00244 A

PROTEIN CRYSTALLIZATION SCREENING THROUGH SCATTERING TECHNIQUES

YANNIS GEORGALIS[a,*], JENS SCHÜLER[a], JOACHIM FRANK[b], MARIO DIKEOS SOUMPASIS[c], WOLFRAM SAENGER[a]

[a]*Institut für Kristallographie, Freie Universität Berlin, Takustr. 6, Berlin, Germany*
[b]*Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4–6, 14195 Berlin, Germany*
[c]*Biocomputing Group, Max-Planck-Institut für Biophysikalische Chemie, P.O. Box 2841, 37018 Göttingen, Germany*

ABSTRACT

We have employed dynamic light scattering and X-ray small-angle scattering for monitoring the cluster formation of hen-egg white lysozyme solutions under crystallization conditions. From time-resolved dynamic light scattering experiments we can obtain a set of observables that exhibit clear extrema, when simultaneously plotted as functions of protein and electrolyte molarity. Close to these extrema, lysozyme crystals appear rapidly and with high probability. Both dynamic light scattering and small-angle X-ray scattering retrieve, under optimal crystallization conditions, particles 12 to 15 times larger than monomeric lysozyme; they could correspond to stable critical nuclei. Scattering amplitude analysis provides also useful information that correlates well with the size observables. A combination of both techniques can thus be successfully employed for screening supersaturated protein solutions.

INTRODUCTION

Controlled growth of protein crystals is the major obstacle in obtaining atomic resolution structures with trough diffraction methods. Atomic structures are in turn indispensable for establishing structure–function relationships in molecular biology and biotechnology.

The crystallization of biomolecules is a process involving nucleation, crystal growth and cessation of growth. Many parameters influence nucleation and crystal growth, making *a priori* predictions for suitable

---

* To whom correspondence should be addressed. Tel: +30-838-4588; Fax: +30-838-6702; E-mail: yannis@chemie.fu-berlin.de

growth conditions difficult. Very few attempts were made until the early 90s for tackling this problem. Since then, diagnostics of protein crystallization turned out to be a challenging topic and several laboratories are involved in the problem*. Despite the plethora of articles that have appeared on this subject, there are only a few systematic screening attempts.

The crystallization of a macromolecular solution is determined (i) by the effective interactions (or potentials of mean force) between the molecules and (ii) by kinetic factors that control both nucleation and growth. All these quantities depend strongly on the fundamental state variables of the system (e.g. species conformations, temperature and pressure) as well as on structural and physical parameters of its components (e.g. sizes, charges etc.). Derived thermodynamic quantities such as the macromolecular solubility may be useful indicators for finding good crystallization conditions, however low solubility is known as a necessary but often not as a sufficient condition for obtaining crystals.

It has to be stressed that classical crystallization theories are applicable to one component or binary systems and for the case of homogeneous nucleation. The extrapolation of similar concepts to proteins, which crystallize anyway inhomogeneously, has been hardly checked experimentally, due to the enhanced complexity of molecular interactions. The protein concentration at equilibrium, i.e. the solubility, is a complex function of size, shape, charge distribution and initial protein concentration. Furthermore, temperature, pH, size, charge and concentration of the precipitating agent (inorganic salt, polymer or organic solvent) will influence to an unpredictable extent the initial stable cluster formation, nucleation and later growth stages [1–3].

If one wishes to put protein crystallization diagnostics in a sound frame it is essential to proceed with the simplest systems possible. The various empirical recipes chosen to enforce protein crystallization (e.g. salts or organic reagents, or high molecular weight polymers, and water) represent systems whose physicochemical properties are not understood in the first place. Even if the problem is constrained to a simple water–electrolyte–protein system one deals, in the thermodynamic sense, with a very complex situation. Side effects, like non-specific binding or electrolyte induced structural changes, cannot at present be taken into account even by the most advanced simulation schemes.

---

* For references on crystallization diagnostics see the proceedings that appeared in J. Crystal Growth 110 (1991), 122 (1992), as well as in Acta Cryst. D50(4) 1994.

The problem of protein crystallization can be understood as a sub-problem of the wider and very complex protein–protein interactions problem. Our inability to either theoretically predict or experimentally tackle protein nucleation emerges, to a large extent, from our incomplete knowledge of interactions between proteins in the molecular level.

In the present work we address the problem of screening crystallization conditions of lysozyme by dynamic light scattering (DLS) in a higher degree of detail than in previous publications [4–8]. We have monitored the cluster formation kinetics of lysozyme at different protein concentrations while simultaneously varying the ionic strength of the electrolyte NaCl. We have used as observables (i) the mean quasi-stationary hydrodynamic radius of the clusters (ii) the fractal dimension and (iii) the zero time-lag size of the clusters. All three observables can be obtained from time-resolved DLS experiments according to previously described procedures [4]. With some effort, tentative estimates of the numbers of monomers and fractals can also be extracted. These estimates provide additional evidence for the definition of the optimal crystallization regions. We show that all observables, when examined as simultaneous functions of protein and electrolyte molarity, show extrema that coincide with those solution conditions where lysozyme crystals are obtained within less than two days.

Since the aggregation of lysozyme is very fast under supersaturation conditions, the kinetic observations made so far concern postnucleation. It is however mandatory to have estimates of the minimum particle size formed in the solution. In a series of small-angle X-ray scattering (SAXS) experiments we have independently redetermined the mean radius of gyration of the smallest particles present in crystallizing solutions. We find excellent agreement between those sizes and the zero time-lag extrapolated hydrodynamic radii.

DLS and SAXS have been reviewed by several authors [9–13]. For ensembles of monodisperse particles, which are small in comparison with the employed wavelength, DLS provides direct means for determining the particle size in the infinite dilution limit. The formalism of SAXS is very similar to that of static light scattering. The major difference is the superior spatial resolution attained through the shorter wavelength of the X-rays.

It has to be stressed that most physical-chemical approaches, including scattering techniques, have been conceived for working with stationary systems in the infinite dilution regime. Crystallization diagnostics of

biomolecules, even in their simplest form, depart from this ideal picture. At high protein and electrolyte concentrations rapid aggregation and cluster formation, render correct particle sizing difficult.

MATERIALS AND METHODS

The chemicals used in the present work were of analytical grade. Deionised water was obtained from a Milli-Q device (Millipore). Three times crystallized lysozyme was purchased from Sigma Chemical Co. (Deisenhofen, Germany) and was further exhaustively dialysed against water and lyophilized. All experiments were conducted in a 0.1 M Na-acetate, pH 4.2. NaCl p.a. grade, was purchased from Merck (Darmstadt, Germany).

Monodispersity of the preparations was controlled by DLS at several protein concentrations, in the absence of electrolyte (non-aggregation conditions) before each kinetic experiment. Protein and NaCl were rapidly mixed in the appropriate ratio and filtered through Millex sterile filters, 0.22 μm pore size, into standard cylindrical light scattering cells. Experiments were initiated within less than one minute after mixing protein with salt.

DLS — Data Acquisition and Evaluation

DLS measurements were conducted with an ALV/SP-86 spectrogoniometer (ALV, Langen, Germany) and the ALV-FAST/5000 digital autocorrelator boards at a scattering angle of 20° throughout. An Ar⁺ laser (operating wavelength 488 nm) was employed as light source.

DLS provides direct means for determining the free-particle diffusion coefficient for ensembles of dilute, monodisperse particles. This is true only if the particles do not interact with each other and their hydrodynamic radius, R, is small compared with the employed wavelength ($R < \lambda /20$).

The electric field autocorrelation function (ACF), $G^{(1)}(\tau)$, is a quantity that is proportional to the distribution of relaxation times τ and scattering amplitudes of the examined components. The formulation of the field ACF involves an integration over the size distribution function to account for polydispersity and eventual polymodality

$$G^{(1)}(\tau) \propto \int_{R_{min}}^{R_{max}} N(R) \, M^2(R) \, P(q) \, S(q) \, \exp(-mR^{-1} q^2 \, \tau_{rel}) \, dR \qquad (1)$$

where m is a proportionality constant, $N(R)$ and $M(R)$ denote number and mass of particles with radius R in the size range between R_{min} and R_{max}. $P(q)$ and $S(q)$ denote the form and the static structure factors [9] respectively, q is the scattering vector equal with $(4\pi n/\lambda) \sin(\theta/2)$, n the refractive index and θ the scattering angle.

Ideally, for non-interacting particle ensembles, Eq. (1) delivers the z-average hydrodynamic radius R that is associated with the translational diffusion coefficient D_t, via the Stokes–Einstein equation

$$D_t = \frac{k_B T}{6\pi\eta R} \qquad (2)$$

where η is the viscosity of the solvent.

For concentrated suspensions the static structure factor can be obtained from measurements of the angular dependence of the total scattered light or it can be approximated by known approaches [14–16]. If all the terms involved in Eq. (1) are available the ACF can be Laplace-inverted. Thus, polymodality and polydispersity can be resolved with adequate precision for concentrated suspensions as well. Some additional aspects that should not escape attention in microstructure formation, involve asymmetry corrections and deduction of the correct moments of the cluster size distribution. Brief descriptions of these issues are given in Appendix I.

Kinetic experiments were conducted at a constant temperature (20 ± 0.1)°C, and in a 0.1 M sodium-acetate buffer, pH 4.25. The protein concentrations examined were 0.72 mM, 1.55 mM and 2.4 mM whereas the NaCl concentration was varied between 0.2 M and 0.8 M. Every protein concentration was measured at 11 different electrolyte molarities resulting in a total of 33 time-resolved experiments. Spectra were collected every 30 s for nearly one hour. Mean radii of the evolving clusters were calculated by Laplace-inversion with a modified version of the program CONTIN [16,17] implemented in a CONVEX-C220 supercomputer.

SAXS — Data Acquisition and Evaluation

Data collection was accomplished with a Kratky small-angle spectrometer (Anton Paar, Graz, Austria) using an X-ray generator with a sealed-off tube 40 kV/20 mA, CuK_d radiation, wavelength $\lambda = 1.5418$ Å. A position-sensitive detector (4.7 cm length — 85 channels per cm), the

multichannel data processor and the appropriate analysis software were obtained from Braun Co. (Munich, Germany). Absolute calibration of the intensities was made with a standard Lupolen block. Solutions were measured in Marck-capillaries (Müller Co., Berlin, Germany) with a diameter of 1 mm. Spectra were recorded at $(20 \pm 0.1)°C$ for periods ranging between two to ten hours. The intensity pulses received in each channel were plotted as a function of the scattering vector. The parameter of interest in these experiments is the radius of gyration, R_g, that can be regarded as a direct measure of the spatial extent of the particle.

According to the Debye approximation [12,13] the scattered X-ray intensity can be expressed as

$$I(h) = \propto \exp\left[-\frac{h^2 R_g^2}{3}\right] \tag{3}$$

where $h = \frac{4\pi}{\lambda} 2\theta$ denotes the scattering vector. After subtracting the background scattering intensity due to solvent, the scattering curves were plotted in semilogarithmic, $\ln I(h)$ vs. h^2, according to the Guinier procedure [12]. The radius of gyration can then be obtained from the slope of the scattering curve at small angles and at the limit of zero angle.

Better resolution, i.e. detection of larger species present in the solutions, can be accomplished by increasing the distance between the probe and the detector. For distances between 20 cm and 37 cm we could explore cluster sizes up to ca. 60 nm or 120 nm, respectively. The latter implies that fractal structures growing in crystallizing solutions (they exceed 100 nm nearly immediately) will not be "seen" in the SAXS experiment. Therefore the possibility of observing additional dominant cluster populations, of low number density and sizes between those of monomers and fractals, is enhanced.

RESULTS AND DISCUSSION

DLS

In the DLS experiments we can clearly observe two particle populations, the first representing monomers (or dimers) and the second mass fractal clusters [4,5], (Fig. 1). Within less than two hours typical clusters reach sizes above 1 μm. The mean cluster size kinetic records were fitted
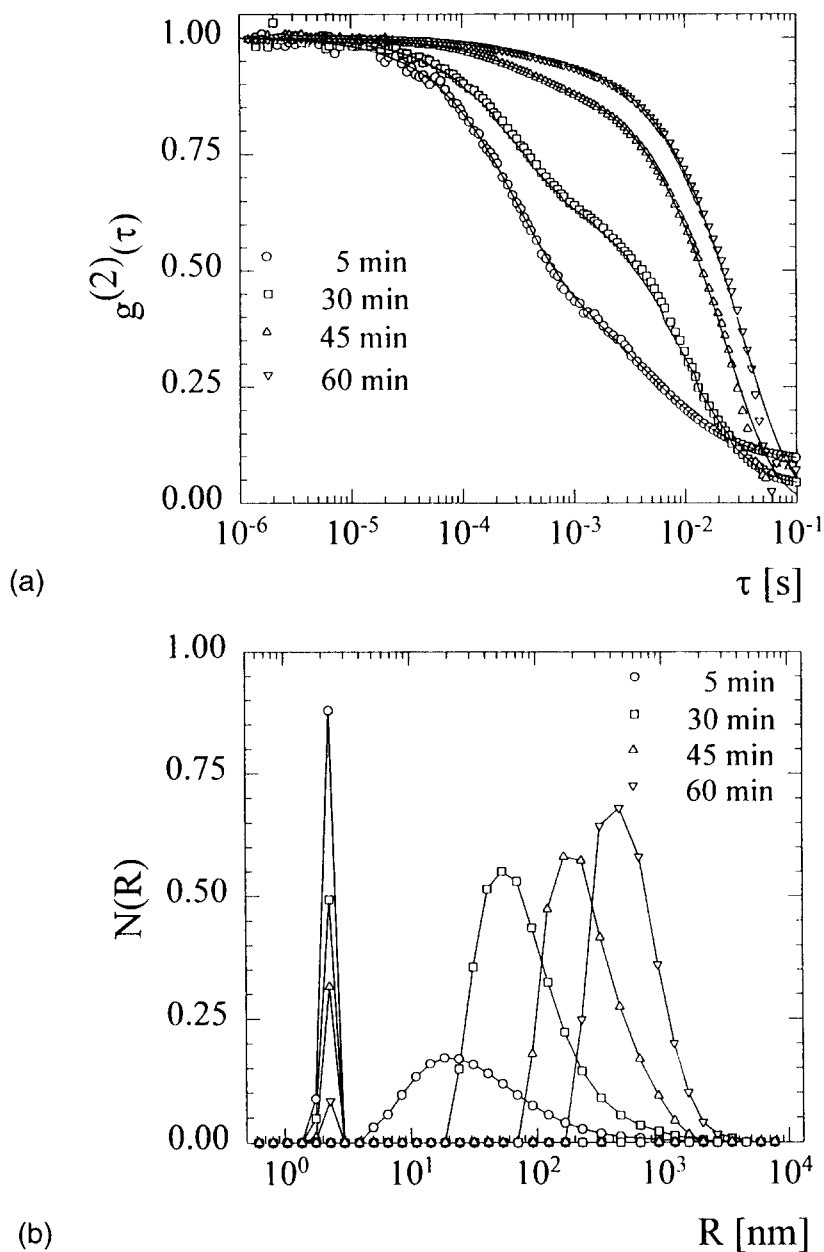
Fig. 1. (a) Typical DLS spectra obtained at different times (inset) with 1.0 M NaCl, in 0.1 M acetate buffer pH 4.2 and 2.1 mM lysozyme. Data were collected at a scattering angle of 20°. The solid line through the data is the CONTIN fit (best solution). (b) Particle size distributions obtained by Laplace inversion of the ACF. The smallest component corresponds, at low supersaturation, to monomeric lysozyme and its size remains stable throughout, R_m = 2.09 nm. The larger components represent fractal clusters; their amplitude and relative peak position evolves with time.
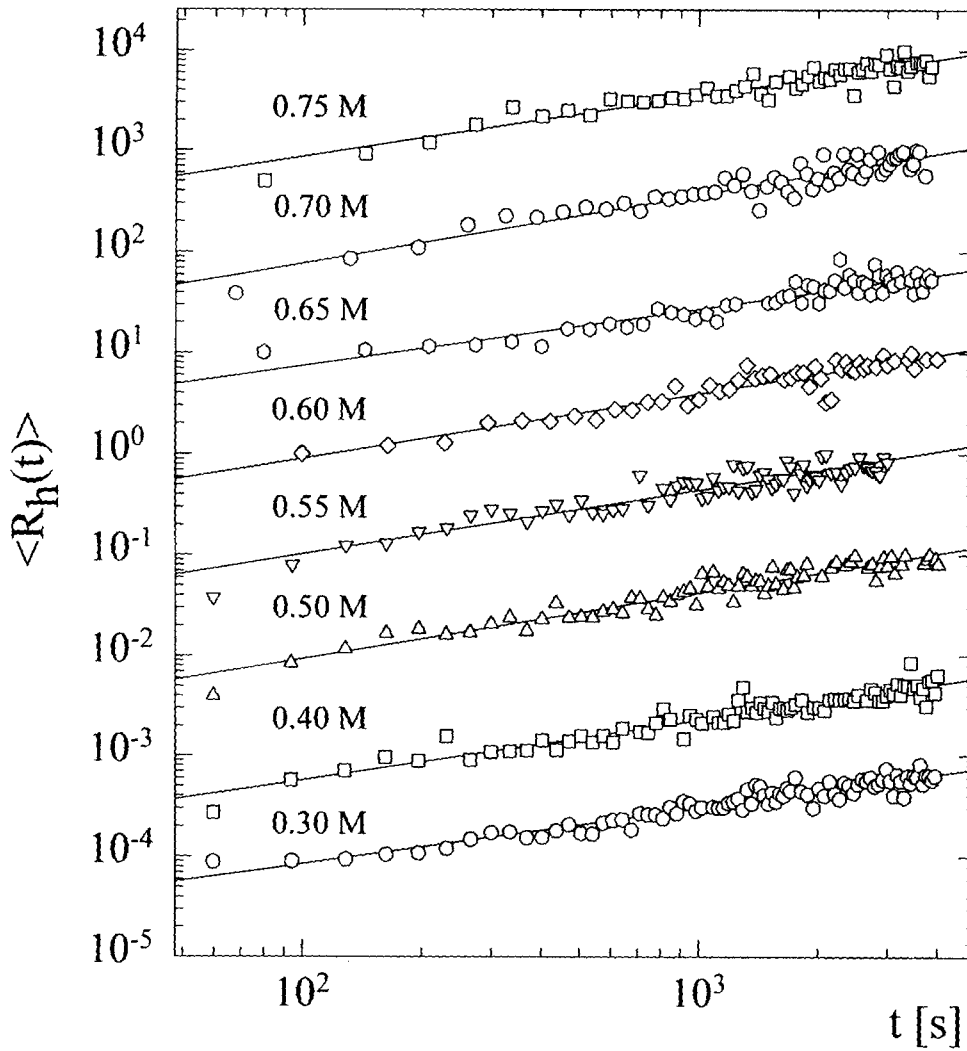
64

Fig. 2. Time-resolved aggregation kinetics of lysozyme under crystallization conditions. Typical traces were obtained by incubating a 1.55 mM lysozyme solution with NaCl and following the cluster size growth by DLS. The lines indicate non-linear least squares fits of Eq. (4) on the data. For the sake of clarity the ordinates are shifted by ascending powers of two between consecutive data sets. The number above each curve denotes electrolyte molarity.

according to a power-law expression [19,20] corresponding to diffusion limited cluster–cluster aggregation (DLCA)

$$R_h(t) = R_h(0) \, [1 + ct]^{1/d_f} \tag{4}$$

where c is a constant involving the Smoluchowski collision rate, $R_h(t)$ denotes the mean cluster radius at time t, $R_h(0)$ can be identified as a zero time-lag extrapolated mean cluster radius and d_f the fractal dimension of the cluster. Typical selected fits of data sets deduced from Eq. (4) are depicted in Fig. 2.

We have grouped data together and recast the DLS results in illustrative three-dimensional surface and spectral plots using standard inverse distance weighting presentation algorithms [21]. The obtained results are displayed in Figs. 3 to 5 as three dimensional plots followed by the respective spectral plots. All three observables, the quasi-stationary radius (Fig. 3), the fractal dimension (Fig. 4) and the zero time-lag extrapolated cluster radius (Fig. 5), display clear maxima. The highlighted regions in the spectral plots indicate that protein and electrolyte molarities where crystals grow in the light scattering vials within less than two days as judged by unaided eye. These maxima define clearly the region above 0.5 M NaCl and 1.55 mM lysozyme as the optimal crystallization region. We expect that more detailed spacing of both electrolyte and protein concentration will result in improved figures. However, the gross screening characteristics allow directly for a first order indexing of the observables.

We will describe below in detail the features of these plots.

The "Quasi-stationary" Cluster Radius

To have a unique measure of every kinetic experiment, we have averaged cluster radii collected between 50 min and 60 min after initiation of the experiments. This average, which involves some twenty points, we have arbitrarily termed as "quasi-stationary cluster radius", $<R_h>$. Strictly speaking the system is always far from stationarity, but for short enough time-scales the approximation is justified. Fractals are expected to grow until they occupy the whole volume available, but in practice they grow only for some 12 h to 24 h before they start sedimenting, unless solutions undergo gelation. Since the DLS experiments cease to render useful data if the mean cluster size exceeds a couple of micrometers, such observations can be better accomplished with static light scattering preferably at low scattering angles [22,23].

Surface and spectral plots of $<R_h>$ are shown in Fig. 3. It is evident that the behaviour of this surface is quite complex and the electrolyte concentration sections indicate a bell-shaped behaviour. In contrast, the
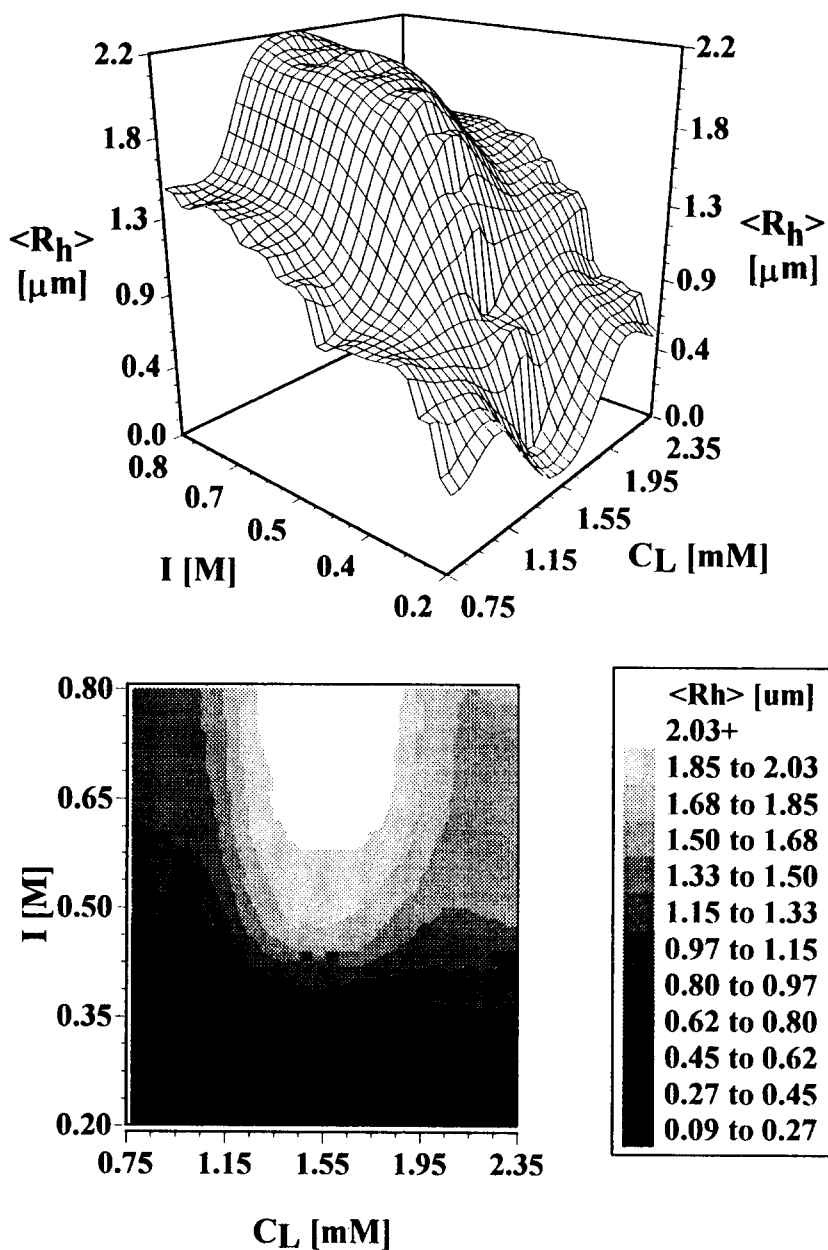
Fig. 3. Surface and spectral plot of the "quasi-stationary" cluster radius $<R_h>$ of lysozyme fractals as a function of lysozyme, (C_L), and NaCl, (I), ionic strength. The estimates were obtained by averaging the cluster radii recorded between 50 min and 60 min. An optimum is confined between 1.25 mM to 1.85 mM protein and above 0.55 M electrolyte (see also [8]). This region correlates very well with microscopic observations on crystal growth within one to two days.

protein concentration sections indicate a parabolic or slightly sigmoidal behaviour in all other regions than at the ends of the surface. In previous studies [4,5], where a more extended electrolyte range was employed, we have shown that this behaviour is bell-shaped for both NaCl and $MgCl_2$.

If the full experiments are plotted, one can obtain a very detailed grid involving all 33 experiments consisting of 120 points each. The computations and data manipulations required for obtaining this very detailed grid are by far more tedious than for the simple "quasi-stationary radius" since the latter involves only a fraction of points (33 vs. 3960 points). Both data sets have been examined in the late stages of the work and the similarity between the three-dimensional surfaces was found to be striking. The absolute values between the two grids did not differ more than 11% at the very most. Therefore, the selection of the "quasi-stationary radius" as an observable is not as arbitrary as it may seem at the first place. These comparisons indicate also that ambiguities due to (i) the choice of weights (ii) data normalization and (iii) specific griding parameters when employing a limited number of data points can be objectively handled.

The Fractal Dimension

Surface and spectral plots of the fractal dimension are shown in Fig. 4. The overall change of d_f is small but this is, on purely theoretical grounds [19,20], not unexpected. The fractal dimension will remain nearly constant as far as one observes aggregation events in a distinct regime without crossing over to another.

For pure DLCA aggregation one would expect a constant fractal dimension of 1.81 but lysozyme fractals exhibit deviations from this figure. The latter may involve some experimental uncertainty since the fractal dimension is less easy to determine. Multiple experiments under identical conditions suggested small standard errors. The fact that the fractal dimension falls below the universal limit of 1.81, is probably due to counterion screening that leads to cluster polarization effects [24,25].

The Zero Lag-time Extrapolated Cluster Radius

It is also interesting to estimate the minimum size of stable clusters present in crystallizing solutions. A stable population of mean size equal to that of the zero time-lag extrapolated cluster size, is assumed to be present.
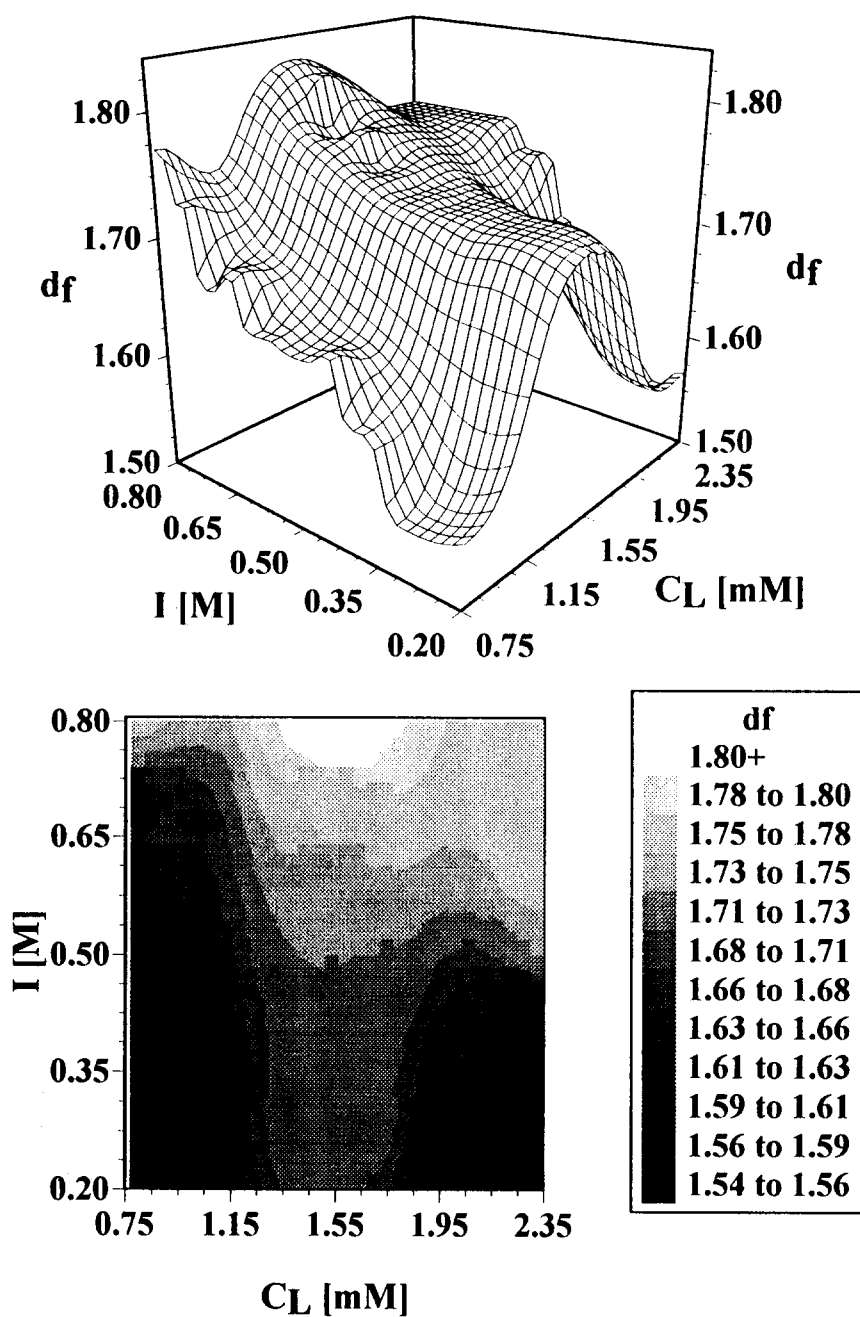
Fig. 4. Surface and spectral plot of the fractal dimension, d_f, of lysozyme clusters obtained from Eq. (4). An optimal region is confined between 1.3 mM and 1.8 mM protein, and above 0.7 M electrolyte. Within these limits the fractal dimension of the clusters is 1.81 indicative of universal DLCA aggregation [19,20].

In general, one faces a problem when attempting to capture the features of intermediate populations in DLS experiments. The scattered intensity is a function of both number and size of the scatterers. In crystallizing (supersaturated) solutions, is still present a very large number of monomers or dimers, 2.09 nm, to 2.65 nm. The scattering process is however dominated by the presence of only a few thousands but comparatively very large, 200 nm to 1 μm, fractal clusters. The two populations can be, in the beginning of the aggregation, decoupled from each other by inverse Laplace transformation, but the resolution of scatterers of intermediate sizes and number densities is in principle, difficult or even impossible. Note also that for stationary systems one can achieve nearly arbitrary precision by sampling for long times. This is however not the case with non-stationary systems, since only meaningless averages would be obtained. Therefore, no more than two components can be resolved with confidence, upon Laplace inversion of the ACF.

From the classical nucleation theory one expects that the free energy of a nucleating cluster increases up to a critical size and is limited by a barrier that has to be surmounted to transform a liquid to a crystalline phase [26]. The size of these clusters is not expected to vary drastically until crystallization is completed. However, the extrapolated $R_h(0)$ should not be identified as the critical size of a nucleus. The typical Ostwald–Miers nucleation regime [27] observed by other investigators [28–30] using viruses and proteins larger than lysozyme is not observed[*] in these experiments. This is due to the rapid diffusion of the lysozyme monomers and the observed kinetics resembles postnucleation events. $R_h(0)$ is expected to be, in general, larger than the hydrodynamic radius of a typical lysozyme nucleus.

The mean size of the observed population varies drastically with lysozyme and electrolyte concentration. In Fig. 5, $R_h(0)$ is plotted as a function of lysozyme and electrolyte molarity. A clear maximum is observed at 1.55 mM protein and around 0.50 M electrolyte. Since at present direct theoretical comparisons of these observations are not available, we have redetermined $R_h(0)$ with SAXS in an attempt to reproduce these results.

---

[*] Occasionally, at low supersaturation the Ostwald–Miers regime may be observed as well. However, such experiments are subject to large uncertainties since the late nucleation events coincide with the times required for the thermostation of the sample. It is therefore difficult to conclude with certainty about the reliability of the first few records.
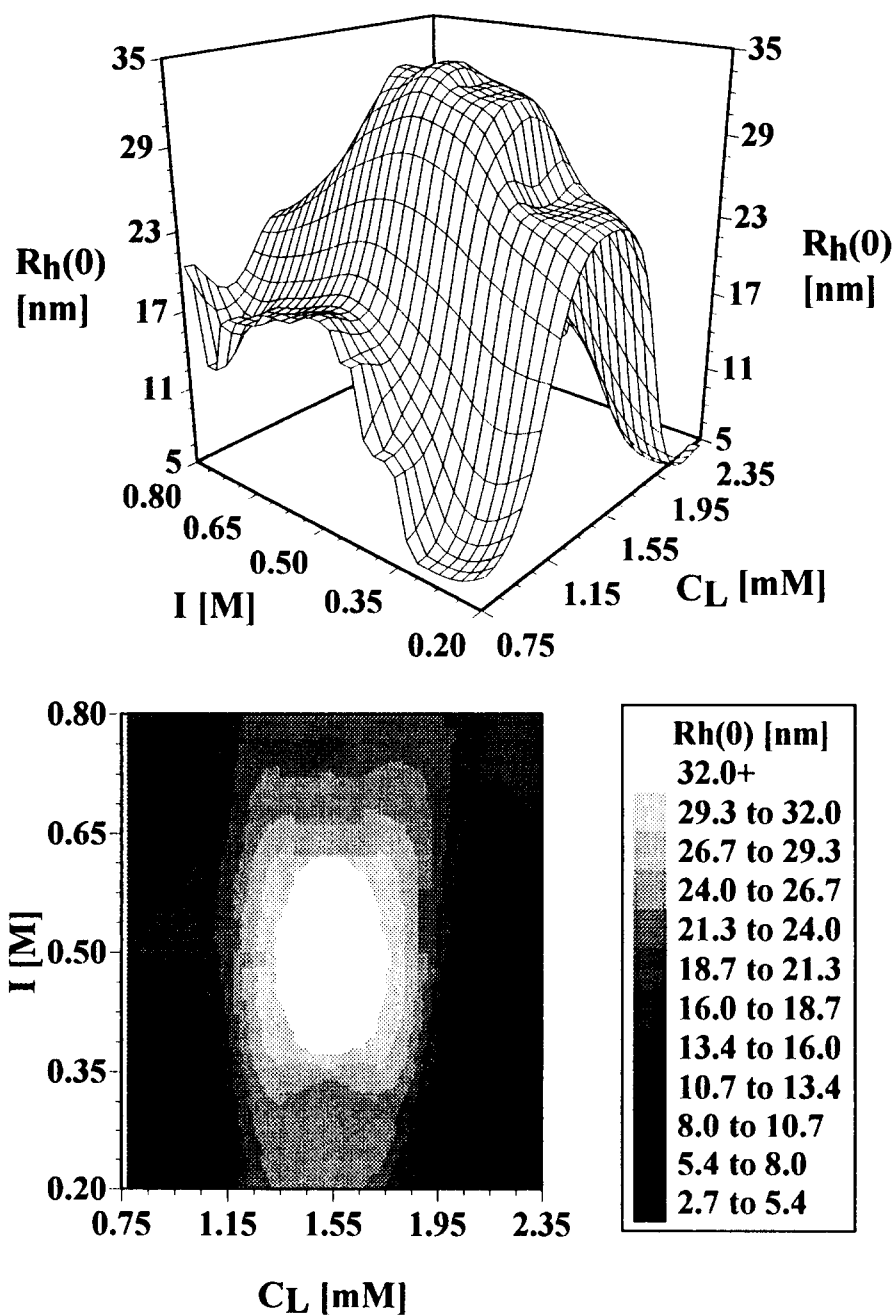
Fig. 5. Surface and spectral plot of the zero-time lag extrapolated mean radius $R_h(0)$ of lysozyme clusters, obtained from Eq. (4). A very clear maximum is observed for 1.35 mM to 1.75 mM protein, and 0.4 M to 0.6 M electrolyte. This region correlates well with microscopic observations on crystal growth within less than two days.

Size Determination of Stable Clusters with SAXS

For identifying the mean size of the lysozyme clusters in solution we have undertaken SAXS measurements at two different ionic strengths and six different protein concentrations. The experiments were conducted under conditions identical to those of the DLS experiment using the same lysozyme preparation.

The Guinier plots were biphasic, indicating the existence of at least two distinct populations in solution. We have decoupled with linear-least squares a low molecular weight population and plotted these data separately as a function of protein and electrolyte molarity. This species does not exhibit any time dependence. Their size, when extrapolated to infinite dilution, is found equal to 1.96 nm, which compares favourably with the DLS result, 2.09 nm, for the hydrated lysozyme monomer. It has to be emphasized that there is no way to decide from SAXS experiments whether the sample is mono- or polydisperse. Therefore, each data set can be evaluated by assuming either a monodisperse or a polydisperse composition. We have analysed the SAXS data by assuming monodisperse ensembles. Because of the complexity of the problem, a more intensive analysis is at this stage, unjustified.

The extraction of size estimates of the clusters by SAXS is by far more complex. For attaining an adequate representation of the data in the inner portion of the scattering curves, the data were fitted with least squares after omitting the first one or two channels that showed systematically upward curvature. Both experiments indicate maximum cluster sizes, R_{SAXS}, up to 16 times larger than monomeric lysozyme. A comparison of this intermediate cluster size, present in crystallizing solutions by both DLS and SAXS, is depicted in Fig. 6. The observed extrema occur again around 1.55 mM lysozyme, and resemble quite closely the zero-lag time extrapolated cluster mean sizes in the respective DLS experiments, Fig. 5.

The Monomer to Dimer Transition

Finally, the mean size of the species that are usually assigned as monomers in a DLS experiment was examined. Careful examination of over 2500 spectra indicated that the mean size of this species obeys a systematic tendency that may include useful information. Since dimerization is the rate limiting step for any kind of association, one expects
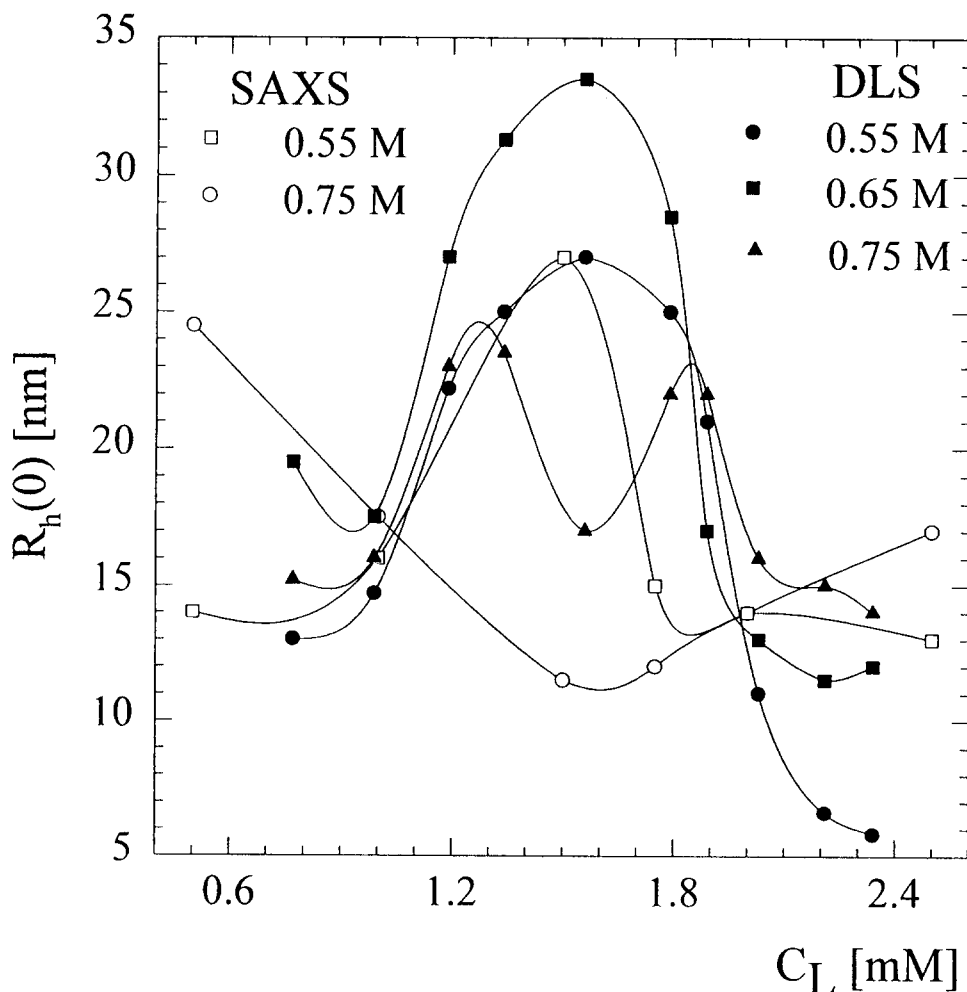
Fig. 6. Comparison of the minimum cluster size, $R_h(0)$, as a function of lysozyme concentration obtained from DLS and SAXS experiments. Note that the extrema shown by SAXS resemble closely those determined by the more detailed DLS experiment.

that these species will behave in a way that reflects bulk properties of the examined solution. Therefore, at some critical protein and electrolyte molarity one would expect the formation of stable oligomeric species having a size larger than that of the hydrated monomer.

The data plotted as surface and spectral plots (Fig. 7) indicate an increment of the mean monomer size, R_0, with increasing protein and electrolyte molarity. These data reproduce accurately those obtained in [8] although the latter experiments covered a wider range of conditions
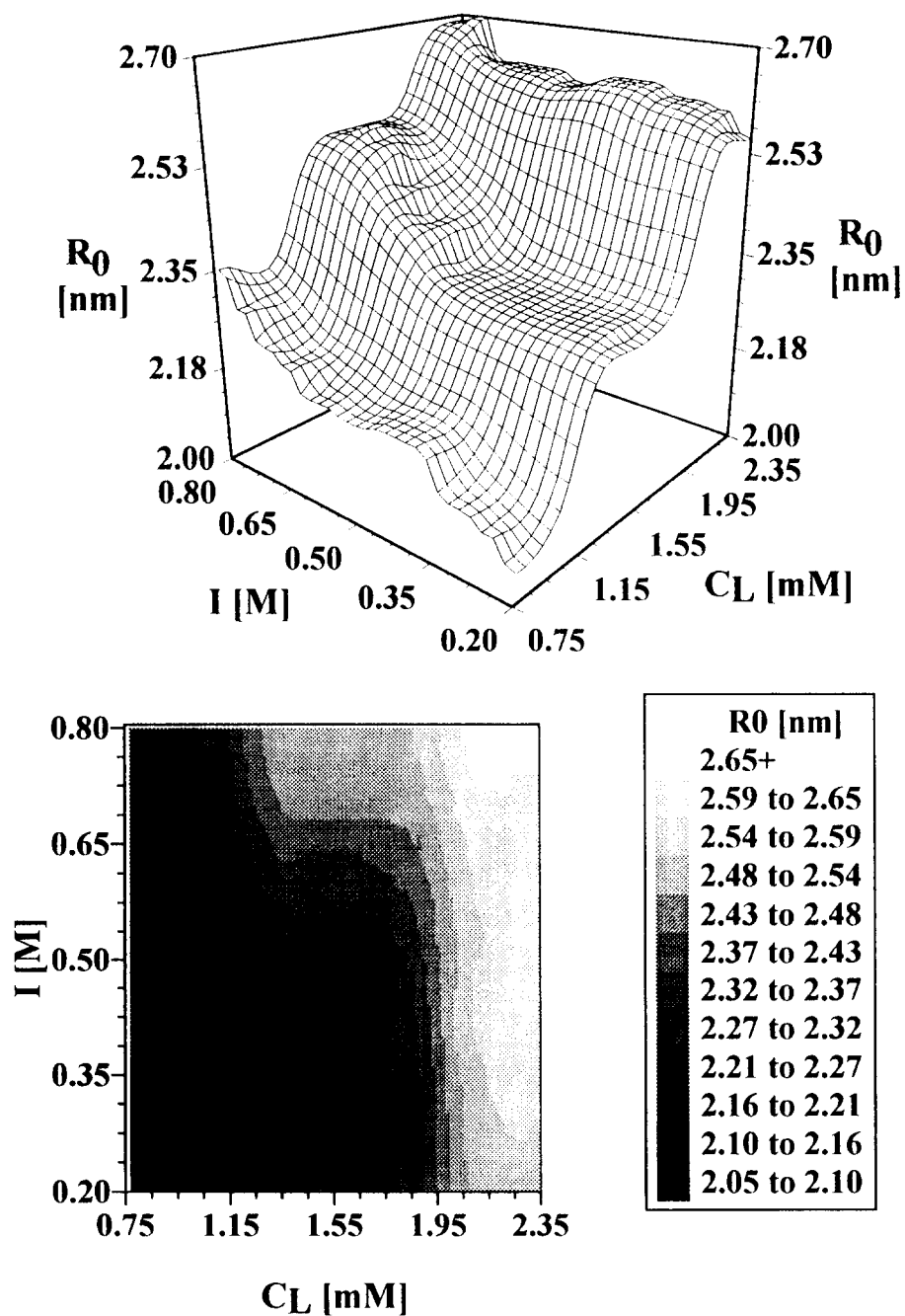
Fig. 7. Surface and spectral plot of the lysozyme monomer to stable lysozyme dimer transition. Stable dimers are formed with higher probability above 1.95 mM lysozyme and 0.65 M NaCl. Intermediate sizes may represent rapidly associating–dissociating species due to incomplete electrostatic screening.

and were purposely designed for studying molecular interactions of lysozyme during the crystallization process.

For free diffusing spheres, the monomer radius is 77.8% of that of the stable dimer. The largest sizes detected in the experiments do not exceed ca. 2.68 nm, a size that could correspond to stable dimers. Larger oligomers, may exhibit sizes that are only slightly higher than the later figure, depending on their mutual orientation. Precise measurements of larger oligomers are however very difficult due to the non-stationary nature of the examined samples.

Particles with such sizes of about 2.65 appear above 0.65 M NaCl and 1.95 mM lysozyme. It is however questionable if the intermediates between pure monomers and dimers represent stable species. The same question has been raised for the dimerization of bovine pancreatic trypsin inhibitor [31, 32]. There, such a behaviour was attributed to an either rapid monomer to dimer association–dissociation or, to the prevalence of van der Waals interactions. In general these concepts are more complex than one may think and their detailed discussion is out of the scope of the present work.

We can conclude that the transition from monomeric species to stable dimers occurs as a rather smooth function of the solution supersaturation. The stable dimer region is again observed within the realm of crystal growth; above 1.95 mM lysozyme and 0.55 M electrolyte.

Amplitude Analysis

One can assume that the total scattered intensity results from two major contributions, namely scattering from the smallest species and scattering from fractal clusters. One may then be tempted to compute tentative estimates of the numbers of scattering species from the above experiments assuming that only "monomers" and fractals are present in the examined solutions. This assumption is rather crude since it neglects the free nucleating species and presumes irreversibility. Neglecting the nucleating species is not expected to influence the relative tendencies since their number is small compared to the number of monomers and fractals. Further, the irreversibility assumption is justified within the frame of fractal cluster formation theory. The tentative estimates obtained from the scattering amplitude analysis can again be used for defining the optimal crystallization region for lysozyme based on criteria similar to those described above.

As can be seen from Eq. (1), the number of scatterers is explicitly included in the kernel of the expression for the ACF. The particle size distribution function of each species can be obtained after Laplace-transformation of the ACF spectra. For small hard spheres, due to the intensity preserved weighting of the ACF, one usually considers the mass-squared weighted average, $P(q) M^2(R) N(R)$, as a function of the radius R. For the "monomeric" species one can simply recast this term as $R^6 N(R)$ whereas for fractal structures as $P(q) S(q) N(R) R^{2d_f}$ due to the mass-size scaling power-law, $M(R) \propto R^{d_f}$ [19,20].

Therefore, estimates of the number of monomers and fractals can be obtained from the kinetic experiment by normalizing the areas under each peak with the determined mean cluster sizes, raised in the respective powers. For either species $P(q)$ can be set equal to 1, since the mean size of the seed particles is smaller than the employed wavelength, for most of the cases encountered in the present analysis. $S(q)$ can be computed from the Fisher–Burford approximation [14] and from the known $R_h(t)$ and d_f of the fractal clusters.

We can focus on the number of monomers and fractals and define as an observable the number of clusters $n(t)$ of size $R_h(0)$, which are incorporated in a fractal cluster at time t. One can then assume that the number of monomers, $N_m(t)$, and fractals, $N_f(t)$, are any time associated with the initial number of lysozyme monomers N_0 through the relationship

$$N_0 \approx N_m(t) + n(t) N_f(t) + \dots \tag{5}$$

Equation (5) should obey the conditions: $N_m(0) = N_0$, $n(0) = 0$ and $N_f(0) = 0$. After infinite time it is theoretically expected that $N_0 = N_m(\infty) = 0$ and $N_f(\infty) = 1$ and $n(t) = N_0$ (if one imagines a huge fractal that fills up all the available space). The latter is in general not true, unless gelation occurs and monomeric species can still be observed in DLS experiments for several days even after the sedimentation of fractals and the appearance of the crystalline phase.

In Fig. 8 the relative number of "monomers" is plotted as a function of protein and electrolyte molarity. The ordinate has been normalized to a convenient scale between 0 and 1. It can be seen that the number of free "monomers" in solution reaches a distinct minimum around 1.55 mM of protein and 0.65 M of electrolyte. If crystallization conditions are close to optimal the number of monomers is indeed expected to be minimized since they have a higher probability to be incorporated onto nuclei than
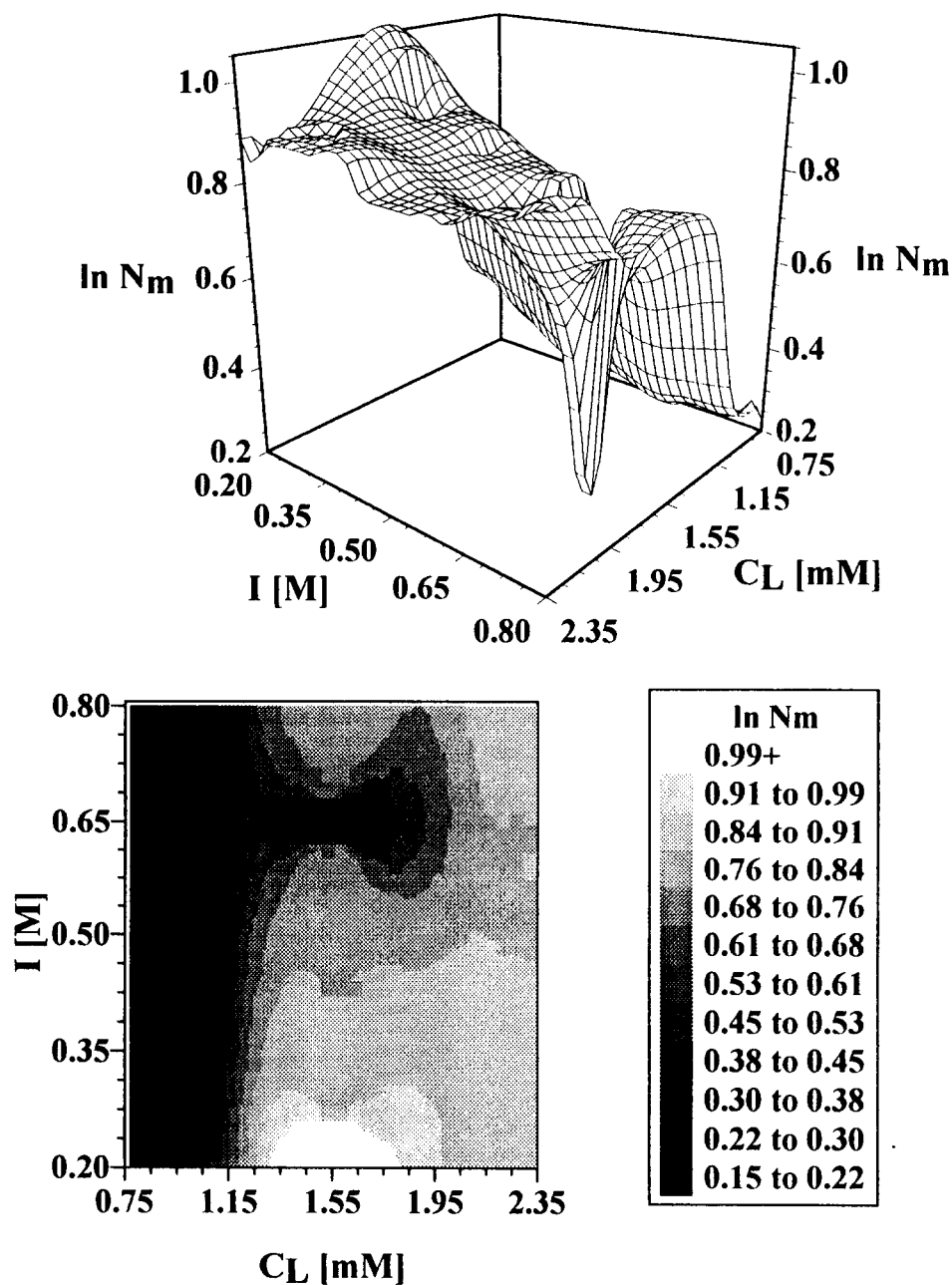
Fig, 8. Surface and spectral semilogarithmic plots of the relative number of monomers, N_m, in lysozyme crystallizing solutions. A clear minimum is observed at 1.55 mM protein and 0.65 M electrolyte. The relative number of free monomers in solution is minimized through incorporation on small nuclei. Please note the changes undertaken at the origin of the x–y axes for the sake of clarity.

move freely in the solution. These figures agree well with the optimum range of observables defined above and in Ref. [8], on the same subject.

The relative number of fractals N_f is plotted in Fig. 9. Whereas the number of observed fractals develops smoothly almost over the entire protein–electrolyte surface, a distinct drop is observed above 1.25 mM protein and 0.30 M electrolyte. This behaviour can be easily understood if one considers fractal and nuclei formation as two competing processes (see discussion below). At optimal regions the number of fractals is expected to be minimized thus favouring nucleation.

Association Between Nucleation and Fractal Formation

Fractal cluster formation is typically observed in systems developing far from equilibrium. Kinetic models are employed for the interpretation of the obtained data that do not allow for structural rearrangements, and therefore minimum energy configurations cannot be deduced. On the contrary, thermodynamic models describe growth phenomena close to equilibrium via compact structure formation. Classical phase transition models (nucleation and growth or spinodal decomposition) are based on thermodynamics. The formation of possibly fractal structures in kinetic growth processes has only been recently considered [33].

For a consistent picture associating nucleation with fractal cluster formation, one can imagine the whole scenario starting with the nucleation burst when protein and electrolyte come into contact. Nucleus–nucleus collisions then trigger catastrophic random aggregation that dominates the light scattering process in the early stages. However a minority of nuclei manages to escape the "Smoluchowski sink" and put crystal formation through.

The question whether the nuclei formed at the early stages are compact or ramified is not easy to answer. The equilibrium shape of a cluster is usually approximated as a smooth sphere. This is certainly not correct since it is known that cubes or polyhedra can represent better lattice forming shapes. In such a case geometrical form factors can account for the deviations from spherical shape [34]. To mention an example, if one assumes cubes instead of spheres for the equilibrium shape, the number of molecules present in the critical nucleus is expected to be twice as large. Nuclei having a mean size of 25 nm will involve up to 1700 monomers, assuming a spherical shape.
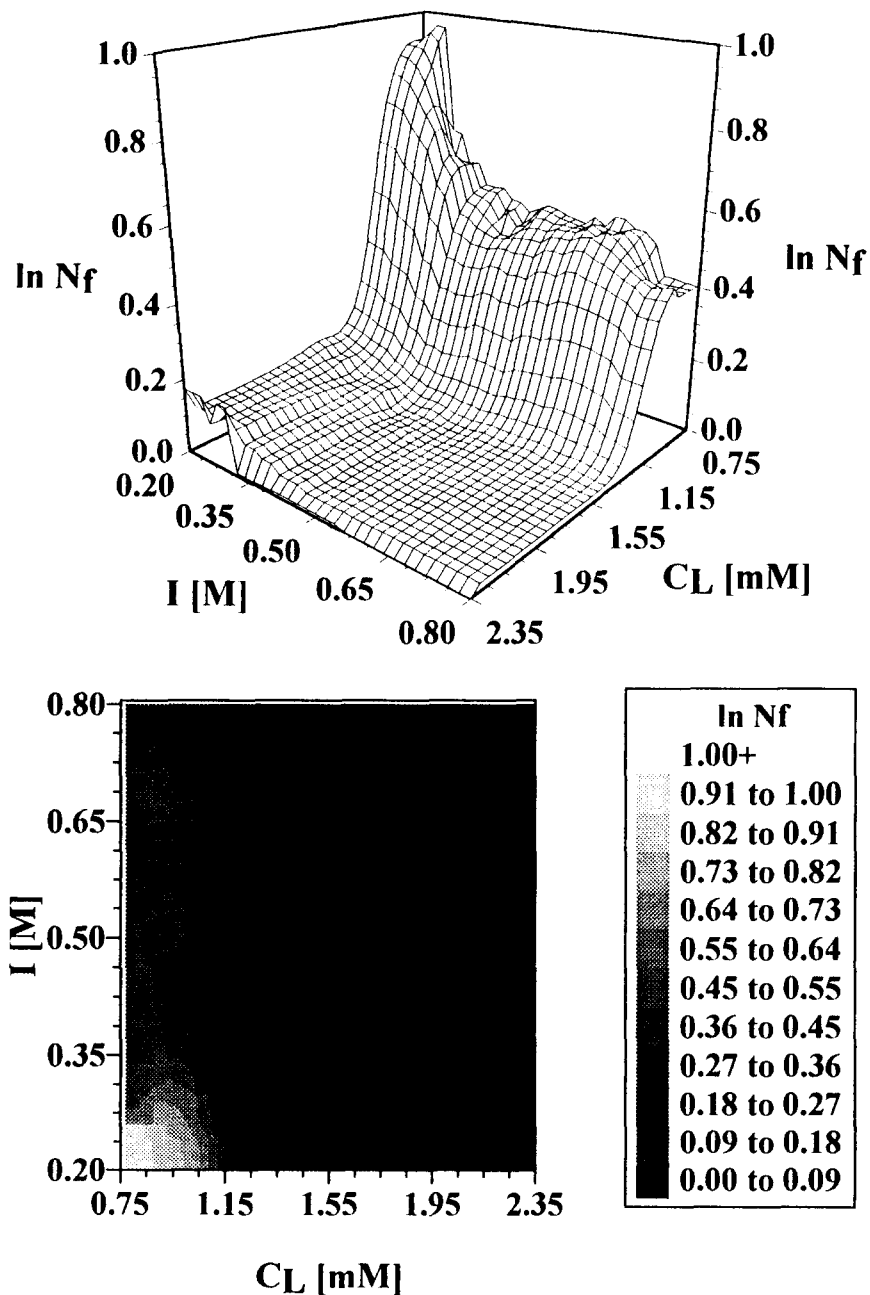
Fig. 9. Surface and spectral semilogarithmic plots of the relative number of fractals, N_f, in lysozyme crystallizing solutions. A broad minimum is observed above 1.55 mM lysozyme and 0.30 M electrolyte. The relative number of fractals is minimized if smaller compact clusters are consumed in nuclei formation. Please note the changes undertaken at the origin of the x–y axes for the sake of clarity.

Further, there is no direct experimental evidence that these nuclei are indeed compact. Theoretical studies [35–37] have shown that deviations from the classical picture are not improbable. When small compact clusters evolve in time, both the number of contact points and the probability of collision will increase. Such clusters will then fail to interpenetrate each other as if they were small. Therefore, the growth of more tenuous structures with lower fractal dimensions, as time proceeds, is not unexpected. Very recently, some interesting observations were made by low- and wide-angle X-ray scattering on zeolites [38]. These experiments indicate that nucleation, in contrast to classical theory, occurs via formation of small fractal structures. These clusters undergo restructuring and after rearrangements a dense microcrystalline phase appears. Whereas it would be rather premature to extrapolate such properties to proteins, these observations render indirect support to our findings.

Although it is difficult to define the exact borders between nucleation and postnucleation, we believe that our light scattering experiments capture well the postnucleation events. Nucleation, in the classical sense, is probably so fast for lysozyme that it escapes observation. Unless special precautions are taken, the information is lost. However, we have recently shown that pertinent, albeit not always easy to interpret, information concerning the fate of protein crystallizing solutions can be extracted with DLS techniques [7,39].

CONCLUSIONS

The combined use of DLS and SAXS allows a first order analysis of lysozyme aggregation during nucleation and growth. The obtained results show that the observed clusters are fractals formed by smaller aggregates. The growth of these clusters can be described by power-law kinetics [4,5,7,8] and the appearance of crystals within short times is reinforced when aggregation takes place in the DLCA regime, at least under the examined conditions. Cases where aggregation takes place in different regimes (reaction limited or cross-over aggregation regimes i.e. while using $(NH_4)_2SO_4$ as a precipitant should be treated carefully keeping always in mind that several notorious effects may modify the value of the fractal dimension. The growth kinetics can be obtained with fair precision; kinetics can be rapidly deduced during the early stages of

the reaction (i. e. the first one to two hours) and can be recast into surface and spectral plots for visualizing the apparent tendencies.

We have persuasively shown that the obtained observables show pronounced extrema that coincide with those solution conditions where tetragonal, well-diffracting, lysozyme crystals are obtained within less than two days. Therefore critical crystallization regions indeed exist, and they can be captured even when observations are made in the post-nucleation phase. The empirical correlation between the occurrence of extrema in cluster formation and crystal growth indicates that the observables derived from the scattering experiment may be identified as useful tools for rapidly screening multiple crystallization set-ups by using DLS instrumentation. Such experiments may help in general to understand metastability that is an intrinsically dynamic process.

Due to the different spatial resolution of the DLS and SAXS experiments, it is easier to determine by SAXS small stable aggregates in nucleating solutions. Similar size estimates are obtained by DLS after extrapolating the growth kinetics to zero time-lag. This is an important aspect since it ensures that one observes cluster–cluster rather than monomer–cluster aggregation. Therefore, the seeding particles that build the fractals are not lysozyme monomers but, with very high probability, nuclei. It is understandable that, due to the small size of lysozyme, capturing the nucleation burst with methods like time-resolved DLS or SAXS may be extremely difficult. The times required for completion of the nucleation burst, may be orders of magnitudes shorter than the dead time of the experiment.

A comparison with experimentally measured solubilities of lysozyme with two different cations Na^+ and Mg^{2+} [2] was also undertaken. The results are displayed in Figs. 10a and 10b as surface and spectral plots. At conditions comparable with ours these plots indicate that the solubility minima, i.e. the regions where the probability for obtaining crystals is increased, are in either case quite close (see also Fig. 3 in Ref. [5]) but they do not coincide with the regions defined by the scattering experiment and the appearance of lysozyme crystals. Solubility maxima and minima alternate in the plots. Whereas such effects can be clearly seen in nearly all data sets measured by the authors, in the plots displayed herein such tendencies are clearer for Mg^{2+} whereas Na^+ show a broader range of low solubilities. The conclusion drawn from such three-dimensional plots is that solubilities are complex functions of both protein and electrolyte concentration (packing and electrostatic screening).
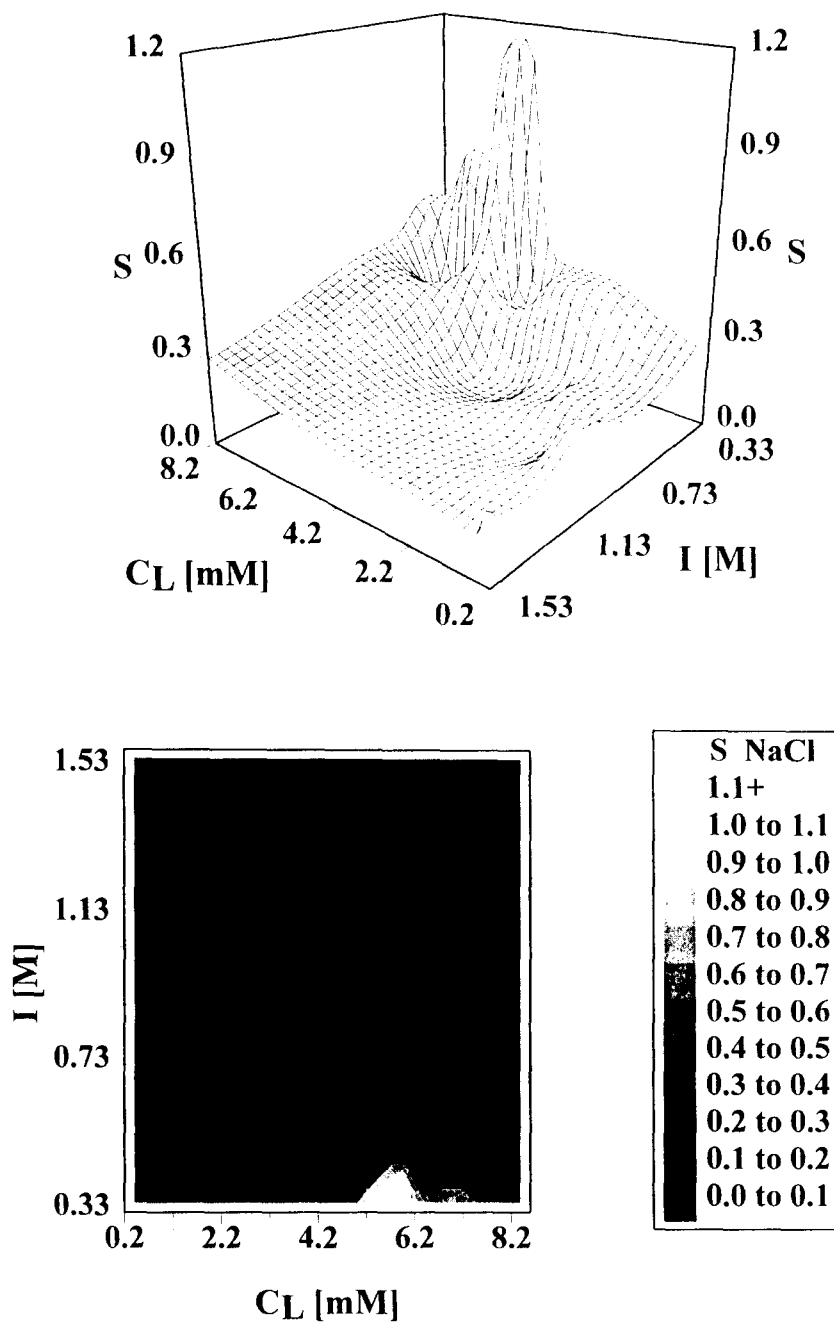
Fig. 10. Solubilities computed from Table I of Ref. [2]. (a) Surface and spectral plots of spectrophotometrically determined solubility S, of lysozyme solutions incubated with NaCl. Minima are identified at 2.01 mM, 4.2 mM and 7.1 mM lysozyme and correspond to 1.07 M, 0.73 M and 0.47 M NaCl, respectively.
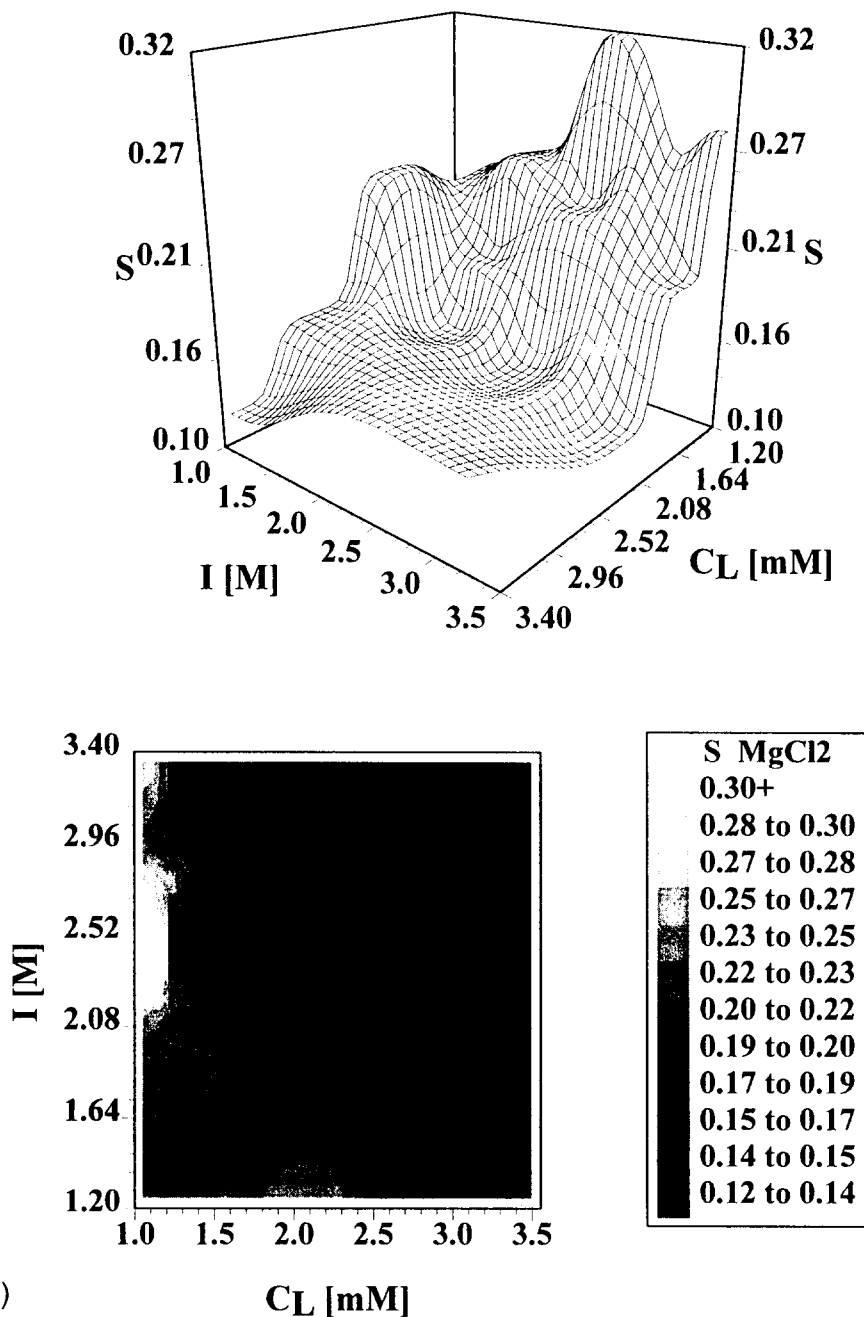
Fig. 10 continued. (b) Surface and spectral plots of the relative solubility S, of lysozyme solutions incubated with $MgCl_2$. Minima are identified at 2.19 mM, 2.27 mM and 3.36 mM lysozyme and correspond to 1.88 M, 3.14 M and 1.39 M $MgCl_2$, respectively. The last value compares favourably with the experiments depicted in Fig. 3 of Ref. [5]. Please note the changes undertaken at the origin of the x–y axes for the sake of clarity.

The extensive experimental data discussed in this work provide a first platform for improving our understanding on protein crystallization and predictive theoretical frameworks. Theoretical descriptions of protein crystallization are highly desirable. There one has to cope with a simultaneous description of nucleation and fractal cluster growth, in a manner similar to that proposed by Binder and Staufer [40], nearly twenty years ago. The first order theoretical modelling of the lysozyme–NaCl system, in the region of thermodynamic states described here will be published in a subsequent communication (Soumpasis and Georgalis, in preparation).

ACKNOWLEDGEMENTS

REFERENCES

1   Z. Kam, H.B. Shore and G. Feher, J. Mol. Biol., 123 (1978) 539.
2   M.M. Riés-Kautt and A.F. Ducruix, J. Biol. Chem., 263 (1989) 745.
3   A.F. Ducruix and R. Giége (Eds.), Crystallization of Nucleic Acids and Proteins. IRL Press, Oxford, 1992.
4   Y. Georgalis, A. Zouni, W. Eberstein and W. Saenger, J. Crystal Growth 126 (1993) 245.
5   Y. Georgalis and W. Saenger, Adv. Colloid Interface Sci., 46 (1993) 165.
6   Y. Georgalis, J. Schüler, W. Eberstein and W. Saenger, in: M.M. Novak (Ed.), Fractals in the Natural and Applied Sciences. IFIP (A-41), pp. 139, Elsevier Science, Amsterdam, 1994.
7   W. Eberstein, Y. Georgalis and W. Saenger, Eur. Biophys. J., 22 (1993) 359.
8   W. Eberstein, Y. Georgalis and W. Saenger, J. Crystal Growth 143, (1994) 71.
9   S.K. Schmitz, An Introduction to Dynamic Light Scattering by Macromolecules. Academic Press, New York, 1990.
10 B. Chu, Dynamic Light Scattering. Academic Press, 1991.
11 W. Brown (Ed.), Dynamic Light Scattering, The Method and Some Applications. Oxford Science Publications, 1993.

12   O. Glatter and O. Kratky (Eds.), Small Angle X-Ray Scattering Academic Press, London, 1982.

13   O. Glatter, in: P. Lindner and Th. Zemb (Eds.), Neutron, X-Ray and Light Scattering. North-Holland, Amsterdam, 1991.

14   M.E. Fisher and R.J. Burford, Phys. Rev. A, 156 (2) (1967) 583.

15   H.M. Lindsay, R. Klein, D.A. Weitz, M.Y. Lin and P. Meakin, Phys. Rev. A, 38 (5) (1988) 2614.

16   H.M. Lindsay, R. Klein, D.A. Weitz, M.Y. Lin and P. Meakin, Phys. Rev. A, 39 (6) (1989) 3112.

17   S.W. Provencher, Comp. Phys. Comm., 27 (1982) 213.

18   S.W. Provencher, Comp. Phys. Comm., 27 (1982) 229.

19   M.Y. Lin, H.M. Lindsay, D.A. Weitz, R.C. Ball, R. Klein and P. Meakin, Nature (London), 339 (1989) 360.

20   M.Y. Lin, H.M. Lindsay, D.A. Weitz, R.C. Ball, R. Klein and P. Meakin, Proc. R. Soc. Lond. A, 423 (1989) 71.

21   Stanford Graphics Presentation and Analysis Package, 3-D Visions.

22   F. Ferri, M. Giglio, E. Paganini und U. Perini, Europhys. Lett., 7 (1988) 599.

23   M. Carpineti, F. Ferri, M. Giglio, E. Paganini and U. Perini, Phys. Rev. A, 42 (12) (1990) 7347.

24   R. Jullien, Phys. Rev. Lett., 55 (1985) 1697.

25   R. Jullien, J. Phys. A, 19 (1986) 2129.

26   D.W. Oxtoby, Adv. Chem. Phys., 70 (1988) 263 .

27   I.M. Lifschitz and V.V. Slyojov, J. Phys. Chem. Solids, 19 (1965) 35.

28   A.J. Malkin, J. Cheung and A. McPherson, J. Crystal Growth, 126 (1993) 545.

29   A.J. Malkin and A. McPherson, J. Crystal Growth, 126 (1993) 555.

30   A.J. Malkin and A. McPherson, J. Crystal Growth, 1232 (1993) 128.

31   P.R. Wills and Y. Georgalis, J. Phys. Chem., 85 (26) (1981) 3978.

32   W.H. Gallagher and C.K. Woodward, Biopolymers, 28 (1989) 2001.

33   M. Carpineti and M. Giglio, Phys. Rev. Lett., 68 (22) (1992) 3327.

34   K.H. Lieser, Angew. Chem. 81 (6) (1969) 206.

35   D.W. Heermann and W. Klein, Phys. Rev. Lett., 50 (14) (1982) 1062.

36   W. Klein and F. Leyvraz, Phys. Rev. Lett., 57 (22) (1986) 2845.

37   J. Yang, H. Gould and R.D. Mountain, J. Chem. Phys., 93 (1) (1990) 711.

38   W.H. Dokter, H.F. van Garderen, T.P.M. Beelen, R.A. van Santen and W. Bras, Angew. Chem., 107 (1) (1995) 122.

39   W. Heinrichs, Y. Georgalis, H.-J. Schönert and W. Saenger, J. Crystal Growth, 133 (1993) 196.

40   K. Binder and D. Staufer, Adv. Phys., 25 (1976) 343.

41   P. Meakin, in: C. Domb and J.L. Lebowitz (Eds.), Phase Transitions and Critical Phenomena, Vol. 12. Academic Press, New York, 1988, 351 pp.

APPENDIX I

Decoupling of Rotational Motions

Fractals are asymmetric structures, and the contributions of both translational and rotational motions to the ACF must be taken into account. The field ACF of anisotropic, inhomogeneous scatterers, has been calculated with the assumption that rotational and translational motions are uncoupled. Computer generated and experimentally produced DLCA aggregates show typical asymmetry ratios of about 1.7, a value that indicates that this assumption is valid.

If the size of the clusters is large[*], rotational motions are expected to contribute through the cluster polydispersity to the ACF because they are q independent.

$$g^{(1)}(\tau) = \frac{\sum N M^2 \exp(-q^2 D_t) \sum S_l(q) \exp[-l(l+1) D_r \tau]}{\sum N M^2 S_l(q)} \tag{6}$$

where D_t and D_r denote the translational and rotational diffusion coefficient of the fractal with radius R respectively, and the summation extends over all clusters characterized by this radius

$$D_r = \frac{k_B T}{8\pi\eta R^3} \tag{7}$$

Each term in the second sum, in Eq. (6), which corresponds to the rotational contribution, is weighted with the factor $S_l(q)$ that can be obtained after a multipole expansion of the structure factor [15,16]

$$S(q) = \sum S_l(q) \tag{8}$$

The latter sum has been computed up to the seventh term from computer generated DLCA aggregates. Equation (8) allows then the reconstruction of the scattered field ACF as follows:

---

[*] For simplification we have assumed that the hydrodynamic radius is equal with the radius of gyration. For DLCA clusters the two radii differ by some 10%.

$$g^{(1)}(\tau) = \frac{\sum N\,M^2 \exp\left[-q^2\,D_{eff}\,\tau\,S(q)\right]}{\sum N\,M^2\,S(q)} \tag{9}$$

where $D_{eff} = D_t\,f(qR)$ denotes an effective diffusion coefficient including both translational and rotational diffusional motions. We have previously shown [4] how an approximate correction can be applied to account for coupling of the rotational motions. One can then obtain estimates of the translational diffusion coefficient, D_t, and therefore R_h, with adequate precision.

While currently, the employed structure factors seem to adequately describe both diffusion and reaction limiting aggregation regimes, there may be a disadvantage in their use. Namely, the large-scale computer simulations where they originate from, are valid for ideal clusters composed of a limited number of particles and they do not consider the effects of internal rotations within the cluster (restructuring), which may modify the fractal dimension. Large-scale computer simulations [41] have indeed shown that changes of the fractal dimension are to be expected when clusters undergo internal rotations.