Acta Cryst. (1994). D**50**, 293-297

Phase Diagram and Dilution Experiments in the Crystallization of Carboxypeptidase G_2

BY EMMANUEL E. G. SARIDAKIS

Blackett Laboratory, Imperial College, London SW7 2BZ, England

PATRICK D. SHAW STEWART

Douglas Instruments Ltd, 25J Thames House, 140 Battersea Park Road, London SW11 4NB, England

AND LESLEY F. LLOYD AND DAVID M. BLOW

Blackett Laboratory, Imperial College, London SW7 2BZ, England

(*Received* 20 *July* 1993; *accepted* 24 *November* 1993)

Abstract

The automated microbatch technique developed at Imperial College has been used to establish a phase diagram for crystallization. The concentrations of the protein (carboxypeptidase G_2) and precipitant (PEG 4000) were varied, while pH and temperature were kept constant. The diagram consists of an undersaturation and a supersaturation zone, the latter being subdivided into the metastable, nucleation and precipitation zones. In the metastable zone, crystals may grow but nucleation of crystals does not occur. It is the best zone for growth of X-ray diffraction quality crystals because of the slower growth rate and the avoidance of uncontrolled nucleation, which uses up protein in the formation of tiny crystals. Nevertheless, in practice, it is rarely well defined or used because nuclei must be introduced artificially into the system. The new method used here consists of setting crystallization droplets at nucleation conditions and later diluting them to conditions where nucleation has not been observed. Single diffracting crystals of typical dimensions $0.3 \times 0.3 \times 0.2$ mm were routinely obtained in the metastable zone, equivalent to the best (very rarely) obtained crystals in the nucleation zone.

1. Introduction

The most frequent limiting factor in the elucidation of protein structure by X-ray crystallography is a supply of good-quality readily diffracting protein crystals.

A phase diagram provides a method for quantifying the influence of the concentrations of protein, precipitant(s) and additive(s), temperature and pH, on the production of crystals. A phase diagram for crystallization is a diagram indicating whether the liquid, crystalline or amorphous solid (precipitate) state is stable under a variety of crystallization parameters. These parameters are usually taken two by two, for the ease of display afforded by a two-dimensional matrix. Here, protein and precipitant concentrations are chosen for the two dimensions. Other dimensions, such as temperature and pH, could be added.

The solubility of a crystal form is defined as the concentration of protein in a solution at equilibrium with crystals, under given conditions. The main feature of a two-dimensional phase diagram is the solubility curve (Ries-Kautt & Ducruix, 1992). This curve divides the two-dimensional space of the matrix into the zones of: (i) undersaturation, in which the protein is always fully dissolved and will thus never crystallize, and (ii) supersaturation, in which the concentration of protein is too high for all the protein to remain in solution, in equilibrium with crystals.

A phase diagram for crystallization is usually established for given values of one parameter, by crystallization trials varying the other, until crystals form (Ries-Kautt & Ducruix, 1989, 1992). The solubility is verified by adding crystals to an undersaturated solution, where they will re-dissolve until saturation is reached. Crystals do not always form spontaneously in the zone of supersaturation. Just above the solubility curve, crystals are stable and may grow but nucleation of crystals may not occur. Nucleation, which is the initiation of a new phase, necessitates overcoming a free-energy barrier. The degree of supersaturation might not be high enough for this barrier to be overcome (Mikol & Giegé, 1992). Under certain conditions, amorphous (non-crystalline) precipitates may be obtained.

The boundary between conditions causing nucleation or precipitation, and conditions yielding a clear solution, is therefore not the solubility curve. It is sometimes called a 'supersolubility' curve (Mikol & Giegé, 1992), and defines the spontaneous nucleation zone, the conditions above the energy barrier for nucleation. Part of the supersaturation region is a subdivision called the metastable zone, in which crystals grow, but no nucleation occurs. The solubility curve forms the lower limit of the metastable zone. There are three subdivisions of the supersaturation region: the metastable, nucleation and precipitation zones (Fig. 1).

ISSN 0907-4449 © 1994

The automated microbatch technique for protein crystallization (Chayen, Shaw Stewart, Maeder & Blow, 1990; Chayen, Shaw Stewart & Blow, 1992) involves small volumes of protein, precipitating agent(s) and additive(s) mixed together at the desired concentrations. Crystals grow directly out of this solution, without any re-concentrating process (such as diffusion or vapour equilibration). The crystallization mixtures are dispensed in depressions (wells) previously covered with paraffin oil. This technique makes the establishment of the phase diagram a simpler task. It allows the screening of numerous crystallization conditions at once, varying both protein and precipitant concentrations simultaneously, by narrow increments.

This screening has been used to define three zones in the parameter space: (i) the region in which the solution has remained clear, comprising the undersaturation and metastable zones; (ii) the nucleation zone, in which crystallization is initiated; and (iii) the precipitation zone, in which excess protein separates from the solution in an amorphous state (Fig. 1).

For protein crystal growth by the conventional 'batch' technique, initial conditions must be set in the nucleation zone. During crystal growth the concentration of protein in solution falls, with negligible change in the other conditions.

In the vapour-diffusion technique, the initial conditions are usually undersaturated. During vapour diffusion the concentration of protein and precipitant both increase (they maintain the same ratio during vapour diffusion only if there is only one volatile component, normally water). In common practice, both concentrations would double if no protein precipitation occurred. Vapour diffusion is often almost complete before significant crystal growth occurs, but during crystal growth, the protein concentration in solution falls (Fig. 1). The precise conditions of nucleation are not usually known in the vapour diffusion method because of the complexity of the rates of vapour diffusion and crystallization. In addition, the conditions in the crystallization drop may have moved well into the nucleation zone before growth

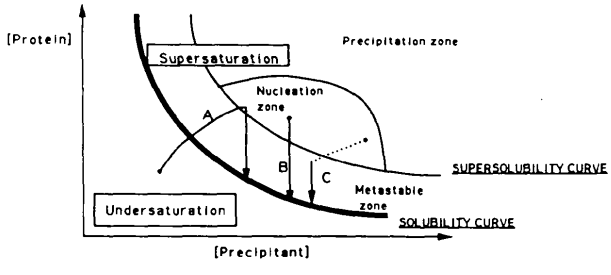

Fig. 1. A theoretical model of a protein crystallization phase diagram, showing the various zones which the protein solution can reach at equilibrium. The three lines *A*, *B* and *C*, show the functional principles of three different crystallization techniques. (*A*) Crystallization by vapour diffusion. (*B*) Batch crystallization. (*C*) Batch crystallization with dilution (the technique presented here).

(and subsequent fall in the protein concentration) starts. This means that excessive nucleation is not necessarily avoided with this technique.

The very slow rate at which supersaturation is reached when PEG is used in a vapour-diffusion system (Mikol, Rodeau & Giegé, 1990), which may be an advantage when large crystals are sought, is certainly a disadvantage in the systematic study of crystallization by vapour diffusion because of the length of the experiments as well as the increased uncertainty about the concentrations in the droplet at the time of nucleation.

If uncontrolled nucleation occurs, crystal growth is likely to lead to the formation of many tiny crystals. In the context of vapour-diffusion experiments, the reservoir has sometimes been diluted to slow crystal growth after nucleation and obtain larger crystals (Yonath, Mussig & Wittman, 1982; Przybylska, 1989). Stura & Wilson (1992) suggest that the proper conditions for growth of large well ordered crystals are in the metastable zone. This proposal is systematically explored in the following experiments, in which crystal growth is carried out in two steps. A solution is initially set by the batch method into the nucleation zone, and the conditions are later adjusted to the metastable zone by direct dilution.

2. Materials and methods

The crystallization trials used carboxypeptidase G_2 (CPG2), a zinc-dependent dimeric enzyme. Purified CPG2 from *Pseudomonas* species, strain RS-16 (Sherwood, Melton, Alwan & Hughes, 1985), was supplied by Dr R. Sherwood of the Public Health Laboratory Service, Porton Down, England. Its concentration was estimated by the absorbance at 280 nm, assuming A_{280} (1 mg ml^{-1}) = 0.5, based on the amino-acid composition (Minton, Atkinson, Bruton & Sherwood, 1984). A single preparation of enzyme, as a stock solution of 6 mg ml^{-1}, in 0.05 *M* Tris–HCl, 0.1 *M* NaCl and 0.1 m*M* $ZnCl_2$ (pH 7.4) was used for these experiments. Before use, samples were concentrated to *ca* 40 mg ml^{-1} in a Centricon C-30 ultrafiltration cell.

Conditions for the crystallization of carboxypeptidase G_2 were reported by Lloyd, Collyer & Sherwood (1991). The presence of zinc acetate as an additive is crucial for obtaining a diffracting crystal form of this protein (Lloyd, Brick, Blow & Lou, 1992).

All the droplets were set up to give a final concentration of 0.1 *M* sodium cacodylate buffer (Aldrich Chemical Co. Ltd), pH 6.3 and 0.2 *M* zinc acetate (Sigma), at 291 K. These concentrations were kept strictly constant in all the droplets. The protein and PEG 4000 (BDH) concentrations were varied. By the use of very wide preliminary screens, the useful range of screening was determined to be between *ca* 6 and 16 mg ml^{-1} of protein, and 8 and 17% PEG.

The Imperial College automated crystallization device (IMPAX) was used (Chayen, Shaw Stewart, Maeder &

Blow, 1990; Chayen, Shaw Stewart & Blow, 1992). Crystallization droplets were dispensed into the wells of tissue-culture plates (Terazaki plates, Sterilin, England) under paraffin oil (BDH). The droplet volumes in the static screenings were set at 2 µl. In the dilution experiments, the initial droplet volumes varied between 1.7 and 2 µl and the final droplet volumes between 2.2 and 2.5 µl.

The dilutions were performed by introducing an additional volume of 0.2 *M* zinc acetate and 0.1 *M* sodium cacodylate buffer solution (in order to keep these concentrations constant) inside the target droplet, using IMPAX. The percentage of dilution (added volume × 100/total final volume of the droplet) was between 10% (for final conditions close to the supersolubility curve) and 30% (for final conditions close to the solubility curve).

When establishing a phase diagram, the quality of the crystals is not important. Ample time delays were allowed before dilutions, in order to ensure formation of nuclei in the droplets destined for dilution. In practice, this meant waiting until the growth of the first visible crystallites, indicating that the crystallization droplet would by then be full of nuclei. Dilution then established which zone the lower concentration condition under examination belonged to. When the aim was to test whether bigger better quality crystals could indeed be obtained in the metastable zone, times prior to the appearance of the first visible crystallites were chosen. Time delays ranged from 5 h to 2 d, according to the initial (nucleation) conditions.

In a further series of experiments, a particular pair of initial and final conditions was selected. The initial and final conditions were 9.24 mg ml^{-1} protein, 12.9% PEG and 7.62 mg ml^{-1} protein, 10.9% PEG, respectively. Dilutions from the initial to the final conditions were performed at various times after setting up (from 1 to 48 h).

The samples were examined under the microscope one, two and three weeks after setup (after dilution, in the case of the dilution experiments). No changes were observed between two and three weeks after setup, but some desiccation of the droplets was evident beyond three weeks.

3. Results and discussion

Fig. 2 presents the results of the static experiments: parts of the supersolubility curve and of the curve separating the nucleation from the precipitation zones are shown, as inferred from the plotted data points.

The accuracy and meaningfulness of determining the limits of the nucleation zone is limited by the relative irreproducibility of the crystallization results: surfaces on which the crystals grow (features of the plastic, contaminants), bacterial action, pH and temperature fluctuations, pipetting errors *etc.*, all mean variations of the borderlines between the different zones of the diagram. The drawing of the curves has therefore involved a certain number of fairly subjective decisions: for example, a monotonic curve shape was implicitly favoured. However, given the high total number of crystallization droplets set up, the number of overlaps of identical conditions giving different results is very low in the regions which are not too close to the zone boundaries. One must bear in mind the usual irreproducibility of crystallization trials, which is not often pointed out due to the limited amount of systematic studies performed. The total amount of protein used in this study is approximately 16 mg. It is unlikely that a larger amount of protein would have led to more decisive results in the areas of greatest irreproducibility, or that it would have significantly altered the shape of the inferred curves. On the contrary, now that a certain confidence with the method has been acquired, a smaller amount of protein should be needed for subsequent similar studies on other proteins.

An obvious characteristic of our protein is the very narrow range of protein and PEG concentrations suitable for successful crystallization. It must be remembered, though, that this phase diagram is only a two-dimensional cross-section of the multidimensional parameter space, at 291 K, using a buffer at pH 6.3.

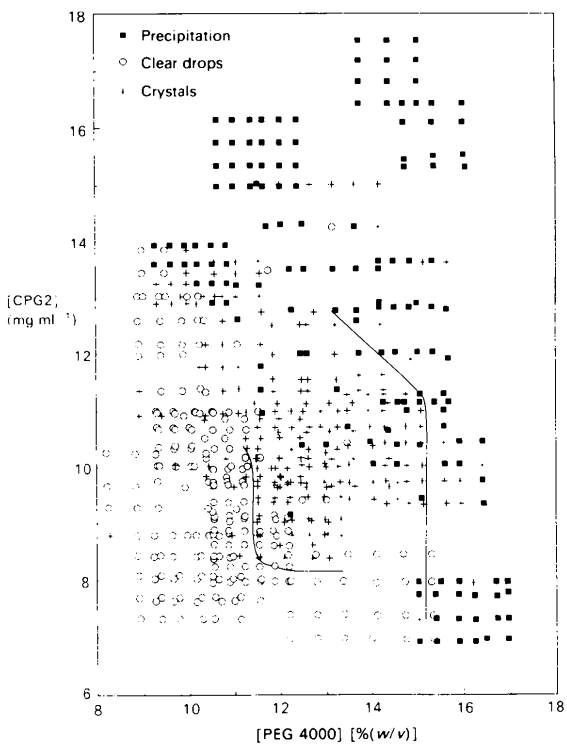

Fig. 2. Phase diagram of CPG2 concentration *versus* PEG concentration derived from the static experiments. Each symbol represents one microbatch droplet and indicates the result of a static experiment. Overlapping data are not shown. Part of the supersolubility curve and of a curve approximately separating the crystalline and amorphous phases are shown.

Below *ca* 8 mg ml^{-1} of protein, either undersaturation or precipitation occur, according to the concentration of PEG. This indicates that, for such low protein concentrations, no ordered aggregation takes place under the chosen conditions.

Two 'triple points', where all phases coexist and maximum irreproducibility of the results follows, are present at two opposite corners of the crystal-yielding region. The equilibrium region at the top (low PEG, high protein) is more extended than the one at the bottom (high PEG, low protein).

Fig. 3 shows the results of dilution experiments, which have been used to define part of the metastable zone, in which crystals grow though they do not nucleate.

The metastable zone corresponds theoretically to the slow growth of crystals without nucleation of new crystals (Ries-Kautt & Ducruix, 1992). Therefore, it represents the ideal conditions for growth of good diffracting crystals. Crystals of smallest dimension 100-200 μm (with largest dimension up to 600 μm) were routinely obtained by dilution into the metastable zone. These are approximately the same size as the best crystals grown from the nucleation zone. Good crystals obtained without dilution are rare and are found on the borderline between the nucleation and metastable zones. In this region most of the droplets do not yield any crystals at all, and crystals occur randomly, with a wide range of dimensions. More than 90% of the crystals from droplets set in the nucleation zone are useless crystallographically. In contrast, dilutions into the metastable zone yield a reliable supply of crystallographically usable crystals (more than 60% of the crystals grown in the metastable zone are of diffracting quality). However, no real improvement in the degree of order of the crystals has been made over previous preliminary crystallographic studies (Lloyd, Collyer & Sherwood, 1991). Diffracted intensities were observed at up to 3.3 Å on film.

In a separate experiment, a particular pair of conditions were chosen, before and after dilution (indicated by the bold symbols on Fig. 3). The solution was incubated at the nucleation conditions for a series of different times before dilution. Fig. 4 presents the number of crystals in each droplet and the largest dimension averaged over all the crystals in the droplet, both against the time of dilution of the droplet in hours. The point for zero time of dilution represents controls set immediately at the final condition. The results show that the shorter the time of dilution (down to 1 h), the fewer and larger the crystals obtained. Many samples, however, gave no crystals. A further study of times of dilution shorter than 1 h gave less reproducible results, leading to no clear conclusion.

'Shock' nucleation, that is, nucleation as a result of the contact of highly concentrated protein stock solution with precipitant stock solution before mixing is complete, is a common problem of many crystallization techniques. Shock nucleation causes the formation of crystal nuclei at conditions which really belong to the metastable region of the phase diagram. The IMPAX microtips dispense all the stock solutions, including protein, simultaneously, and the various solution-containing

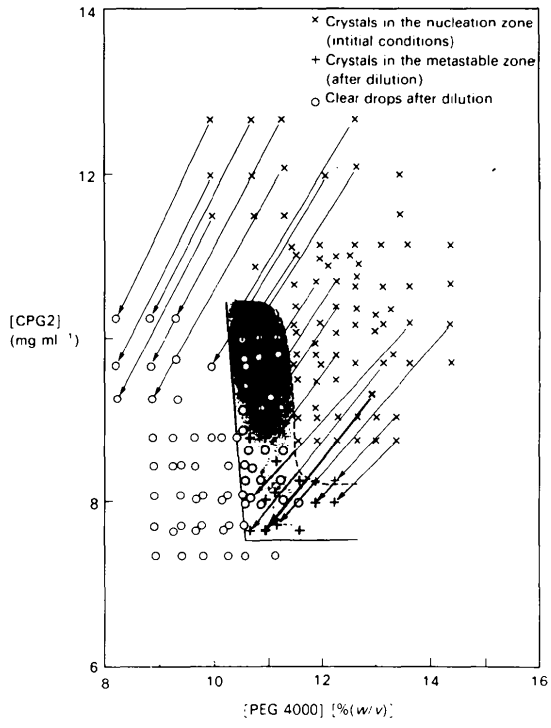

Fig. 3. The results of the dilution experiments. Initial conditions in the nucleation zone are joined in pairs with final conditions. The symbols at the end-point of each vector show the outcome of a dilution experiment. The metastable zone, which is shaded in grey, is bordered by the solubility curve, which is shown as a solid line (the supersolubility curve now shown as a dashed line). For clarity, only a few of the vectors joining the pairs, which represent the dilutions (dilution vectors), are shown. The bold symbols indicate the dilution employed in the timed dilution experiments.

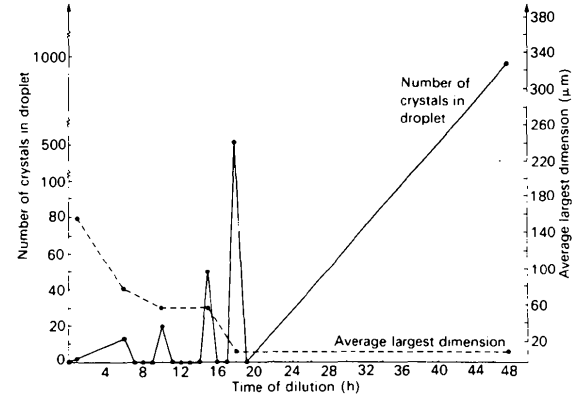

Fig. 4. Graphs of the number of crystals in each droplet and the largest dimension averaged over all the crystals in the droplet, both against the time before dilution of the droplet, in hours. Graphs are for a set of identical drops, corresponding to the dilution illustrated by the bold line in Fig. 3.

channels emerge within 100 μm of each other. Diffusive mixing is rapid, and the dispensed sample is stirred within a few seconds. If stable nuclei are formed during the short time before complete mixing has taken place, no batch method can avoid the problem. However, Fig. 4 shows that at least in certain cases, nucleation proceeds more slowly than this, and shock nucleation is not a crucial parameter. However, shock nucleation may be the main reason why 9% of the static experiments within the zone later identified as 'metastable' resulted in nucleation. It could also mean that the area identified here as 'metastable' is smaller than the true metastable zone.

A systematic study of the various areas of a phase diagram, as presented here, is very time consuming. Batch dilutions at selected conditions, however, can provide a general method for improving the average quality of crystals, when growth at tightly controlled conditions is necessary. Carboxypeptidase G_2 has only rarely given large crystals grown by conventional means. The method of batch dilution provides a reproducible procedure which regularly gives large crystals.

We thank Roger Sherwood of the Division of Biotechnology, PHLS Centre for Applied Microbiology and Research, Porton Down, Salisbury, England, for providing a generous supply of purified protein, and we thank a referee for a useful comment on 'shock' nucleation.

References

CHAYEN, N. E., SHAW STEWART, P. D. & BLOW, D. M. (1992). *J. Cryst. Growth*, **122**, 176-180.

CHAYEN, N. E., SHAW STEWART, P. D., MAEDER, D. L. & BLOW, D. M. (1990). *J. Appl. Cryst.* **23**, 297-302.

LLOYD, L. F., BRICK, P., BLOW, D. M. & LOU, M.-Z. (1992). *J. Cryst. Growth*, **122**, 355-359.

LLOYD, L. F., COLLYER, C. A. & SHERWOOD, R. F. (1991). *J. Mol. Biol.* **220**, 17-18.

MIKOL, V. & GIEGÉ, R. (1992). In *Crystallization of Nucleic Acids and Proteins*, edited by A. DUCRUIX & R. GIEGÉ, pp. 219-239. New York: Oxford Univ. Press.

MIKOL, V., RODEAU, J.-L. & GIEGÉ, R. (1990). *Anal. Biochem.* **186**, 332-339.

MINTON, N. P., ATKINSON, T., BRUTON, C. J. & SHERWOOD, R. F. (1984). *Gene*, **31**, 31-38.

PRZYBYLSKA, M. (1989). *J. Appl. Cryst.* **22**, 115-118.

RIES-KAUTT, M. & DUCRUIX, A. (1989). *J. Biol. Chem.* **264**, 745-748.

RIES-KAUTT, M. & DUCRUIX, A. (1992). In *Crystallization of Nucleic Acids and Proteins*, edited by A. DUCRUIX & R. GIEGÉ, pp. 195-218. New York: Oxford Univ. Press.

SHERWOOD, R. F., MELTON, R. G., ALWAN, S. M. & HUGHES, P. (1985). *Eur. J. Biochem.* **148**, 447-453.

STURA, E. A. & WILSON, I. A. (1992). In *Crystallization of Nucleic Acids and Proteins*, edited by A. DUCRUIX & R. GIEGÉ, pp. 99-126. New York: Oxford Univ. Press.

YONATH, A., MUSSIG, J. & WITTMAN, H. G. (1982). *J. Cell. Biochem.* **19**, 145-155.