Nucleation of protein crystals in a wide continuous supersaturation gradient

A. Penkova,[a] N. Chayen,[b] E. Saridakis[b] and Chr. N. Nanev[a]*

[a]*Institute of Physical Chemistry, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria, and* [b]*Biological Structure and Function Section, Division of Biomedical Science, Sir Alexander Fleming Building, Imperial College of Science, Technology and Medicine, London SW7 2AZ, UK. E-mail: nanev@ipchp.ipc.bas.bg*

By using a supersaturation gradient along a protein solution contained in a glass capillary tube, we modified the classical double pulse technique, thus substantially accelerating the procedure of measurement of nucleation parameters. Data for the number of crystal nuclei, **n** *vs* nucleation time, **t**, were obtained for hen-egg-white lysozyme, chosen as a model because of the availability of reliable solubility data in the literature. The stationary nucleation rate and the nucleation time lag have been measured. Quantitative data for the work required for nucleus formation ($\mathbf{A_k} = 4.3\text{x}10^{-13}$ erg) and the size of the critical cluster (three molecules) were also obtained. Besides, it was observed that Ostwald ripening seems to play an important role for nucleation times longer than 150 min. Using the same technique, semi-quantitative investigations were performed with porcine pancreatic trypsin.

Keywords: protein crystal nucleation; direct measurements with hen-egg-white lysozyme; HEWL, trypsin

1. Introduction

Nucleation is a key step in the crystallization process. Being its first stage, it determines to a great extent many features of the subsequent crystal growth. The nucleation rate, for instance, determines the number of crystals that may grow.

Unfortunately, due to the inherent difficulties in monitoring the nuclei themselves, our present knowledge about the nucleation stage in protein crystallization is insufficient. A single successful attempt has been reported (Yau & Vekilov, 2000) to visualize directly, by Atomic Force Microscopy, the shape of near-critical-size clusters formed during the crystallization of apoferritin. The shape of the critical nucleus has been interpolated on the basis of the coincidence in shape of sub-critical and super-critical clusters. Rather surprisingly, the results indicated that the critical nucleus was by no means nearly spherical. It was shaped like a raft and contained some 20 to 50 apoferritin molecules (Yau & Vekilov, 2000).

Light scattering is widely used in the study of homogeneous protein crystal nucleation (Kam *et al*., 1978; Malkin *& McPherson*, 1994). Although indirect, it is a powerful investigation technique. Unfortunately, "for data interpretation it is heavily dependent on assumptions about the interactions between the molecules" (Galkin & Vekilov, 2000).

By separating the nucleation and growth stages, Galkin and Vekilov (2000) have recently developed a novel technique that allows direct determinations of homogeneous nucleation rates. The same classical principle of stage separation had been used earlier in the study of the rate of heterogeneous protein crystal nucleation (Tsekova *et al*., 1999; Nanev & Tsekova, 2000). The results of both studies were interpreted by applying the classical theory of crystal nucleation, developed originally for inorganic small-molecule crystals. Fundamental quantities, such as the work required for nucleus formation, the number of molecules in the critical nucleus, etc. were determined from data for stationary nucleation rates (Tsekova *et al*., 1999; Nanev & Tsekova, 2000; Galkin & Vekilov, 2000).

Crystal nucleation depends exponentially on the supersaturation. Moreover, the most appropriate conditions for protein crystal nucleation frequently lie within a rather narrow interval of supersaturation levels (McPherson, 1999). In basic research however, (e.g. Tsekova *et al*., 1999; Nanev & Tsekova, 2000; Galkin & Vekilov, 2000), the supersaturation is usually changed in discrete and rather wide steps. This approach is justified only with hen-egg-white lysozyme (HEWL) and other intensively studied protein crystals, e.g. ferritin. In general, however, and especially with yet-to-be crystallized proteins, the conditions have to be scanned more carefully since the step-wise supersaturation change is fraught with the risk of completely overlooking the nucleation zone.

In the present work we used a continuous supersaturation gradient in order to probe into some important details of protein crystal nucleation. During the nucleation stage, the supersaturation was varied continuously along a glass capillary tube filled with protein solution at crystallization conditions. The continuous variation of supersaturation was achieved by applying a continuous temperature gradient along the capillary (see also Luft *et al*., 1999). After a given nucleation time, the supersaturation throughout the capillary was brought to a lower value (corresponding to growth without additional nucleation) by placing the capillary at constant temperature. Using this method, quantitative studies on nucleation can be performed, provided that the temperature dependence of the protein solubility at given conditions (protein and precipitating agent concentration, pH, etc.) is precisely known. Since exact data on the temperature dependence of the solubility are available for HEWL, it was chosen as a model.

In the present work we are investigating the stationary rate of nucleation of (tetragonal) lysozyme crystals. In addition to the stationary nucleation rate, another fundamental quantity which characterizes nucleation kinetics, namely the nucleation time lag (Kashchiev, 2000), has also been studied. Our most general aim was to assess the extent to which the theory of crystal nucleation, developed initially for small molecule crystals, is quantitatively valid for protein crystallization, in the case of HEWL.

In order to demonstrate the applicability of the new technique to optimizing protein crystals, the nucleation and growth of porcine trypsin crystals have been studied semi-quantitatively, in a relatively wide continuous supersaturation range.

2. Experimental

2.1. Set up

Strict separation in time of the nucleation and growth stages lies at the root of the classical approach for determining nucleation rates at a given supersaturation. The principle is very simple (Fig. 1a). During the first stage, of higher supersaturation, the crystals are (predominantly) nucleated. By keeping this period relatively short, the crystals do not have enough time to grow. If it is too long Ostwald ripening may occur (see below). After the (expected) onset of nucleation, the supersaturation is rapidly lowered to below the threshold which is necessary for nucleation, into the so-called metastable zone, so that no further nuclei may appear during that second (growth) stage. (Simply, the presence or absence of crystals determines the limits of the metastable zone in preliminary experiments.) Only the existing nuclei grow into crystals during the second stage, which is made long enough for crystals to become

visible. The crystals are then counted under an optical microscope. The stationary nucleation rate is obtained simply by plotting the number **n** of crystals grown to visible sizes *vs* the nucleation time, **t**. The technique is usually called *double pulse technique*. It is in use since the 1930s and enables the measurement of rates of nucleation experimentally, without ever actually seeing the nuclei themselves. By analogy, the image on a photographic film is invisible before its development.

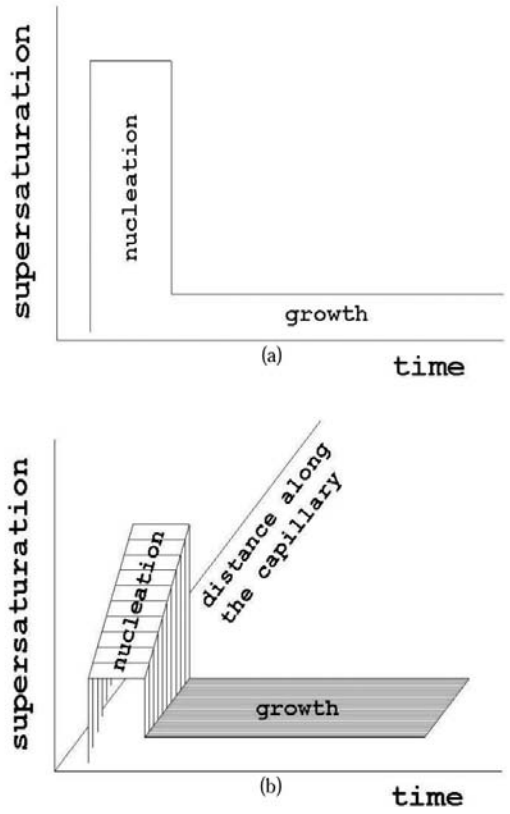

Figure 1

(a) Principle of the classical double pulse technique. (b) A supersaturation gradient is applied along the capillary tube during the nucleation stage.

In the present work we have modified the double pulse technique as follows (the principle is depicted schematically in Fig. 1b). Instead of moving from one starting (nucleation) value for the supersaturation to an end (growth) value, we have produced a continuum of starting (nucleation) supersaturation levels, using a temperature gradient. Thus, several measurements of nucleation rates were performed at once, for different values of the supersaturation. The supersaturation gradient is applied along a (cylinder of) protein solution contained in a glass capillary tube (the "nucleation" area in Fig. 1b). At every point of the capillary tube the supersaturation is kept constant throughout the entire nucleation time period, **t**, and in all experiments of a given series. Temperatures were kept fixed, at 6° C and 20° C, at the two ends of the capillary and thus a nearly linear temperature gradient was established along the tube. Good thermal insulation ensured that the temperature at each point on the capillary remained virtually constant during each stage of the experiment, i.e. that each "slice" of crystallization solution was moved from a single nucleation temperature, depending on its position along the gradient, to the unique growth temperature.

A glass capillary tube with an internal diameter of 550 μm was used. The capillary was graduated in 5 mm segments (18 segments in all). It was then filled with HEWL crystallization solution (described under 2.1) and placed horizontally in the temperature gradient for various time intervals **t**, found appropriate for nucleation during preliminary experiments (we used nucleation times **t** = 10-240 min). Two minutes for capillary pre-tempering have been added before the nucleation time, for every measurement procedure.

At the end of the pre-set nucleation time, the capillary was exposed to a constant temperature of 20 ± 0.3° C, in a constant-temperature room. The supersaturation was thus substantially lowered, and kept constant both in time and along the capillary during the growth stage (Fig. 1b). This temperature had been found in preliminary experiments to be sufficient for crystal growth, but not for nucleation, at the conditions described under 2.1. Growth for approx. one day yielded visible crystals. The number of crystals in each segment was then counted and plotted as a function both of t and of the (initial, i.e. nucleation) supersaturation. This procedure was performed repeatedly for each **t** in order to yield reliable statistics, since the stochastic character of the nucleation process means a high inherent data scatter. Thus, more than five dozens of measurement runs were performed in order to plot **n** *vs* both **t** and supersaturation. The data presented in Fig. 2 were thus averaged over more than 450 single results. The maximum deviation of **n** for a given supersaturation depends on the nucleation time **t**, ranging from ± 40% for longer **t** up to two to three times for the shortest **t** used here (where numbers of crystals are very low).

One advantage of this technique is that a relatively wide range of supersaturations is investigated in a single experiment, ensuring that differences in nucleation rates are not due to the inherently imperfect reproducibility of the nucleation process.

2.2. HEWL solution preparation

Seikagaku 6x crystallized lysozyme was used without additional purification. 30 mg/ml protein was dissolved in a solution containing 3% (w/v) NaCl and 50 mM sodium acetate, pH = 4.5. This HEWL solution is known to be metastable at 20 ± 0.5° C. The thermodynamic supersaturation is: $\Delta\mu = k_B T\ \boldsymbol{\sigma}$, where k_B is Boltzmann's constant, T the absolute temperature and **σ** depends on the solute activities. Neglecting solution non-ideality, **σ** may be expressed as the logarithm of the ratio of the concentration, **c**, over the concentration at equilibrium, $\mathbf{c_e}$, in mg/ml. Thus, in this case $\boldsymbol{\sigma} = \ln(30/\mathbf{c_e})$. Data for $\mathbf{c_e}$ as a function of temperature are taken from Rosenberger *et al*., 1993; Ries-Kautt & Ducruix, 1992; Howard *et al*., 1988, Cacioppo & Pusey, 1991.

2.3. Crystallization investigations with porcine trypsin

In an analogous experiment, this time aimed only at determining an optimum range of supersaturation for crystallization, the glass capillary tube was filled with porcine pancreatic trypsin solution (Sigma, Steinheim, Germany, cat. No. T-0134) and placed in a temperature gradient, where it was held for a time long enough for crystals to nucleate and grow.

The concentrations were chosen by preliminary screening around published conditions (Christofer *et al*., 1998), using the microbatch technique. The chosen conditions were 20 mg/ml protein in 32% (sat) ammonium sulfate, 100 mM Tris, pH 8.4. Weeks after the start of the experiment, the capillary was removed from the gradient and examined under a standard stereoscope.

3. Results and discussion

3.1. Quantitative investigations with HEWL

Crystal nucleation only took place in the parts of the capillary where the supersaturation was greater than that corresponding to the upper limit of the metastable zone. Heterogeneous nucleation is assumed: recent experiments, performed with the capillary in vertical position, have shown that more than two thirds of the (tetragonal) lysozyme crystals are sticking to the glass surface (Nanev *et al*., 2002). We have to keep in mind that the heterogeneous crystal nucleation is easier from theoretical point of view.

In agreement with general expectations, the number n of nucleated crystals increases exponentially with supersaturation, along the capillary. This was verified in all measurement runs, and can be considered as an indication that our method worked reliably.

A fan of eight **n - t** curves is presented in Fig. 2, the number **n** of HEWL crystals per cm^2 being plotted against the mean supersaturation of the corresponding capillary segment. Although we sometimes observed crystals in 12-13 of the 18 segments, only eight **n - t** curves are shown. The nucleation rates being too low for the lowest supersaturations (i.e. from the ninth **n - t** curve onwards), the statistical error on the corresponding curves would be unacceptably high.

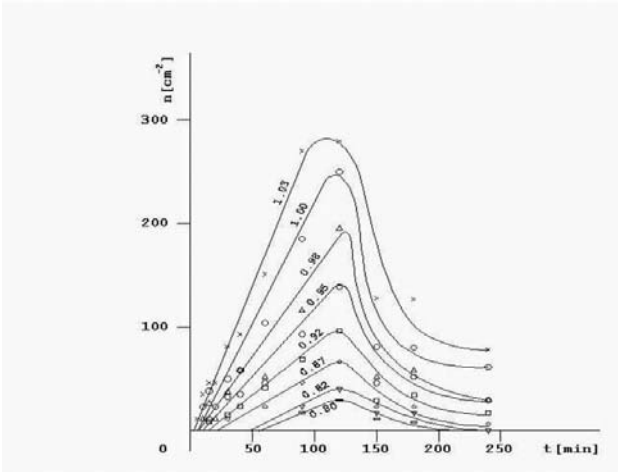

Figure 2

A fan of **n** [number of crystals per square centimeter] *vs* **t** [minutes] curves. The numbers on the curves are showing mean supersaturations, $\Delta\mu x10^{13}$, for every corresponding segment in which the measurement has been done.

The **n - t** curves go through a maximum, at approx. 100-120 min. A linear increase of **n** is observed for shorter times, whereas for longer times the number of crystals (nuclei) falls drastically. Three important features are addressed:

1. The linear **n - t** dependence for nucleation times shorter than 100-120 min, which enables the determination of the stationary nucleation rates, **I**.

2. The distinct time lag τ, which depends on the supersaturation. Due to this time lag, the **n** *vs* **t** curves do not start from the origin.

3. The substantial decreases in **n,** especially for nucleation times longer than 150 min, which may be discussed as a case of Ostwald ripening.

Let us discuss these features in greater detail:

1. From the linear part of the **n** *vs* **t** curves the stationary nucleation rates **I** [nuclei/cm^2 sec] have been calculated. On the basis of the stationary nucleation rates, the work for nucleus formation $\mathbf{A_k}$ and the number **n*** of protein molecules in the critical nucleus have been estimated.

1a. The work $\mathbf{A_k}$, was calculated according to the classical approximation, on the basis of Folmer's expression for the dependence of the nucleation rate on the supersaturation (e.g. see Tsekova *et al*., 1999; Nanev & Tsekova, 2000). Plotting ln **I** *vs* $(\Delta\mu/k_BT)^{-2}$, the slope of the straight line yields a coefficient which enables the calculation of $\mathbf{A_k}$ (Fig. 3). The value obtained for $\mathbf{A_k}$ is $4.3x10^{-13}$ erg. This figure is in good agreement with the values obtained for small molecule crystals as well as the value previously obtained for HEWL crystal nucleation (Tsekova *et al*., 1999; Nanev & Tsekova, 2000).

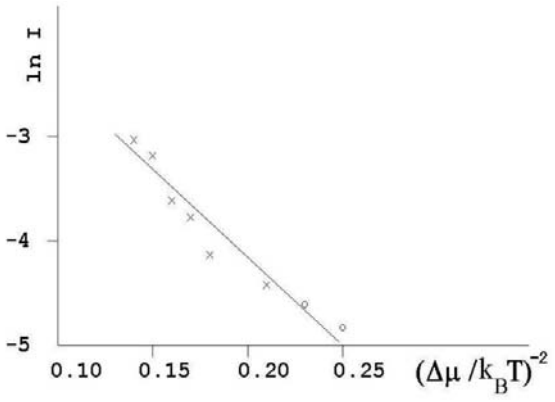

Figure 3

Plot of ln **I** *vs* $(\Delta\mu/k_BT)^{-2}$. The slope of the straight line are yielding $\mathbf{A_k}$, for the notations see text**.** (The circles are denoting points of insufficient statistical accuracy.)

1b. Since the temperature ^{o} K does not change significantly (only by about 4%) along the capillary during the nucleation stage, we are working in a quasi-isothermal approximation. The so-called nucleation theorem (Kashchiev, 1969 & 2000; see also Stoyanov, 1973; Milchev *et al* 1974a, 1974b; Milchev & Stoyanov, 1976; Milchev, 1991) can be used. Its derivation does not rely on the classical approximation and applies down to the smallest nuclei of only a few molecules. According to this theorem, the number **n*** of protein molecules in the critical nucleus is obtained from the plot of ln**I** *vs* $\Delta\mu/k_BT$ (Fig. 4). This plot has already being used (e.g. Nanev & Tsekova, 2000; Galkin & Vekilov, 2000) for determining the size of the critical nucleus:

$$\mathbf{n^*} = d(k_BT \ln \mathbf{I})/d(\Delta\mu)$$

The result, **n*** = 3 molecules, is in full agreement with earlier results for nucleation on a glass support (Nanev & Tsekova, 2000). Small critical nuclei are typical for heterogeneous crystal nucleation. On the other hand, however, dimers, trimers etc are present in protein solutions. Does this mean that those trimers do not have the required bonding/structural properties of a crystalline state although they may be of sufficient size?

2. No nuclei appear at nucleation times shorter than a distinct nucleation time lag τ. We observed that τ is supersaturation-dependent. The higher the supersaturation, the shorter the time that appears to be sufficient for the creation of at least one cluster of critical size (i.e. a critical nucleus). Conversely, the critical supersaturation threshold for nucleus formation depends on the nucleation time, ranging from $\Delta\mu = 0.64x10^{-13}$ erg/molecule (for

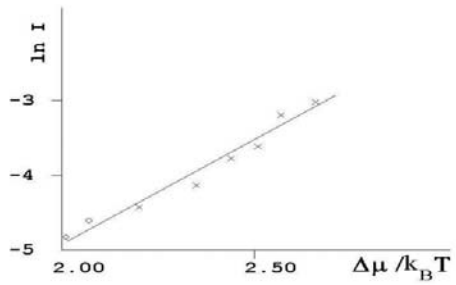

Figure 4

Plot of ln**I** *vs* $\Delta\mu/k_BT$, which is yielding **n***, with **I** [nuclei/cm^2 sec], for the remaining notations see text. (The circles are denoting points of insufficient statistical accuracy.)

t=120 min) to $\Delta\mu = 1.03\text{x}10^{-13}$ erg/molecule (for t=5 min). The supersaturation dependence of the time lag has been determined independently in every measurement, for each nucleation time. The capillary was scanned and the supersaturation below which nucleation did not take place during the given time was recorded. All data (also for the ninth segment onwards) are included in Fig. 5, their inherent scatter being taken into account. The τ-data from Fig. 2 can be fitted to a linear plot, which is shown on the right-hand side of the curve in Fig. 5. In spite of the large data scatter, the remaining experimental results do not fit a straight line.

The direct measurement of τ is another advantage of our technique.

According to calculations performed within the framework of the nucleation theory (Kashchiev, 1969 & 2000) the straight-line τ *vs* $(\Delta\mu)^{-1}$ should start from the origin. This is obviously not the case, since τ already diminishes for finite $\Delta\mu$ values. This result shows a discrepancy in the time lag in the case of protein crystallization from that predicted by the nucleation theory developed initially for small molecule crystals. Let us discuss the possible reasons.

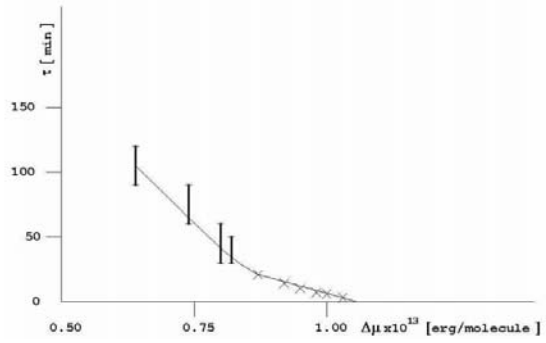

Figure 5

Plot of τ *vs* $\Delta\mu$ is shown. Only the data obtained from the **n** - **t** curves can be fitted to a straight line. (They are supplemented by other experimental results.)

The physical explanation for the time lag in the case of small molecule crystals relies on the fact that steady state is established after a given time. The nucleation process cannot become stationary immediately after the conditions for steady state have been imposed. A certain time elapses before the establishment of the entire set of stationary populations of sub-critical and super-critical clusters in the system, even under conditions ensuring steady state. One may have thought that this explanation remains also valid in the case under consideration, especially since the big protein molecules move more slowly. With a critical nucleus of only three protein molecules, however, this explanation does not appear to be sufficient. An additional explanation lies in the fact that τ is very sensitive to the mechanism of monomer attachment to the nucleus: the difficulties in steric accommodation of the protein molecules into the clusters may thus play an important role. The time required to reach the size of a critical nucleus should also be added (for a comprehensive discussion of the nucleation lag time, see Kashchiev, 2000, pp. 249-270).

Another reason for the observed differences may be the high volume fraction of protein in a solution at typical crystallization conditions, resulting in smaller intermolecular distances. The large size of the macromolecules should also be kept in mind.

3. The decrease in the number of crystals grown to visible sizes for nucleation times longer than 150 min may be explained by Ostwald ripening. A coalescence of nuclei is less probable since under the moderate nucleation rates chosen, the number of critically sized crystals is relatively small and they are thus likely to be sufficiently far from each other to avoid contact.

Ostwald ripening is a well-known process connected with the competitive growth of ensembles of clusters. Due to the Gibbs-Thomson effect, smaller crystals have higher solubilities. Thus, larger crystals grow faster. As these crystals grow, the solution becomes protein-depleted and the supersaturation in the system drops. This leads to an increase in the critical cluster size, making the smaller crystals sub-critical. The latter shrink and dissolve, supplying the larger crystals with monomers for further growth. Only the largest supercritical crystals thus survive. Theoretically the process is completed only after the largest crystal has won the competition: this may be the case here with the lowest supersaturations in Fig. 2.

Something must be said here about the accuracy of this experimental approach. By considering segments of capillary, and defining the supersaturation in each of them to be constant and equal to the mean supersaturation of the segment, an approximation is introduced, which is minimized my making the segments as short as is practicable. Since the supersaturation segments remain exactly the same throughout the investigation, however, only the nucleation time **t** changes. Moreover, by using the logarithm of the nucleation rate **I**, the inaccuracy in **n**, which is considerable due to its exponential dependence on the supersaturation, become less important. Furthermore, by plotting ln**I** *vs* $\ln(\Delta\mu/k_BT)$ any systematic error is eliminated. In conclusion, this experimental approach ensures adequate accuracy, especially in view of the inherent data scatter due to the statistical character of the nucleation process. Although perhaps a complication from the theoretical point of view, the use of a thermal gradient in this study offers practical advantages: it accelerates the measurement procedure, makes it easier and uses limited amounts of protein.

3.2. Results of the investigation with porcine pancreatic trypsin

Applying the same experimental approach with porcine pancreatic trypsin, semi-quantitative results on its nucleation and growth have been obtained. We found the most appropriate temperature for the growth of relatively big crystals of this model protein. Large single crystals suitable for X-ray diffraction were only observed in a small part of the capillary, corresponding to a narrow supersaturation range. The solution at the side of the capillary corresponding to lower supersaturation remained clear. At higher supersaturation, micro-crystals, clusters and finally precipitate appeared, in that order. It is thought that the solution that is subjected

to supersaturation levels insufficient for nucleation may be acting as a source of protein molecules that feeds the crystals growing in the adjacent part of the capillary. This offers an additional advantage in screening a continuous supersaturation gradient in a single experiment, which is counterbalanced in our case by the fact that the smaller crystals growing on the other side of the optimal "slice" act as a sink of molecules. We suggest that this effect may be used for further optimization of crystal size. After finding the optimal supersaturation region, the temperature gradient may be reset so that the optimal region is at one end of the capillary, whilst the rest is at a supersaturation insufficient for nucleation. A large molecular feeding source for the crystals would thus be obtained, without a sink.

4. Conclusion

Quantitative agreement, in some respects, of protein crystal nucleation with predictions of the classical nucleation theory has been shown, as well as some deviations. More specifically, the nucleation time lag behaves only qualitatively (insofar as it is supersaturation-dependent) in accordance with the classical theory.

The number of growing protein crystals (e.g. in a batch method) depends on the nucleation rate. Obviously, Ostwald ripening contributes in decreasing the number of nuclei at long nucleation times. Therefore, if the nucleation and growth stages are been uncoupled, it is advantageous to use either short or substantially long nucleation pulses, in order to obtain a smaller number of larger crystals.

This project has been completed with the partial financial support of the Committee for Science at the Council of Ministers of Bulgaria under contract X-804/98. The valuable remarks of two anonymous reviewers on the manuscript are appreciated.

References

Cacioppo, E. & Pusey, M. L. (1991). *J. Cryst. Growth*, **114**, 286-292.
Christofer, G.K., Phipps, A.G. & Gray, R.J. (1998). *J. Cryst. Growth*, **191**, 820-826.
Galkin, O. & Vekilov, P. (2000). *J. Am. Chem. Soc.* **122**, 156-163.
Howard, S. B.. Twigg, P. J., Baird, J. K. & Meehan, E. J. (1988). *J. Cryst. Growth*, **90**, 94-104.
Kam, Z., Shore, H. B. & Feher, G. (1978). *J. Mol. Biol*. **123**, 539-555.
Kashchiev, D. (1969). *Surf. Sci*. **14**, 209-220.
Kashchiev, D. (2000). *Nucleation, Basic Theory and Application*. Oxford: Butterworth-Heinemann.
Luft, J. R., Rak, D. M. & DeTitta, G. T. (1999). *J. Cryst. Growth*, **196**, 447-449.
Malkin, A. J. & McPherson, A. (1994). *Acta Cryst*. D**50**, 385-395.
McPherson, A. (1999). *Crystallization of Biological Macromolecules*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
Milchev, A., Stoyanov, S. & Kaischew, R. (1974a). *Thin Solid Films*, **22**, 255-265.
Milchev, A., Stoyanov, S. & Kaischew, R. (1974b). *Thin Solid Films*, **22**, 267-274.
Milchev, A., Stoyanov, S. (1976). *J. Electroanal. Chem.* **72**, 33-43.
Milchev, A. (1991). *Contemp. Phys.* **32**, 321-332.
Nanev, C. N., Penkova, A. N. & Chayen, N. E. (2002). To be published.
Nanev, C. N. & Tsekova, D. (2000). *Cryst. Res. Technol.* **35**, 189-195.
Ries-Kautt, M. & Ducruix, A. F. (1992). *Crystallization of nucleic acids and proteins: A practical approach*, edited by A. F. Ducruix & R.Giegé, pp. 195-218. Oxford: University Press.
Rosenberger, F., Howard, S. B., Sowers, J. W. & Nyce, T. A. (1993). *J. Cryst. Growth*, **129**, 1-12.
Stoyanov, S. (1973). *Thin Solid Films*, **18**, 91-98.
Tsekova, D., Dimitrova, S. & Nanev, C. N. (1999). *J. Cryst. Growth*, **196**, 226-233.
Yau, S.-T. & Vekilov, P. (2000). *Nature (London)*, **406**, 494-497.