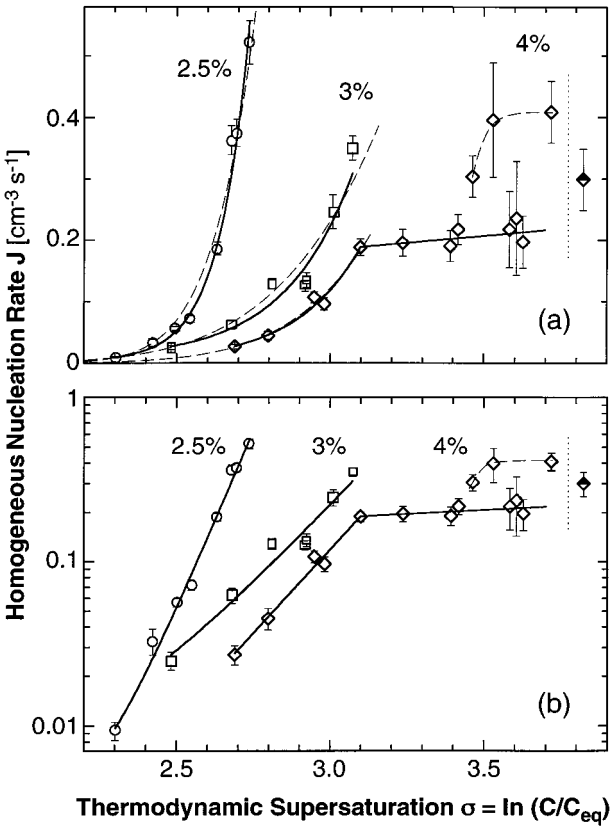

Figure 7. Dependencies of homogeneous nucleation rate J on thermodynamic supersaturation $\sigma \equiv \Delta\mu/k_BT$ at $T = 12.6$ °C and at the three precipitant concentrations indicated on the plots. Solid lines: fits with exponential functions. Dashed lines: fits with the classical nucleation theory expression, eq 3. Data points with higher J values at $C_{NaCl} = 4\%$ and the datum point obtained in a cloudy solution at $C_{NaCl} = 4\%$ and lysozyme $C = 68$ mg/mL, $\sigma = 3.8$, were not used in the fitting procedures, see text for details. Vertical dotted lines at $\sigma = 3.78$ indicate the liquid−liquid demixing boundary at this T and $C_{NaCl} = 4\%$. (a) Linear coordinates; (b) semilogarithmic coordinates.

theorem of Kashchiev and Oxtoby,[72,73] a universal, model-independent nucleation law. Since the nucleation work ΔG can be estimated from the logarithm of the nucleation rate J, in terms of J and n^* the nucleation theorem becomes

$$n^* - n_0 = k_B T \frac{\partial \ln J}{\partial \Delta\mu} + \alpha \qquad (2)$$

where α is a correction that takes values between 0 and 1.[73]

To present the dependencies of the nucleation rates on the concentrations of the protein and precipitant in Figure 4 in the variables of the nucleation theorem, we re-plotted in Figure 7 the data in terms of functions of thermodynamic supersaturation $\sigma \equiv \Delta\mu/k_BT$. We calculate σ as $\ln(C/C_{eq})$, where C is the protein concentrations of the solution and C_{eq} is the concentration at equilibrium with large crystals at the temperature T and precipitant concentration of the experiment.[74,75] This definition of supersaturation does not account for solution nonideality.[76,77] Following the logic of ref 77, it is easy to show that nonideality

(68) Gibbs, J. W. *The Scientific Papers*; Dover: New York, 1961; Vol. 1.
(69) Volmer, M. *Kinetik der Phasenbildung*; Steinkopff: Dresden, 1939.
(70) Oxtoby, D. W. *J. Phys.: Condens. Matter* **1992**, *4*, 7627−7650.
(71) Oxtoby, D. W. *Acc. Chem. Res.* **1998**, *31*, 91−97.
(72) Kashchiev, D. *J. Chem. Phys.* **1982**, *76*, 5098−5102.
(73) Oxtoby, D. W.; Kashchiev, D. *J. Chem. Phys.* **1994**, *100*, 7665−7671.
(74) Cacioppo, E.; Pusey, M. L. *J. Cryst. Growth* **1991**, *114*, 286−292.
(75) Rosenberger, F.; Howard, S. B.; Sowers, J. W.; Nyce, T. A. *J. Cryst. Growth* **1993**, *129*, 1−12.
(76) Ross, P. D.; Minton, A. P. *J. Mol. Biol.* **1977**, *112*, 437−452.
(77) Guo, B.; Kao, S.; McDonald, H.; Wilson, W. W.; Asanov, A.; Combs, L. L. *J. Cryst. Growth* **1999**, *196*, 424−433.

Table 1. Characteristics of the Nucleation Process Determined from Fits of Eqs 2−5 to Data[a]

C_{NaCl}	n^*-n_0	A [s^{-1}]	B	γ [mJ/m^2]	n^*_{CNT}
2.5%	9.6 ± 0.2	$(9 \pm 5) \times 10^{-16}$	65 ± 4	0.64	11−7
3%	4.2 ± 0.2	$(4 \pm 2) \times 10^{-18}$	33 ± 4	0.51	4−2
4%	4.7 ± 0.3	$(1.4 \pm 0.7) \times 10^{-17}$	44 ± 4	0.56	5−3
4%	0.2 ± 0.3				

[a] For definitions and details, see text.

leads to a correction in the above expression for the supersaturation $\Delta\mu/k_BT = \ln(C/C_{eq}) + 2B_2M(C - C_{eq}) + O(B_3C^2)$, where B_i are the virial coefficients and M is the protein molecular mass. At a NaCl concentration of 2.5% $B_2 = -2.1 \times 10^{-4}$ mol cm^3/g, and the higher order virial coefficients have negligible affects on the solution behavior.[35] At 75 mg/mL, $\sigma = 2.7992$ the correction $2B_2M(C - C_{eq}) \cong -0.4$, i.e., ∼15%. Similar estimates at the higher precipitant concentrations require data on the solution virial coefficients that are not available. Hence, for consistency, we employ the simplified expression for supersaturation for all data in Figure 7.

Figure 7b indicates that at $C_{NaCl} = 2.5$ and 3% n^* does not change throughout the respective supersaturation ranges, while at $C_{NaCl} = 4\%$ the nucleus size changes abruptly at $\sigma = 3.1$, corresponding to $C = 33.5$ mg/mL. The values of $n^* - n_0$ extracted from the four linear segments in Figure 7b are shown in Table 1. Since supersaturation is defined as the logarithm of the ratio of C to C_{eq}, the exact value of C_{eq} does not affect the slope of the straight lines in Figure 7b. Hence, the values of $n^* - n_0$ are independent of possible experimental errors of the solubility measurements.

We attribute the deviations of $n^* - n_0$ from whole numbers to the correction factor α in eq 2. To roughly evaluate n_0, we compare the molecular diameter, ∼30 Å, to the distance between the molecular centers $n_1^{-1/3}$, where n_1 is the protein *molecular* concentration. At $C \approx 70$ mg/mL, $n_1 = 2.9 \times 10^{18}$ cm^{-3} and the distance is ≈70 Å. At 2.5% NaCl and at protein concentrations close to 70 mg/mL, in the volume occupied by a crystal consisting of ∼10 molecules, there may be $n_0 \approx 1$ solute molecule. At the other precipitant concentrations, the $n^* - n_0$ and the volume occupied by the nuclei are smaller, hence, n_0 is smaller and the correction it introduces in the nucleus size is comparable to or smaller than α. Keeping this in mind, we can extract the n^* values corresponding to the $n^* - n_0$ values in Table 1: for $C_{NaCl} = 2.5\%$, n^* is 10 or 11, at $C_{NaCl} = 3\%$, $n^* = 4$ or 5, at $C_{NaCl} = 4\%$, $n^* = 4$ or 5 and then 1 or 2 molecules. Critical clusters consisting of one molecule have been encountered before in investigations of electrochemically driven nucleation of new phases under high overvoltages/supersaturations.[59] Under those conditions, the nucleation rate is determined by the kinetics of attachment of molecules to this critical cluster.

With these small numbers of molecules in the nuclei, the discrete character of the n^* changes of one or a few whole molecules becomes apparent. This partially explains the constant n^* values within certain σ intervals. One may still wonder why does the nucleus not cover all sizes between 10 and 1−2, 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.[59] Calculations of ΔG of clusters of various sizes assuming compact cluster shapes revealed that the molecular configurations consisting of a different number of molecules have different stability.[78,79] As a result, the ΔG dependence on the cluster size is not monotonic and the locus of 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[80] 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.

Comparison with Predictions of Classical Nucleation Theory. One of the basic assumptions of classical nucleation theory (CNT) is of continuous cluster size changes that is a good approximation to reality only for large nuclei. Thus, the size of the nuclei determined above precludes application of CNT to our data. We compared our experimental results to the predictions of this theory only for the sake of completeness. Within the framework of this theory, the dependence of the nucleation rate on supersaturation σ and protein molecular concentration n_1 is[46,81]

$$J = An_1 \exp(-B/\sigma^2) \tag{3}$$

The coefficient A is a complicated function of the molecular-level attachment-kinetics parameters. There have been attempts to theoretically derive an expression for this coefficient for nucleation from solution.[82−85] In all cases, the final formulas for A contain variables that are often impossible to determine independently.

The parameter B is related to the thermodynamic barrier for creation of the critical cluster ΔG^* and for a spherical cluster can be written as

$$B = \frac{16\pi}{3} \frac{\Omega^2\gamma^3}{(k_BT)^3} \tag{4}$$

with Ω the protein molecular volume in the crystal and γ the surface free energy of the critical cluster. Two-parameter fits of our data with eq 3 show reasonably good correspondence, Figure 7a. The best-fit values of A and B are shown in Table 1. From the values of B we estimate the surface free energy γ. The values of γ shown in Table 1 are lower by about half than a previous determination in ref 66. That previous work used expressions similar to eqs 3 and 4 to fit crystallization kinetics data in which the nucleation stage was not separated from growth. Hence, the estimate for γ given there is of limited reliability.

Classical nucleation theory also allows determinations of the nucleus size as

$$n^*_{CNT}{}^* = \frac{2B}{\sigma^3} \tag{5}$$

This size continuously changes in the supersaturation ranges of the experiments. The nucleus sizes determined using eq 5 straddle the more accurate determinations based on the nucleation theorem, eq 2. This correspondence seems to support the general belief that CNT provides a fair approximation for the nucleation barrier,[86−88] while spectacularly failing to predict the preexponential factors.

(78) Stoyanov, S. In *Current Topics in Materials Science*; North-Holland: Amsterdam, 1974.
(79) Milchev, A.; Malinovski, J. *Surf. Sci.* **1985**, *156*.
(80) Yau, S.-T.; Vekilov, P. G. *Nature.* Submitted for publication.
(81) Vekilov, P. G.; Monako, L. A.; Thomas, B. R.; Stojanoff, V.; Rozenberger, F. *Acta Crystallogr.* **1996**, *D52*, 785−798.
(82) Turnbull, D.; Fisher, J. C. *J. Chem. Phys.* **1949**, *17*, 71−73.
(83) Kahlweit, M. In *Physical Chemistry*; Eyring, H., Ed.; Academic Press: New York, 1969; Vol. VII, pp 675−698.
(84) Neilsen, A. E. In *Crystal Growth*; Peiser, S., Ed.; Pergamon: Oxford, 1967; pp 419−426.
(85) Walton, A. G. In *Nucleation*; Zettlemoyer, A. C., Ed.; Marcel Dekker: New York, 1969; pp 225−307.

Heterogeneous Nucleation on Foreign Particles. In terms of the classical nucleation theory, heterogeneous nucleation, i.e., nucleation on a foreign surface, is enhanced if the new phase wets the available surface. Then, the thermodynamic work of formation of a nucleus of the new phase can be reduced up to several times, with the reduction factor determined by the wetting angle.[89] This is the reason for the significantly faster rates and shorter times of this nucleation mode. The observations in Figure 6 of more intense heterogeneous nucleation at higher precipitant and protein concentrations roughly agree with this trend and with previous experimental observations.[90] However, the weak, linear increase in the number of heterogeneously nucleated crystals with protein concentration does not fit this simplified scenario. This should not be surprising, since (see above) we suspect that heterogeneous nucleation in our system occurs on the surface of submicron particles. The sizes of these particles likely cover the whole range below 0.22 μm, and their surfaces may be very rough. Such peculiarities of the substrates have not been considered by theory.

Our technique of nucleation rates determinations was specially designed to differentiate between homogeneous and heterogeneous contributions. However, if a technique only records the sum number of crystals for a given time, the resulting nucleation rates will be biased by the presence of heterogeneously nucleated crystals. Since the dependencies of the rates of two nucleation modes on the system parameters are similar, the data will not allow distinctions between them. The introduced bias may be propagated far into the conclusions based on such results.

Heterogeneous nucleation of protein crystals may soon become an important technological process for the fabrication of, for instance, coupled protein−semiconductor sensors, data storage devices, etc. Despite the above caveat, the observation of the general correspondence of the heterogeneous nucleation trends to the prediction of nucleation theory may provide guidelines in these new exciting areas.

Liquid−Liquid Separation and Its Effects on Crystal Nucleation. The liquid−liquid (L−L) coexistence boundaries in the system lysozyme−NaCl−water at pH 4.5 maintained by acetate buffer have been determined in our laboratory for a few concentrations of NaCl using the same materials as those used in the present study.[67] In the (C,T) plane the phase separation curve $T_{L-L}(C)$ is best described by

$$T_{L-L} = T^{\mathrm{crit}}\left\{1 - A^{L-L}\left|\frac{C^{crit} - C}{C^{\mathrm{crit}}}\right|^{1/\beta}\right\} \qquad (6)$$

where $C^{\mathrm{crit}} = 255 \pm 30$ mg/mL is the critical lysozyme concentration, β is the critical exponent, and $A^{\mathrm{L-L}}$ and T^{crit} are adjustable parameters, with T^{crit} being a roughly linear function of C_{NaCl}. Using the parameter values from ref 67 we calculated the liquid−liquid separation boundary for $C_{\mathrm{NaCl}} = 4\%$. The result is plotted in Figure 8 together with the solubility curve at these conditions[74] and the respective L−L and solubility curves for $C_{\mathrm{NaCl}} = 3\%$.[74,75]

The line A−A in Figure 8 shows the locations in the (C,T) plane of all experiments discussed here. We see that the experimental conditions are far from the L−L separation curve

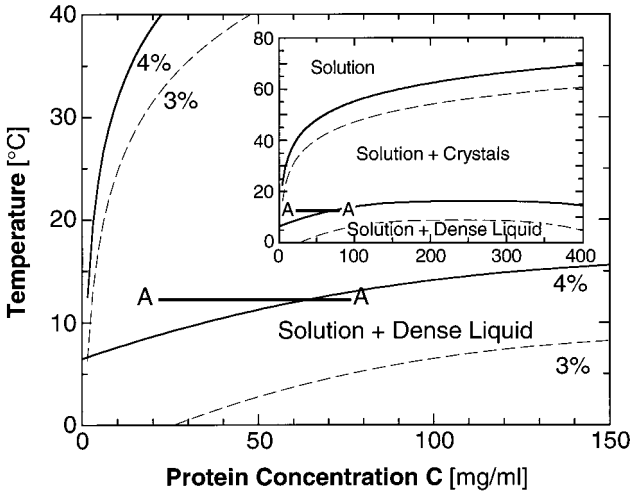

Figure 8. A section of the phase diagram of lysozyme solutions (whole phase diagram is shown in the inset) in the presence of 3% and 4% NaCl. The 3% liquid−liquid coexistence curve is according to experimental determinations in ref 67; the 4% curve is calculated using eq 6 and parameters determined in the same study.[67] Solubility curves for tetragonal crystals (upper pair) at both NaCl concentrations were calculated using empirical formulas from ref 74. The horizontal line A−A marks conditions used in the studies reported here.

for $C_{\mathrm{NaCl}} = 3\%$ and should be even further away for $C_{\mathrm{NaCl}} = 2.5\%$. Thus, we do not expect any effects of the L−L demixing on the nucleation of crystals at these two precipitant concentrations. Correspondingly, no solution clouding or spherulitic crystals were found in these series of experiments. However, the line A−A crosses the L−L separation curve for $C_{\mathrm{NaCl}} = 4\%$. The data point at this precipitant concentration and $C = 68$ mg/mL in Figure 4 was recorded below this curve in the region of liquid−liquid demixing. Correspondingly, as discussed above, during nucleation, the solution was cloudy. Crystal nucleation occurring below the L−L phase separation boundary has higher energy barriers than at the boundary, or slightly above it.[45] We attribute the slow nucleation at these experimental condition (slower that the nucleation at a lower protein and precipitant concentration) to this effect. Note that since three phases, dilute solution, concentrated liquid, and crystals are present, the nucleation theorem does not apply to nucleation below the L−L phase boundary.

The liquid−liquid separation boundary is the location where the compressibility $(\partial V/\partial p)_{T=\mathrm{const}} = (\partial^2\Delta G/\partial p^2)_T$ reaches a maximum, i.e., the fluctuations in the system are highest. At the point on that curve (C^{crit},T^{crit}), $(\partial^2\Delta G/\partial p^2)_T$, and the fluctuation amplitudes even reach ∞ (or the container size). This has been predicted to enhance the nucleation of a new phase and sharply increase the crystal nucleation rate in the vicinity of the phase boundary.[44,45,91] Minor variations of the solution composition that cause small changes in the locations of the L−L curve may have a strong effect on the nucleation rate. We attribute the bifurcation in $J(C)$ for $C > 48$ mg/mL to such variations. One would expect the highest deviation between the branches to occur about the demixing concentration of ~66 mg/mL. Quite surprisingly, the strongest deviation is observed at $C = 56$ mg/mL. We tentatively correlate this shift with the result of a phase field computation of the nucleation barrier[45] which predicts a shift of the minimum from the binodal point by about +1 K. Figure 8 shows that around the experimental conditions this shift corresponds to a −(10−15) mg/mL change in protein concentration.

(86) Laaksonen, A.; Telaquer, V.; Oxtoby, D. W. *Annu. Rev. Phys. Chem.* **1995**, *46*, 489−524.
(87) Schenter, G. K.; Kathmann, S. M.; Garrett, B. C. *Phys. Rev. Lett.* **1999**, *82*, 3483−3487.
(88) McGraw, R.; Laaksonen, A. *Phys. Rev. Lett.* **1996**, *76*, 2754−2757.
(89) Chernov, A. A. *Modern Crystallography III: Growth of Crystals*; Springer: Berlin, 1984.
(90) Tsekova, D.; Dimitrova, S.; Nanev, C. N. *J. Cryst. Growth* **1999**, *196*, 226−233.
(91) Wolde, P. R. t.; Oxtoby, D. W.; Frenkel, D. *Phys. Rev. Lett.* **1998**, *81*, 3695−3698.

It is appealing to view in a unified way the sequence $n^* =$ (4 or 5) $\rightarrow$ $n^* =$ (1 or 2) $\rightarrow$ L−L separation occurring with increasing σ at $C_{NaCl} = 4\%$. In this sequence, the L−L phase transition appears as an extension of crystal nucleation for $n^* = 0$. There are arguments against such interpretation. The nucleus size is determined by the distance in the (C,T) plane in Figure 8 between the point corresponding to the experiment conditions and the solubility curve, while the location of the L−L boundary may be independent of the solubility curve. Although the two boundaries may be related, as in recent simulations with a spherically symmetric intermolecular potential,[42,44] the relation should not necessarily result in the discussed sequence. For instance, in the case discussed in refs 42 and 44, the crystal nucleus size in the direct proximity of the L−L boundary was about 50 molecules. On the other hand, spherically symmetric models fail to correctly predict the locations of the solubility and the L−L boundaries.[92] Only models based on asymmetric potentials have recently been able to quantitatively reproduce a typical protein phase diagram.[92] Thus, if such advanced models are applied to simulations of the crystal nucleation close to the L−L boundary, an answer could emerge as to whether the L−L phase transition can be viewed as an extension of crystal nucleation at high supersaturations.

Conclusions

We found that for the chosen model protein system, crystal nucleation is an intrinsically stochastic process. In this respect, protein nucleation is similar to nucleation of simple liquids or water-soluble inorganic materials. Multiple repetitions of a nucleation experiment under identical conditions allow reproducible determinations of its characteristic rates. Variations of the time allocated for nucleation allow differentiation between homogeneous and heterogeneous nucleation. Although the dependencies of the homogeneous nucleation rate on protein and precipitant concentration are similar to those found in small-molecule systems, the nuclei consist of only a few molecules. Strictly speaking, this precludes direct applications of classical nucleation theory to the studied system. With increasing supersaturation, imposed by increasing protein or precipitant concentrations, the nucleus size takes discrete values of 10 or 11, then 4 or 5, then 1 or 2 molecules. This leads to a broken dependence of the nucleation rate on supersaturation that is beyond the predicting capabilities of classical nucleation theories. Furthermore, if crystals nucleate under conditions that are close to the liquid−liquid separation boundary in the phase diagram of the protein solution, crystal nucleation kinetics are affected. In the region of liquid−liquid separation, nucleation rate is lower than expected for the given protein concentrations and temperature. On the other hand, experiments in the area immediately above this phase boundary yielded nucleation rates that vary by a factor of up to 2 in identical experiments. Theory predicts a sharp nucleation rate maximum in this area of the phase diagram. Hence, we attribute this kinetic instability to minor shifts of the experimental conditions that affect the exact distance of the experimental conditions to the phase boundary and the location of maximum rate.

Acknowledgment. We thank D. W. Oxtoby and D. N. Petsev for helpful discussions, M. D. Serrano for critical reading of the manuscript, and L. Carver for expert graphics work. Support by the National Blood, Hearth and Lung Institute, NIH (Grant No. HL RO1 58038), the Life and Microgravity Sciences and Applications Division of NASA (Grant Nos. NAG8 1354 and 97-HEDS-0250), and the State of Alabama though the Center for Microgravity and Materials Research at the University of Alabama in Huntsville is gratefully acknowledged.

JA9930869

(92) Lomakin, A.; Asherie, N.; Benedek, G. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9465−9468.