Abstract

This paper deals with experimental investigation, mathematical modeling and numerical simulation of the crystallization processes induced by counter diffusion method of a precipitant agent in a lysozyme protein solution. Novel mathematical strategies are introduced to simulate the experiments and in particular to take into account the kinetics of the growth process and the motion of the crystals due to the combined effect of gravitational force and viscous drag if the sedimenting process is allowed (protein chamber free of gel). Comparison between experimental observations and numerical simulations in the presence of convection and sedimentation and without them provides a validation of the model. The crystal formation in gel results modulated in space. If the gel matrix is not present, convective cells arise in the protein chamber due to local inversions in the density distribution associated to nucleation phenomena. As time passes, these vortex cells migrate towards the top of the protein chamber exhibiting a different wave number according to the distance from the gel interface. The sedimentating particles produce a wake due to depletion of protein from the surrounding liquid. The models and the experiments may represent a useful methodology for the determination of the parameters and conditions that may lead to protein crystallization.

1. Introduction

During last decades, great interest has been directed towards crystals of biological macromolecules and towards their crystallization processes. In fact, single crystals with good diffraction properties and structural quality are needed to achieve atomic resolution data on protein structure, needed for progress in biology and drug design. Typically the crystals are obtained by precipitation from super-saturated solutions using a number of techniques (McPherson [1], Rosenberger [2]). Producing protein crystals of adequate size, however, is often the ‘bottleneck’ for three-dimensional structure studies of protein molecules. Superimposed on this is the poor state of our current understanding of many factors that may influence the crystallization process (sedimentation, convection, etc.) and the disappointing lack of mathematical models and numerical methods able to provide a comprehensive description of the overall process. For this reason further investigation is required.

It is worthwhile to stress how many authors have focused the attention on "what happens close to the crystal". Several studies have shown that crystal growth rate is sensitive to very large amount of parameters: e.g. temperature, solution composition, supersaturation, impurities, convection etc. (see e.g. Otálora and García-Ruiz [3]). Moreover it usually depends on the crystallographic orientation (Coriell, Chernov, Murray and McFadden [4]) and on crystal habit changes with the level of supersaturation (Pusey, Snyder and Naumann [5], Monaco and Rosenberger [6]). Morphological stability of crystal facets and thus crystal quality also depend on supersaturation and impurities [4]. A comprehensive review of the previous fundamental protein crystal growth and morphology experimental studies has been given for instance in Monaco and Rosenberger [6], Otálora et al. [7,8] and Vekilov and Chernov [9].

A big effort has been made also to understand nucleation kinetics (see e.g. Otálora et al. [8], Haas and Drenth [10], Galkin and Vekilov [11] and references therein). Interface kinetics, and onset of convection were analysed in Lee and Chernov [12].

Macromolecular crystallization is a matter of searching, as systematically as possible, the ranges of the individual parameters that impact upon crystal formation, finding a set or
multiple sets of these factors that yield some kind of crystals, and then optimizing the variable sets to obtain the best possible crystals for X-ray analysis. However fundamental methods for identifying favourable protein crystallization conditions remain still poorly developed (Rosenbaum and Zukoski [13], Giege and Ducruix, [14], McPherson [15]).

In light of the arguments given above the present paper aims to analyze “ensemble behaviour” during crystal formation and growth, with the objective of pointing out "macroscopic" aspects. In fact, crystallization is characterised by the interplay of different phenomena: transport in liquid phase, nucleation and crystal growth, sedimentation and convection, etc. Understanding of this interplay requires a global analysis that would consider all the relevant phenomena simultaneously in order to track system evolution. Such global analysis may support optimisation of growth techniques i.e. it may help to discern interrelations between various "macroscopic" parameters of interest, e.g. crystal number and size distribution in space and time.

The counter diffusion method is investigated. This choice is motivated by the fact that this technique can be particularly suitable to produce macromolecular crystals, because: a) the nucleation and growth phenomena self-limit maximal supersaturation that is actually reached during the crystallization, thus establishing better conditions for obtaining relatively large single crystals; b) a range of supersaturations is established in different locations along the protein solution during the process, allowing to scan different crystallization conditions in the same experiment. Both these features are of interest for growth of macromolecular crystals.

Both the conditions of gellified and non-gellified protein solution are investigated in order to discern by comparison the effect of convection and sedimentation.

The experimental investigation is supported by novel mathematical models and numerical methods. The computational model takes into account the surface incorporation kinetics. The approach used relies directly on the application of an integral form of the mass balance for each protein crystal. Shape of the crystals is ignored. This approach comes through major modification and extension of the kinetic model used by Lappa [16] (able to elucidate distribution of the local growth rate along crystal face as well as to predict shape instabilities,
onset of surface depressions due to diffusive and/or convective effects, etc, i.e. all those factors dealing with the "local" history of the shape).

In addition to the above-mentioned theoretical modelling and numerical investigation, a novel numerical strategy is introduced to investigate motion of the crystals due to the combined effect of gravitational force and viscous drag and interaction of this motion with the concentration field. The study is carried out first under the assumption that there is no buoyancy convection in protein chamber. Then the model is further improved taking into account the effect of convection due to solutal buoyancy forces.

Comparison between experimental observations and numerical simulations in presence of convection and sedimentation and without them provides a validation of the model.

The ultimate aim of the present analysis is to show how the proposed novel mathematical model and the experiments may represent a “combined” useful methodology giving insights on the investigated phenomena as well as leading to rational guidelines that can increase the probability of success in crystallizing protein substances.

2. System configuration

The crystallization technique used in the experiments is based on diffusion of a precipitant agent. Usually, this technique involves two chambers containing two different solutions (protein solution and precipitant agent solution) that are put in contact directly or by a membrane. In our case, the two solutions are put in contact, one above the other, inside a single crystallization cell. In order to prevent convective mixing and to obtain a stable interface between the two solutions, gel is added to the lower solution. For this reason the configuration under investigation simply consists of a protein solution and a precipitant agent solution placed one above the other and separated by the gel interface. The crystallization cell is a chamber whose length and height are L and H respectively; the interface is placed at \( y=h \).

At the initial time, both solutions are at constant concentration respectively. Experimentally, the shape of the gel interface cannot be horizontal due to the occurrence of a meniscus which is caused by surface tension effects. The function
\[ \xi(x) = h \left[ 1.1 - 0.1 \left( \sin \left( \frac{\pi}{4} + \frac{\pi}{2} \frac{x}{L} \right) - \frac{\sqrt{2}}{2} \right) \right] \left( 1 - \frac{\sqrt{2}}{2} \right) \]  

(1)

is used to model interface curvature in order to have a minimum protruding in the lower chamber at the mean point along the horizontal length of the chamber \((\xi(x=0)=\xi(x=L))\) and the minimum at \(x=L/2\). The gel interface is supposed to be impermeable to the protein.

3. Experimental technique

The precipitant agent is NaCl. Hen egg white (HEW) lysozyme is used as model protein, being a well-characterised molecule. A 1 % agarose solution is prepared by stirring agarose powder in a 50 mM acetate buffer solution (pH=4.5); then lysozyme (case 1) or salt (case 2) is added to agarose solution. This solution is injected (before gel solidification occurs) in the lower half part of the cell. After gel formation we introduce the precipitant agent (case 1) or lysozyme (case 2) solution in the upper part of the cell, like a supernatant liquid above the gel. Crystallization occurs by diffusion of salt in the part of the cell containing the protein solution: lower part in the case (1) and upper part in the case (2). In fact, lysozyme diffuses more slowly than salt. Sedimentation and convection in the protein chamber are prevented by the use of gel for case (1).

A Mach Zehnder interferometer powered by a 632 [nm] laser is used to monitor \textit{in situ} the transport dynamics in the fluid phase by observing the compositional field. This field is magnified by a microscope objective and observed simultaneously by a CCD digital camera (1280 x 1024 pixels). The recorded images are in turn transferred to a computer.

4. Mathematical model and numerical method:

4.1 Governing field equations:

The model is based on the mass balance equations, without consider cross-coupled coefficients, which are two order of magnitude lower at the considered concentration levels. Therefore, in absence of convection (gelified protein solution), the diffusion of lysozyme is governed by the equation

\[ \frac{\partial c}{\partial t} = D \nabla^2 c + J_N + J_K \]

(2)
D being the diffusion coefficient of lysozyme, \( J_N \) the protein depletion due to the production of new solid mass (nucleation) and \( J_K \) the protein production due to the growth (or dissolution) of an already existing crystal. As described in major detail afterwards, "analytical" expressions for \( J_N \) and \( J_K \) derived by existing models cannot be used; in fact such expressions include kinetic parameters that are not available for most of proteins and, for nucleation results are not handable for a global description of the process at macroscopic scale (for the present analysis the nucleation process will be managed with the application of evolution rules, as described in the next sections).

The initial values of protein and salt concentration are denoted by \( C(o) \) and \( C_{NaCl(o)} \) (expressed in [g/cm\(^3\)]) respectively. The walls are supposed to be impermeable to protein and salt. The gel interface is supposed to be impermeable to the protein.

4.2 The phase field variable

In the specific case of mass crystallization from a supersaturated solution one must generally accomplish at least two things simultaneously: (a) determine the concentration fields of organic substance and precipitant in the liquid phase and (b) determine the position of the interface between the solid and liquid phases. According to the technique used to address (a) and (b), in principle the numerical procedures able to solve these problems can be divided into different groups.

If the size of the crystals is negligible with respect to the size of the reactor i.e. if the seeds are small and undergo only small dimensional changes with respect to the overall dimensions of the cell containing the feeding solution, the only information associated to each grain is its position and mass. These data can be stored in special arrays. However a more elegant approach consists in introducing a non-dimensional phase-field variable \( \phi \).

This approach accounts for the solid mass stored in the generic computational cell by assigning an appropriate value of \( \phi \) to each mesh point (\( \phi=1 \) for a computational cell filled with solid mass, \( \phi=0 \) liquid and \( 0<\phi<1 \) for a computational cell containing both liquid and solid phases [16]). The key element for the method here implemented is its technique for
adjourning $\phi$. Upon changing phase, the $\phi$-value of the cell containing the crystal is adjusted to account for mass release or absorption, this adjustment being reflected in the protein concentration distribution as either a source or sink.

The volume [cm$^3$] of the crystal mass $M$ stored in a grid cell can be computed as:

$$\Delta V_{\text{stored}} = \frac{M}{\rho_p} \rightarrow \phi = \frac{\Delta V_{\text{stored}}}{\Delta \Omega}$$

where $\rho_p$ is the protein mass density in the crystal and $\Delta \Omega$ is the volume of the computational cell.

### 4.3 Nucleation

The degree of supersaturation $\sigma$ is defined by $\sigma = \frac{C}{S}$ where $S$ is the solubility (its value is function of the local concentration of the precipitant agent). The dependence of $S$ of lysozyme on $C_{\text{NaCl}}$ has been determined experimentally. In this work the expression reported by Otálora and García Ruiz [3] is used.

$$S = \exp\left[\log(0.97617) - 30.3755 \ C_{\text{NaCl}}\right] + \exp\left[\log(33.847728) + 15.13065 \sqrt{C_{\text{NaCl}}} - 260.84 \ C_{\text{NaCl}}\right]$$

(4)

When there is no pre-existing deposit, it is generally found that the concentration of protein has to be greater than $S$ to create one spontaneously, say

$$\sigma \geq \eta$$

(5)

where $\eta$ is the supersaturation limit (it is obvious that $\eta > 1$); once nuclei are created, precipitation can continue if $1 < \sigma \leq \eta$ and vice-versa material can come back to the solute condition if $0 < \sigma < 1$. Further details on the well known physics of these phenomena can be found in the excellent theoretical analysis of McPherson [1].

With regard the kinetics of the nucleation phenomena, however it is important to stress how the problem is very complex and far to be well understood. Recently, much progress has been made in the understanding of the protein crystal growth mechanism. By comparison, insight into protein nucleation processes is still very limited. Many detailed studies of the nucleation process appeared in literature described crystallization in terms of engineering mass-crystallization models, rather than crystal nucleation theories (Galkin and Vekilov [11]).
Several authors spent some effort to interpret their results in terms of the classical nucleation theories and found appreciable differences. Other authors argued that nuclei possess fractal structure and that such structure could significantly alter the nucleation mechanism and kinetics with respect to classical theories. Up to now there is still a disappointing lack for a theory of nucleation able to provide a definitive explanation of the nucleation process kinetics as well as mathematical model to be applied to predict "a priori" the experimental behaviours. In conclusions the state of the art does not allow to handle the phenomena of nucleation kinetics from a "computational" point of view, in particular at the macroscopic scale.

Fortunately, it is well known that "at the supersaturations required for nucleation, crystals typically grow extremely rapidly" (Galkin and Vekilov [11]). For this reason in the present analysis the kinetics of the nucleation process are neglected while taking into account the kinetics of the ensuing growth process. If the supersaturation limit is exceeded, the amount of supersaturated protein in liquid phase is posed in solid without any time delay. This assumption, even if being very "crude", however is supported by the very good agreement between present numerical and experimental results.

According to the crystallization criteria introduced by Henisch and García Ruiz [18] and Henisch [19] (algebraic model), once the solid particles are formed it is assumed that they are in equilibrium with protein and salt in liquid phase. Thus the solute content of the solution will have to be decremented from the original C value to the solubility value.

If $\phi^n = 0$ and $\sigma^n = (C/S)^n \geq \eta \rightarrow$

$$M_{\text{grain(o)}} = (C - S)^n \Delta \Omega, \quad \phi^{n+1} = \frac{1}{\Delta \Omega} \frac{M_{\text{grain(o)}}}{\rho_p} = \frac{S^n}{\rho_p} (\sigma^n - 1), \quad C^{n+1} = S^n$$

Formally this corresponds to the following expression for $J_N$: $J_N = -\frac{(C-S)^n}{\Delta t}$ (subjected to the condition $\sigma > \eta$) where $n$ indicates subsequent steps and $\Delta t$ is the computational time step.

4.4 The kinetic conditions

The modelling of mass exchange between the already existing solid mass and the nutrient solution leads to the introduction of a group of differential equations, strictly related, from a
mathematical point of view, to the ‘kinetic conditions’ used to model mass transfer at the crystal surface. Surface attachment kinetics at the crystal surface depend on the local value of solubility and on a coefficient $\lambda$ (kinetic coefficient) having the dimension of a velocity [cm/s]; using mass balance (see e.g. Rosenberger [2] and Pusey et al. [5]), and assuming a linear dependence of the growth rate by the interface supersaturation (see e.g. Lappa [16] and Lin et al.[17]), one obtains:

$$\left( \frac{D}{\rho_p - \rho_C C_{cs} / \rho_L} \right) \frac{\partial C}{\partial n_{cs}} = \lambda \left( \frac{C_{cs} - S}{S} \right)$$

(7)

where $C_{cs}$ is the concentration of the protein at the crystal surface, $D$ is the related diffusion coefficient, $\rho_p$ and $\rho_c$ are the protein mass density and the total mass density in the crystal, $\rho_L$ is the total density of the solution, and $\lambda$ is the kinetic coefficient. The concentration gradient can be expressed as: $\frac{\partial C}{\partial n} = \nabla C \cdot \hat{n}$ where $\hat{n}$ is the unit vector perpendicular to the solid/liquid interface pointing into the feeding solution.

Whenever protein in solute phase and solid crystal co-exist in equilibrium (saturation condition):

$$C_{cs} = S$$

(8)

Thus ‘crystals do not grow from a saturated solution’. The system must be in a non-equilibrium, or supersaturated state to provide the thermodynamic driving force for crystallization (McPherson [1]). As long as $C_{cs} < S$, more solid material will dissolve if any. If, on the other hand, $C_{cs} > S$, material will condense on any material already existing and augment its size.

It is well known (Coriell et al. [4], Lappa [16]) that the surface attachment kinetics (eq. (7)) are essential in determining the “local” evolution (surface growth rate distribution for each crystal and shape morphology). Such analysis however is out the scope of the present work.

The approach used here relies directly on the application of an integral form of the kinetic balance law for each protein seed. The shape of the crystals is ignored. A local detailed description of the morphology evolution of all the particles would require in fact very dense
grids (of the order of $10^7$ points) and therefore a prohibitive computational time. Moreover it
is out the scope of the present analysis that aims to show ensemble behaviours (i.e.
macroscopic evolution of the system under investigation).

For the sake of completeness the Reader is remanded to reference [16] for some numerical
results dealing with the morphology evolution of interacting protein seeds (able to elucidate
the distribution of the local growth rate along the sides of the crystal as well as to predict
shape instabilities, the onset of surface depressions due to diffusive and/or convective effects,
etc., i.e. all those factors dealing with the "local" history of the shape that are not taken into
account in the present work).

As pointed out in paragraph 4.2, each solid particle is considered in terms of its mass and
position. Accordingly, to account for the kinetic condition, eq. (7) is re-written as an integral
condition to be applied along the frontier $\partial \Omega$ of the computational cell containing the solid
particle. The concentration at the crystal surface ($C_{cs}$) is assumed to be equal to the value of
the concentration $C$ computed along the frontier of the computational cell containing the
crystal. Of course $C_{cs}$ is not constant along the frontier; it has a different value according to
the side where it is computed i.e. the integral form of eq. (7) takes into account the relative
direction of the different faces, which induces varying conditions for each face. In particular,
for each face $C_{cs}$ is expressed as average value ($\langle C_{i}^{\text{average}} \rangle$) with respect to the concentration
$\tilde{C}$ of the cell accommodating the crystal and that of the neighbour cell sharing the face in
question $C_{i}^{\text{neighbour}}$.

Integrating over the frontier of the computational cell equation (7) reads:

$$\int_{\partial \Omega} \nabla C \cdot \hat{n} \, dA = \int_{\partial \Omega} \lambda \left( \rho_p - \rho_c \frac{C_{cs}}{\rho_L} \right) \left( \frac{C_{cs}}{S} - 1 \right) \, dA$$

(9a)

now $\hat{n}$ is the unit vector perpendicular to the computational cell frontier pointing into the
generic neighbour cell. In discretized form:

$$\sum_{i} D \frac{C_{i}^{\text{neighbour}} - \tilde{C}}{\Delta n} A_i = \sum_{i} \lambda \left( \rho_p - \rho_c \right) \left( \frac{C_{i}^{\text{neighbour}}}{2 \rho_L} + \frac{\tilde{C}}{2S} - 1 \right) \, A_i$$

(9b)
The concentration \( \tilde{C} \) satisfying the integral form of eq. (7) is obtained from eq. (9b), then the mass stored in the computational cell is updated according to the equation:

\[
\frac{\partial M}{\partial t} = \int_{\Omega} \lambda \left( \rho_p - \rho_c \frac{C_{cs}}{\rho_L} \right) \left( \frac{C_{cs}}{S} - 1 \right) dA \to \frac{\partial \phi}{\partial t} = \frac{1}{\Delta \Omega} \int_{\Omega} \lambda \left( 1 - \frac{\rho_c C_{cs}}{\rho_p \rho_L} \right) \left( \frac{C_{cs}}{S} - 1 \right) dA
\] (10a)

And in discretized form:

\[
\phi^{n+1} = \phi^n + \Delta t \sum_i \frac{1}{\Delta \Omega} \left( 1 - \frac{\rho_c C_{cs}^{\text{neighbour}}}{\rho_p} + \tilde{C} \right) \left( \frac{C_{cs}^{\text{neighbour}} + \tilde{C}}{2S} - 1 \right) A_i
\] (10b)

with \( \tilde{C} \) satisfying eqs. (9b). Formally, these balance differential equations correspond to the following expression for \( J_K \) in eq (2): 

\[
J_K = -\frac{1}{\Delta \Omega} \frac{\partial M}{\partial t}
\]

According to equations (9) and (10) the growth velocity is not directly imposed but it results from internal conditions related to solute transport and incorporation kinetics. If the protein concentration is locally depleted, correspondingly, the solid mass stored in the computational cell grows and the phase variable is increased; on the other hand if mass stored in the cell begins to re-dissolve, protein is released in solute phase and the local value of protein concentration is increased.

The shape of the crystals is ignored while sacrificing little in accuracy for the macroscopic description of the spatio-temporal behaviour. Evidence of the reliability of the present assumptions is provided by the good agreement between experimental and numerical results.

4.5 Sedimentation:

This paragraph deals with the modelling of the sedimentation process under the assumption that solutal buoyancy convection cannot establish in the protein chamber. The model for the combined sedimentation-convection will be discussed in the next paragraph.

Crystal sedimentation will occur due to a difference between the gravitational force and the viscous resistance. The viscous resistance acting on a solid particle (the crystals are assumed to be spheres) moving with a velocity \( V_{\text{grain}} \) can be given by Stokes law as

\[
F_D = 6\pi R \mu V_{\text{grain}}
\]

where \( \mu \) is the coefficient of dynamic viscosity of the solution, \( V_{\text{grain}} \) the
velocity of the solid particle of radius $R$ and non-slip conditions are assumed at the particle surface. The gravitational force is given by $F_g = \frac{4}{3} \pi R^3 (\rho_c - \rho_L) g$.

The equation of motion of sedimenting crystals reads:

$$\frac{4}{3} \pi \rho_c \frac{d(R(t)^3 V_{\text{grain}})}{dt} = \frac{4}{3} \pi R^3 (\rho_c - \rho_L) g - 6\pi R \mu V_{\text{grain}}$$  \hspace{1cm} (11)$$

where $\rho_L = \rho_{H_2O} (1 + \beta_{\text{lys}} C + \beta_{\text{NaCl}} C_{\text{NaCl}})$ ($\beta_{\text{lys}}$ and $\beta_{\text{NaCl}}$ are the solutal expansion coefficients related to lysozyme and salt respectively); eq. (11) reads:

$$\frac{dV_{\text{grain}}}{dt} = g \left( 1 - \frac{\rho_L}{\rho_c} \right) - \frac{9}{2} \nu \frac{\rho_L}{\rho_c} \frac{V_{\text{grain}}}{R(t)^2} - 3 \frac{V_{\text{grain}}}{R(t)} \frac{dR}{dt}$$  \hspace{1cm} (12)$$

where, $R(t) = \sqrt[3]{\frac{3}{4\pi} \frac{M_{\text{grain}}(t)}{\rho_p}} = \sqrt[3]{\frac{3}{4\pi} \phi(t) \Delta \Omega}$,

$M_{\text{grain}}(t)$ is the mass of protein associated to each grain. In eqs. (11)-(12) the dependence of the radius of the solid particle on time has been highlighted in order to point out that the size of a sedimenting particle may change during sedimentation due to further absorption (or release) of mass ($M_{\text{grain}}$ is a function of time) as explained in the previous paragraph. Note that the viscous force has been computed according to the Stokes formula that is valid only when the velocity does not change in time (i.e. at steady conditions for the motion of the solid particle); since the time required by the generic solid particle to reach the bottom of the protein chamber is of the order $O(10^2 \ [s])$ whereas the characteristic time (the diffusion time of salt) of the phenomena under investigation is of the order $O(10^5 \ [s])$, it is reasonable to suppose that the approximations introduced in eq. (12) do not alter significantly the phenomena under investigation.

Eq. (12) has been solved for each crystal assuming $V_{\text{grain}}=0$ and $y=y_{\text{nucl}}$ as initial conditions, where $y_{\text{nucl}}$ is the position at which nucleation occurs. Protein crystals are denser than the feeding solution thus on Earth they settle in the computational cells on the bottom of the growth chamber.
4.6 Sedimentation-Convection:

The sedimentation model discussed in the previous paragraph is now improved and further refined by taking into account the effect of convection arising in the protein chamber due to buoyancy forces. The flow is governed by the continuity, Navier-Stokes and species equations, that in non-dimensional conservative form read:

\[ \nabla \cdot \mathbf{V} = 0 \quad (13) \]

\[ \frac{\partial \mathbf{V}}{\partial t} = -\frac{1}{\rho} \nabla p - \nabla \cdot [\mathbf{V} \mathbf{V}] + \nu \nabla^2 \mathbf{V} + g \beta_{\text{lys}} \left( C - C_{(o)} \right) \mathbf{g} + g \beta_{\text{NaCl}} \left( C_{\text{NaCl}} - C_{\text{NaCl(o)}} \right) \mathbf{g} \]

\[ \frac{\partial C}{\partial t} = -\nabla \cdot \left[ \mathbf{V} C \right] + D \nabla^2 C \quad (15) \]

\[ \frac{\partial C_{\text{NaCl}}}{\partial t} = -\nabla \cdot \left[ \mathbf{V} C_{\text{NaCl}} \right] + D_{\text{NaCl}} \nabla^2 C_{\text{NaCl}} \quad (16) \]

The non-dimensional Rayleigh numbers are defined as

\[ Ra_{\text{lys}} = \frac{g \beta_{\text{lys}} L^3 C_{(o)}}{\nu D} \] and

\[ Ra_{\text{NaCl}} = \frac{g \beta_{\text{NaCl}} L^3 C_{\text{NaCl(o)}}}{\nu D_{\text{NaCl}}} \]. Non-slip conditions are imposed on the walls.

Eqs. (13-16) subjected to the initial and boundary conditions were solved numerically in primitive variables by a finite-difference method. The domain was discretized with a uniform mesh and the flow field variables defined over a staggered grid. Forward differences in time and central-differencing schemes in space (second order accurate) were used to discretize the partial differential equations.

For further details on the numerical method (MAC) used to compute the velocity see Refs.[20-22]. The protein and salt distributions at time (n+1) are obtained from Eqs. (15) and (16) after the velocity calculation. The nucleation and kinetic incorporation phenomena are taken into account according to the technique described in paragraphs 4.3 - 4.4. The equation of motion of sedimenting crystals has to be modified in order to take into account the velocity field due to the buoyancy forces

\[ \frac{d V_{\text{grain}}}{dt} = g \left( 1 - \frac{\rho_{L}}{\rho_{\text{c}}} \right) - \frac{9}{2} \nu \frac{\rho_{L}}{\rho_{\text{c}}} \frac{V_{\text{grain}}}{R(t)^2} - 3 \frac{V_{\text{grain}}}{R(t)} \frac{dR}{dt} \quad (17) \]

which in discrete form reads:
\[ V_{\text{grain}}^{n+1} = V_{\text{grain}}^n + \Delta t g(1 - \frac{\rho_L}{\rho_C}) - \Delta t \frac{9}{2} V_j \frac{\rho_L}{\rho_C} \frac{V_{\text{grain}}^{n+1} - V_{\text{grain}}^n}{(R_{n+1})^2} - 3 \frac{V_{\text{grain}}^{n+1}}{R_{n+1}} (R_{n+1} - R^n) \] (18)

i.e. it is assumed that, due to the small size of the sedimenting crystals, the velocity field \( V \) is not affected by the motion of these particles and that, vice-versa, the sedimenting particles may be accelerated or decelerated by the effect of \( V \).

The numerical technique discussed from paragraphs 4.1 to 4.4 has been validated through comparison with the numerical results of Henisch [19]. The numerical method described in paragraph 4.6 requires the application of subroutines which have been already used by the authors in previous works and have been widely validated (see e.g. Refs.[20-22]).

5. Results and discussion

The following properties and operating conditions have been used for the numerical simulations: \( D = 10^{-6} \text{ [cm}^2/\text{s}] \), \( D_{\text{NaCl}} = 10^{-5} \text{ [cm}^2/\text{s}] \), \( \nu = 8.63 \times 10^{-3} \text{ [cm}^2/\text{s}] \), \( \rho_C = 1.2 \text{ [g/cm}^3] \), \( \rho_p = 0.82 \text{ [g/cm}^3] \), \( \lambda = 3 \times 10^{-7} \text{ [cm/s]} \), \( \beta_{\text{lys}} = 0.3 \text{ [(g/cm}^3)\text{-1]} \), \( \beta_{\text{NaCl}} = 0.6 \text{ [(g/cm}^3)\text{-1]} \), \( C_{\text{NaCl(o)}} = 14 \times 10^{-2} \text{ [g/cm}^3] \), \( L = 1 \text{ [cm]} \), \( H = 4 \text{ [cm]} \), width = 0.1 [cm], average height of the protein chamber \( h = 1.95 \text{ [cm]} \), 60 computational points along \( x \) and 190 computational points along \( y \). Figures 2-7 show the results for the crystallization in gel. Figures 8-12 deal with the sedimentation process.

5.1 Simulation of protein precipitation in gel

For the case of gellified protein chamber (case (1), diffusive transport regime), different values of the initial protein concentration have been considered, in order to investigate the effects of supersaturation level on the system dynamics.

Simulations show that some "bands" of crystals are produced in the gel i.e. discrete nucleation events separated by certain time (and thus space ) intervals occur in the protein chamber. These bands are not spatially uniform (see Figs. 4 for \( C(o) = 4 \times 10^{-2} \text{ [g/cm}^3] \); note that in these figures the salt chamber is located on the top of the protein chamber). According to
the evolution of the concentration field (Figs.3) nucleating particles deplete their surroundings of protein which causes a drop in the local level of supersaturation such that the nucleation rate falls in the neighbourhood, leading naturally to a spacing between regions of nucleation that gives rise to the alternate presence of depleted zones and crystals. This behaviour exhibits interesting analogies with the well-known so-called "Liesegang patterns" (see e.g. Garcia-Ruiz [23], Henisch [24] and Lappa et al. [25]).

The spacing among different solid particles and the size of the particles changes according to the distance from the origin of the imposed concentration gradient (for the present case the gel interface between the salt and protein chambers). The space distance among near solid particles, in fact, tends to decrease uniformly towards the interface of gel. Correspondingly the size of these particles is minimum at the interface and increases downwards up to a constant value in the bulk. According to the results discussed above, the “density of particles” defined as ratio of the number of particles in a fixed reference volume and the reference volume, is maximum near the interface and almost constant in the bulk. These behaviours are shown in Figs. 5a and 5b.

According to figures 4 the crystallization process is characterized by a certain degree of periodicity in space. Lysozyme precipitates to produce deposits spaced in both the x and y directions. The term “periodic”, however, should be interpreted with caution and tolerance. Strictly speaking, it demands a period (i.e. a constant time or space interval), something that the phenomena here under discussion fail to show. The periods are not constant throughout the crystallization process, but they are not random either. An attempt made to measure the wavelength has shown in fact that a certain time interval can be highlighted where this factor exhibits a somehow constant value. This interval is $3 \cdot 10^4 \text{ [s]} < t < 7 \cdot 10^4 \text{ [s]}$. In this range, the wavelength $\gamma$ (distance between two consecutive crystals along y) is almost constant ($\gamma = 0.6 \text{ [cm]}$).

Further to the previous case, the precipitation from super-saturated solutions with $C_{(o)} = 6 \cdot 10^{-2} \text{ [g/cm}^3\text{]}$ and $C_{(o)} = 8 \cdot 10^{-2} \text{ [g/cm}^3\text{]}$, has been analyzed (see Figs. 6 for the case $C_{(o)} = 6 \cdot 10^{-2} \text{ [g/cm}^3\text{]}$, the case $C_{(o)} = 8 \cdot 10^{-2} \text{ [g/cm}^3\text{]}$ is not shown for the sake of brevity; an exhaustive
description of this case can be found in Carotenuto et al. [26]). According to the results shown in Figs. 6, the spacing among different solid particles is different with respect to the previous case. Particles are closer to each other. The average distance between the particles distributed along x and y is lower if compared with the case $C_{(o)} = 4 \cdot 10^{-2}$ [g/cm$^3$] ($\gamma = 0.35$ [cm] for $C_{(o)} = 6 \cdot 10^{-2}$ and $\gamma = 0.20$ [cm] for $C_{(o)} = 8 \cdot 10^{-2}$). Moreover the solid particles are characterized by a smaller diameter.

The average size of particles is a decreasing function of the initial amount of protein available in solution (compare Figs. 5a and 6c). Figs. 5b, and 6d show the density of particles distribution. The decrease of the average size of the particles is coupled to an increase of the number of crystals.

5.2 Comparison with the experiments on precipitation in gel

Comparison between experimental observation and numerical simulations provides a validation of the present hypotheses, numerical models and strategies and gives insights on the crystallization process, providing, for instance, an evaluation of supersaturation threshold for nucleation.

The results discussed above have been obtained for particular values of the ratio between the supersaturation limit and the solubility ($\eta$). Regarding this ratio, it should be pointed out that it is affected by a large uncertainty. This uncertainty is overcome through the comparison between the experimental and numerical results obtained for different values of this parameter. A parametric analysis has been carried out ($3 \leq \eta \leq 5$) for the case $C_{(o)} = 4 \cdot 10^{-2}$ [g/cm$^3$]. The computations show that the problem is quite sensitive to $\eta$. It can influence the number of locations where nucleation occurs (i.e. the number of solid crystals produced due to nucleation) and their spatial distribution too, i.e. the density of particles. Figs. 7 show the size of particles versus the distance from the gel interface and their density as function of the parameter $\eta$. According to these behaviours, the value of $\eta$ should be assumed in the range [3-4] in order to be compatible with experimental observations. Finally, in fact, the value $\eta = 3.6$ has been found for this case. The same parametric analysis carried out for the cases
C(o)=6·10^{-2} [g/cm^3] and C(o)=8·10^{-2} [g/cm^3] has shown that the best agreement between numerical and experimental results is achieved for \( \eta=2.3 \) and \( \eta=1.6 \) respectively.

The numerical results are in good quantitative agreement with the experimental ones both for the particles size and density distributions (see Fig.4a,b and 6c,d).

According to this study, the value of \( \eta \) is a decreasing function of the initial value of protein concentration, i.e., the larger the amount of protein available in liquid phase, the lower the supersaturation limit. This trend is not unexpected and allows to explain the different values of \( \eta \) experimentally found by the different investigators. There is no physical reason forcing \( \eta \) to be a constant. On the contrary one may suspect that a larger initial amount of protein available in liquid phase can lower activation energy of the nucleation process.

5.3 Discussion for the results in gel

The present effort demonstrates both the potential and challenges of mathematical modeling of protein crystal growth. Very interesting information can be gathered from the comparison between numerical results and experimental observations, such as: "physical" parameters still affected by large uncertainty, cause-and-effect relationships, intrinsic mechanisms of the process, counteracting behaviours, etc.

For instance, according to the computations, the increase of the solid particles size is due to the delicate balance between two counteracting effects: 1) protein condenses on any solid crystal already existing augmenting its size; this depletes the protein concentration and leads to lower values of C; 2) salt continues to diffuse through the interface so that S is reduced due to solubility modulation. If the second effect prevails over the first one, i.e. the protein concentration is larger than the "solubility" then further protein precipitation in the considered region is possible.

The growth law is driven by the surface attachment kinetics according to equations (9) and (10). Therefore the model developed is also able to provide a description of crystal growth as a function of local conditions established in the protein chamber. As an example Fig. 5c shows the growth law for a reference crystal located in the bulk of the protein chamber (x
This seed is formed at $t \approx 7 \times 10^4$ [s] and reaches its final size after almost $10^5$ seconds. A "measure" of the average growth rate is also possible. The growth rate (computed over the time interval required to reach the 90% of the final mass of the crystal) is $1.27 \times 10^{-9}$ [g/s].

From a "global" point of view, as time passes the extension of the region affected by protein depletion increases towards the bottom of the protein chamber (Figs 3a-c). This is due to the combined effect of the nucleation and ensuing growth phenomena. The propagation towards the bottom of the nucleation process depletes of lysozyme the protein chamber and further depletion occurs due to the phenomenon of subsequent precipitation (Figs. 3 and 4). As time passes new regions are involved in the nucleation and further growth processes whereas previous regions are no longer characterized by these phenomena.

The computations show that the macroscopic space-modulated dynamics of the above process are not very sensitive to the incorporation kinetics. In fact both size and crystal number distributions are similar to those obtained in previous investigations (see e.g. Carotenuto et al. [26]) in which local equilibrium conditions were assumed. This can be explained by the fact that the evolution is mainly driven by phenomena such as the competition for growth of different crystals and the intersection and overlapping of the related depletion zones. It is worthwhile to stress how due to the presence of many growing crystals, the process is mainly controlled by these aspects. Some differences between kinetic and equilibrium models are expected in the early phase of the growth when depletion zones are separated; in our case due to the large number of crystals the dynamics become rapidly limited by solute availability, reducing the relative importance of kinetic barrier.

5.4 Sedimentation

Consider the case where gel matrix is not used in the protein chamber (see Fig. 1b). The case $C_{(o)}=4 \times 10^{-2}$ [g/cm$^3$] has been investigated. The buoyancy-driven convection has been neglected for the first group of simulations.
Since protein crystals are denser than the feeding solution, sedimentation occurs. Due to sedimentation, the phenomena under investigation do not show the patterns described in the previous paragraphs. The solid particles in fact settle in the growth chamber. The lower boundary of the protein chamber being the interface of gel used to fill the salt chamber, the crystals sediment on this interface. This is clearly shown in Figs. 9 where the thickness of the layer of crystals sedimented on the bottom is an increasing function of time due to the continuous mass incoming from the bulk. Figs. 9a show that the first nucleation occurs at the free interface. Solid particles accumulate on the interface and then, if the local value of concentration is larger than the “solubility”, they grow due to further precipitation of protein. As time passes and salt diffuses in the protein chamber, other nucleation occurs far from the interface. Similarly to the configuration investigated in paragraph 5.1 (i.e. crystals locked on the gel matrix), each nucleation occurring in the protein chamber gives rise to a depletion zone around the location where the solid particle has condensed. For the present case, however the solid particles are not locked on their initial position and for this reason they leave the position where nucleation has occurred. During the fall to the bottom, the solid particles deplete of protein the surrounding liquid. For this reason a wake is produced beyond the sedimenting particles and a small depletion spot highlights their initial position (Figs. 8, note that in these figures the salt chamber is located on the bottom of the protein chamber). As time passes, due to the nucleation process, a depleted “band” is created starting from the gel interface and protruding in the protein chamber. The amplitude of the band is an increasing function of the time. This effect behaves as a "disturbance" in the concentration field propagating towards the top of the protein chamber (Figs. 8).

The propagation towards the top of the nucleation events depletes of lysozyme the protein solution. This behaviour is similar to that described in paragraph 5.1. For the present case, however, the “rate of depletion” is lower. In this case, in fact, further depletion due to the phenomenon of subsequent precipitation after nucleation is negligible (it simply applies to the wakes beyond the sedimenting particles). Due to the sedimentation process, solid particles leave the regions of high protein concentration where further precipitation could occur. The
crystals reach the bottom of the protein chamber where the concentration is low due to previous protein absorption. Since all the solid particles tend to be clustered and accumulated in a same zone of low protein concentration, the possibility of further precipitation and consequent size augmenting is very reduced with respect to the case of locked crystals (gel matrix). For the reasons discussed above, the value of the protein concentration in the bulk of the chamber tends to be higher with respect to the case of gellified protein solution. This trend is confirmed by the experimental results.

5.5 Sedimentation-Convection

Fig. 10 (at each station y the density, computed as \( \rho_L = \rho_{H,0} \left( 1 + \beta_{by} C + \beta_{NaCl} C_{NaCl} \right) \), has been averaged along x) shows the time-evolution of the density profile along y in the protein chamber obtained under the assumption of protein chamber free of solutal buoyancy convection. For a fixed instant, the density has a maximum on the gel interface and a minimum on the top of the protein chamber. This behaviour is the sum of two counteracting effects: lysozyme is depleted in regions close to the gel interface (due to the nucleation process and subsequent solid particles accumulation due to sedimentation) thus reducing the density; salt diffuses in the protein chamber through the gel interface thus increasing the density. The latter effect prevails since salt has a larger diffusion coefficient. For this reason the density has a maximum on the gel interface and then tends to decrease towards the top of the protein chamber. The shape of the density profile however is not always decreasing along y. Some local inversions occur in this profile due to nucleation phenomena in the bulk of the protein chamber. Nucleation phenomena, in fact, absorb protein; this leads to a sink in the density distribution and correspondingly a ripple appears in the shape of the density profile. The position of this ripple is not fixed as time passes. It migrates along y according to the propagation towards the top of the nucleation process. For \( t=1.6 \cdot 10^5 \) [s] the sink reaches the top of the protein chamber and then it disappears for further increase of the time.

These results suggest that, for \( t<1.6 \cdot 10^5 \) [s], onset of convection is possible due to buoyancy forces associated to the inversion in the density profile. For this reason simulations have been
carried out taking into account the solution of the non-linear and time-dependent Navier-Stokes equations (Figs. 12). The numerical results show that vortex cells arise. These rolls however do not extend over the entire bulk of the protein chamber and are confined to the ripple zone of the density profile.

As time passes, the vortex rolls migrate towards the top of the protein chamber (see Figs. 12). This behaviour is strictly related to the propagation along y of the zone where the inversion occurs. Moreover, the migration of the cells towards the top is coupled to a change of their number (i.e. the value of the wave number m associated to the convective instability). This number changes according to the distance from the gel interface. Close to the gel interface m is about 6 (Fig.12a), but this number is reduced when, due to the propagation towards the top, the distance from the gel interface increases. The wave number becomes 5 in Figs 12b and 12c and is reduced to m=2 in Fig. 12d.

It is known that discrete wavenumbers of convective disturbances are selected out of the full spectrum of disturbances because the convection rolls are closed in a special zone geometry. The selection rule is given simply by the constraint that the wavelength must be an aliquot of the length along x of the chamber (L) and by the fact that the convection rolls are confined to the extension along y of the ripple zone (L_i). According to this theory the critical wave number should be related to the length along x of the chamber and to the extension along y of the ripple zone (L_i), i.e. it should scale according to the law

\[ m = \frac{L}{L_i} \quad (19) \]

Fig.10 shows that the extension along y of the ripple zone (L_i) is an increasing function of the distance from the gel interface. This explains why the wave number is reduced as time passes. Since the diameter of the vortex rolls is almost equal to L_i (the force driving convection acts along y over a length equal to L_i) and the space available along x is fixed, the number of convective cells decreases if L_i is increased.

Regarding the lysozyme concentration distribution, the simulations point out that the effect of the vortex rolls is very weak. They simply add some “sinusoidal noise” to the concentration contour-lines and make the crystals distribution on the bottom of the protein chamber non
symmetric with respect to the point x=L/2. The wave number in Fig. 12a (m=6) is in good agreement with the experimental results shown in Fig. 11. Fig. 11 shows sinusoidally distorted fringes (fringes represent sites of equal refractive index (n) that is dependent on protein and salt concentrations). The sinusoidal distortion of the fringes close to the gel interface (t=2 10^4 [s]) is due to convection and the number of waves is 6-7.

6. Conclusions

A counter-diffusion technique and different values of the initial protein concentration have been considered in order to investigate the effect that the initial amount of protein available in solution (i.e. the effect of different supersaturation conditions) may have on the spatio-temporal behaviour of the crystallization in the case of gellified configuration. Results show that increase of the initial protein concentration decreases the average size of the solid particles, increases their number and “accelerates” propagation of the nucleation process towards the bottom of the protein chamber.

Without gel (second configuration studied), the patterns of crystals are destroyed. The solid particles in fact settle in the growth chamber. The thickness of the layer of crystals sedimented on the bottom increases with time due to the continuous mass incoming from the bulk. As time passes and salt diffuses into the protein solution, nucleation occurs further from the interface. During the fall to the bottom, the solid particles deplete the surrounding liquid of dissolved protein. For this reason a wake is produced beyond the sedimenting particles. If convection due to buoyancy forces is taken into account, a very complex convection pattern arises which dynamics depend on the nucleation in the bulk of the protein chamber. The results show that convective cells arise due to local inversions in the density distribution. These vortex cells migrate towards the top of the protein chamber exhibiting a size dependent on the distance from the gel interface and affecting the nucleation dynamics in the bulk. Sedimentation inhibits growth of crystals since they tend to be clustered in a region where the concentration of protein in liquid phase is small.
The proposed models and methods exhibit heretofore unseen capabilities to predict and elucidate experimental observations at macroscopic level and to identify cause-and-effect relationships. Both experimental observation and numerical modelling show comparable features of the phenomena under investigation. The consistency of model predictions with experimental data suggests that the proper rate-controlling steps have been taken into account, and that simplifications do not influence much the system behaviour. It has been shown how the proposed novel mathematical model and the used experimental technique provide a “combined” useful methodology that gives insights on the crystallization process.

7. Acknowledgements
The present work has been supported by the European Space Agency (ESA) and the Italian Space Agency (ASI). The authors would like to thank Professor Alexander A. Chernov (USRA) for the precious suggestions and for the guidelines provided, in particular for the discussion about the effect of a large number of interacting crystals and the related transition from kinetic to diffusion-controlled mode.

8. References
evidence of the stability of the depletion zone around a growing crystal under microgravity",
1681.
1-147
[12] C. P. Lee, A. A. Chernov, "Solutal convection around growing protein crystals and
Approach" (Practical Approach Series, 210) 2nd ed. Oxford University Press, New York
(1999.).
[15] A. McPherson, "Procedures for Macromolecular Crystallization, Crystallization of
(1999).
conditions”, Phys. Fluids 15 (4), (2003), 1046. This article has been also selected by the
American Physical Society for the 15 March 2003 issue of the Virtual Journal of Biological
Physics Research (Volume 5, Issue 6).


Fig.1: sketch of the configurations: (a) no sedimentation allowed (b) Lysozyme sedimenting on the gel interface

Fig.2: Experimental distribution of the crystals – (gellified configuration, \(C_{(o)}=4\cdot10^{-2}\, [\text{g/cm}^3]\))
Time=1.73\cdot10^5\, \text{s}
Non-dimensional protein concentration contour lines (3a) and distribution of the crystals (4a) – (gellified configuration, $C_{(o)}=4 \cdot 10^{-2} \text{ [g/cm}^3\text{]}$) Time=8.00 $\cdot 10^3$ [s]

Non-dimensional protein concentration contour lines (3b) and distribution of the crystals (4b) – (gellified configuration, $C_{(o)}=4 \cdot 10^{-2} \text{ [g/cm}^3\text{]}$) Time=7.2 $\cdot 10^4$ [s]

Non-dimensional protein concentration contour lines (3c) and distribution of the crystals (4c) – (gellified configuration, $C_{(o)}=4 \cdot 10^{-2} \text{ [g/cm}^3\text{]}$) Time=1.08 $\cdot 10^5$ [s]
Figs. 5: comparison between numerical and experimental results (gellified configuration, \(C_0=4 \times 10^{-2} \text{ [g/cm}^3\text{]}\)): (a) solid particle size distribution, (b) density of particles distribution, (c) computed "growth law" of a reference seed (\(x = 0.4 \text{ [cm]}\), \(y=0.9 \text{ [cm]}\)) located in the bulk of the protein chamber.
Fig. 6a: Experimental distribution of the crystals—Fig. 6b: Numerical distribution of the crystals
(gellified configuration, $C_{(o)} = 6 \cdot 10^{-2}$ [g/cm$^3$]) Time = $1.7 \cdot 10^5$ [s]

Fig. 6c: Solid particle size distribution: comparison between numerical and experimental results (gellified configuration, $C_{(o)} = 6 \cdot 10^{-2}$ [g/cm$^3$])
Figs. 6: Comparison between numerical and experimental results (gellified configuration, \(C_{(o)} = 6 \cdot 10^{-2} \text{[g/cm}^3]\)): (d) density of particles distribution.

Figs. 7: Effect of the parameter \(\eta\) – comparison between numerical and experimental results (gellified configuration, \(C_{(o)} = 4 \cdot 10^{-2} \text{[g/cm}^3]\)): (a) the solid particle size distribution, (b) density of particles distribution.
Non-dimensional protein concentration contour lines (8a), distribution of the crystals (9a) – $(C_o=4\cdot10^{-2} \,[g/cm^3])$ Time=$7.35\cdot10^4 \,[s]$

Non-dimensional protein concentration contour lines (8b), distribution of the crystals (9b) – $(C_o=4\cdot10^{-2} \,[g/cm^3])$ Time=$1.365\cdot10^5 \,[s]$

Non-dimensional protein concentration contour lines (8c), distribution of the crystals (9c) – $(C_o=4\cdot10^{-2} \,[g/cm^3])$ Time=$1.68\cdot10^5 \,[s]$
Fig. 10: Density distribution in the protein chamber versus time ($C_0 = 4 \cdot 10^{-2} \text{ [g/cm}^3\text{]}$)

Fig. 11: interferogram, $t=2.0 \cdot 10^4$ [s]

Figs. 12: velocity field– ($C_0 = 4 \cdot 10^{-2} \text{ [g/cm}^3\text{]}, R_{\alpha S} = 1.36 \cdot 10^9, R_{NaCl} = 9.5 \cdot 10^8$):
(a) Time $= 1.8 \cdot 10^4$ [s], (b) Time $= 3.9 \cdot 10^4$ [s], (c) Time $= 6.3 \cdot 10^4$ [s], (d) Time $= 9.3 \cdot 10^4$ [s]