

Protein Crystallization

Labwork Nanoanalytics and Interfaces
WS 2021/22

Supervisor: Hadra Sioux Banks-Machado

I hereby declare that my elaboration (or that of my group) was made independently and is not a copy (even partially) of an existing protocol. I am aware that violation can lead to an exclusion from the labwork.

Date: 04.01.2022

Signature:

Content

1	INTRODUCTION.....	1
1.1	Motivation.....	1
1.2	Relevant Theories for Protein Crystallization	1
1.2.1	Classical Nucleation Theory	1
1.2.2	Two-Step Nucleation Theory.....	3
1.2.3	DLVO Theory	3
2	MATERIALS AND METHODS.....	5
2.1	Sample Preparation	5
2.2	Microscope Images	5
3	RESULTS AND DISCUSSION	6
3.1	Two-Step Crystallization	6
3.2	One-Step Crystallization and Crystal Growth Process.....	7
3.3	Crystal Aspect Ratio Analysis	9
4	CONCLUSION	11
5	LITERATURE	12

1 Introduction

The aim of this experiment was to investigate the protein crystallization of the enzyme lysozyme under varying conditions. Therefore, seven samples with different concentrations of the protein (lysozyme), the salt (NaCl) and the buffer solution were prepared. The samples were then observed and evaluated under a microscope by applying the software ZEN 2010.

1.1 Motivation

Proteins play a crucial role for humans as they are the basis for many relevant functions in the human body, e.g. DNA replication, digestion or the immune system. However, it is assumed that protein interaction or aggregation can lead to diseases like Parkinson or Alzheimer [1]. An important group of proteins are enzymes to which the lysozyme, that is investigated in this experiment, belongs to. Lysozyme is a part of the innate immune system of humans [2]. Due to its antibacterial effect the cell wall of peptidoglycan is attacked by cleaving glycosidic bonds [2]. Since there is a correlation between function and structure of a protein, it is important to know the exact structure of a protein, to better understand its role in the human body. The commonly used X-ray structure analysis requires crystallized proteins [3]. Therefore, it is crucial to know under which conditions a protein crystallizes the best.

1.2 Relevant Theories for Protein Crystallization

There exist a few theories that explain the crystallization of proteins in a supersaturated solution. These theories differ in their complexity regarding the assumptions that have been made previously. The theories have in common that a supersaturated solution creates a nonequilibrium condition [4]. As a result, molecules dissolved in the solution start aggregating what leads to the creation of nuclei and subsequent crystal growth [4].

1.2.1 Classical Nucleation Theory

The classical nucleation theory (CNT) is considered being a simple theory to explain protein crystallization and is therefore widely used [4]. This “simplicity” is due to several theoretical assumptions that are made when using it [4]. First, it is assumed, that the clusters are spherical with a homogeneous density inside that is independent of the size of the cluster [4]. Second, the specific surface tension is independent of size and shape of a cluster [4]. Third, the cluster grows by addition of single growth units [4]. Fourth, the nucleation rate is time independent [4]. The last assumption is that the cluster can be seen as incompressible [4].

Due to collisions of growth units (molecules or oligomers) in the supersaturated solution, these units adhere to each other and form small clusters [5].

Cluster formation can be described by the change of the Gibbs free energy ΔG . It is on the one hand described by the energy change in the cluster ΔG_v that supports the formation of the cluster, and on the other hand by the energy change of the surface ΔG_s that supports the dissolution of the cluster [4]. So, these two terms compete against each other during a potential formation of a cluster [4]. ΔG_{crit} is the energy barrier that has to be overcome forming a cluster (see fig. 1) [4].

$$\Delta G = \Delta G_v + \Delta G_s \quad (1)$$

ΔG_v describes the energy difference when new particles of the supersaturated solution deposit to the cluster [4]. The value can be determined by the following equation [5]:

$$\Delta G_v = n(\mu_L - \mu_S) = -nRT \ln \left(\frac{c}{c_{eq}} \right) = -\frac{4\pi r^3}{3\nu} RT \ln \left(\frac{c}{c_{eq}} \right) \quad (2)$$

It is dependent on the difference of chemical potentials of the dissolved (μ_L) and the solid phase (μ_S) of a particle, that is multiplied with the number of growth units n within the cluster [5]. Therefore, this term decreases the free energy, as the solid phase is more stable than the liquid one, what results in a negative term for ΔG_v [4].

The protein concentration at the beginning is denoted by c , whereas c_{eq} describes the concentration after the terminated crystal growth, thus when there exists an equilibrium in the solution [5]. The volumetric particle density is described by ν [5].

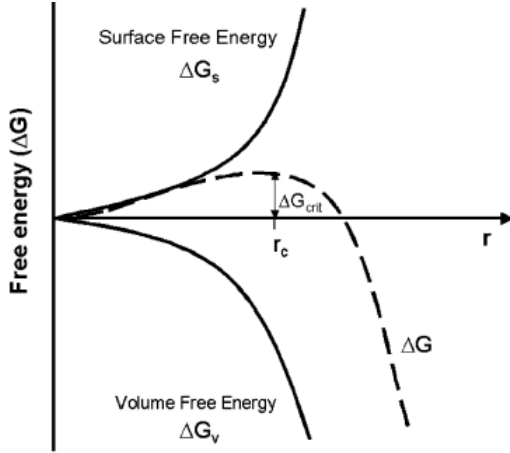


Figure 1: Free energy ΔG between two molecules and its components (surface and volume free energy) dependent on molecule distance r [4]

With increasing solid/liquid interface ΔG_s increases, and thus the total Gibbs free energy [4]. The change of the surface energy ΔG_s can be calculated by [5]:

$$\Delta G_s = A \cdot \gamma = 4\pi r^2 \cdot \gamma \quad (3)$$

with A as surface area of the cluster, and γ as specific surface tension [5].

Adding the terms for ΔG_v and ΔG_s , the following equation is obtained [5]:

$$\Delta G = -\frac{4\pi r^3}{3\nu} RT \ln \left(\frac{c}{c_{eq}} \right) + 4\pi r^2 \cdot \gamma \quad (4)$$

By deriving equation (eq.) 4 after r , the so-called critical radius r^* of a cluster, required to build a stable cluster, is reached when the derivative $\frac{d\Delta G}{dr} = 0$ [5].

$$r^* = \frac{2\gamma\nu}{RT \ln \left(\frac{c}{c_{eq}} \right)} \quad (5)$$

As described above, the interplay between ΔG_v and ΔG_s is crucial for a potentially stable formation of a cluster [4]. This leads to the conclusion that if the radius of a cluster is larger than r^* , the interparticle bonds are stronger than the free bonds at the surface of the cluster [5]. In this case, a stable cluster can be formed [5].

1.2.2 Two-Step Nucleation Theory

The two-step nucleation theory is more complex than the classical one. Thus, it is better suitable to describe the “real” nucleation process [4]. In contrast to the CNT assuming that a cluster is built directly, the two-step nucleation theory consists of – as it is named – two steps forming a cluster (see fig. 2) [4]. The first step comprises the so-called liquid-liquid phase separation (LLPS). This means that due to an extremely supersaturated solution droplets are formed that have a higher protein concentration than the surrounding protein concentration in the solution [5]. So, there is an increased “dense” local protein concentration. With regard to the crystal phase, the dense phase is metastable [5]. In the second step, out of these droplets with the higher protein concentration, the crystals are formed [4].

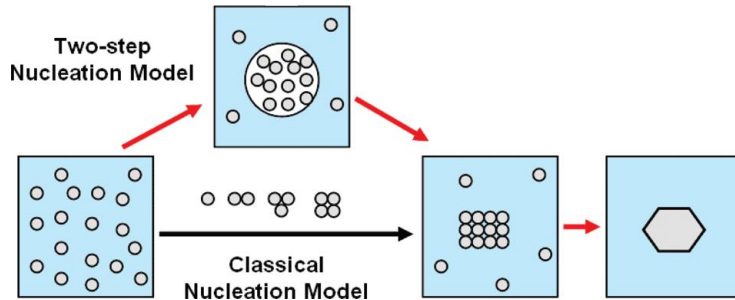


Figure 2: Model of CNT and two-step nucleation theory [4]

The LLPS can be influenced by depleting agents, e.g. NaCl [6]. By adding NaCl to the solution, the charge of the protein molecules is reduced. Resulting from this, the repulsive interactions between the protein molecules are depleted and the neutral growth units can closely approach. Furthermore, the LLPS can be influenced by a change in temperature [6]. The crystallization process is either induced below or above a critical temperature [6]. The starting temperature for LLPS is called cloud point temperature, because the solution seems to be cloudy [6].

1.2.3 DLVO Theory

The DLVO Theory describes the influence of different forces on the interaction of colloidal particles, e.g. the interaction of two protein molecules when they come close to each other [7]. The overall interaction potential $V_{DLVO}^{(r)}$ of the molecules is determined by three different forces [7].

$$V_{DLVO}^{(r)} = V_{EE}^{(r)} + V_{VDW}^{(r)} + V_{HS}^{(r)} \quad (6)$$

The so-called “Hard Sphere” interaction $V_{HS}^{(r)}$ is the potential of repulsive forces due to overlapping electron clouds of two molecules [7]. However, this potential is vanishingly small and can therefore be neglected [7].

$V_{EE}^{(r)}$ describes the potential of the repulsive Coulomb interaction [7]. Around a protein, there is the electric double-layer (EDL) with its thickness λ_D (Debye length). The Debye length describes the distance from an electrically charged point along which the charge decays by the factor $1/e$ [7]. Around the EDL, there is the diffusive layer with randomly charged ions in it [7].

$$\lambda_D = \sqrt{\frac{k_B T \epsilon \epsilon_0}{e^2 I}} \quad (7)$$

With I as ionic strength:

$$I = \frac{1}{2} \sum z_i^2 \cdot c_i \quad (8)$$

The Coulomb interaction potential $V_{EE}^{(r)}$ can be calculated by [8]:

$$V_{EE}^{(r)} = \frac{Q}{\epsilon (1+\kappa a)} \cdot e^{\frac{-\kappa(r-\lambda)}{r}} \quad (9)$$

with Q as the total charge of the molecule, ϵ as dielectric constant, κ as Debye screening length ($\kappa = \lambda_D^{-1}$), a as radius of the molecule and λ as the diameter of the molecule ($\lambda = 2a$).

$V_{VDW}^{(r)}$ describes the potential of the attractive van der Waals interaction that occurs between two molecules with an electric dipole moment [7]. It can be calculated by [8]:

$$V_{VDW}^{(r)} = -\frac{A_H}{12} \left(\frac{\lambda^2}{r^2} - \frac{\lambda^2}{r^2 - \lambda^2} - 2 \ln \left(\frac{r^2 - \lambda^2}{r^2} \right) \right) \quad (10)$$

A_H is the Hamaker constant, r is the distance between the two molecules.

Fig. 3 shows the net interaction potential $V_{DLVO}^{(r)}$ (green curve) in dependence of the distance between two molecules. At a certain point when two molecules come close to each other, the attractive van der Waals forces are stronger than the repulsive Coulomb forces. So, the net interaction potential can be overcome, and the molecules remain next to each other. By adding NaCl or other ions to a protein solution, the net interaction potential is decreased because the NaCl leads to an increase in the ionic strength what weakens the repulsive Coulomb forces (see eq. 7-9).

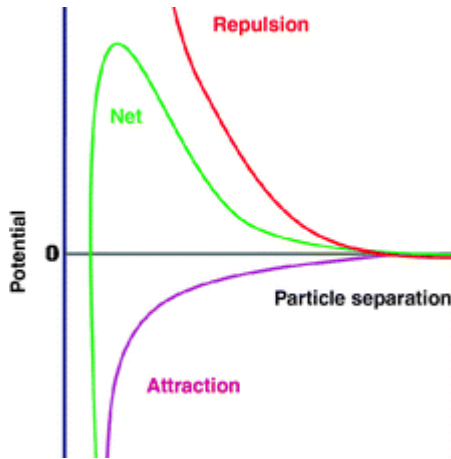


Figure 3: DLVO theory, net interaction potential between repulsive and attracting potentials [9]

2 Materials and Methods

2.1 Sample Preparation

First, the samples were prepared. In the following order, the buffer solution, the protein stock solution, and the sodium chloride stock solution were pipetted into Eppendorf tubes. In total, seven samples with varying volumes of buffer, protein stock and sodium chloride stock solutions were prepared. The composition for each of the seven samples can be seen in tab. 1. The samples were mixed using a Vortex mixer afterwards. As buffer, CH₃COOH (pH 5.3) was used. For the salt stock solution 2 M NaCl was used. Lysozyme is the protein used in this experiment.

Table 1: Composition of samples

Sample	Volume buffer [μ l]	Volume protein [μ l]	Volume salt [μ l]	Protein conc. [$\frac{mg}{ml}$]	Salt conc. [$\% \frac{w}{v}$]
1	101.7	40	58.3	40	3.4
2	93.3	40	66.7	40	3.9
3	85.0	40	75.0	40	4.4
4	91.7	25	83.3	25	4.9
5	46.7	70	83.3	70	4.9
6	38.3	70	91.7	70	5.4
7	30.0	70	100.0	70	5.8

While preparing the samples, glass slides were prepared on which each sample was pipetted afterwards. First, the glass slides were cleaned with ethanol. Then, three adhesive tapes were fixed on each glass slide next to each other. Into each of the tapes a hole was made, using a hole puncher before. In total, seven glass slides were prepared in the same way.

After that, droplets of 10 μ l each were placed into the holes of the tape ring on the glass slide. So, there were three approaches for each of the seven samples. Finally, a cover slide was placed on each tape ring with the sample in it.

2.2 Microscope Images

The images were taken under a light microscope and evaluated by using the software ZEN 2010. All images were observed with a magnification of 20x.

The two-step-crystallization was analysed in the samples 5 to 7. From each of these samples, images were taken immediately after preparing them. Then, images of the samples 1 to 3 were taken. These samples were for investigating the one-step crystallization. After that, sample 3a was chosen to be suitable doing an image series to observe crystal growth. The images were automatically taken for a time period of 30 min in an interval of 5 min each. So, in the end seven images for observing the crystal growth were available. Another series of images was tried to make, but it was unsuccessful due to either lacking or excessive crystal growth. Therefore, an already finished series of another group was used for further evaluation of the image series (sample 2c). Three hours after preparation, images of all samples (one- and two-step crystallization) were taken to compare the initial results with the progress in crystal growth.

3 Results and Discussion

3.1 Two-Step Crystallization

In the following, the two-step crystallization is investigated on sample 7a. The samples 5 and 6 could not be used for investigation because the two-step crystallization was less obvious in the respective images than for sample 7a. Fig. 4 depicts sample 7a right after its preparation. It can be observed that crystal formation is already in progress. Since the lysozyme concentration, as well as the concentration of the salt (that is used as depleting agent) are relatively high for sample 7a, this fast formation of the crystals seems to be logical. There are many crystals of a few μm in size and some, that have already grown to almost 500 μm in diameter. It can be seen that the crystals vary in their shape. Between the structures that can already be identified as crystals, there can be seen areas with small structures that cannot be identified as single crystals yet. These areas seem to be still cloudy. Here, also the LLPS can be observed.

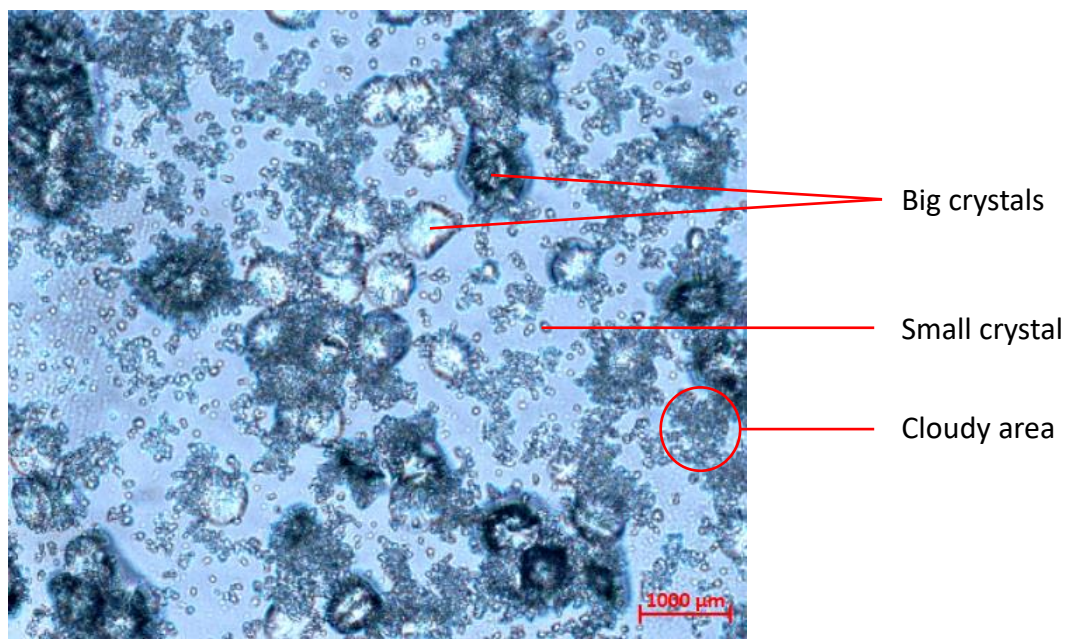


Figure 4: Sample 7a right after preparation of the sample

Fig. 5 depicts the sample about 3 h after its preparation, thus 3 h after the microscopy image, shown in fig. 4, was recorded. Compared to fig. 4, the crystals in fig. 5 seem significantly more structured. In addition, several crystals with a diameter of 500 μm or bigger can be observed, still varying in their shape. Many vanishingly small crystals seem to still have the potential to grow. Though, there seem to be less small crystals or even unstructured areas than right after the start of the crystallization, as it is depicted in fig. 4.

The crystallization process for sample 7a is a two-step crystallization due to the high concentration of lysozyme. It makes sense that 3 h after sample preparation the formation of the crystals is still in progress as the two-step process takes more time than the one-step process.

Although this crystal growth represents the two-step crystallization, the so-called “sea urchin” crystals that are typically growing during this type of crystallization cannot be detected in our experiment. This might be either due to the reason that 3 h of crystal growth are not enough time for growing of these structures, or to the reason that these structures do not always

appear, even during the two-step crystallization. The samples 5 and 6 do also not show “sea urchin” crystals.

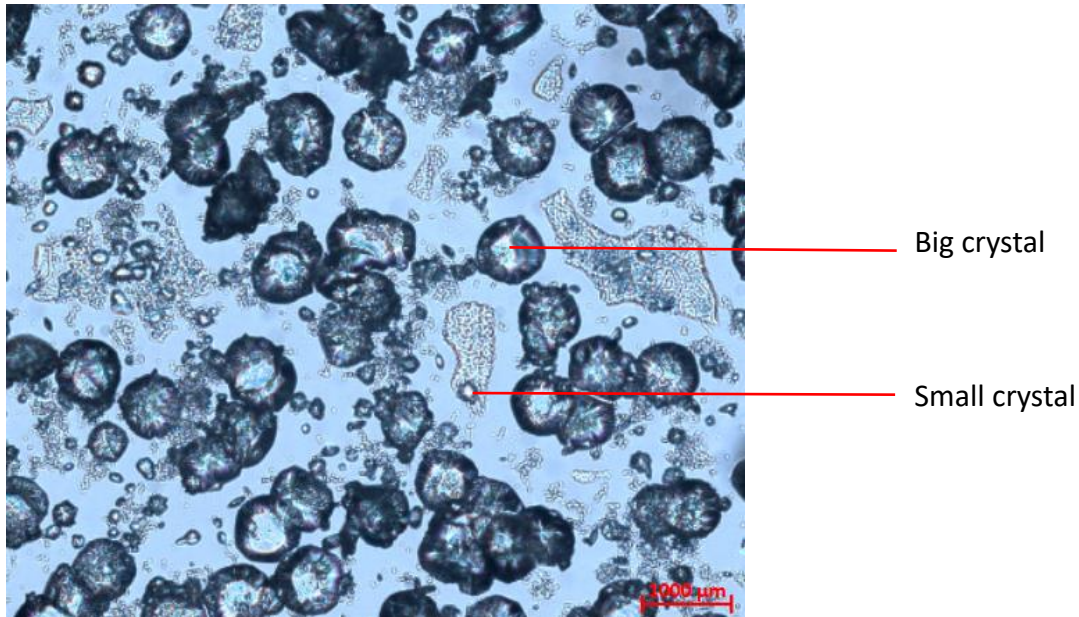


Figure 5: Sample 7a 3 h after preparation

3.2 One-Step Crystallization and Crystal Growth Process

The one-step crystallization was investigated for the samples 3a and 2c. To observe the crystal growth, the length and the width of one crystal for each sample were measured over time.

Fig. 6 shows crystals from the image series of 3a (left) and 2c (right) at $t = 20$ min each.

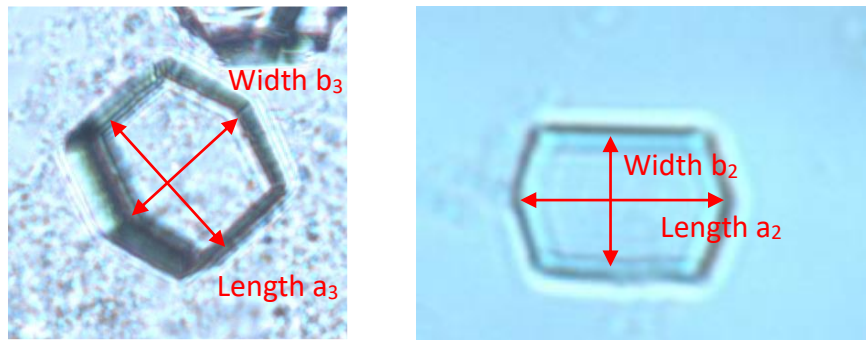


Figure 6: Measuring of length and width of a hexagonal crystal of sample 3a (left) and 2c (right)

The measured values for length and width of the two investigated crystals at the respective time points are shown in tab. 2.

Table 2: Measured values for length and width of the crystals over time

Sample	Parameter	0 min	5 min	10 min	15 min	20 min	25 min	30 min
3a	length [μm]	380	409	464	501	525	569	601
	width [μm]	308	359	408	447	496	529	565
2c	length [μm]	255	291	320	341	375	414	452
	width [μm]	169	192	212	234	250	301	338

The measurement values are plotted in dependency of the time (start of image series $t = 0$ min, end of image series $t = 30$ min) which is depicted in fig. 7 (length) and fig. 8 (width). A linear regression was performed with Excel. Measurement data of sample 3a are depicted in blue, whereas data for sample 2c are depicted in orange. Since the R^2 values (coefficients of determination of the linear regression) of all regression lines are close to 1, it can be deduced that the evaluated crystals grow almost linearly in length and width.

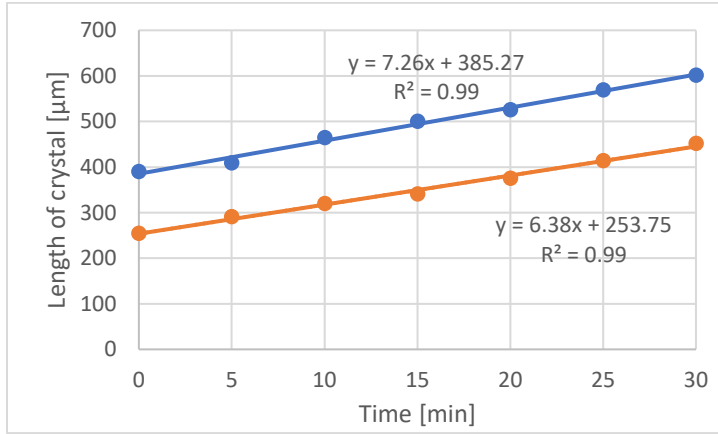


Figure 7: Growth of the length of the crystals from sample 3a (blue) and 2c (orange)

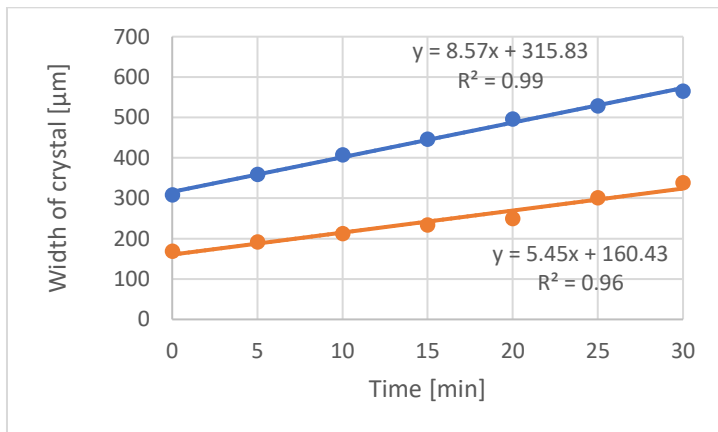


Figure 8: Growth of the width of the crystals from sample 3a (blue) and 2c (orange)

In tab. 3 the calculated growth rates for the length and width of each sample are described. The growth rates can be determined as the slopes of the linear regressions, whose formulas are depicted in Fig. 7 and 8. The parameter t_c describes the induction time (t_c (a) for length and t_c (b) for width) which provides information about the starting point of the crystal growth, and thus additionally about the time between sample preparation and starting crystal growth. The induction time can be calculated out of the respective regression formula. y must therefore be 0, because the crystal has a negligible expansion in length or width when it first appears. A representative calculation of t_c related to the length of sample 3a can be seen in the following:

$$0 = 7.26x + 385.27$$

$$x = \frac{-385.27}{7.26} = -53.0 = t_c$$

Table 3: Growth rate and t_c of crystals calculated from image series via Excel

Sample	$\frac{da}{dt} \left[\frac{\mu m}{min} \right]$	$\frac{db}{dt} \left[\frac{\mu m}{min} \right]$	t_c (a) [min]	t_c (b) [min]
3a	7.3	8.6	- 53.0	-36.9
2c	6.4	5.5	-39.8	-29.4

The time obtained by calculating t_c that way has to be subtracted from $t = 0$. It shows how many minutes before the first image of the series has been taken, the crystal has started to grow. When knowing the exact time of starting the image series (e.g. sample 3a $t = 0$ was taken at 2.15 μm) the exact time of the start of the crystal growing can be determined. So, for sample 3a, it is 1.22 μm . Since sample 3a was prepared at 1.17 μm , it can be deduced that the time until the crystal started to appear was 5 min. This can be seen in fig. 9.

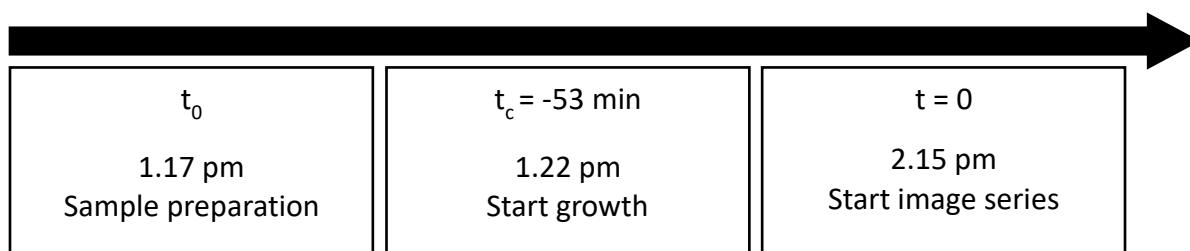


Figure 9: Timeline of sample preparation, start of growth and start of the image series of sample 3a (length)

As the image series of sample 2c was not our own, the exact points of time for preparation and start of the image series are unknown. Therefore, the induction time for length and width could only be determined relatively.

The period of time of 5 min (induction time) between t_0 and t_c , that sample 3a needs until the crystal starts to grow in length, is quite low. It can be assumed that for sample 2c the time span between t_0 and t_c would be higher by some minutes because the crystals needing more time to start growing due to less NaCl in the sample at the same protein concentration.

In tab. 3 it can be seen, that the growth rates for length and width of sample 3a are higher than for sample 2c. This is due to the higher concentration of the depleting agent NaCl in sample 3a that leads to a faster crystal growth as it is explained in section 1.2.2. With a higher concentration of NaCl, the potential $V_{EE}^{(r)}$ of repulsive forces is reduced, as the Debye screening length becomes higher with increasing ionic strength of the solution (see section 1.2.3).

Furthermore, it can be observed that the crystals in both samples start to grow in length first. This seems to make sense because all crystals showing higher values for the length a than for the width b , and the growth rates being not (significantly) higher for the width than for length. Thus, the crystal starts to grow in one direction (length) first, and after a few minutes it additionally starts to grow in width.

3.3 Crystal Aspect Ratio Analysis

The dependence of the average crystal length to width ratio on the salt concentration is analysed in the following for the one-step crystallization.

Fig. 10 shows the average aspect ratio of length a to width b in dependence of the salt concentration for sample 2 in orange and for sample 3 in blue. The average aspect ratio of

length to width and its standard deviation were calculated out of the length and width of seven crystals each. Sample 1 could not be used for this evaluation because no crystals had been appeared three hours after the preparation.

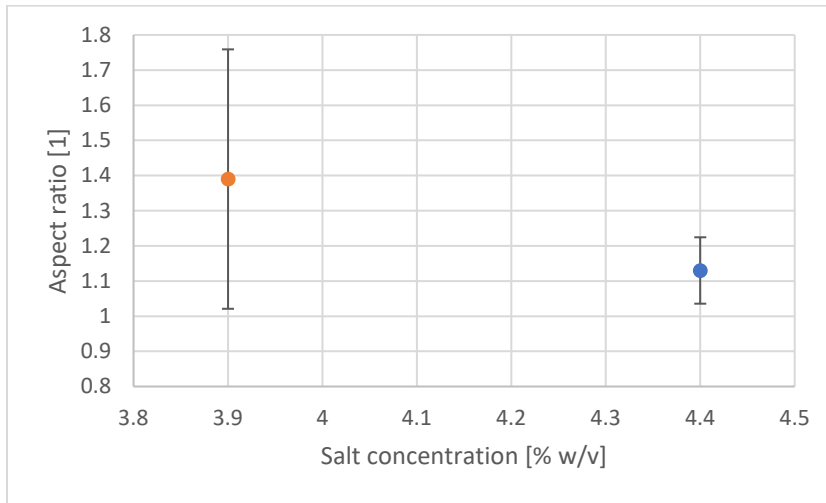


Figure 10: Average aspect ratio of length/width of crystals in dependence of salt concentration (sample 2 orange, sample 3 blue)

The average aspect ratio of length to width is lower for sample 3 than for sample 2. The standard deviation is also smaller. Due to the higher salt concentration in sample 3 compared to sample 2, it can be assumed that the crystals grow more homogeneously in length and width. The salt, that acts as depleting agent, neutralizes the initially charged protein and increases therefore the Debye screening length. Therefore, the repulsive interaction potential $V_{EE}^{(r)}$ between two protein molecules is depleted. Thus, less energy is needed for a new growth unit to assemble at any place on the crystal and not just in length because of less repulsive interaction.

Table 4: Data of diagram (fig. 10)

Sample	Average aspect ratio [1]	Standard deviation [1]	Salt concentration [% w/v]
2	1.39	0.37	3.9
3	1.13	0.09	4.4

4 Conclusion

Proteins play a key role during many processes in the human body. Since their structure is closely connected to their function, it is important being able to identify these exact structures. *Via* protein crystallization it is possible to observe the structure of a protein by X-ray analysis.

Therefore, in the present experiment the influence of protein and salt concentration on the crystallization process was investigated. The protein used in this experiment was lysozyme. It could be shown that a high concentration of lysozyme and NaCl leads to two-step crystallization. In this two-step process LLPS occurred, and the crystals grew fast and in large numbers. When using less lysozyme, one step crystallization could be observed. In that case, a higher salt concentration led to the crystals growing more homogeneously in length and width. This could be explained by the function of salt as a depleting agent, which reduces the repulsive forces between the respective protein molecules. It could also be observed that in one step crystallization, less crystals appeared compared to the two-step crystallization.

5 Literature

1. Hashimoto, M., et al., *Role of protein aggregation in mitochondrial dysfunction and neurodegeneration in Alzheimer's and Parkinson's diseases*. NeuroMolecular Medicine, 2003. **4**(1): p. 21-35.
2. Callewaert, L. and C.W. Michiels, *Lysozymes in the animal kingdom*. Journal of Biosciences, 2010. **35**(1): p. 127-160.
3. Papageorgiou, A.C. and J. Mattsson, *Protein structure validation and analysis with X-ray crystallography*. Methods Mol Biol, 2014. **1129**: p. 397-421.
4. Erdemir, D., A.Y. Lee, and A.S. Myerson, *Nucleation of Crystals from Solution: Classical and Two-Step Models*. Accounts of Chemical Research, 2009. **42**(5): p. 621-629.
5. Manuel García-Ruiz, J., *Nucleation of protein crystals*. Journal of Structural Biology, 2003. **142**(1): p. 22-31.
6. Muschol, M. and F. Rosenberger, *Liquid-liquid phase separation in supersaturated lysozyme solutions and associated precipitate formation/crystallization*. The Journal of Chemical Physics, 1997. **107**(6): p. 1953-1962.
7. Lauth, G.J. and J. Kowalczyk, *Stabilität von Suspensionen*, in *Einführung in die Physik und Chemie der Grenzflächen und Kolloide*, G.J. Lauth and J. Kowalczyk, Editors. 2016, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 285-330.
8. Gunton, J.D., A. Shiryayev, and D.L. Pagan, *Protein Condensation: Kinetic Pathways to Crystallization and Disease*. 2007: Cambridge University Press.
9. Pujala, R.K., *Introduction*, in *Dispersion Stability, Microstructure and Phase Transition of Anisotropic Nanodiscs*, R.K. Pujala, Editor. 2014, Springer International Publishing: Cham. p. 1-16.