Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Optimization of the critical nuclear size for protein crystallization: a note

Emmanuel Saridakis†

The Biochemistry Laboratory, Biology Department, University of Athens, Panepistimiopolis, Athens 15 701, Greece

† Present address: Biophysics Section, Physics Department, Imperial College of Science, Technology and Medicine, Prince Consort Road, London SW7 2AZ, England.

Correspondence e-mail: e.saridakis@ic.ac.uk

Received 2 April 1999
Accepted 20 October 1999

It was observed that for some proteins the best crystals for X-ray diffraction have been obtained at supersaturation ratios of *ca* 2.5–3 (in experiments without seeding). It was then noticed that under certain conditions specific to the protein such values are close to a local minimum of the critical radius for nucleation. A relation between the two observations is proposed.

1. Introduction

Nucleation is being increasingly recognized as a crucial stage in macromolecular crystal growth (Rosenberger *et al.*, 1996). Its excessive onset leads to untimely protein depletion in the crystallization solution and the protein moving towards the formation of many tiny crystals instead of being used in the growth of a few large ones. By the same token, useful space in the drop is being taken up by these small crystals, not leaving enough space for large ones to grow unimpeded.

It has been shown (Boistelle & Astier, 1988) that at a given temperature and supersaturation ratio (defined as the ratio of protein concentration to the solubility of the protein) of a crystallization system, there is a critical radius r^* at which a crystal aggregate is stable (a nucleus). It is given by

$$r^* = 2\Omega\gamma/[kT\ln(C/S)], \qquad (1)$$

where Ω is the volume of the protein molecule inside the crystal, γ is the interfacial free energy per unit area between nucleus and solution, k is Boltzmann's constant, C is the protein concentration and S is the protein concentration at saturation (the solubility).

Theoretically, if one molecule is withdrawn from a nucleus of radius r^*, the latter dissolves spontaneously; if one molecule is added, it grows spontaneously. The same happens if the supersaturation is slightly decreased or increased, respectively.

r^* obviously tends towards zero, a situation corresponding to excessive nucleation, as the supersaturation ratio C/S increases and, conversely, tends to infinity as C/S approaches 1 – that is, saturation. It is well known, however, that it is not sufficient to be just above $C/S = 1$ for nucleation to take place spontaneously. Rather, there is a metastable zone in any crystallization phase diagram where the supersaturation of the solution is sufficient for growth but not for nucleation (Miers & Isaac, 1907).

Many fairly successful efforts have been devoted, by nucleation–growth uncoupling (Yonath *et al.*, 1982; Stura & Wilson, 1992; Saridakis *et al.*, 1994; Blow *et al.*, 1994), to the growth of protein crystals in the metastable zone which, owing to the slowness of the growth process and the absence of further nucleation, is thought to be a zone of optimum growth conditions (Stura & Wilson, 1992). The large majority of crystallization experiments leading to useful crystals are, however, carried out without the help of nucleation–growth uncoupling techniques and here we will concentrate on these.

2. Results

Because of the arguments advanced earlier against excessive nucleation (that is, against a too small r^*) one might intuitively assume that the best supersaturation conditions for crystallization would be just above the metastable conditions, that is, at the low end of the 'labile zone' of the corresponding crystallization phase diagram.

However, in a much discussed paper by Ataka & Tanaka (1986), it was found that the largest tetragonal crystals of lysozyme grew at a C/S of between 2.5 and 3 and that these ratios were not at the low end of the labile zone but much higher. Crystals grown at the low end of the labile zone were smaller and less well diffracting.

Further published data were sought by the author in order to investigate a possible similarity with other proteins' crystallization conditions. There is only a handful of proteins for which both solubility and crystallization data at comparable conditions (identical precipitant, pH, temperature, additives) are available. Furthermore, only two or three different initial protein concentrations are usually explored in crystallization experiments and sometimes the initial protein concentrations are not even reported. This makes

Table 1
Comparison of available data for a few proteins and assessment of agreement with theoretical result.

Protein (molecular weight, kDa)	Crystal form	Supersaturation ratio for optimum crystals	Source of solubility data	Source of crystallization data	Agreement	Comments
Hen lysozyme (14.5)	Tetragonal	2.5–3	Ataka & Tanaka (1986)	Ataka & Tanaka (1986)	Yes	
Hen lysozyme (14.5)	Orthorhombic	1.5–4	Ataka & Tanaka (1986)	Ataka & Tanaka (1986)	Yes	Less stringent condition than previous
Jack bean canavalin (49)	Trigonal	∼3	DeMattei & Feigelson (1991)	McPherson & Spencer (1975) (from BMCD)	Yes	
Arthrobacter glucose isomerase (43)	Trigonal	26	Chayen *et al.* (1988)	Akins *et al.* (1986)	No	
Calf γ-crystallin II (21)	Tetragonal	∼2.5	Berland *et al.* (1992)	Carlisle *et al.* (1977) (from BMCD)	Yes	
Calf γ-crystallin IIIB (21)	Orthorhombic	See comments	Berland *et al.* (1992)	Chirgadze *et al.* (1977) (from BMCD)	Yes	Solubility curve is too steep for accurate reading, but shows close agreement
Pseudomonas sp. carboxypeptidase G_2 (41.8)	Monoclinic	1.1–1.3	Saridakis *et al.* (1994)	Saridakis *et al.* (1994)	No	
Rhodobacter sphaeroides photoreaction centre (not given in paper or BMCD)	Orthorhombic	∼2.5	Odahara *et al.* (1994)	Odahara *et al.* (1994)	Yes	Membrane protein
Clostridium cellulolyticum endoglucanase A (43)	Orthorhombic	∼2.6	Budayova *et al.* (1999)	Zou *et al.* (1993) (from BMCD)	Yes	
Aspergillus niger acid proteinase A (22.3)	Orthorhombic	∼7	Kudo *et al.* (1996)	Kudo *et al.* (1996)	No	Solubility is extremely sensitive to the additive (DMSO) concentration

comparison of crystallization conditions with standard two-dimensional phase diagrams (in which one of the parameters is protein concentration) difficult. Data for which the optimum supersaturation range for crystallization encompasses the aforementioned range, but is too wide or imprecise for meaningful argument, were excluded. Published data useful for our purpose are presented, with some succinct comments, in Table 1. Out of the ten proteins for which such data have been compared (lysozyme was counted twice to account for two distinct crystal forms), seven follow the result closely. Five of the presented cases have both crystallization and solubility data from the same paper. For the other five, crystallization conditions were researched, in an unbiased way, in the Biological Macromolecular Crystallization Database (Gilliland *et al.*, 1994).

In trying to explain this observation, the author initially thought that C/S = 2.5–3 might correspond to a finite maximum for r^*,

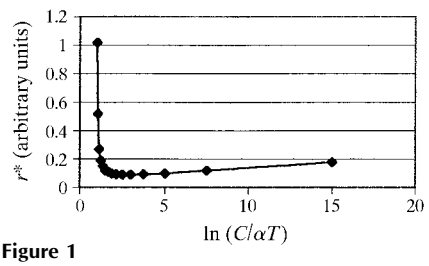

Figure 1
r^* *versus* supersaturation ratio around the area of the local minimum, using arbitrary units for r^* and an arbitrary (but realistic) value of α.

because of the arguments advanced earlier against excessive nucleation. The usual way to find local critical points of a function with respect to one of its variables is to set to zero its first derivative with respect to that variable. Thus, $dr^*/dx = 0$. Ω and γ are assumed constant for a given protein and crystal form and k is a universal constant. We are therefore left with C, T and the variables on which S depends: the concentration of precipitating agent G (assuming this agent has, in some way, been already chosen), the pH and, again, T. Since C, G and the pH appear only once and are part of the argument of the monotonic function 'natural logarithm', no critical points can be expected to arise with respect to any of those. We are left with the temperature, which appears in the denominator of r^* both by itself and as part of S.

Let $S = \alpha T$, where α is a variable depending in any possible way on G, the pH and any other crystallization parameter except temperature. Then

$$\frac{dr^*}{dT} = 2\Omega\gamma[\ln^{-1}(C/\alpha T) - 1]/[kT^2 \ln(C/\alpha T)], \quad (2)$$

which, apart from the trivial zeros as T or C become infinite, is zero for $\ln(C/\alpha T) = 1$ or $C/\alpha T = 2.7$. However, since we have defined $C/\alpha T = C/S$, this is our frequently obtained optimum supersaturation ratio. However, as can be seen in Fig. 1, this critical point is not a local maximum, as was originally thought by the author, but a local minimum. This minimum r^* is quite far from the asymptotic minimum which corresponds to the excessive nucleation characteristic of very high supersaturations.

The assumptions which are made are that (i) the protein solubility is approximately linearly dependent on temperature in the regions of pH and precipitant concentration which are of interest for crystallization; (ii) the protein solubility can be fairly correctly described by an equation of the form $S = \alpha T$, if we let all other relevant parameters be incorporated into the temperature-independent variable α.

The first assumption is by no means valid for all proteins, but it is approximately the case for many. As for the second, which of course requires (i) to be true, it can only be seen as a mathematical simplification rather than a realistic picture of the actual complex interplay of the solubility parameters.

3. Conclusions

An explanation for this observation may be sought in work by Forsythe *et al.* (1994), which provides evidence for the hypothesis that growth proceeds by the addition of preformed ordered aggregates from the solution. These 'growth units' are not stable nuclei; they associate and dissociate, but it is helpful if the association–dissociation process proceeds more slowly than their attachment. Arguably, if these 'growth units' do not have enough time to attach themselves in the right steric orientation before dissociating, premature cessation of growth will occur. The supersaturation ratio

should therefore be of such an order that the difference between r^* and the average size of these aggregates is not very large: therefore, r^* should be fairly low.

It is well known that protein crystallization is a multi-parametric process which can be influenced by hundreds of factors. It is therefore to be expected that the proposed usefulness of minimum r^* conditions will often be offset by other factors which, depending on each macromolecule's specific characteristics, may prove much more important in the given process. A very narrow (or very wide) zone of metastable conditions will of course influence the results. In the case of carboxypeptidase G_2, for example, which was found to disagree with the minimum r^* indication, we have a narrow metastable zone above which heavy homogeneous nucleation takes place (Saridakis *et al.*, 1994). The importance of staying as close as possible to the supersolubility curve, in order to prevent the deleterious effect of excessive nucleation, is therefore overwhelming. Aiming for a minimum r^* in this case would mean than no protein at all would be left to attach onto the growing nucleus, nor would there be any free space in the drop for the crystal to grow. On the contrary, glucose isomerase, the other protein in Table 1 which is at odds with the proposed condition, presents a very wide metastable zone: the theoretical local r^* minimum of 2.7 is by far insufficient for nucleus formation and the supersaturation must be pushed a long way towards the asymptotic r^* minimum. Acid proteinase A is a case where highly specific interactions with dimethylsulfoxide (DMSO) seem to be crucial (Kudo *et al.*, 1996) and this factor is clearly too complex to be tackled with such a limited purely solubility-based approach such as the one presented here.

Furthermore, not only the supersaturation ratio itself, but also the way and speed by which it is reached, are crucial. It is also generally recognized (Stura & Wilson, 1992) that optimum crystal growth necessitates much lower supersaturations than nucleation and it is thus arguable whether one should not try to get closer to optimum crystal growth rather than optimum nucleation conditions when performing crystallization trials without nucleation–growth uncoupling. All the above qualifications show the severe limitations of our approach.

However, given the wide range of possible supersaturation ratios leading to the growth of useful crystals (this very limited study already shows ratios ranging from 1.1 to 26) the observations of Table 1 (five out of ten crystal forms grown at ratios between 2.5 and 3 and an additional two under less stringent conditions centered on the same region) are statistically significant. Therefore, it is suggested that when some solubility information is available and the solubility of the protein is temperature-dependent (in a non-retrograde way), aiming for a supersaturation ratio of approximately 2.5–3 might prove a fruitful route to try, especially when the more intuitive and frequently correct approach of staying as close as possible to the metastable zone fails.

References

Akins, J., Brick, P., Jones, H. B., Hirayama, N., Shaw, P.-C. & Blow, D. M. (1986). *Biochim. Biophys. Acta*, **874**, 375–377.

Ataka, M. & Tanaka, S. (1986). *Biopolymers*, **25**, 337–350.

Berland, C. R., Thurston, G. M., Kondo, M., Broide, M. L., Pande, J., Ogun, O. & Benedek, G. P. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 1214–1218.

Blow, D. M., Chayen, N. E., Lloyd, L. F. & Saridakis, E. (1994). *Protein Sci.* **3**, 1638–1643.

Boistelle, R. & Astier, J. P. (1988). *J. Cryst. Growth*, **90**, 14–30.

Budayova, M., Astier, J.-P., Veesler, S., Czjzek, M., Belaich, A. & Boistelle, R. (1999). *J. Cryst. Growth*, **196**, 297–304.

Carlisle, C. H., Lindley, P. F., Moss, D. S. & Slingsby, C. (1977). *J. Mol. Biol.* **110**, 417–419.

Chayen, N., Akins, J., Campbell-Smith, S. & Blow, D. M. (1988). *J. Cryst. Growth*, **90**, 112–116.

Chirgadze, Y. N., Nikonov, S. V., Garber, M. B. & Reshetnikova, L. S. (1977). *J. Mol. Biol.* **110**, 619–624.

DeMattei, R. S. & Feigelson, R. S. (1991). *J. Cryst. Growth*, **110**, 34–40.

Forsythe, E., Ewing, F. & Pusey, M. (1994). *Acta Cryst.* D**50**, 614–619.

Gilliland, G. L., Tung, M., Blakeslee, D. M. & Ladner, J. (1994). *Acta Cryst.* D**50**, 408–413.

Kudo, N., Ataka, M., Sasaki, H., Muramatsu, T., Katsura, T. & Tanokura, M. (1996). *J. Cryst. Growth*, **168**, 118–123.

McPherson, A. & Spencer, R. (1975). *Arch. Biochem. Biophys.* **169**, 650–661.

Miers, H. A. & Isaac, F. (1907). *Proc. R. Soc. London Ser. A*, **79**, 322–336.

Odahara, T., Ataka, M. & Katsura, T. (1994). *Acta Cryst.* D**50**, 639–642.

Rosenberger, F., Vekilov, P. G., Muschol, M. & Thomas, B. R. (1996). *J. Cryst. Growth*, **168**, 1–27.

Saridakis, E. E. G., Shaw-Stewart, P., Lloyd, L. F. & Blow, D. M. (1994). *Acta Cryst.* D**50**, 293–297.

Stura, E. A. & Wilson, I. A. (1992). *Crystallization of Nucleic Acids and Proteins*, edited by A. Ducruix & R. Giegé, pp. 99–126. Oxford University Press.

Yonath, A., Mussig, J. & Wittman, H. G. (1982). *J. Cell. Biochem.* **19**, 145–155.

Zou, J., Flocco, M. M. & Mowbray, S. L. (1993). *J. Mol. Biol.* **233**, 739–752.