ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Structural Biology 142 (2003) 22–31

Journal of Structural Biology

www.elsevier.com/locate/yjsbi

Nucleation of protein crystals

Juan Manuel García-Ruiz*

Laboratorio de Estudios Cristalográficos (IACT), CSIC-Universidad de Granada, Facultad de Ciencias, Av. Fuenteneuva s/n, 18002 Granada, Spain

Received 5 February 2003

Abstract

This paper introduces nucleation theory applied to crystallizing protein solutions. It is shown that the classical approach explains the available nucleation data under most conditions used for growing protein crystals for structural studies and for industrial crystallization. However, it fails to explain most experimental data on the structure of the critical clusters. It is also shown that for open systems working out of equilibrium, such as hanging-drop and counterdiffusion techniques, the geometry of the Ostwald–Myers protein solubility diagram and the number, size, and quality of the forming crystals depend not only on supersaturation but also on the rate of development of supersaturation.

Keywords: Crystals; Kinetics; Protein; Nucleation

1. Introduction

The birth of a crystal—the so-called nucleation process—is one of the most fascinating problems in protein crystallization. As will be shown in this review, the mechanisms leading to the formation of clusters of molecules displaying translational and rotational order have their own rules completely different from those of crystal growth. Nucleation theory was developed during the first half of the past century for the case of condensation of a drop from its vapor (Becker and Döring, 1935; Volmer, 1939; Volmer and Weber, 1926) and it was later extended to crystallization from melts (Turnbull, 1950; Turnbull and Fisher, 1949) and also from solutions (see Myerson, 1993). There are several reviews on the subject (Chernov, 1984; Mutaftschiev, 1993; Nielsen, 1964) but the "must read" book is probably the one edited by Zettlemoyer (1969). As for the application to protein crystallization, the pioneering work of Feher and Kam (1985) and the review of Veesler and Boistelle (1999) are recommended reading.

Specific information on nucleation kinetics of protein molecules is not abundant, probably because the value of the nucleation stage was not realized by the protein crystallization community until recently. To those looking for micro- to millimeter crystals for structural studies, it is not fair to promise that understanding nucleation will avoid the exploration of the phase space of protein systems and will provide an ab initio determination of crystallization conditions. This is nowadays impossible even for small molecules. But certainly, a better general understanding of the nucleation process will help to set a protocol reducing the variables currently assumed to induce nucleation. In addition, it is well demonstrated that the nucleation step controls the structure of the crystallizing phase and the number of particles (and thus the crystal size) appearing in a crystallization system. This information is critical for mass (large-scale) protein crystallization, an activity with increasing economical interest. Finally, protein molecules will be useful in the future for fundamental studies on nucleation because of their large size, which will permit visualization of the growth of crystals at the very beginning, when indirect information inferred from scattering studies is not very accurate. Last but not least, the in vivo nucleation process in the liquid–liquid region is also interesting for detection and understanding at the molecular level of illnesses related to protein aggregation (Ansari and Datiles, 1999; Galkin et al., 2002; Pande et al., 2001; Tardieu, 1998).

* Fax: +34-958-243384.
E-mail address: jmgruiz@ugr.es (J. Manuel García-Ruiz).

1047-8477/03/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S1047-8477(03)00035-2

This contribution aims to introduce the nucleation theory for solutions and also to provide an overview of the current state of knowledge on protein nucleation studies. To fit the framework of this volume, the paper focuses on the classical nucleation approach, which accounts satisfactorily for most of the conditions under which crystals for structural studies are obtained.

2. Intuitive approach to nucleation

Since the 19th century, it has been known that crystals grow from solutions by accretion. Growth units freely moving in the solution flow toward the crystal and attach to its surfaces under appropriate conditions. Certainly, we have to learn much more about how to tune the crystallization scenario to control growth rate, growth mechanisms, and crystal size and quality, but the process is relatively well known and easy to grasp intuitively. The scenario is not so easy to imagine at the very beginning, when a crystal is born from its mother solution. For instance, it is clear and well tested that, in the absence of impurities blocking the crystal surfaces, a crystal soaked in its supersaturated solution will continue to grow until the solution reaches the equilibrium concentration. Thermodynamics also dictates that a supersaturated solution must return to equilibrium by segregating a solid phase until equilibrium is achieved. However, this goes against observations, as it is easy to demonstrate experimentally that a threshold value of supersaturation must be exceeded for precipitation to occur spontaneously. Let us see the origin of this paradox.

We can imagine a protein solution as a certain number of protein molecules moving freely among the molecules of the solvent. These molecules, being either monomers or oligomers, are the growth units to build the crystal. Let us assume that the growth units are little balls which, by successive aggregation, may form a cluster with a primitive cubic structure (Fig. 1). Let us assume also that the growth units have six bonds arranged perpendicularly with cubic symmetry. The growth units move randomly in the solution, colliding inelastically with each other as billiard balls do, so that from time to time a number of them converge, forming clusters U_i of growth units whose location is subjected to fluctuations. The lifetime of these clusters is governed by the balance between the cohesion forces that tend to maintain the cluster and those that tend to pull apart the growth units, depending on the number of bonds shared by the growth units forming the cluster and the number of unshared bonds pointing toward the solution. The force (F_A) that maintains a cubic cluster of eight growth units (i.e., a cluster U_2 of edge $L = 2$ growth units) is proportional to the number of bonds saturated between them (namely 12), and the force F_S that tends to separate them is proportional to the number of unsaturated bonds on the surface (namely 24). It is clear that F_A (the number of shared bonds in the cluster) is proportional to the volume, while F_S (the number of unshared bonds in the cluster) is proportional to the cluster's surface, so that, for cubic clusters, the energy balance can be written as

$$\Delta G = -\Delta G_v + \Delta G_s \propto -L^3 + 6L^2.$$

Note that for clusters of edges $L < 3$, the force pulling out the surface to break up the cluster is larger than the cohesion force between the growth units forming the cluster. Therefore, these clusters have a larger probability of being dissolved than of continuing to grow. Hence there is a critical size (in the case of this simple model, the cluster of edge 3 growth units) for which F_A equals F_S. This cluster, termed the nucleus of critical size or critical nucleus, has the same probability of growth as of dissolution. All those clusters with a size larger than the size of the critical nucleus will be likely to grow spontaneously. Plotting F_A and F_S contributions to free energy (Fig. 1) allows us to illustrate the main variables affecting the nucleation problem. It is clear that there is an energetic barrier ΔG^* (the nucleation barrier) that must be crossed to induce the formation of stable nuclei. This is the reason why to form crystals, the solution must be supersaturated beyond a threshold value. The value of the free energy of the forming cluster depends

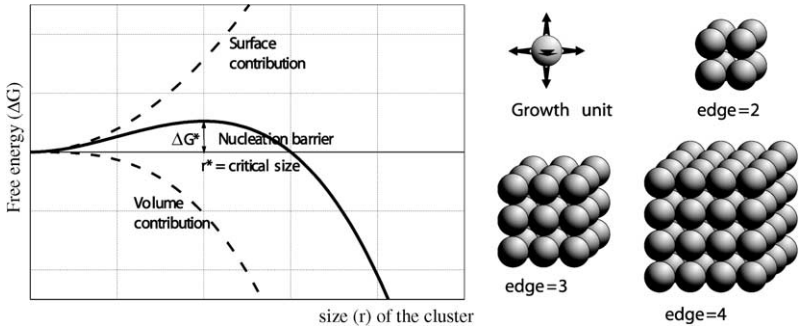

Fig. 1. Energetics of cubic cluster formation as a function of the cluster's size.

on its size, the critical size r^* corresponding to the nucleation barrier. For a given driving force of the system, any cluster with a size below this critical size will tend to dissolve, while clusters with a size larger than the critical size will grow to form crystals. This intuitive approach also emphasizes that the nucleation process is a phase transition that results from fluctuations leading to local variations in the solution density whose lifetime is related to interaction between solute molecules.

3. Formalizing nucleation

Formalization of this simple intuitive model was performed during the first half of the 20th century for the case of the condensation of a drop from its vapor phase (Becker and Döring, 1935; Turnbull, 1950; Turnbull and Fisher, 1949; Volmer, 1939; Volmer and Weber, 1926). The surface term ΔG_s can be expressed as the product of the surface area A of the cluster and the specific energy γ of its surface (to visualize γ, think of surface tension),

$$\Delta G_s = A\gamma.$$

The volume energy is given by the difference in chemical potential of the n molecules in solution μ_α and in the solid phase μ_β:

$$\Delta G_v = -n(\mu_\alpha - \mu_\beta). \tag{1}$$

The chemical potentials of the molecules in the solution and in the solid can be expressed as a function of their respective activities a_α and a_β,

$$\mu_\alpha = kT \ln a_\alpha$$

$$\mu_\beta = kT \ln a_\beta,$$

and, defining supersaturation S as the ratio of the actual activity (or concentration) of the solution divided by the activity (or concentration) at equilibrium, we obtain

$$\mu_\alpha - \mu_\beta = kT \ln S. \tag{2}$$

Substituting into Eq. (1), we have

$$\Delta G = -n[kT \ln S] + A\gamma. \tag{3}$$

Considering for simplicity that the nuclei are spherical (surface energy γ does not change with direction) and making v the molar volume occupied by a growth unit (molecule) in the crystal the ratio between the unit cell volume and the number Z of molecules per unit cell, we can express the energy as a function of the radius of the spherical nucleus,

$$\Delta G = \frac{\frac{4}{3}\pi r^3}{v} kT \ln S + 4\pi r^2 \gamma, \tag{4}$$

known as the Gibbs–Thompson expression for the free energy of a liquid drop of radius r that condenses from its vapor. The formalism is also useful for the case of a crystal growing from its solution, if we include the crystal anisotropy, a property that is not relevant for a liquid drop. To consider the crystalline anisotropy, so that surface energy takes different values for different faces i of a crystal, γ is replaced by an expression of the type $\Sigma A_i\gamma_i$.

Fig. 2 shows graphically the Gibbs–Thompson expression for different values of supersaturation. Note that ΔG increases with r until a critical value ΔG^* is reached and then decreases. Any cluster larger that this critical size experiences a lowering of free energy when it grows, i.e., the cluster becomes more stable as it grows. To reach clusters of that critical size the solution needs to overcome this activation barrier. To calculate the value of the critical radius r^*, we need only to differentiate Eq. (4) with respect to r and set it to zero. Thus, the expression for the critical nucleus is obtained:

$$r^* = \frac{2v\gamma}{kT \ln S}. \tag{5}$$

The value ΔG^* of the activation barrier is immediately given by substituting r^* into Eq. (4):

$$\Delta G^* = \frac{16\pi v^2 \gamma^3}{3[kT \ln S]^2}. \tag{6}$$

Note that the values of r^* and ΔG^* vary inversely with supersaturation (Fig. 2), r^* tends to infinity as S tends to 1. The existence of this energy barrier explains why a solution that should experience precipitation under thermodynamic conditions does so only if a certain value of supersaturation is exceeded.

It should be now clear that nucleation is a probabilistic phenomenon. Density fluctuations, whose spatial and temporal distributions are unpredictable, occur in the bulk of the solution and play the critical role. According to Boltzman's distribution law, the probability of a fluctuation of magnitude W is given by $e^{(W/kT)}$.

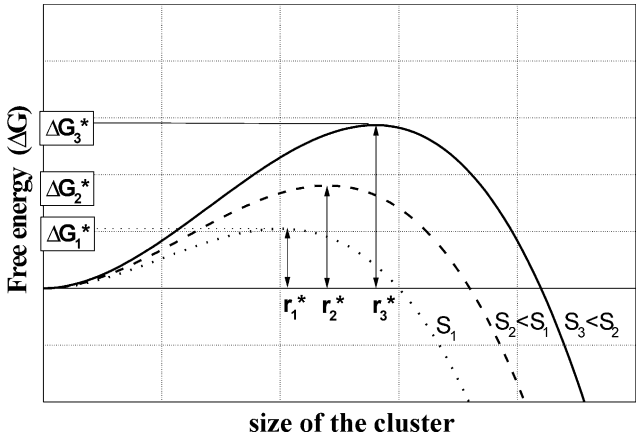

Fig. 2. Variation in the activation barrier as a function of supersaturation. At very high supersaturation, when the size of the critical nucleus becomes smaller than the smallest structural unit, the activation barrier disappears and amorphous phases form.

Thus, the probability for a nucleus of size $i+1$ to be formed from a nucleus of size i is $e\ (-\Delta G_i/kT)$, where ΔG_i is the change in free energy associated with the addition of a growth unit to a cluster of size i. Since $n\Delta G_i = \Delta G$, we have

$$\frac{N_n}{N_1} = \exp\left(\frac{-\Delta G}{kT}\right), \tag{7}$$

or, in another way,

$$N_n = N_1 \exp = \exp\left(\frac{-\Delta G}{kT}\right). \tag{8}$$

Clearly, the nucleation frequency J, i.e., the number of nuclei per unit volume and per unit time that exceed the critical size, can be expressed as

$$J = \kappa_0 \exp\left(\frac{-\Delta G^*}{kT}\right). \tag{9}$$

Substituting the value of G^*, we finally obtain

$$J = \kappa_0 \exp\left(-\frac{16\pi v^2\gamma^3}{3(kT)^3[\ln S]^2}\right). \tag{10}$$

The preexponential term κ_0 is difficult to derive theoretically. It is related to the kinetics of attachment of growth units to the forming cluster. It depends, for instance, on the viscosity of the solution, the molecular charge, the molecular volume, and the density of the solution. The exponential term is related to the activation barrier for nucleation and in principle is more reliable for gaining information about the nucleation process. Note that the factor $16\pi/3$ is a shape factor valid only assuming the cluster to be spherical.

It is straightforward to make use of nucleation kinetics to understand typical protein solubility (Ostwald–Myers) diagrams. The nucleation frequency J is governed by the exponential term of Eq. (10). As shown in Fig. 3, the nucleation flow J is very low at low supersaturation values and then J increases rapidly after a certain critical value of supersaturation S^* is achieved. It explains the existence in the solubility diagram of a region between $S = 1$ (saturation) and S^*, termed the metastable zone, where there is a very low probability of nucleation. The metastable zone is bounded by two curves of different natures. The lower boundary is the solubility curve, defined as the loci of solutions at equilibrium concentration. The location of this curve is fixed and it is of thermodynamic nature. A solution located on this curve will be stable forever. The probability for a nucleation event to occur is 0 and the induction time for nucleation is infinite. The upper limit of the metastable zone is termed the metastable limit or supersolubility curve and, unlike the solubility curve, it is kinetic in nature. It is defined by the loci of the solubility diagram where the probability for a nucleation event to occur is 1, so that the induction time for nucleation can be considered 0 (instantaneous nucleation). Any solution inside the metastable zone will nucleate spontaneously given enough time, the induction time being a function of supersaturation: the higher the supersaturation, the shorter the induction time.

This term of induction time for nucleation deserves further discussion because it is an important parameter describing nucleation kinetics. The induction time τ is defined as the time elapsed between the creation of supersaturation and the formation of nuclei of critical size. For a given volume of solution, the induction time τ is the inverse of the frequency of formation of a critical nucleus J:

$$\tau_N = \frac{\kappa^{te}}{J}. \tag{11}$$

Taking logarithms and expressing τ as a function of S:

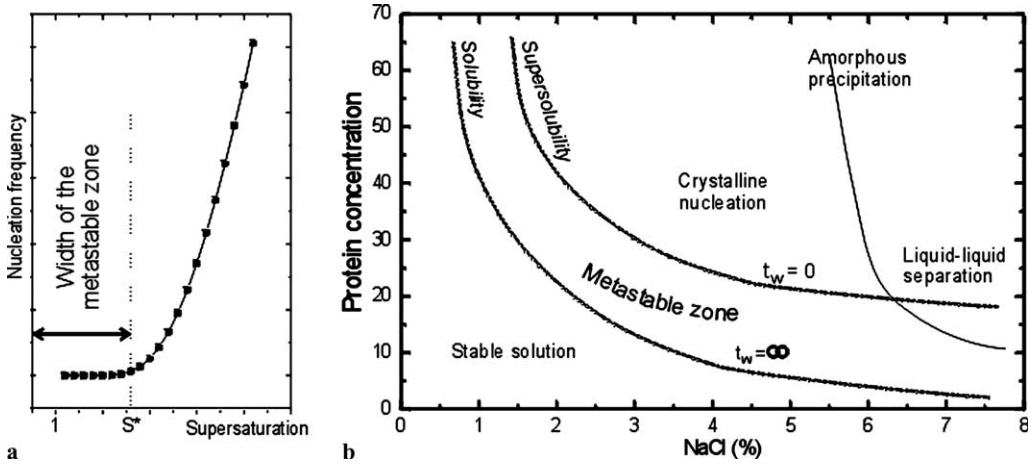

Fig. 3. (a) Nucleation flow as a function of supersaturation. (b) The solubility diagram showing the different phases. Note that the location of the supersolubility (defined for waiting time $t_w = 0$) depends on the rate of development of supersaturation. The location of the liquid–liquid boundary depends strongly on temperature.

$$\log \tau = \kappa_0 + \left[\frac{16\pi v^2\gamma^3}{2.3^3 \cdot 3k^3T^3}\right] / \log^2 S. \quad (12)$$

Therefore, when plotting $\log \tau$ versus $1/\log^2 S$, a linear relationship is obtained whose slope m is given by

$$m = \frac{16\pi v^2\gamma^3}{2.3^3 \cdot 3k^3T^3}. \quad (13)$$

The nucleation time τ is difficult to measure experimentally but the waiting time for nucleation is not. The waiting time for nucleation is the time interval elapsed between the setting of a protein solution at a given supersaturation value and the detection of the first nuclei to form. This waiting time is the sum of three factors (Mullin, 1993):

$$t_w = t_d + t_n + t_g = \tau + t_g. \quad (14)$$

The first term t_d is the time needed to achieve a stationary size distribution of precritical clusters, the second term t_n is the time required to form nuclei of critical size, and the third term t_g accounts for the time required for those nuclei to grow to a size large enough to be experimentally detected. Thus, t_g depends on the technique used for detection of nucleation (Martín-Calleja et al., 1991). For a data set obtained with the same technique, the value of τ can be replaced by t_w in Eq. (14).

The induction time is one of the few parameters related to nucleation that can be assessed experimentally, detaching nucleation from growth being the trickiest part of the experiments. It has been done many times for small molecules (Liu, 2000; Mullin, 1993), but only limited data are available for protein crystallization systems. According to Eq. (12) a plot of $\log \tau$ versus $1/\log^2 S$ should yield a linear relationship, which slope is given by relation (13). At constant temperature, it is then possible to obtain the value of the interfacial tension γ. Then using this value of γ in Eq. (5) the value of the critical nuclei is estimated. Galkin and Vekilov (2000, 2001) found values for γ about $0.6\,\mathrm{mJ\,m^2}$ corresponding to a cluster of about 5–10 molecules. Similar values have been also estimated by Drenth et al. (2003). The linear relationship arising from Eq. (12) has been validated for many small molecule compounds grown from solution. However, in most cases, rather than to one straight line, data on $\log \tau$ versus $1/\log^2 S$ are better fitted to two linear relationships. The linear relation for high supersaturation accounts for homogeneous nucleation while the one at low supersaturation values accounts for heterogeneous nucleation (see below).

It is worth noting that the experimental measurement of the nucleation process depends on the volume of the system, as soon realized by Turnbull (1950) and tested for ice nucleation (Ning and Liu, 2002). In fact, Eq. (10) can be also expressed as the nucleation rate, i.e., the number of nuclei forming per unit time,

$$\frac{\partial N}{\partial t} = V\kappa_0 \exp\left(-\frac{16\pi v^2\gamma^3}{3(kT)^3[\ln S]^2}\right), \quad (15)$$

where V is the volume of the system. This also has implications for protein nucleation studies. Nowadays, so-called high-throughput protein crystallization tends to use large-scale screening for the search of crystallization conditions with the help of robotics. In order to minimize the waste of protein, this strategy implies necessarily reducing the drop size in batch or hanging-drop experiments down to the nano- and picoliter scales (Kuhn et al., 2002; Uber et al., 2002). According to Eq. (15) the number of crystals appearing in a crystallizing drop is proportional to its volume V, which means that scale-up of crystallization conditions from "nanodrops" to "microdrops" must be performed carefully (Bodenstaff et al., 2002).

Back to the solubility diagram, the location of the supersolubility curve (and therefore the width of the metastable zone) is not fixed but depends on the rate at which supersaturation is created. The reason is related again to the stochastic nature of the nucleation phenomenon. This concept is crucial to understand some "magic" behavior of protein crystallization experiments. To catch it, let us use a simile and suppose that we want to cross a number of parallel streets on which cars run with a given frequency (Fig. 4), that frequency increasing from the first street to the last. Evidently, the number of streets that we will be able to pass without being run over depends on the speed at which we cross the streets: the greater the speed, the larger the number of streets that we will be able to cross. This analogy is interesting particularly for the case of nonequilibrium crystallization techniques such as vapor diffusion and counterdiffusion techniques. For a hanging-drop experiment, the value of supersaturation changes continuously due to the evaporation of the drop. Therefore, in the solubility diagram,

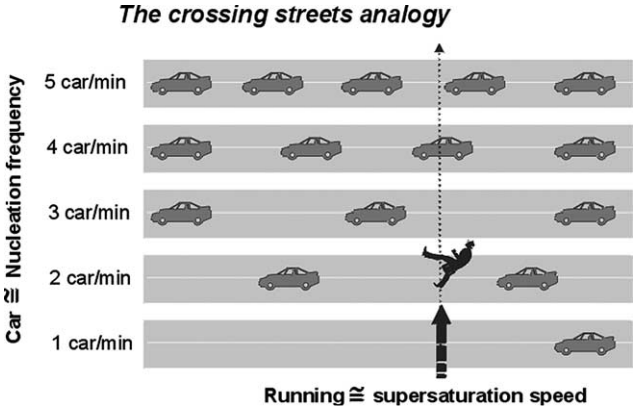

Fig. 4. The number of streets that the pedestrian will be able to cross without being run over depends on his speed. By analogy, the faster an open system moves across the solubility diagram the higher the supersaturation at which nucleation will take place (see text).

the experiment describes a trajectory from the undersaturated to the supersaturated region (Fig. 4). In the metastable region, the experiment moves away from equilibrium as supersaturation increases at a rate $\delta\sigma/\delta t$ that depends on the starting concentration values in the drop and the well and on the geometry of the experimental set-up (Diller and Hol, 1998; Luft and DeTitta, 1997). We can divide the metastable zone into a number of intervals σ_i–σ_{i+1} similar to the streets of our analogy. For a given value of supersaturation rate $\delta\sigma/\delta t$, the residence time in the interval σ_i–σ_{i+1} is t_R. If the induction time τ for a given interval is shorter than the induction time corresponding to the supersaturation value at that interval, a nucleation event will take place when the experiment crosses that interval. However, if the induction time τ for a given interval is longer than the induction time, the supersaturation will continue to increase without nucleation. In other words, if the experiment crosses the metastable region quickly, it will be able to reach higher values of supersaturation before nucleation starts. Thus, the location of the metastable limit depends not only on the technique used to detect nucleation but also on the rate of supersaturation. This dependence of the waiting time for nucleation on supersaturation rate has been previously observed for inorganic compounds for which the width of the metastable zone has been found to depend on cooling rate (Nývlt, 1983).

This last discussion is currently forgotten in classical nucleation approaches, which deal with closed systems such as batch experiments in which supersaturation is immediately achieved upon mixing of protein and precipitating agent solutions. In these closed systems crystals appear after a certain waiting time $t_w(\sigma)$ corresponding to the actual value of supersaturation. However, it is important in open systems moving toward equilibrium, such as most usual protein crystallization techniques. In these cases, the location of the metastable limit depends on how fast the system moves toward the precipitation zone, the nucleation probability being an inverse function of the rate of development of supersaturation. In summary, it means that the nucleation density, the final crystal size, and the structural quality of the crystal are a function not only of the supersaturation σ at which precipitation occurs but also of how fast the supersaturation evolves ($d\sigma/dt$). For an evaluation of the power of a crystallization technique to scan for the best crystallization conditions, the plot of supersaturation versus rate of development of supersaturation is the most reliable one (García-Ruiz, 2003; García-Ruiz et al., 1999).

4. Heterogeneous nucleation

Up to now, we have considered that the protein solution is homogeneous and therefore the probability of a given fluctuation occurring is identical over the whole volume of the system. This is what is termed homogeneous nucleation, which takes place fundamentally inside extremely pure solutions when the supersaturation is sufficiently high to overcome the activation barrier. Under laboratory conditions this is unlikely to occur because of the existence of foreign surfaces, such as the container surface, dust particles, large impurities, etc. Note that the energetics of nucleation concerns mainly the work to create a surface. If there is already a hydrophobic/hydrophilic surface in the system this will decrease the work required to create critical nuclei and will increase locally the probability of nucleation with respect to other locations in the system. This is termed heterogeneous nucleation. Considering the interaction between solute and substrate in terms of the contact angle that the nucleus forms with the substrate, the reduction of the activation energy is given by the equation

$$\Delta G_{het} = \Delta G_{hom}\left(\frac{1}{2} - \frac{3}{4}\cos\alpha + \frac{1}{4}\cos^3\alpha\right).$$

Fig. 5 shows graphically the above expression. Note that if the nucleus wets the substrate completely ($\alpha = 180$), $G_{het} = G_{hom}$; when the contact angle $\alpha = 90$, $G_{het} = 1/2G_{hom}$, and the smaller the contact angle α, the smaller the value of the activation energy for nucleation, which turns to be zero for $\alpha = 0$. Hetereogeneous nucleation has been studied for several hydrophobic surfaces and glasses (Nanev and Tsekova, 2000) and has been also proposed to control the nucleation rate. The idea originally proposed by McPherson (1999) was to induce nucleation by the presence of different mineral surfaces. A more systematic study was performed for the case of lysozyme by Kimble et al. (1998) and for other proteins by Paxton et al. (1999). While the original aim was to look for epitaxial control of nucleation, it appears that the protein nucleation is rather controlled by electric charges. We are still far from a complete

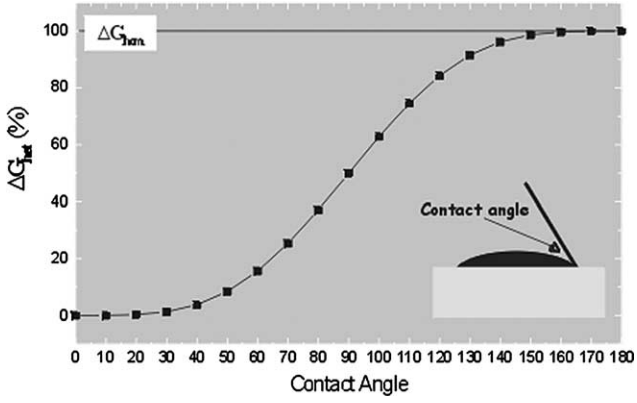

Fig. 5. Reduction in free energy of the nucleation barrier due to heterogeneous nucleation as a function of the contact angle with the substrate.

understanding of the mineral–protein interaction, a problem that is also interesting for the reverse case of biomineralization, i.e., how protein affects the nucleation of minerals (Jimenez-Lopez et al., 2003). The existence of heterogeneous nucleation has been also suggested for protein crystals grown in agarose gels (Vidal et al., 1998) but observed shorter nucleation time can be explained by the increment of actual supersaturation due to uptake of water by agarose molecules.

It is under discussion if homogeneous nucleation can be achieved under typical laboratory conditions (Liu, 2000). As anticipated by Turnbull (1950), the use of small volumes reduces the probability of heterogeneous nucleation. Therefore for most protein crystallization systems, which work at the scale of microliters, nucleation should be homogeneous but this guess has not been yet proven experimentally. In fact, when plotting experimental data in $\log \tau$ versus $1/\log^2 S$ plots, rather than one linear relationship, data are better fitted to two or three different linear relationships (Mullin, 1993; Nývlt, 1983). Even for the case of small molecules, which were much more studied than macromolecules, there are strong differences in the size of the critical nucleus, suggested values ranging from tens to thousands of molecules.

5. The so-called nonclassical approach

The above classical approach is based on two main concepts: (a) the existence of local density fluctuations and (b) the existence of a large structural difference between the degree of order in the solution and the nuclei. In other words, nucleation is a probabilistic phenomenon and to form nuclei implies using energy in creating a surface. This classical approach loses its power of explanation at very high supersaturation values. First, at high supersaturation, the size of the stable cluster is so small that there is no way to differentiate between the surface and the bulk of the cluster. The important feature, however, is that the stable clusters are small enough to contain a number of molecules smaller than the number required to build the repeatable unit cell. Thus, upon aggregation of these little clusters of molecules, a crystalline structure does not necessarily follow, the most probable configuration being an amorphous material lacking translational order. This is the origin of the region of amorphous precipitation in the solubility diagram.

An interesting complementary approach to protein nucleation can be made considering protein solutions as colloidal systems made of particles with a highly anisotropic interaction potential. It is well-known that these systems display (protein solutions also do it) a liquid–liquid immiscibility region (Broide et al., 1991; Darcy and Wiencek, 1999; Ishimoto and Tanaka, 1977; Muschol and Rosenberger, 1997; Tardieu et al., 1999). At high supersaturation, when protein–protein interactions are of very short range (compared to the protein molecular size) in addition to being highly anisotropic, the nucleation process turns out to be different in the vicinity of the metastable critical point for protein solutions (Haas and Drenth, 2000). ten Wolde and Frenkel (1997) have shown using numerical simulations that density fluctuations in a protein solution located near the liquid–liquid boundary may induce the formation of a high-density protein drop surrounded by low-concentration solutions. It is evident that the nucleation process in these drops is easier because of the lower activation barrier than in classical nucleation. So here rather than one step, characteristic of the classical approach, nucleation proceeds by a two-step mechanism. The first is the separation of high protein concentration drops from the bulk of the solution at low protein concentration. The second is the formation of tiny crystals inside these high-concentration drops, which will grow surrounded by a high-protein-concentration liquid film. Although this two-step mechanism has been claimed to be connected with the formation of high-quality crystals, it is a very unlikely scenario for most crystallization trials (see Fig. 3b).

6. Detection of the nucleation process

Several techniques have been used for detecting nucleation in protein solutions, including optical microscopy (Galkin and Vekilov, 1999; Mühlig et al., 2001), atomic force microscopy (Yau and Vekilov, 2000), static (Umbach et al., 1998) and dynamic light scattering (Georgalis and Saenger, 1993; Juárez-Martínez et al., 2001), small-angle X-ray or neutron scattering (Ducruix et al., 1996; Finet et al., 1998; Tardieu et al., 1999, 2000), turbidimetry (Dao et al., 2000; Hu et al., 2001), neutron magnetic resonance (Drenth and Haas, 1998) and Raman spectroscopy (Schwartz and Berglund, 1999), electron microscopy (Michinomae et al., 1999; Yonath et al., 1982), and differential scanning calorimetry (Igarashi et al., 1999).

Classical nucleation assumes that the formation of nuclei occurs by the addition of single growth units, either monomer or oligomers, as postulated in the pioneering work of Feher and Kam (1985). Malkin and McPherson (1994) also suggest an interpretation of their dynamic light scattering data from four different proteins in terms of classical nucleation. While the few kinetic data available suggest that this view is valid (Galkin and Vekilov, 1999; Drenth and Haas, 1998), the information from direct imaging and from the study of the protein interactions is not so clear. For Tardieu and co-workers (Finet et al., 1998) it appears that an undersaturated lysozyme solution (at the metastable zone)

is made of monomeric lysozyme and shows overall repulsive interactions. Moving the solution toward the metastable zone by adding salt and/or lowering the temperature progressively leads to attractive interactions under which crystallization occurs. The same results were also found by Muschol and Rosenberger (1995, 1996) using dynamic light scattering. While these authors find no sign of crystallization intermediates between monomers and actual crystals, other authors have reported the existence of polymeric clusters prior to crystal nucleation. Niimura et al. (1995) reported the existence of two types of clusters: (a) 20–60 nm and (b) less than 4 nm in size. The growth units appear to be monomers leaving from small clusters, which is also suggested by Azuma et al. (1989). Georgalis et al. (1993) infer that these clusters forming at the very beginning of the process have a fractal dimension characteristic of diffusion-limited-aggregation (DLA) processes. It seems that the process starts with the formation of clusters of protein molecules induced by attractive interactions that Rosenberger et al. (1996) interpreted as not real clusters. These labile clusters, which have a very low energetic barrier, have a branched geometry characteristic of DLA processes. They either rearrange to form ordered tetrameric clusters (about 65 nm in size) through hydrophobic interactions, as proposed by Igarashi et al. (1999), or just dissolve to form monomers.

Dynamic light scattering, and small-angle X-ray, and small-angle neutron diffraction mostly inform on the interactions in supersaturated (crystallizing) solutions, but light scattering data are difficult to interpret unambiguously (Galkin and Vekilov, 1999). To gain some insight into the actual nucleation mechanism direct visualization of the nucleation will be extremely useful. However, catching crystals at birth is not an easy matter even for large macromolecules like proteins. In a few cases direct imaging of the nucleation process was achieved with electron and atomic force microscopy. Yonath et al. (1982) observed the crystallization of ribosomal subunits starting with the formation of an amorphous phase that later rearranged to form ordered nuclei. Yau and Vekilov (2000) reported that the nuclei of apoferritin (a globular molecule with a molecular mass of about 500 kDa that crystallizes in the cubic system) are unexpectedly flat rather than spherical, the critical size ranging between 20 and 50 apoferritin molecules. Michinomae et al. (1999) studied the crystallization of hen egg white lysozyme by transmission electron microscopy. They observed the formation of short threads as well as larger amorphous structures with lack of translational and rotational order, something similar to rice balls. It seems that some of these "*sushi*"-type structures later self-rearrange to form crystalline nuclei that continue to grow. What switch converts these *sushi* structures into a crystal lattice arrangement is unknown. If these unstructured clusters are confirmed to exist, then the approach considering the cluster energy to depend on the spherical density profile rather than on the critical radius would be more realistic (Oxtoby, 1998). In summary, the physical reality of the critical nucleus, its geometry, and its internal texture are far from being well understood and certainly the problem of visualizing the earliest stages of protein crystal formation is open to further investigation.

Acknowledgments

I thank Dr. Julyan Cartwright for critically reading the manuscript. I also acknowledge the financial support of the Spanish Ministerio de Ciencia y Tecnología.

References

Ansari, R.R., Datiles, M.B., 1999. Use of dynamic light scattering and Scheimpflug imaging for the early detection of cataracts. Diabetes Tech. Ther. 1, 159–168.

Azuma, T., Tsukamoto, K., Sunagawa, I., 1989. Clustering phenomenon and growth units in lysozyme aqueous solution as revealed by laser light scattering method. J. Crystal Growth 98, 371–376.

Becker, R., Döring, W., 1935. Kinetische Behandlung der Keimbildung in übersättigten Dämpfen. Ann. Phys. 24, 719–752.

Bodenstaff, E.R., Hoedemaeker, F.J., Kuil, M.E., de Vrind, H.P.M., Abrahams, J.P., 2002. The prospects of protein nanocrystallography. Acta Crystallogr. D 58, 1901–1906.

Broide, M.L., Berland, C.R., Pande, J., Ogun, O., Benedek, G.B., 1991. Binary-liquid phase separation of lens protein solutions. Proc. Natl. Acad. Sci. USA 88, 5660–5664.

Chernov, A.A., 1984. Modem Crystallography. Crystal Growth, vol. III. Springer, Berlin.

Dao, L.H., Nguyen, H.M., Mai, H.H., 2000. A fiber optic turbidity system for in-situ monitoring protein aggregation, nucleation and crystallisation. Acta Astronaut. 47, 399–409.

Darcy, P.A., Wiencek, J.M., 1999. Identifying nucleation temperatures for lysozyme via differential scanning calorimetry. J. Cryst. Growth 196, 243–249.

Diller, D.J., Hol, W.G.J., 1998. An accurate numerical model for calculating the equilibration rate of a hanging-drop experiment. Acta Crystallogr. D 55, 656–663.

Drenth, J., Haas, C., 1998. Nucleation in protein crystallization. Acta Crystallogr. D 54, 867–872.

Drenth, J., Haas, C., Leppert, J., Ohlenschläger, O., 2003. Effect of molecular anisotropy on the nucleation of lysozyme, J. Phys. Chem. B., in press.

Ducruix, A., Guilloteau, J.P., Riès-Kautt, M., Tardieu, A., 1996. Protein interactions as seen by solution X-ray scattering prior to crystallogenesis. J. Cryst. Growth 168, 28–39.

Feher, G., Kam, Z., 1985. Nucleation and growth of protein crystals: general principles and assays. Method Enzymol. 114, 77–112.

Finet, S., Bonneté, F., Frouin, J., Provost, K., Tardieu, A., 1998. Lysozyme crystal growth, as observed by small angle X-ray scattering, proceeds without crystallization intermediates. Eur. Biophys. J. 27, 263–271.

Galkin, O., Vekilov, P.G., 1999. Direct determination of the nucleation rates of protein crystals. J. Phys. Chem. B 103, 10965–10971.

Galkin, O., Vekilov, P., 2000. Control of protein crystal nucleation around the metastable liquid–liquid phase boundary. Proc. Natl. Acad. Sci. USA 97, 6277–6281.

Galkin, O., Vekilov, P.G., 2001. Nucleation of protein crystals: critical nuclei, phase behavior, and control pathways. J. Cryst. Growth 232, 63–76.

Galkin, O., Chen, K., Nagel, R.L., Hirsch, R.E., Vekilov, P.G., 2002. Liquid–liquid separation in solutions of normal and sickle cell hemoglobin. Proc. Natl. Acad. Sci. USA 99, 8479–8483.

García-Ruiz, J.M., 2003. Counter-diffusion methods for protein crystallisation. Methods Enzymol., in press.

García-Ruiz, J.M., Novella, M.L., Otálora, F., 1999. Supersaturation patterns in counter-diffusion crystallisation methods followed by Mach–Zehnder interferometry. J. Cryst. Growth 196, 703–710.

Georgalis, Y., Saenger, W., 1993. Time-resolved light scattering studies on protein precrystallization fractal clusters. Adv. Colloid Interface Sci. 46, 165–183.

Georgalis, Y., Zouni, A., Eberstein, W., Saenger, W., 1993. Formation of protein precrystallization fractal clusters. J. Cryst. Growth 126, 245–260.

Haas, C., Drenth, J., 2000. The interface between a protein crystal and an aqueous solution and its effects on nucleation and crystal growth. J. Phys. Chem. B 104, 368–377.

Hu, H., Hale, T., Yang, X., Wilson, L.J., 2001. A spectrophotometer-based method for crystallization induction time period measurement. J. Cryst. Growth 232, 86–92.

Igarashi, K., Azuma, M., Kato, J., Ooshima, H., 1999. The initial stage of crystallization of lysozyme: a differential scanning calorimetric (DSC) study. J. Cryst. Growth 204, 191–200.

Ishimoto, C., Tanaka, T., 1977. Critical behavior of a binary mixture of protein and salt water. Phys. Rev. Lett. 39, 474–477.

Jimenez-Lopez, C., Rodriguez-Navarro, A., Dominguez-Vera, J.M., Garcia-Ruiz, J.M., 2003. Influence of lysozyme on the precipitation of calcium carbonate: a kinetic and morphologic study. Geochim. Cosmochim. Acta, in press.

Juárez-Martínez, G., Garza, C., Castillo, R., Moreno, A., 2001. A dynamic light scattering investigation of the nucleation and growth of thaumatin crystals. J. Cryst. Growth 232, 119–131.

Kimble, W.L., Paxton, T.E., Rousseau, R.W., Sambanis, A., 1998. The effect of mineral substrates on the crystallization of lysozyme. J. Cryst. Growth 187, 268–276.

Kuhn, P., Wilson, K., Patch, M.G., Stevens, R.C., 2002. The genesis of high-throughput structure-based drug discovery using protein crystallography. Curr. Opin. Chem. Biol. 6, 704–710.

Liu, X.Y., 2000. Heterogeneous nucleation or homogeneous nucleation. J. Chem. Phys. 112, 9949–9955.

Luft, J.R., DeTitta, G.T., 1997. Kinetics aspects of macromolecular crystallization. Methods Enzymol. 276, 110–131.

Malkin, A.J., McPherson, A., 1994. Light scattering investigations of nucleation processes and kinetics of crystallization in macromolecular systems. Acta Crystallogr. D 50, 385–395.

Martín-Calleja, J., Alcántara, R., García-Ruiz, J.M., 1991. The detection of salting-out: a comparative study. Cryst. Res. Technol. 26, 35–42.

McPherson, A., 1999. Crystallization of Biological Macromolecules. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Michinomae, M., Mochizuki, M., Ataka, M., 1999. Electron microscopic studies on the initial process of lysozyme crystal growth. J. Cryst. Growth 197, 257–262.

Mühlig, P., Klupsch, T., Schell, U., Hilgenfeld, R., 2001. Observation of the early stage of insulin crystallization by confocal laser scanning microscopy. J. Cryst. Growth 232, 93–101.

Mullin, J.W., 1993. Crystallisation. Butterworth–Heinemann, Oxford.

Muschol, M., Rosenberger, F., 1995. Interactions in under- and supersatured lysozyme solutions: static and dynamic light scattering results. J. Chem. Phys. 103, 10424–10432.

Muschol, M., Rosenberger, F., 1996. Lack of evidence for prenucleation aggregate formation in lysozyme crystal growth solutions. J. Cryst. Growth 167, 738–747.

Muschol, M., Rosenberger, F., 1997. Liquid–liquid phase separation in supersaturated lysozyme solutions and associated precipitate formation/crystallisation. J. Chem. Phys. 107, 1953–1962.

Mutaftschiev, B., 1993. Nucleation theory. In: Hurle, D.T.J. (Ed.), Handbook of Crystal Growth. North-Holland, Amsterdam, pp. 187–247.

Myerson, A.S., 1993. The structure of supersaturated solutions. In: Hurle, D.T.J. (Ed.), Handbook of Crystal Growth. North-Holland, Amsterdam, pp. 247–306.

Nanev, C.N., Tsekova, D., 2000. Heterogeneous nucleation of hen-egg-white lysozyme–molecular approach. Cryst. Res. Technol. 35, 89–195.

Nielsen, A.E., 1964. Kinetics of Precipitation. Pergamon Press, Oxford.

Niimura, N., Minezaki, Y., Ataka, M., Katsura, T., 1995. Aggregation in supersaturated lysozyme solution studied by time-resolved small angle neutron scattering. J. Cryst. Growth 154, 136–144.

Ning, D., Liu, X.Y., 2002. Controlled ice nucleation. Appl. Phys. Lett. 81, 445–447.

Nývlt, J., 1983. Induction period of nucleation and metastable zone width. Coll. Czech. Chem. Commun. 48, 1977–1983.

Oxtoby, D.W., 1998. Nucleation of first-order phase transitions. Acc. Chem. Res. 31, 91–97.

Pande, J., Asherie, N., Lomakin, A., Ogun, O., King, J., Benedek, G.B., 2001. Crystal cataracts: Human genetic cataract caused by protein crystallization. Proc. Natl. Acad. Sci. USA 98, 6116–6120.

Paxton, T.E., Sambanis, A., Rousseau, R.W., 1999. Mineral substrates as heterogeneous nucleants in the crystallization of proteins. J. Cryst. Growth 198, 656–660.

Rosenberger, F., Muschol, M., Thomas, B.R., Vekilov, P.G., 1996. Nucleation and crystallization of globular proteins–What do we know and what is missing. J. Cryst. Growth 168, 1–27.

Schwartz, A.M., Berglund, K.A., 1999. The use of Raman spectroscopy for in situ monitoring of lysozyme concentration during crystallization in a hanging drop. J. Cryst. Growth 203, 599–603.

Tardieu, A., 1998. Alpha-crystallin quaternary structure and interactive properties control eye lens transparency. Int. J. Biol. Macromol. 22, 211–217.

Tardieu, A., Finet, S., Bonneté, F., 2000. Structure of the macromolecular solutions that generate crystals. J. Cryst. Growth 232, 1–9.

Tardieu, A., Le Verge, A., Malfois, M., Bonneté, F., Finet, S., Riès-Kautt, M., Belloni, L., 1999. Proteins in solution: from X-ray scattering intensities to interaction potentials. J. Cryst. Growth 196, 193–203.

ten Wolde, P.R., Frenkel, D., 1997. Enhancement of protein crystal nucleation by critical density fluctuations. Science 277, 1975–1979.

Turnbull, D., 1950. Formation of crystal nuclei in liquid metals. J. Appl. Phys. 21, 1022–1028.

Turnbull, D., Fisher, J.C., 1949. Rate of nucleation in condensed systems. J. Chem. Phys. 17, 71–73.

Uber, D.C., Cornell, E.W., Nordmeyer, R.A., Kolbe, W.F., et al., 2002. An approach to rapid protein crystallisation using nanodroplets. J. Appl. Crystallogr. 35, 278–281.

Umbach, P., Georgalis, Y., Saenger, W., 1998. Time-resolved small angle static light scattering on lysozyme during nucleation and growth. J. Am. Chem. Soc. 120, 2382–2390.

Veesler, S., Boistelle, R., 1999. Diagnostic of pre-nucleation and nucleation by spectroscopic methods and background on the physics of crystal growth. In: Ducruix, A., Giegé, R. (Eds.), Crystallisation of Nucleic Acids and Proteins. Oxford Univ. Press, Oxford, pp. 313–340.

Vidal, O., Robert, M.C., Boué, F., 1998. Gel growth of lysozyme crystals studied by small angle neutron scattering: case of agarose gel, a nucleation promotor. J. Cryst. Growth 192, 257–270.

Volmer, M., 1939. Kinetik der Phasenbildung. Steinkopf, Leipzig.

Volmer, M., Weber, A., 1926. Keimbildung in übersättigten Gebilden. Z. Phys. Chem. 119, 277–301.

Yau, S.-T., Vekilov, P.G., 2000. Quasi-planar nucleus structure in apoferritin crystallization. Nature 406, 494–497.

Yonath, A., Khavitch, G., Tesche, B., Muessig, J., Lorenz, S., Erdmann, V.A., Wittmann, H.G., 1982. The nucleation of crystals of the large ribosomal subunits from *Bacillus stearothermophilus*. Biochem. Int. 5, 629–636.

Zettlemoyer, A.C. (Ed.), 1969. Nucleation. Dekker, New York.

Further reading

Carter Jr., C., 1999. Experimental design, quantitative analysis and the cartography of crystal growth. In: Ducruix, A., Giegé, R. (Eds.), Crystallisation of Nucleic Acids and Proteins. Oxford Univ. Press, Oxford, pp. 313–340.