Acta Cryst. (1994). D**50**, 355–360

Influence of Polydispersity on Protein Crystallization: a Quasi-Elastic Light-Scattering Study Applied to α-Amylase

BY S. VEESLER, S. MARCQ, S. LAFONT, J. P. ASTIER AND R. BOISTELLE

Centre de Recherche sur les Mécanismes de la Croissance Cristalline, CRMC2-CNRS, Campus de Luminy, Case* 913, *F*-13288 *Marseille CEDEX* 09, *France*

(*Received* 21 *October* 1993; *accepted* 22 *December* 1993)

Abstract

The early stages of the crystallization process of porcine pancreatic α-amylase were investigated by quasi-elastic light scattering. It is shown that at 288 and 293 K the diffusion coefficient does not monotonically change with increasing protein concentration but passes through a maximum at 10 mg ml^{-1}. In supersaturated solutions, prior to nucleation, the protein is strictly monodisperse. Nucleation induces the formation of aggregates and a polydispersity of, for example, 18% for an initial supersaturation $C/C_e = 5.8$. Monodispersity is restored after the nuclei have grown and partially consumed the solute. On the other hand, polydispersity increases up to 20% at 298 K if the protein concentration decreases to 3–4 mg ml^{-1}, values at which the solutions are undersaturated. When the protein concentration exceeds 5–6 mg ml^{-1} the protein becomes monodisperse again. These results, confirmed by those of another system we are studying (bovine pancreatic trypsin inhibitor), are at variance with the statements that supersaturation is always at the origin of aggregation and polydispersity, and that in undersaturated solutions the diffusion coefficient should remain constant for obtaining crystals once the solutions are supersaturated.

1. Introduction

Porcine pancreatic α-amylase with its multidomain organization together with structure–function studies (Pasero, Pierron, Abadie, Chicheportiche & Marchis-Mouren, 1986) is an attractive model for understanding more clearly the nucleation and growth processes of protein crystals. Actually, there are two isoenzymes, I and II, which can be separated by ion-exchange chromatography. Both enzymes have the same molecular weight (55 kDa) and crystallize under two polymorphic modifications named *A* and *B*. Solubility, phase transition, kinetic ripening and growth rates of *A* and *B* have been reported in a previous paper (Boistelle, Astier, Marchis-Mouren, Desseaux & Haser, 1992). Here, we are interested in the very early stages of the crystallization process and investigate the behaviour of the molecules in solutions slightly under- or supersaturated. During this study, we have used the quasi-elastic light-scattering (QELS) technique in order to determine whether the molecules in solution are monomers or *n*-mers and give an evaluation of the polydispersity of the aggregate size distribution at the different stages of the crystallization process.

2. Experimental

In the present study, we have worked with amylase I. It is the only molecular form whose amino-acid residues (496) have been sequenced (Pasero, Pierron, Abadie, Chicheportiche & Marchis-Mouren, 1986). In addition, the crystals of the polymorph *AI* were used for the determination of the crystal structure (Buisson, Duée, Haser & Payan, 1987).

Amylase I was delivered in aqueous solutions at pH 8 (10 m*M* Tris–HCl) containing 6 m*M* NaCl, 1 m*M* $CaCl_2$ and 3 m*M* NaN_3, in the presence of 0.1 m*M* phenylmethylsulfonate fluoride, a protease inhibitor. The protein concentrations ranged from 5 to 30 mg ml^{-1}. The ionic strength of the solution is about 40 m*M*.

The apparatus we used for the QELS experiments consisted of an SEM 633 goniometer (Sematech, Nice), a real-time RTG correlator (Sematech, Nice) with 12 channels on a logarithmic scale covering all ranges of delay times normally required for these experiments, and a Spectra Physics 2017 5 W argon-ion laser. The laser was running at a power ranging from 50 to 500 mW on the 514.5 nm line, depending on the protein concentration. The samples were contained in a 12 mm diameter cylindrical glass cuvette with a flat bottom, immersed in an 80 mm diameter index matching bath filled with 0.22 μm filtered *m*-xylene and thermostated at 278, 288, 293 and 303 K. The samples were filtered through a 0.5 μm millex LCR single-use membrane (Millipore), the analysis volume being about 300 μl. The data were collected

* Laboratory associated with the Universities Aix-Marseille II and III.

at a scattering angle of 30° with a sample time of 0.8 μs. There are two major reasons for working at 30°. First, at this angle, the technique is more sensitive than at larger angles in detecting the presence of large particles and, second, the autocorrelation function has a longer decay rate. Viscosity measurements of the different solvents were performed using a Cinevisco™ temperature-controlled capillary standard viscometer (Sematech, Nice).

Since 1978, when the quasi-elastic light-scattering technique was first used to study the early stages of protein crystallization (Kam, Shore & Feher, 1978), many studies dealing with light scattering and protein solutions have been published. The principles of QELS have been discussed in several books (Cummins & Pike, 1973; Berne & Pecora, 1976; Pecora, 1985). In short, the technique consists of measuring the translational diffusion coefficient D of molecules dispersed in a solvent undergoing Brownian motion. This coefficient depends on several parameters: temperature, pH, concentration in protein, nature and concentration of precipitating agent, interactions between molecules, aggregation *etc.* By extrapolation of the diffusion coefficient to zero concentration of the protein, 'free molecule' or 'free particle' properties can be obtained, together with the particle size. The experiments consist of analysing the time-dependent fluctuations of the light intensity scattered at a scattering vector $\mathbf{q}$ [$q = (4\pi n/\lambda)\sin(\theta/2)$, where n is the refractive index of the solution and θ the scattering angle]. These fluctuations are described by the intensity autocorrelation function. For a solution of monodisperse molecules the autocorrelation function is given by

$$G(\tau) = A[1 + \beta' \exp(-2\Gamma\tau)],$$

where $G(\tau)$ is the intensity autocorrelation function, τ the delay time, A the baseline value, Γ the decay rate ($\Gamma = Dq^2$) and β' an instrument constant. When the system becomes more complex, as in the case of a mixture of several species having different diffusion coefficients, $G(\tau)$ becomes a squared sum of single exponentials with decay rates directly related to the different species. The mean diffusion coefficient D and the polydispersity v of the system were directly determined from the cumulant analysis (Koppel, 1972). When the polydispersity was high ($v > 6\%$), the cumulant analysis gave only qualitative information on the molecules in solution. Accordingly, for a better data analysis we used software based on the singular system and exponential sampling method (Ostrowsky, Sornette, Parker & Pike, 1981; Bertero & Pike, 1991*a*,*b*). The algorithm directly determines a particle size distribution from the QELS data. However, it must be born in mind that this problem, *i.e.* the inversion of the Laplace transform in photon correlation spectroscopy, is an ill posed problem so that a single solution does not exist.

3. Results

3.1. *Diffusion coefficients*

Fig. 1 shows the variation of the diffusion coefficient measured at 293 K as a function of protein concentration, Fig. 2 being a typical particle size distribution diagram. The measurements were performed using three protein batches prepared and purified independently. We see that the three samples behave in the same way. At the lowest concentration (5 mg ml^{-1}), the diffusion coefficient ranges from 4.8 to 5.5 $\times 10^{-7}$ cm^2 s^{-1}. Then it slightly increases with increasing protein concentration passing through a maximum ranging from 6.4 to 7.1 $\times 10^{-7}$ cm^2 s^{-1} at a protein concentration of 10 mg ml^{-1}, before

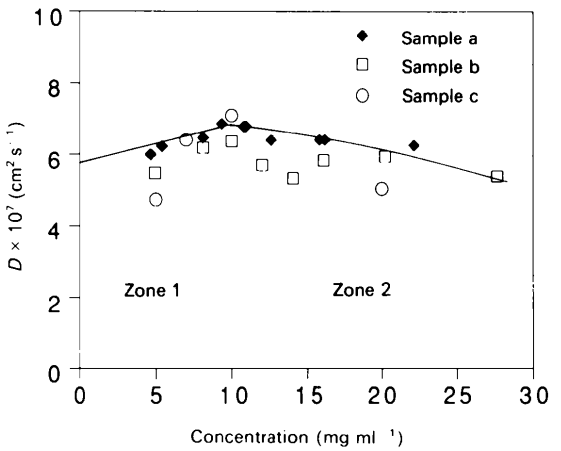

Fig. 1. Diffusion coefficients measured on three different batches of α-amylase *versus* protein concentration at 293 K. The continuous line is only a guideline.

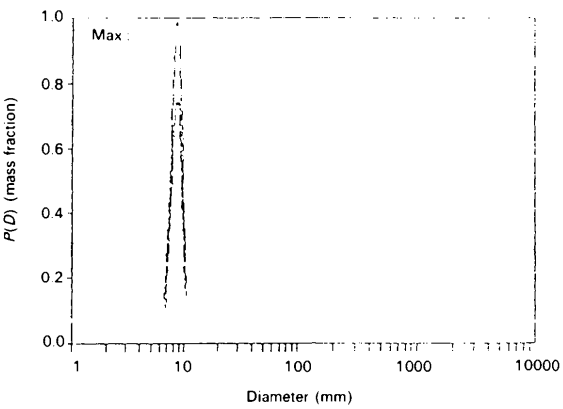

Fig. 2. Size distribution of α-amylase molecules in a supersaturated solution at 16.2 mg ml^{-1} and 293 K, several hours prior to nucleation ($D = 6.24 \times 10^{-7}$ cm^2 s^{-1}, $v = 1.3\%$).

decreasing again, down to 5.4×10^{-7} cm^2 s^{-1} at 28 mg ml^{-1}. The line we have drawn on Fig. 1 must only be considered as a guideline. However, in our opinion, the shape of the curve cannot be attributed to experimental uncertainties of the diffusion coefficient, which, for the same batch, are $\pm 0.2 \times 10^{-7}$ cm^2 s^{-1}. In zone 2, beyond 10 mg ml^{-1}, the translational diffusion coefficient changes in the same way as those observed in other systems such as lysozyme (Skouri, Munch, Lorber, Giegé & Candau, 1992) and bovine pancreatic trypsin inhibitor (BPTI; Gallagher & Woodward, 1989). Let us point out here that for the highest concentrations, higher than 15 mg ml^{-1}, the measurements were performed within a few hours of the beginning of the experiments in order to avoid nucleation problems.

If we now look at the polydispersity of the solutions (Fig. 3) we see that there is no clear relationship between polydispersity, diffusion coefficient and sample number. At a given protein concentration, polydispersity is low, less than 6%. Accordingly, the diffusion coefficients discussed previously are not significantly affected by aggregation.

In Fig. 4 we show the variation of the diffusion coefficient at different temperatures. The same maximum at 10 mg ml^{-1} exists at 288 K but it does not exist at all at 278 K and is only outlined at 303 K. In the latter case, we think that thermal motion has screened the molecular interactions. By extrapolation to zero protein concentration, using the Stokes–Einstein equation $D_0 = k_B T/6\Pi\eta R_h$ where k_B is the Boltzmann constant, T the absolute temperature and η the solution viscosity, we find $D_0 = 8.0$, 5.8, 5.0 and 3.8×10^{-7} cm^2 s^{-1} at 303, 293, 288 and 278 K. The corresponding R_h values are 34.8, 36.9, 37.0 and 34.6 Å, respectively. These latter values of the hydrodynamic radii correspond to the equivalent radius of the monomer obtained from crystallographic studies (Payan, Haser, Pierrot, Frey, Astier, Abadie, Duée & Buisson, 1980): considering the volume occupied by one molecule in the unit cell and assuming that the molecule is a sphere of same volume, its radius is about 31.1 and 38.4 Å for polymorphs *A* and *B*, respectively. The presence of monomers in solution was confirmed by small-angle X-ray scattering measurement in a study which is in progress.

3.2. *Polydispersity*

From all the experiments which were carried out, it was found that polydispersity increases with decreasing protein concentration. Polydispersity becomes very high even when the concentration drops below the solubility curve of polymorph *A*, which is the less soluble polymorph above 291 K whereas *B* is the less soluble form below 291 K. It is unlikely that this point has a real physical significance and there has been no proof until now that polydispersity depends more on the solubility of one polymorph than on the solubility of another. In our opinion, in the present case, polydispersity results from the low protein concentration regardless of the value of the solubility of polymorphs *A* and *B*.

In order to support this idea, we investigated the behaviour of the protein at three temperatures around the solubility curves of *A* and *B*. The principle of the experiments was to start with supersaturated solutions and to dilute them progressively by addition of the mother liquor until they became undersaturated. Conversely, undersaturated solutions were concentrated until they became supersaturated. The results are displayed in Fig. 5 where we have reproduced the solubility curves of *A* and *B* determined previously (Boistelle, Astier, Marchis-Mouren, Desseaux & Haser, 1992). Instead of giving the diffusion coefficients, we have indicated directly

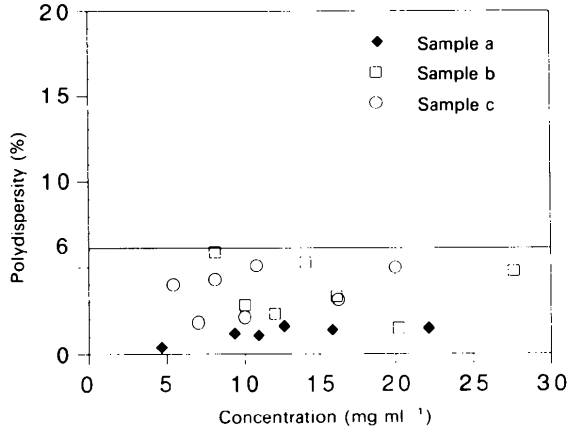

Fig. 3. Polydispersities measured on the same samples as those of Fig. 1.

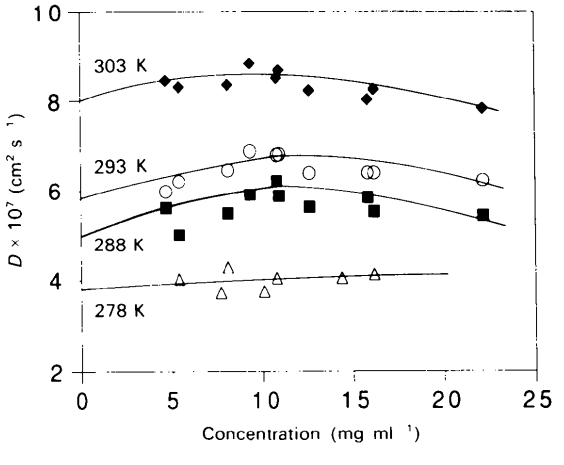

Fig. 4. Influence of temperature on the diffusion coefficient of α-amylase. Continuous lines are only guidelines.

the values of polydispersity measured for each concentration. At both temperatures the solutions become polydisperse with decreasing concentration and monodisperse with increasing concentration. Moreover, polydispersity drastically increases below the solubility curves and we can emphasize that polydispersities as large as 22% cannot result from experimental uncertainties. The aggregation and dis-aggregation process is reversible and follows the dilution and concentration cycles.

3.3. *Pre-nucleation and nucleation*

In the third and final stage of our investigations, we followed the evolution of the supersaturated solutions *versus* time. The experiments were carried out at 278 K with the protein concentration fixed at 9.3 mg ml^{-1} so that the supersaturations, expressed as the ratio of actual concentration over equilibrium concentration $\beta = C/C_e$, were $\beta = 1.9$ and 5.8 with respect to polymorphs *A* and *B*, respectively. Under these conditions, from previous experiments, we know that *B* occurs preferentially to *A* within a reasonable induction time for nucleation. In Fig. 6 we have plotted the evolution of the diffusion coefficient *versus* time, together with the evolution of polydispersity. The figure can be divided into three zones.

In the prenucleation zone 1 ($0 < t < 10$ d), D slowly decreases from 5.0 to 3.8×10^{-7} cm^2 s^{-1}. Polydispersity is low, ranging around 5% with two unexplained surges at 10%.

In the nucleation zone 2 ($10 < t < 13$ d) several particle-size distributions are observed. Polydispersity is around 10–12%.

In the growth zone 3 ($t > 13$ d) amylase crystals are visible. They slowly deposit on the bottom of the glass cuvette so that the QELS measurements are not disturbed by their presence. At 14 d, $D = 3.6 \times 10^{-7}$ cm^2 s^{-1} while polydispersity decreases to 5%.

From this experiment it was found that (i) the solution contains only monomers for a very long period of time, until the induction period for nucleation is exceeded; (ii) nucleation induces the formation of aggregates of different sizes which turn into crystals when they overcome the critical size or dis-aggregate again due to the decrease of the supersaturation; and (iii) growth takes place in a solution where the molecules are monomeric and monodisperse, larger particles not being detected.

However, unlike the study made on lysozyme (Georgalis, Zouni & Saenger, 1992; Georgalis, Zouni, Eberstein & Saenger, 1993) we cannot speculate here on the aggregation mechanism of α-amylase.

4. Discussion

In the concentration range 5–28 mg ml^{-1}, the solutions of α-amylase are supersaturated but perfectly monodisperse provided that the measurements of the diffusion coefficients and polydispersities are made before the induction period for nucleation has elapsed. In the present case the time delay for nucleation lasts several days so that nucleation does not hinder the measurements. At room temperature there are two concentration zones where the molecules behave in different ways, so that the diffusion coefficient passes through a maximum for a protein concentration of about 10 mg ml^{-1}. In both zones polydispersity is low, generally less than 5%. In undersaturated solutions polydispersity increases with decreasing protein concentration whereas in supersaturated solutions the protein remains perfectly monodisperse for several days until nucleation occurs.

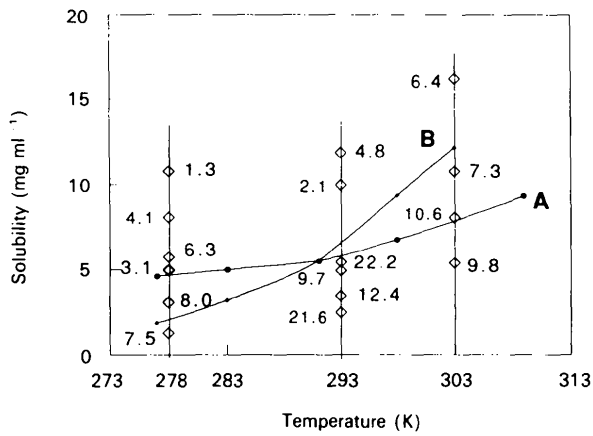

Fig. 5. Polydispersities (%) measured by concentrating or diluting α-amylase solutions with respect to the solubility curves of the polymorphs *A* and *B*.

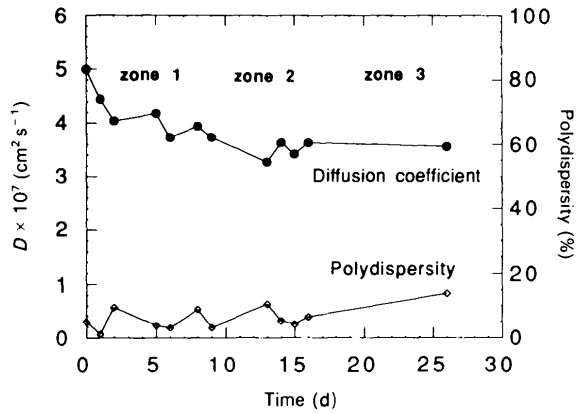

Fig. 6. Evolution of the diffusion coefficient and polydispersity *versus* time in a supersaturated solution where nucleation occurs after 11–13 d.

The decrease of the diffusion coefficient with time at constant protein concentration prior to nucleation (Fig. 6) has already been observed and explained by Bishop (Bishop, Fredericks, Howard & Sawada, 1992). It is due to strong molecular interactions leading to crystallization.

However, some of our results are partially in variance with data collected on other systems. In the case of hen egg-white lysozyme and concanavalin (Mikol, Hirsch & Giegé, 1990; Mikol, Vincendon, Eriani, Hirsch & Giegé, 1991) the authors observed that the variation of the diffusion coefficient with protein and salt concentration in the undersaturated zone could be correlated with the inability of the protein to crystallize once the solution became supersaturated. Conversely, it was claimed that with other precipitants leading to crystallization, there was no variation of the diffusion coefficient up to the saturation point, the protein remaining essentially monodisperse. These statements are not valid in the case of α-amylase where the protein becomes and remains monodisperse with increasing concentration. However, it could be argued that α-amylase solutions deposit good large crystals without any precipitant and that the absence of precipitant is an essential difference with the lysozyme and concanavalin systems. In our opinion, this argument is not totally correct because, in work which is in progress on the crystallization of BPTI in concentrated NaCl solutions, we have observed the same trend as for α-amylase: both the diffusion coefficient and polydispersity significantly decrease with increasing salt or protein concentration. In the supersaturated zone the protein is perfectly monodisperse prior to nucleation and the solutions also deposit good large crystals. From this standpoint our conclusions are partially in agreement with those of Thibault (Thibault, Langowski & Leberman, 1992). According to these authors the solutions deposit crystals on the condition that the percentage of large aggregates stays small and that the diffusion coefficient of the monomer peak stays constant until very close to the precipitation point. Actually, the authors refer to a 'monomer peak' regardless of whether the native protein is in fact monomeric or polymeric. In our case the first condition is fulfilled whereas in our opinion the second one might be expressed in other terms: 'the solution deposits crystals if the protein becomes and stays monodisperse when going from slightly undersaturated to supersaturated solutions regardless of whether the protein is really monomeric or consists of small aggregates.'

The behaviour of α-amylase around the concentration of 10 mg ml^{-1} is not well understood. However, similar behaviour has already been observed on micellar systems. For instance, the mutual diffusion coefficient of tetradecyltrimethylammonium bromide in KBr solution increases with increasing concentration up to a maximum before decreasing again (Walrand, Belloni & Drifford, 1986). The maximum is more pronounced in solutions at low ionic strength and disappears if the KBr concentration exceeds 150 mmol l^{-1}. At low protein concentration (zone 1 in Fig. 1), we directly observe the influence of the electrostatic repulsive interactions. At increasing protein concentration (zone 2 in Fig. 1), the hydrodynamic interactions are preponderant and are responsible for the decrease of D, all solutions behaving nearly as uncharged hard-sphere solutions. Accordingly, the decrease of D cannot be attributed to molecular aggregation. Moreover, we can try to correlate this behaviour of the protein with the solubility and growth rates of the polymorphs which were studied previously. At 293 K, the solubilities of polymorphs A and B are 5.5 and 6.2 mg ml^{-1}, respectively (Fig. 5). Except for the solutions at 5 and 6 mg ml^{-1}, which are slightly undersaturated (Fig. 1), all other solutions are supersaturated. At 10 mg ml^{-1}, the supersaturations are $\beta = 1.8$ and 1.6 for A and B, respectively. We do not attach too much importance to these values but it is worth noting that the crystals hardly grow for $\beta < 1.5$ and very slowly for $1.5 < \beta < 2.0$. This may be due to several causes including the repulsive interactions between molecules in weakly concentrated solutions. This behaviour at low protein concentration could explain why the time delay for nucleation exceeds a few hours.

Finally, concerning this maximum of the diffusion coefficient we can also remark that the surface charges of α-amylase are not screened by a precipitant which implies that repulsive forces may play a role at low protein concentration. As previously suggested (Bishop, Fredericks, Howard & Sawada, 1992), attractive interactions are detectable only above some threshold concentration corresponding to an intermolecular distance of 150 Å for lysozyme, a value to be compared with 240 Å in the case of α-amylase for the concentration of 10 mg ml^{-1}.

5. Concluding remarks

In the present study we have measured the diffusion coefficients of porcine pancreatic α-amylase and paid special attention to the evolution of polydispersity as a function of temperature and protein concentration. The diffusion coefficient passes through a maximum at 10 mg ml^{-1}, this peculiarity being clearly observed in the 288–303 K temperature range. Moreover, the protein does not exactly behave like other proteins with increasing concentration, mainly because the

repulsive interactions are not screened in our system. When the solution is undersaturated polydispersity is high, the growth units associate and dissociate to form monomers and *n*-mers. When the solution is supersaturated but remains in a metastable state, the protein is monodisperse and monomeric. When supersaturation is high enough to induce nucleation, the protein becomes polydisperse. Monodispersity is restored during the growth of the crystals. From these observations and those made in work in progress on bovine pancreatic trypsin inhibitor (in NaCl solutions) and on ornithine carbamoyltransferase (in $MgSO_4$ solutions) we draw the following conclusions: (i) polydispersity of a protein can decrease with increasing protein concentration; (ii) supersaturation does not systematically induce a detectable protein aggregation and an important polydispersity, at least as long as the solution remains in the metastable state; (iii) after the nucleation stage, growth is easy only under the condition that the protein again becomes monodisperse. This implies, in the present case, that the lifetime of under-critical aggregates, which do not turn into crystals, is short.

These conclusions are at variance with the statement that increasing concentrations always favour aggregation and that the diffusion coefficient of the protein should remain constant, or nearly constant in the undersaturated zone for obtaining crystals once the solution is supersaturated. Actually, monodispersity is probably the real pre-requisite for obtaining good crystals.

The authors are indebted to the CM2AO program for financial support, and to M. C. Toselli for technical assistance.

References

BERNE, B. J. & PECORA, R. (1976). *Dynamic Light Scattering: with Applications to Chemistry, Biology and Physics*. New York: John Wiley/Interscience.

BERTERO, M. & PIKE, E. R. (1991*a*). *Inverse Probl.* **7**, 1–20.

BERTERO, M. & PIKE, E. R. (1991*b*). *Inverse Probl.* **7**, 21–41.

BISHOP, J. B., FREDERICKS, W. J., HOWARD, S. B. & SAWADA, T. (1992). *J. Cryst. Growth*, **122**, 41–49.

BOISTELLE, R., ASTIER, J. P., MARCHIS-MOUREN, G., DESSEAUX, V. & HASER, R. (1992). *J. Cryst. Growth*, **123**, 109–120.

BUISSON, G., DUÉE, E., HASER, R. & PAYAN, F. (1987). *EMBO J.* **6**, 3909–3916.

CUMMINS, M. Z. & PIKE, E. R. (1973). *Photon Correlation and Light Beating Spectroscopy*. New York, London: Plenum Press.

GALLAGHER, W. H. & WOODWARD, C. K. (1989). *Biopolymers*, **28**, 2001–2024.

GEORGALIS, Y., ZOUNI, A., EBERSTEIN, W. & SAENGER, W. (1993). *J. Cryst. Growth*, **126**, 245–260.

GEORGALIS, Y., ZOUNI, A. & SAENGER, W. (1992). *J. Cryst. Growth*, **118**, 360–364.

KAM, Z., SHORE, H. B. & FEHER, G. (1978). *J. Mol. Biol.* **123**, 539–555.

KOPPEL, D. E. (1972). *J. Chem. Phys.* **57**, 4814–4820.

MIKOL, V., HIRSCH, E. & GIEGÉ, R. (1990). *J. Mol. Biol.* **213**, 187–195.

MIKOL, V., VINCENDON, P., ERIANI, G., HIRSCH, E. & GIEGÉ, R. (1991). *J. Cryst. Growth*, **110**, 195–200.

OSTROWSKY, N., SORNETTE, D., PARKER, P. & PIKE, E. R. (1981). *Optica Acta*, **28**(8), 1059–1070.

PASERO, L., PIERRON, Y., ABADIE, B., CHICHEPORTICHE, Y. & MARCHIS-MOUREN, G. (1986). *Biochim. Biophys. Acta*, **869**, 147–157.

PAYAN, F., HASER, R., PIERROT, M., FREY, M., ASTIER, J. P., ABADIE, B., DUÉE, E. & BUISSON, G. (1980). *Acta Cryst.* B**36**, 416–421.

PECORA, R. (1985). *Dynamic Light Scattering: Applications of Photon Correlation Spectroscopy*. New York, London: Plenum Press.

SKOURI, M., MUNCH, J. P., LORBER, B., GIEGÉ, R. & CANDAU, S. (1992). *J. Cryst. Growth*, **122**, 14–20.

THIBAULT, F., LANGOWSKI, J. & LEBERMAN, R. (1992). *J. Cryst. Growth*, **122**, 50–59.

WALRAND, S., BELLONI, L. & DRIFFORD, M. (1986). *J. Phys. (Paris)*, **47**, 1565–1576.