Optimization of crystal growth of membrane proteins for advanced diffraction techniques

Elham Vahdatahar

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : Physique pour les Sciences du Vivant

Arrêté ministériel : 25 mai 2016

Présentée par

Elham VAHDATAHAR

Thèse dirigée par **Monika SPANO**,Maître de Conférence, Université Grenoble Alpes

préparée au sein du **Laboratoire Institut de Biologie Structurale** dans **l'École Doctorale Physique**

Optimisation de la croissance cristalline des protéines membranaires pour les techniques de diffraction avancées

Optimization of crystal growth of membrane proteins for advanced diffraction techniques

Thèse soutenue publiquement le **28 juin 2021**, devant le jury composé de :

Monsieur CLAUDE SAUTER
DIRECTEUR DE RECHERCHE, CNRS DELEGATION ALSACE, Rapporteur
Madame FRANCOISE BONNETE
CHARGE DE RECHERCHE HDR, CNRS DELEGATION PARIS CENTRE, Rapporteure
Monsieur WIM BURMEISTER
PROFESSEUR DES UNIVERSITES, UNIVERSITE GRENOBLE ALPES, Président
Monsieur ESKO OKSANEN
DOCTEUR EN SCIENCES, European Spallation Source, Examinateur
Madame CATHERINE VENIEN-BRYAN
PROFESSEUR DES UNIVERSITES, SORBONNE UNIVERSITE, Examinatrice

Acknowledgments

This thesis would have never been finished without support, friendship and advice from a number of people, who I would like to take this opportunity to thank.

First and foremost, I would like to thank Dr. Monika Budayova Spano for supervising this project and teaching and supporting me through these years. It was a great opportunity to work with you and learn from you. Thank you for giving me the opportunity to be a member of the RAMP ITN network and your workgroup. I also want to thank you for entrusting me with various assistantships in several workshops in Grenoble.

I would like to thank the jury members of this thesis, Claude Sauter, Françoise Bonneté, Wim Burmeister, Esko Oksanen, and Catherine Vénien-Bryan who took their precious time to evaluate this work and give me their precious advice.

Many thanks go to all colleagues at the GSY and metalloprotein groups at the Institut de Biologie Structurale, Grenoble who created such an outstanding working atmosphere. In particular, I want to mention Dr. Jean-Luc Ferrer for all his advice and support during my PhD. Dr. David Cobessi with whom I worked closely together during the first year of my Ph.D. studies and from whom I learned how to purify the membrane proteins, which I later used for crystallization, and crystallographic data analysis. Dr. Frank Borrel, with whom I learned how to collect X-ray data in the synchrotron and who patiently answered all of my questions and Antoine Royant, who supported all the GSY members especially during the covid crisis.

I wish also to thank Lydie Martin from whom I learned a lot in the biochemistry lab, Wai-Li Ling for her help with microED studies, and Eva Pebay-Peyroula for kindly giving us the genetic material for AcrB.

I am grateful to Leon Jenner for helping me in the biochemistry lab and for all his help especially during the last year of my Ph.D. I also would like to thank all my friends at IBS. I owe special thanks to Sofia Jaho, Niels Junius, Rémi Ruedas, Roman, Hamida, Haiyan, Shaghayegh, Tu Quynh, and Soumiya who helped me and supported me during my Ph.D.

Thank you to the members of the metalloprotein group Patricia, Juan, Yvain, Anne, and Eve for their support and kindness. My thanks also go to Aline Schwoob, Marta Elzo-Aizarna, and Caroline Barathon, for their help with administrative procedures for the university and RAMP network.

I want to thank also GSY members at ESRF, Yoann Sallaz-Damaz, Eric Rive-Mathieu, Christophe Berzin, and Philippe Jacquet for all their technical help and good memories. Special thanks go to all people from ESRF French CRG for a friendly environment that made me feel at home and for all their help and support during my stay at ESRF. It was really great to meet you all.

I had four secondments during my PhD and I want to take this opportunity to thank the supervisors and students who introduced me to new methods and helped me during my stay in their laboratories. First I want to say thanks to Quentin Mortier, Christophe Coué, Melodie Bedu

for the great productive month at NatXray. I would like to thank Dr. Poul Nissen, Sofia Trampari, Joseph Lyons, and Carolina Neumann for helping me and teaching me the HiLiDe method. I want to say thanks to Dr. Jennifer McManus, Thi Thanh My Nguyen, Alessandro Strofaldi and all of the other lab members for their help, guidance, and teaching me light scattering techniques (DLS and SLS) and also for making my third secondment enjoyable. I'd like to say a big thank you to Dr. Richard Sear and Virginia Apostolopoulou for helping me during my secondment at Surrey university. I also want to thank Dr. Arwen Pearson and Diogo Melo for our collaboration for detergent studies and the good and productive discussions we had during the meetings. It was really good to work with you.

I also would like to say thank you to the European Union for funding this project under Marie Skłodowska-Curie Actions H2020-MSCA-ITN 2017-2021.

Last but certainly not least, I would like to thank my family and friends especially my brother who helped me and supported me all these years. Thank you very much Amin, I am really grateful to have a brother like you. I also want to thank Maryam, Sana and Rashmi for all the good times we shared.

Abstract

Membrane proteins account for around 30% of the proteins in eukaryotic cells. These proteins are potential targets for drug design due to their vital biological functions including their involvement in signaling, transport or reception mechanisms. Investigating the 3D structure of proteins is essential to understand their function. Macromolecular X-ray crystallography is the most widely used technique to determine the structure of proteins. However, obtaining high diffraction quality crystals of membrane proteins, whilst maintaining their native conformation during the production, purification and crystallization steps, is challenging. Screening and optimization of crystallization conditions as well as finding optimal methods for growing high-quality crystals with appropriate size distribution is another bottleneck in modern structural studies of membrane proteins. Neutron protein crystallography requires crystals with a volume larger than 0.1 mm^3, whereas micro- or nano-sized crystals are sufficient for Microcrystal Electron Diffraction (MicroED).

Knowledge of the protein crystallization phase diagram allows the quality of the crystals to be controlled and subsequently optimized. Thanks to these tools, it is possible to adjust crucial physico-chemical parameters such as the temperature or the chemical composition of the crystallization solutions to optimize the crystallization process. Moreover, the automated control of these parameters has been achieved using a recently designed crystallization bench developed in our laboratory. Automated dialysis-based crystallization with variation of the temperature and composition of the crystallization solution by this instrument allows to screen different crystallization conditions using a single sample as well as to optimize the crystal growth in a reversible manner. Therefore, crystallization optimization and phase diagram exploration are possible using a small amount of protein. The crystallization of model soluble and membrane protein has been demonstrated using this instrument to obtain crystals suitable for downstream structural studies by crystallography.

This rational approach to crystallization was applied to the membrane proteins Acriflavine resistance protein B (AcrB) from *Escherichia coli* and TonB-dependent heme outer membrane transporter (ShuA) from *Shigella dysenteriae.* The purification of these proteins for crystallization was optimized. The purified proteins were then crystallized using the different crystallization techniques, identifying the dialysis method as the most successful and consistent although comparable crystals were obtained with almost all techniques, but not always reproducible. AcrB crystals co-crystallized with rifampicin were used for systematic X-ray diffraction studies and the best statistics for the AcrB-rifampicin crystal from each method were compared. By co-crystallization of ShuA with europium (III) chloride, the 3D structure of this protein was solved at 2.49 Å. The results also showed that the purity of the protein sample and well-chosen additives improve the crystal quality.

MicroED is an ideal method for studying membrane proteins and protein complexes that do not readily form crystals large enough for other diffraction techniques. The Coulomb potential maps

determined by MicroED also provide information on the electrostatic environment and charge states of amino acids, which can complement the structural information obtained by this and other techniques. Preliminary studies applying this approach have been performed for our two model membrane proteins co-crystallized with additives.

For further characterization of the crystallization process using the dialysis method, the mass transport of detergents across semi-permeable membranes was studied. This information, combined with previously determined knowledge of the diffusion of salts and polymeric precipitants across semi-permeable membranes, can be used to establish effective strategies for optimizing the crystallization of membrane proteins using this technique with automated control.

Résumé

Les protéines membranaires représentent environ 30 % des protéines des cellules eucaryotes. Ces protéines sont des cibles potentielles pour la conception des médicaments en raison de leurs fonctions biologiques vitales, notamment leur implication dans les mécanismes de signalisation, de transport ou de réception. L'étude de la structure 3D des protéines est essentielle pour comprendre leur fonction. La cristallographie macromoléculaire aux rayons X est la technique la plus largement utilisée pour déterminer la structure des protéines. Cependant, obtenir des cristaux de protéines membranaires de haute qualité de diffraction, tout en maintenant leur conformation native pendant les étapes de production, de purification et de cristallisation, est un défi. Le criblage et l'optimisation des conditions de cristallisation ainsi que la recherche de méthodes optimales pour la croissance des cristaux de haute qualité, avec une distribution de taille appropriée, constituent un autre goulot d'étranglement dans les études structurales modernes des protéines membranaires. La cristallographie neutronique des protéines nécessite des cristaux d'un volume supérieur à 0,1 mm^3, alors que des cristaux de taille micro- ou nanométrique sont suffisants pour la diffraction électronique des microcristaux (MicroED).

La connaissance du diagramme de phase de la cristallisation des protéines permet de contrôler la qualité des cristaux, puis de les optimiser. Grâce à ces outils, il est possible d'ajuster des paramètres physico-chimiques cruciaux tels que la température ou la composition chimique des solutions de cristallisation pour optimiser le processus de cristallisation. De plus, le contrôle automatisé de ces paramètres a été réalisé grâce à un banc de cristallisation récemment conçu dans notre laboratoire. La cristallisation automatisée par dialyse avec variation de la température et de la composition de la solution de cristallisation par cet instrument permet de cribler différentes conditions de cristallisation en utilisant un seul échantillon ainsi que d'optimiser la croissance des cristaux de manière réversible. Par conséquent, l'optimisation de la cristallisation et l'exploration du diagramme de phase sont possibles en utilisant une petite quantité de protéine. La cristallisation d'un modèle de protéine soluble et membranaire a été démontrée à l'aide de cet instrument afin d'obtenir des cristaux adaptés aux études structurales avancées par cristallographie.

Cette approche rationnelle de la cristallisation a été appliquée aux protéines membranaires, la protéine de résistance à l'acriflavine B (AcrB) d'*Escherichia coli* et le transporteur d'hème de membrane externe TonB dépendant (ShuA) de *Shigella dysenteriae*. La purification de ces protéines pour la cristallisation a été optimisée. Les protéines purifiées ont ensuite été cristallisées à l'aide des différentes techniques de cristallisation. La méthode de dialyse a été identifiée comme la plus réussie et la plus cohérente, bien que des cristaux comparables aient été obtenus avec presque toutes les techniques, mais pas toujours reproductibles. Des cristaux d'AcrB co-cristallisés avec de la rifampicine ont été utilisés pour des études systématiques de diffraction des rayons X et les meilleures statistiques pour le cristal d'AcrB-rifampicine de chaque méthode ont été comparées. Par la co-cristallisation de ShuA avec du chlorure d'europium (III), la structure 3D de cette protéine a été résolue à 2,49 Å. Les résultats ont également montré que

la pureté de l'échantillon de protéine et des additifs bien choisis améliorent la qualité des cristaux.

La MicroED est une méthode idéale pour étudier les protéines membranaires et les complexes protéiques qui ne forment pas facilement des cristaux suffisamment grands pour les autres techniques de diffraction. Les cartes de potentiel de Coulomb déterminées par MicroED fournissent également des informations sur l'environnement électrostatique et les états de charge des acides aminés, qui peuvent compléter les informations structurales obtenues par cette technique et d'autres. Des études préliminaires appliquant cette approche ont été réalisées pour nos deux protéines membranaires modèles co-cristallisées avec des additifs.

Pour caractériser davantage le processus de cristallisation par la méthode de dialyse, le transport de masse des détergents à travers les membranes semi-perméables a été étudié. Ces informations, combinées aux connaissances précédemment déterminées sur la diffusion des sels et des polymères à travers les membranes semi-perméables, peuvent être utilisées pour établir des stratégies efficaces pour optimiser la cristallisation des protéines membranaires en utilisant cette technique avec un contrôle automatisé.

Preface

Understanding proteins' functions is important in order to understand the biochemical processes in living cells. Such information can be used for different pharmaceutical and biotechnological studies, such as drug design or protein engineering. Since there is a correlation between the function and structure of a protein, studying protein structure contributes greatly to understanding protein function. There are several techniques to study protein structure but crystallography is used the most. Depending on the method used for crystallography (X-ray, neutron, and electron diffraction), protein crystals with different sizes are needed. As a result optimization of crystallization is one of the most important steps in structural studies using crystallography. This thesis is the result of a nearly four-year research project, undertaken as part of the EU Rationalizing Membrane Protein crystallization (RAMP) Innovative Training Network (ITN). RAMP is a consortium of academic institutions and industrial partners focused on advancing the field of membrane protein crystallization. Within the context of RAMP, this specific project aimed to optimize the crystallization of two model membrane proteins in order to obtain high-quality crystals with suitable sizes for different crystallization techniques. One of these proteins is the TonB-dependent heme outer membrane transporter (ShuA) of the gram-negative bacteria *Shigella dysenteriae*. The second protein is Acriflavine resistance protein B (AcrB), which is embedded in the inner membrane of *Escherichia coli*, a gram-negative bacteria. This research project has received funding from the European Union's Horizon 2020 Research and Innovation program under the Marie Skłodowska-Curie grant agreement no 722687. This PhD was completed under the supervision of Monika Budayova Spano at the Institut de Biologie Structurale, Université Grenoble-Alpes, France.

In order to apply our approach, and as part of the RAMP ITN, substantial time was spent away on multiple secondments at several institutions. The first involved a one-month stay at NatXray, one of the industrial partners in the RAMP network, to participate in the industrialization of crystallization equipment with the engineers at this company. The second secondment involved a one-month stay at Aarhus University, where the major aim was learning the HiLiDe crystallization method to study the structure of our two model membrane proteins. The third secondment involved a two month stay at Maynooth university to learn and use light scattering techniques (SLS and DLS) to study the protein-protein interaction and to characterize protein solutions to the prior crystallization process. The last secondment was a six week stay at the University of Surrey to learn how to make a theoretical model of mass transport of detergent through the semipermeable membrane based on experimental data.

This thesis contains 5 chapters. It starts with an introduction followed by purification and optimization of this process for the two model membrane proteins. Then results from the crystallization of model soluble and membrane proteins are presented. Finally, the quality of the crystals studied using crystallographic techniques is discussed. In chapter 5, all findings are summarized and conclusions are made, followed by the discussion of recommendations for future works.

In the first chapter, model soluble and membrane proteins that are used in this study are introduced. Then protein crystallization and variables that affect this process are briefly discussed. Followed by the introduction of conventional and non-conventional crystallization techniques that are employed in this study. After introducing the concept of the protein crystallization phase diagram and understanding how to use this knowledge to grow high-quality crystals for different crystallography techniques, these crystallography techniques are briefly introduced. Then detergents and the challenges of using them in purification and crystallization of membrane proteins are introduced.

The focus of the second chapter is on the purification of the two model membrane proteins, ShuA and AcrB. Protocols for purification and an optimization of this process are presented in this chapter. Specifically, one and two-step purification methods with affinity chromatography in the presence and absence of Anion exchange chromatography or size exclusion chromatography are detailed.

In chapter three, initial crystallization experiments with model soluble (Lysozyme and thaumatin) and membrane proteins (AcrB and ShuA) with conventional and non-conventional techniques are shown. Then the optimization of crystal growth for different crystallographic techniques (X-ray, neutron, and electron diffraction) is presented. In the second part of this chapter, the effect of different variables such as temperature, pH, salt, and protein concentrations on protein crystallization are demonstrated. Moreover, the effect of varying these parameters is illustrated and rationalized via the kinetic path taken in the schematic phase diagram to reach the nucleation zone for each crystallization method. Also after introducing the two crystallization benches that were designed in our laboratory previously, this chapter shows how by controlling the temperature and exchanging crystallization solutions in an automatic and semi-automatic manner it is possible to control the size of the crystals and optimize the crystallization process reversibly. Finally, the mass transport of several detergents through dialysis membranes is studied. This study is important for the optimization of membrane protein crystallization as detergents are nearly always used to preserve membrane protein structure during the purification and crystallization process.

After optimizing the sizes of crystals for different crystallography techniques, the quality of these crystals is compared in chapter four. For that, first AcrB crystals obtained from one and two-step purifications and hanging drop vapor diffusion are compared. Then AcrB crystals from hanging drop vapor diffusion, dialysis, and batch methods are compared using X-ray crystallography. Following this, the 3D structure of ShuA obtained using molecular replacement is reported. Comparative preliminary results from electron diffraction studies of ShuA and AcrB are then presented.

Finally in chapter 5 summary and the conclusion of this study are discussed along with some suggestions for future work

LIST OF ABBREVIATIONS:

CMC: Critical micelle concentration

CYMAL6: 6-Cyclohexyl-hexyl-beta-D-maltoside

DDM: n-Dodecyl-β-D-Maltopyranoside

DMSO: Dimethyl sulfoxide

DOPC: 1,2-dioleoyl-sn-glycero-3-phospha-tidylcholine

EDTA: Ethylenediaminetetraacetic acid

EGTA: Ethylene glycol tetraacetic acid

HEWL: Hen egg white lysozyme

HiLiDe: High concentration of Lipid and Detergent

IPTG : Isopropyl-β-D-1-thiogalacto pyranoside

kDa: kilo Dalton

MAD: Multi-wavelength anomalous diffraction

MAGs: Monoacylglycerols

MIR: Multiple isomorphous replacement

MPD: 2-Methyl-2,4-pentanediol

NIP: Neutron–sensitive image-plate

OG: Octyl glycoside

OPOE: n-Octylpolyoxyethylene

PDB: Protein Data Bank

PDC: protein detergent complex

PEG: Polyethylene glycol

PEG MME: Polyethylene glycol monomethyl ether

PSDs: Position-sensitive detectors

SAD: Single-wavelength anomalous diffraction

SDS: Sodium dodecyl sulfate

SDS-PAGE: Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis

SIR: Single isomorphous replacement

TEM: Transmission electron microscopy

XDS: X-ray detector software

Table of contents:

Table of figures:

List of tables:

Chapter 1: Introduction

1.1 Introduction

Knowledge of the three dimensional structure of a protein molecule is arguably one of the most powerful sources of information in understanding the function of that protein. Protein structures contribute greatly to the understanding of biochemical processes in living cells and are therefore important for medical and biotechnological studies. This information can be used by the pharmaceutical, biotechnological and chemical industries for protein engineering, drug design, and other applications [1].

Various methods are used to study protein structure, the most prevalent of which are crystallography (X-ray, neutron and electron), nuclear magnetic resonance (NMR) spectroscopy, and electron microscopy [1], [2]. This thesis focuses on rational methods which can be used to make the crystallization of proteins (especially membrane proteins) more predictable and reproducible for structural studies by crystallography. When successful, crystallographic techniques can provide atomic-level information about the structure of proteins or other biomacromolecules at the atomic level.

For biomacromolecular crystallography, protein crystals of sufficient sizes and qualities are needed. To date obtaining crystals with these features is largely based on trial-and-error screening methods [3]. For a systematic approach to the crystallization process (including nucleation and growth [1]), both crystallization methods and experimental variables (physical, chemical and biochemical) should be considered [4].

Membrane proteins are encoded by approximately 20 to 30% of the human genome and are involved in a diverse range of functions [5], [6]. Misfolding or mutation of these proteins is implicated in various diseases such as cancer, obesity, depression, Parkinson's and Alzheimer's diseases [7]. In their native environment, these proteins partially or fully exist within the phospholipid membranes of cells. The stability of any protein is often a precarious balance of different factors [8]. As such, membrane proteins are seldom stable in aqueous solution alone and crystallization of membrane proteins is therefore highly challenging as they readily precipitate outside of their native environment. Illustrating this, the first protein was crystallized in 1840 but the first membrane protein only in 1985 [9]. To preserve native membrane protein structure during purification and crystallization, surfactants such as detergents are used, increasing the complexity of crystallization.

This chapter introduces the model soluble and membrane proteins used for this study, the theoretical and practical aspects of protein crystallization, the various techniques employed in crystallization experiments and the main experimental variables to be considered for these. The concept of the protein crystallization phase diagram is introduced, along with how knowledge of this can be used in order to grow high-quality crystals with different sizes suitable for different crystallographic techniques. Detergents are discussed and in particular the challenges involved in using them for purification and crystallization of membrane proteins. Finally, the chapter

concludes with the objective of this study, namely the optimization of crystal growth for different advanced diffraction techniques.

1.1.1 Proteins

For this work, the soluble proteins lysozyme and thaumatin and membrane proteins *Shigella dysentriae* TonB-dependent heme outer membrane transporter (ShuA) and *Escherichia coli* acriflavine resistance protein B (AcrB) were used.

1.1.1.1 Model soluble proteins

Well-characterized model proteins whose crystallization conditions are already known can be used in the first step of exploring, characterizing and controlling the crystallization process [10].Lysozyme is the most widely used model protein that is commercially available in large quantities at an affordable price [10]. Another model protein used in our study is Thaumatin.

1.1.1.1.1 Lysozyme

Lysozyme is an enzyme that breaks glycosidic bonds [11]. It hydrolyses the β (1, 4) glycosidic bond between N- acetylmuramic acid and N-acetylglucosamine of the peptidoglycan in cell walls of bacteria [12], [13]. Because of this muramidase activity it is also called N-acetylmuramide glycanhydrolase [12].

There are three categories of the lysozymes in the animal kingdom, c-type (for the conventional or chicken type), g-type (for goose type) and i-type (for invertebrate type) [13]. In this study hen-egg white lysozyme (HEWL), which is of c-type, was used. Figure 1 depicts the structure of the this protein.

HEWL has a molecular mass of 14.6 kDa, has 129 amino acid residues and four disulfide bridges [14]. This basic protein has a high isoelectric point (pI ≈10.5) [15]. The three-dimensional structure of this protein consists of two domains (one domain mainly with β-sheet structures and other more helical) and its active site is located in a deep cleft [13].

1.1.1.1.2 Thaumatin

Thaumatin (see Figure 1) is a globular protein extracted from the fruit of a plant called *Thaumatococcus daniellii* that belongs to the Marantaceae family [16]. This protein is used as a sweetener in the food industry. The molecular mass of this single-chain protein is 22 kDa [16]. It has 207 amino acid residues, eight disulfide bridges and a high isoelectric point (pI ≥ 11) [10]. Thaumatin I and II are the two main forms of this protein [8] that only differ by five amino acids

[10]. This protein is widely used as a model protein because crystals of thaumatin can be rapidly formed with the addition of L-tartrate ions to the protein solution [10]. This has enabled previous experiments to examine the crystallization process under a variety of conditions such as crystallization in gel or under pressure. Additionally, thaumatin has been used as a standard for optimization of diffraction data collection and to check for radiation damage in crystals [10].

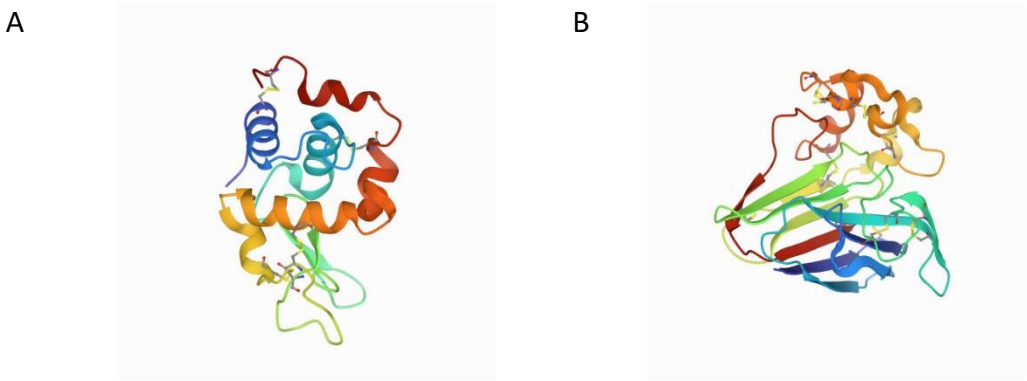


Figure 1. Structures of A) hen egg white lysozyme at 1.33 Å. PDB code: 193L and B) Thaumatin at 1.05 Å. PDB code: 1RQW [17].

1.1.1.2 Model membrane proteins

Membrane proteins are coded by around one-third of the human genome. The two main categories of membrane proteins are transmembrane or integral and peripheral, assigned according to how they associate with the lipid bilayer [7], [18]. Peripheral proteins interact with the surface of the membrane by non-covalent interactions [7]. Thus, when the ionic strength and pH of these proteins changes they can dissociate from the membrane [7]. Integral proteins, bound by hydrophobic forces partly or fully within the membrane, are further sub-divided according to how they are located in the membrane. Three examples of integral membrane proteins are illustrated in Figure 2.

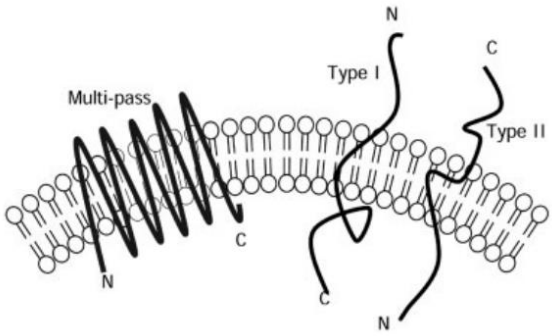


Figure 2. Different types of integral membrane proteins [18].

Type I (monotopic) integral membrane proteins have their N-terminus located outside the membrane and their C terminus inside. Type II (monotopic) is the opposite with their N-terminus

inside and their C-terminus outside. Multi-pass (or polytopic) integral membrane proteins instead cross the membrane multiple times. Multi-pass proteins can also be structurally divided into α-helical bundles, which are found in cytoplasmic and subcellular compartment membranes and β-barrels, which are found in the outer membranes of Gram-negative bacteria, mitochondria and chloroplasts [7]. In contrast, anchored proteins are integrated into one side of the membrane (single-pass α-helices and integral monotopic membrane proteins are exceptions in this group of membrane proteins) [7]. Consequently, to dissociate the integral membrane proteins from the membrane for crystallization studies use of a surfactant (e.g. detergents) is necessary [7].

Understanding the three-dimensional structure of membrane proteins is one of the hot topics in structural biology due to the important biological roles they perform. The first membrane protein structure solved in 1985 was the 3 Å structure of the photosynthetic reaction centre of *Rhodopseumonas viridis* [19]. In addition to photosynthetic and respiratory proteins, many membrane proteins have roles in transporting ions and metabolites across membranes or acting as signal receptors [20]. As such, around 50% of drug targets in humans are membrane proteins [6] underscoring the importance of structural studies of them.

In this study, two model membrane proteins, ShuA from *Shigella dysenteriae* and AcrB from *Escherichia coli* are employed. Although these two proteins are both membrane proteins, they have completely different structural features as might be expected from this diverse category of proteins. ShuA is a beta-barrel protein located in the outer membrane of gram-negative bacteria, whereas AcrB contains 12 transmembrane α- helices and is located in the inner membrane of gram-negative bacteria. In the next two sections, each model membrane protein is briefly discussed.

1.1.1.2.1 AcrB

Bacteria and other cells use four different strategies for antibiotic drug resistance: mutation of a target enzyme of a specific antibiotic, degradation of an antibiotic by other enzymes, exclusion of antibiotics from the cell by modifying different membrane transporters and pores, or exportation of an antibiotic from the cell. There are different methods to export antibiotics and chemicals outside the cell. These processes are called multidrug resistance (MDR). There are five major groups of MDR transporters found in all kingdoms of life [21]:

1. ATP binding cassette or ABC transporters: the export of the substrate is coupled with the hydrolysis of adenosine triphosphate (ATP) [21], [22]

2. Major facilitator superfamily (MFS) family [21], [22]: export of substrate is coupled to the movement of protons across the membrane.

3. Small multidrug resistance (SMR) family [21], [22]

4. Resistance nodulation division (RND) family transporters [21], [22]

5. Multi antimicrobial extrusion (MATE) family: substrate export is coupled with a sodium or proton gradient across the membrane [21], [22]

One of the proteins in RND family transporters is the acriflavine resistance protein B or AcrB, each subdomain of this protein is 114 kDa and contains 1049 amino acid residues [22], [23]. This protein is a homotrimer and each protomer contains periplasmic and membrane parts. The periplasmic part contains the TolC docking domain and pore domain. The membrane part contains 12 transmembrane α- helices [23], [24]. As illustrated in Figure 3, the TolC docking domain contains parts denoted as DN and DC and the Pore domain contains PN1, PN2, PN3, and PN4. Each part plays a role in the overall structure of the protein [25].

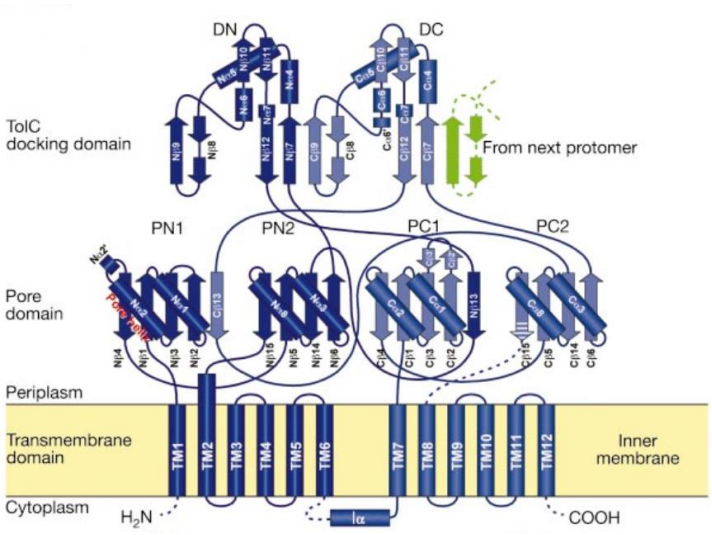


Figure 3. Structure of the AcrB protomer [25]. TolC docking domain and Pore domain are located in the periplasm and the Transmembrane domain is located in the inner membrane.

At the junction of the protomers outside the inner membrane (shown with a star in Figure 4), three vestibules (open areas) are seen. Substrates are believed to be captured through these vestibules and catalytic outflow then removes them from the periplasmic area [23].

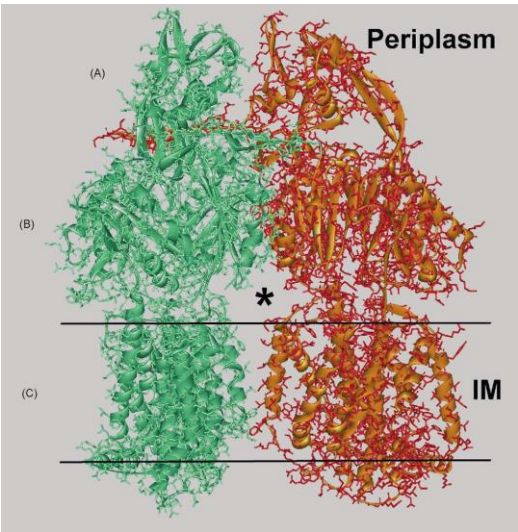


Figure 4. Sectional image of two protomers (green and orange) of AcrB and a vestibule (*) at the junction of these protomers [23]. The third protomer has been removed for better visualization.

AcrB, which has a wide variety of substrates among multidrug resistance pumps [24], is part of the AcrAB-TolC complex [23]. This complex transfers detergents, antibiotics, dye, disinfectants, and simple solvents out of the *E. coli* cell [24].

1.1.1.2.2 ShuA

Iron is essential for all living organisms [26]. The human body uses the proteins hemoglobin, ferritin, transferrin and lactoferrin for iron transport and storage [26]. Two thirds of all iron within the body is incorporated in the heme cofactors of hemoglobin [26]. After hemoglobin, ferritin is used to store iron inside cells and regulate intracellular iron levels [26]. Iron in serum can also bind to transferrin [26]. In addition, lactoferrin binds to free iron. Bacteria like other organisms need iron and they use different mechanisms to get it [26]. They can secrete small molecules called siderophores that have a high affinity to iron and scavenge free iron from the surrounding environment. In pathogenic or commensal bacteria these can compete with transferrin from host cells to bind iron [26]. Siderophores can also extract iron directly from that already bound to lactoferrin, transferrin or hemoglobin [26].

Gram-negative bacteria like *Escherichia coli* or *Shigella dysenteriae* have inner and outer membranes. Iron transfer through these membranes requires specific import mechanisms. Across the outer membrane, the TonB system is used. For the inner membrane, a heme-binding protein, a channel and an ATPase are needed [26]. TonB-dependent transporters (TBDTs) have a high affinity for siderophores and interact with inner membrane protein complex (TonB, ExbB, and ExbD) to acquire the energy needed to transfer heme-bound iron into the cell. TBDTs have a 22-stranded β-barrel and a plug domain [27]. At the extracellular side of the membrane the plug domain binds to the siderophores and at the periplasmic side, it interacts with TonB of the TonB-

ExbB- ExbD protein complex [27]. Therefore, to combat bacteria like *Shigella dysenteriae* one approach is to target the ability of these bacteria to uptake heme, which is essential for their survival [28]. ShuA is a TBDT in *Shigella dysenteriae*, which interacts directly with hemoglobin. Figure 5 shows the structure of ShuA. This protein has 22 antiparallel β-strands that are connected by eleven turns on the periplasmic side and eleven loops on the extracellular side. The N-terminus of the protein creates a plug domain [28]. The molecular weight of this protein is approximately 70 kDa and it contains 640 amino acid residues. For hemoglobin recognition and heme extraction, two conserved residues, His 86 (in the plug domain) and His 420 (in loop 7) (see Figure 5) [28]–[30] and extracellular loops are required [28]. It seems that transient interactions of short regions of these loops with hemoglobin lead to the correct orientation of this protein on the pore opening region, so heme can be extracted by the His residues [28]. The mechanism of interaction between heme and the extracellular loops is still not fully understood and more information is needed. Using neutron crystallography to reveal the position of H atoms (see 1.1.5.2) [31], especially those involved in ligand binding, could be useful for this purpose.

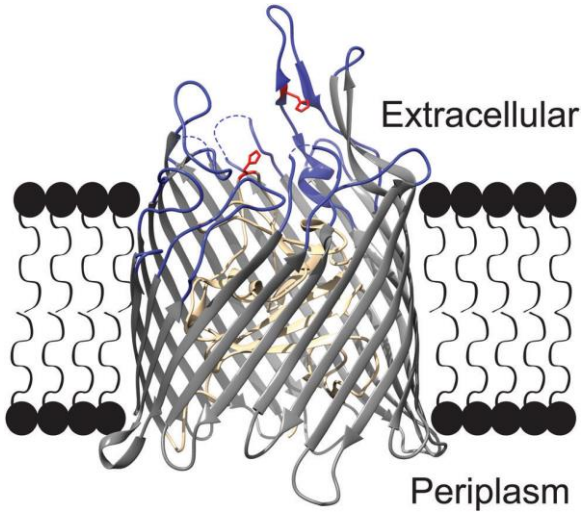


Figure 5. ShuA structure from *Shigella dysenteriae* [28]. Residues 136 to 632 create a transmembrane β-barrel (grey) that is filled by a plug domain (residues 1 to 135). Eleven extracellular loops and two histidine residues necessary for heme import are shown in blue and red, respectively.

1.1.2 Protein crystals and challenges for obtaining them

Before discussing crystallization and the different crystallographic techniques employed for this work, it is worth mentioning that biomacromolecular crystals have some features that make them different from other kinds of crystals. First, they contain between 25 to 90% solvent [4] making them very soft and fragile. Secondly, they have a gel-like structure that permits small molecules, additives and solvents to pass through them easily. Biomacromolecular crystals are limited in size and they diffract X-rays poorly in comparison to other kinds of crystals where lattice interactions cause an internal order which leads to better diffraction [4], [32], [33].

One of the difficulties in studying protein structure by crystallography is therefore obtaining high-quality crystals with appropriate size for each technique. Another difficulty is related to membrane proteins. In order to keep them in their native structure detergent is required. The next section briefly reviews the methods and variables used for crystallization of soluble and membrane proteins.

1.1.3 Rational protein crystallization: theoretical and practical approaches

Crystallization of proteins and generally biomacromolecular crystallization started around a century ago. The first protein crystal was obtained from hemoglobin [4], [34], [35]. After hemoglobin, hen egg white albumin and series of plant proteins were crystallized in different laboratories [35]. In 1926 urease from jack beans was crystallized by J.B. Sumner [36] which was followed by the crystallization of pepsin and series of other proteolytic enzymes. In 1988 H. Michel obtained the Nobel prize for crystallization of the first membrane protein [35]. Macromolecular crystallization is not limited to proteins, in 1935, Stanley crystallized the tobacco mosaic virus [37]. In 1934, the first X-ray diffraction of a protein was recorded using pepsin and protein crystals started to be used as a source for structural study [35] instead of just a method for obtaining purified protein [4]. As the methods related to protein crystallography were developed, the need for better crystallization techniques was felt. In the 1960s crystallization techniques like dialysis and vapor diffusion were developed and made crystallization of a limited amount of protein easier [38]. In the 1980s the availability of recombinant DNA technology led to a fundamental change in structural biology [4]. Overexpression of recombinant proteins finally lets researchers obtain a sufficient amount of protein for large crystallization trials [32]. Having reached this stage, finding the best conditions for crystallization is the most important next step in developing crystallography as a technique. This can be achieved by rationally studying the relationship between the protein and the variables used for crystallization, namely the phase diagram.

1.1.3.1 Phase diagram

The protein crystallization phase diagram (Figure 6) is used to describe different states of protein (liquid or solid) as a function of different physical or chemical variables. Such diagrams can be constructed using empirical data to describe protein solubility in a way that can be used to find optimal conditions for nucleation and growth [39]. Both of these are essential stages in the crystallization of proteins or any other molecules [32], [40]. During nucleation, which is the first phase transition, molecules change from a fully disordered state to an ordered state. In order for nucleation and growth to happen, supersaturation should first occur [32].

In Figure 6, undersaturated and supersaturated regions are separated by the solubility curve. When the solution is in thermodynamic equilibrium the concentration of the protein is equal to the solubility, which is dependent on the solution conditions. The undersaturated region is where the protein concentration is below the solubility curve, whereas the supersaturated region is where protein concentration is higher than solubility [39] and protein crystals can form or grow. However, this does not necessarily mean that crystals will form in the entire supersaturated region. This region is divided into three zones: metastable, labile or nucleation, and precipitation. In the metastable zone, supersaturation is not yet high enough, therefore, nucleation cannot happen in a reasonable time, however, crystal growth can take place in this zone. In the nucleation zone, supersaturation is high enough to overcome the energy barrier required for spontaneous nucleation to occur. In the precipitation zone, aggregates and precipitate are formed by excess supersaturation instead of ordered crystal growth [41].

Different parameters affect the supersaturation, for example, increasing ionic strength, decreasing temperature and increasing protein concentration can increase the supersaturation [40].

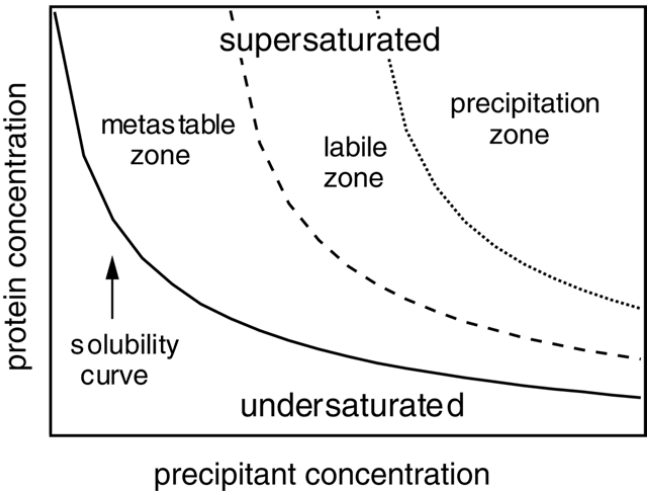


Figure 6. Schematic protein phase diagram as a function of precipitate concentration [41].

Phase diagrams can be determined by the following methods.

A. Equilibrium technique

For this, the solubility of the protein is determined by measuring the protein concentration at equilibrium at the constant temperature. One way to determine solubility curve is by using a solution of interest, placing protein crystals in it, and monitoring the protein concentration over time. If crystals grow, it means that the solution is supersaturated. In the undersaturated solution, crystals dissolve, and finally, at equilibrium, crystals are stable [39], [41]. Measuring the protein concentration at equilibrium gives one point of the solubility curve. Depending on the

solution volume, measuring one point on the solubility curve can take between several days to a week [42].

B. Michaelson interferometry

Michaelson interferometry is another tool to obtain protein solubility and requires less material and time in comparison to the equilibrium technique [39], [43]. To check if a crystal is growing or dissolving when temperature is changed, the concentration gradient near crystals is monitored [38]. By changing the temperature, fringes around the crystals change - in crystals growth they bend upward and in crystal dissolution, they bend downward- or remain flat (crystal at equilibrium). This observation helps to determine if the solution is undersaturated or supersaturated [44].

1.1.3.2 Variable affecting crystallization

Although many different crystallization techniques have been developed, optimization of macromolecular crystallization, especially for membrane proteins, is still an open area of research. One way to improve the optimization process is to understand and identify the impact of variables that affect crystallization. In the next section, some of these variables like protein nature and concentration, the nature of crystallization agents and their concentrations, physical and chemical variables like temperature, pH, ionic strength, and buffers are briefly discussed [45]. Some of these variables are related to the characteristics of the macromolecules, for example surface features. Others are different parameters of the crystallization environment itself.

Table 1 is a summary of them. These variables can also potentially interfere with each other. Consequently, it is hard to predict the conditions that should be explored or the importance of each variable in the crystallization process without some empirical data [46].

Table 1: Different variables affecting crystallization [32].

Physical	Chemical	Biochemical
1.Temperature/temperature variation 2.Surfaces/heterogeneous nucleants 3.Methodology/approach to equilibrium 4. Mother-liquor volume 5.Geometry of chamber or capillary 6. Gravity 7. Pressure 8. Time 9.Vibrations/sound/mechanical perturbations	1. pH 2. Precipitant type 3.Final precipitant concentration 4. Ionic strength 5. Cation type and concentration 6. Anion type and concentration 7. Degree of supersaturation 8. Reductive/oxidative environment 9.Concentration of the macromolecule 10. Metal ions 11. Initial precipitant concentration 12. Cross-linkers/polyions	1.Purity of the macromolecule/nature of impurities 2.Ligands, inhibitors, effectors 3.Aggregation state of the macromolecule 4.Post-translational modifications 5. Source of macromolecule 6. Proteolysis/hydrolysis 7. Chemical modifications 8. Genetic modifications 9. Inherent symmetry of the macromolecule 10. Degree of denaturation

10. Electrostatic/magnetic fields 11.Dielectric properties of the medium 12.Viscosity of the medium 13. Rate of equilibration	13. Detergents/surfactants/amphophiles 14. Non-macromolecular impurities 15. Chaotropes	11. Isoelectric point 12. Unstructured regions 13. His tags, purification tags 14. α-Helix content 15.Conformational states 16. Thermal stability 17. Allowable pH range 18. History of the sample

1.1.3.2.1 Physical variables

Although several physical variables affect the crystallization of biomacromolecules, the approach adopted in this thesis is to predominantly focus on temperature and rate of equilibration.

1. Temperature

Temperature plays an important role in crystallization studies. It affects protein solubility [47], supersaturation, and crystal nucleation and growth [48], [49]. It is easy to control during the experiment and as long as the sample is not damaged, the crystallization process induced by temperature variations is reversible [50].

Generally, the solubility of protein samples increases with increasing temperature [51]. If the solubility of protein is low or independent of temperature, it is possible to decrease the concentration of the crystallization agent because at low ionic strength solubility becomes more dependent on temperature [50], [52].

Changing the temperature also can be used to control the size of crystals. To grow larger crystals, the temperature can be changed to keep the crystal in the metastable zone of the phase diagram [42]. An automated crystallization bench, OptiCrys, (the instrument that was developed in our laboratory) uses this as the basis of an optimization workflow for the growth of large crystals. This process is fully detailed in the published paper “Optimization of Crystal Growth for Neutron Macromolecular Crystallography” inserted in appendix I. This instrument can also be used to obtain a large number of small crystals by using control of temperature [53]. Details of this complementary experiment are discussed in chapter 3.

Other than incubation temperature during the crystallization process, the mixing temperature, i.e. the temperature that sample preparation is performed, during crystallization solution preparation can affect the success of the crystallization and reproducibility of the experiment [48]. When the mixing temperature is lower than the incubation temperature, supersaturation is higher and nuclei may appear in the crystallization plate before it is transferred to an incubator or similar for a temperature-controlled experiment. This phenomenon can be observed in the case of proteins with direct solubility [48]. At higher temperatures, nucleation starts after the plate is transferred into an incubator or temperature-controlled room [48]. There are other temperature related phenomena such as kinetic ripening and crystal polymorphism. In kinetic ripening, fluctuation of temperature is used to dissolve smaller crystals of the same phase for the

benefit of the larger crystals [42], [53]. In crystal polymorphism, different forms of crystal grow from the same molecules. In the case of crystal polymorphism, temperature variation can be used to dissolve the undesirable polymorph [42], [54].

2. Rate of equilibration

Rate of equilibration is another physical variable that affects crystal quality. It can be optimized according to the purpose of the study. For example in initial screening, fast equilibration is more favourable to reduce the screening time by identifying conditions where nucleation can occur rapidly [55]. However, for obtaining high quality crystals, in most cases decreasing rate of equilibration results in better-ordered crystals with better diffraction properties [55].

Control of the equilibration rate to obtain high-quality crystals is a more complex process. The rate of equilibration depends on multiple variables such as temperature, pH, protein concentration, and etc. [56]. However generally, when crystallization rate is high, protein molecules in the solution join growing crystals more quickly, increasing the probability of heterogeneous molecular orientation which can hinder crystal growth or introduce lattice heterogeneity [57].

1.1.3.2.2 Chemical variables

This section briefly discusses the effects of pH and chemical composition of crystallization solutions on the protein crystallization process.

The pH of the crystallization solution can affect the solubility of proteins within it, which affects crystallization. However, as shown in Figure 7, protein crystallization can still happen over a large pH range.

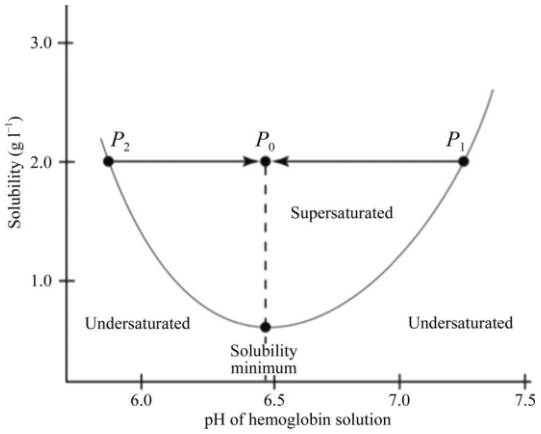


Figure 7. Solubility curve of hemoglobin as function of the pH. Changing the pH has an effect on the solubility of the protein but crystallization can still occur across a range of pH values [4].

The pH of the protein environment determines protonation states of the ionizable groups on the macromolecular surface [52]. It can also change the net charge and charge distribution, conformation, dipole moment and often the oligomeric state of macromolecules [46]. Thus, screening of pH is an essential analysis and is carried out routinely in screening assays [46].

Another chemical variable is the concentration of crystallizing agents. These are necessary in any crystallization experiment to bring a protein sample into a supersaturated state so that crystallization can occur. According to their features, they are divided into four groups: (I) salts, (II) organic solvents, (III) long-chain polymers, and (IV) low-molecular-mass polymers and non-volatile organic compounds[4], [46].

1. Salts

Salt ions and protein macromolecules compete to bind the water molecules that maintain the solubility of both [4]. At a high enough salt concentration, macromolecules cannot interact properly with water molecules leading to self-association. This can cause amorphous precipitate or ordered crystals to form. Conversely, at very low ionic strengths, there are not enough cations to maintain macromolecule solubility, which may also lead to crystal formation. In Figure 8 these salting-in and salting-out effects are depicted [4].

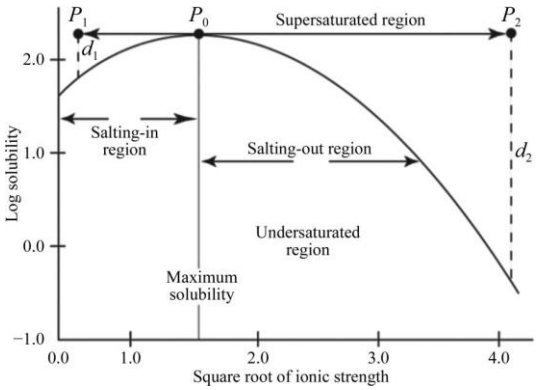


Figure 8. Solubility curve of a protein in different concentrations of salt [4]. As salt concentration increases in a protein solution, protein solubility increases (salting-in region) until the maximum solubility, after that it decreases (salting-out region) [58].

Multivalent ions, like citrates, phosphates, and sulphates are the most efficient at competing to bind water molecules and thereby dehydrating proteins [4] because their ability is proportional to the square of the valences of the salt's ion[46].

In addition to the hydration effect, protein-ion interaction is also important. For example, lithium sulphate and ammonium sulphate can have different effects on protein samples. They can create different crystal morphologies and/or diffraction qualities [4], [46], [52]. Therefore, during the

crystallization optimization process for finding the best conditions, screening a wide range of salts can be helpful.

2. Volatile Organic solvents

Organic solvents affect macromolecular interactions in aqueous solutions by changing (reducing) the dielectric of the medium. Increasing the concentration of organic solvents increases the attraction between macromolecules making the solid-state (crystal formation) favourable [4]. Methanol, ethanol, acetone, isopropanol, and dioxane are the most common organic solvents [46]. Because nucleic acids are more sensitive to dielectric effects, organic solvents are more often used to crystallize them compared to proteins [46].

3. Long-chain polymers

Long-chain polymers like polyethylene glycol (PEG) cause separation of macromolecules from solution due to volume-exclusion effects. As they occupy more space, other existing macromolecules have less available space, which causes them to form solid-states like aggregates or crystals [4]. Since PEG is not volatile, it can be used like salts and equilibrated with protein during the crystallization process. PEG competes for water and leads to dehydration in a similar manner to salts. This is important because it leads to stronger affinities for ligands. PEGs also reduce the dielectric properties of the medium (like organic solvents). Since they have low electron density, they create less noise in structures obtained by X-ray diffraction [46].

4. Low-molecular-mass polymers and non-volatile organic compounds

PEG (molecular weight ≤ 1000), MPD (2-Methyl-2,4-pentanediol) [4] and glycerol increase the viscosity of the crystallization solution. Glycerol, for example, has structure-ordering potential that makes it suitable for crystallizing flexible proteins [46]. MPD with properties between PEG and organic solvents was used successfully in the crystallization of both proteins and oligonucleotides. Salts and long-chain PEGs are used with compounds on this group [46].

1.1.3.2.3 Biochemical variables

The state of the protein itself is the most important variable in the crystallization process [3], [45]. Therefore, when there is a problem in the crystallization process, further purification or modification of the protein can increase the chance of crystallization. Proteins can be deliberately modified to try to increase the likelihood of crystallization through the use of genetic engineering and recombinant DNA techniques (e.g. single or multiple point mutation) or chemical reactions (e.g. using limited proteolysis to produce truncated proteins) [46].

The chemical purity of the protein sample is essential for having well-diffracting crystals [59]. Ideally, chemical purity as well as physical homogeneity and monodispersity of protein samples can be checked before crystallization [40]. There are various methods that can be used to evaluate sample quality, one popular example is light scattering [45].

Macromolecule concentration can be varied from 2 to 100 mg/mL [4], however, a concentration between 8 to 20 mg/mL is the optimal concentration range for crystallization of most proteins [52].

The mother liquor in a crystallization experiment typically contains solvent and a precipitating agent. It is also possible to add compounds or components (additives) in an attempt to improve the success rate of crystallization. Different types of additives are used in macromolecular crystallization. Some of them bond to the protein and change its conformation or physical-chemical properties. Some, on the other hand, make crucial interactions between macromolecules [4].

McPherson *et al*. and Sauter *et al*. list the following additives [4], [46]:

1. Physiological or biochemical ligands like coenzymes, prosthetic groups, substrate analogues, metal ions, inhibitors, etc. that bind to the active sites of enzymes or specific sites of the macromolecules. These can change macromolecule conformation, promote homogeneity or stabilize conformations.

2. Chemical protectants. Compounds like 2-mercaptoethanol, glutathione or dithiothreitol, heavy metal ion scavengers like ethylenediaminetetraacetic acid (EDTA) and ethylene glycol tetraacetic acid (EGTA) are in this group and prevent oxidation. Other compounds such as phenol or chlorobutanol and sodium azide prevent the ageing of protein samples.

3. Detergents and solubilizing agents. Different surfactants, detergents, quaternary ammonium salts, sulfobetaine, chaotropic agents as urea, are examples of components in this group.

4. Poisons. These are used to reduce twinning however; the exact role is currently not well defined. Some examples are dimethyl sulfoxide, ethanol, dioxane, acetone, butanol or MPD.

5. Co-solvents, kosmotropes, and osmolytes. These compounds stabilize protein structure by changing surface interaction of protein with water. Among them molecules like sugars (sucrose, trehalose) proline, trimethylamine N-oxide, glycine, betaine, taurine and sarcosine can be mentioned which apply their effect at high concentration (1 M or more).

6. Some small molecules such as diamino-containing or dicarboxylic acid-containing molecules can stabilize protein conformations or reduce their dynamic nature. This can occur if such agents reversibly cross-link carboxyl and amino groups on the surface of the proteins or form intramolecular hydrogen-bonding networks with surface polar groups.

7. Multivalent charged compounds that stabilize proteins through noncovalent bonds by forming intermolecular cross-links. These molecules typically have greater length in comparison to the previous group.

8. Nucleation enhancer materials and compounds. PEG with low concentrations or other polymeric materials such as jeffamine in high salt concentration solutions are in this group and provide surfaces for nucleation.

9. Ionic liquids like 1-ethyl-3-methylimidazolium tetrafluoroborate, in the concentration range 0.05–0.15 M, are also additives. They improved the crystallizability of lysozyme, trypsin, and a Fab complex without changing the structures of those crystals.

1.1.3.3 Crystallization methods

In addition to physical, chemical, and biological parameters, the crystallization method used also influences crystal nucleation and growth. Although there are different methods for crystallization of biomacromolecules, they all have the same objective, which is bringing the solution to a supersaturated state [38]. The methods used in this study are briefly described here.

1.1.3.3.1 Vapor diffusion

Vapor diffusion is the most commonly used technique for crystallization [1], [60]. This technique is generally used for small volume samples. It is worth mentioning that understanding principles of water equilibration in vapor diffusion technique has been established by Raoult's rules [61]. As shown in Figure 9, hanging drops, sitting drops, and sandwich drops are three different modes of vapor diffusion [38]. They all feature the same mechanism of protein crystallization.

In this technique (Figure 9), a drop containing the protein sample mixed with a crystallization agent is equilibrated with crystallization agents in a larger-volume reservoir [39], [46], [56], [60], [62]. The concentration of the protein and crystallization agents in the drop is lower than the concentration required for crystallization. Over time, equilibration between the drop and the reservoir will occur through the vapor phase that leads to an increase in the concentration of the ingredients in the drop [39] and decrease the size of the drop [46]. The increase of concentration in the drop drives the system toward the supersaturation and crystal growth [46], [60], [62].

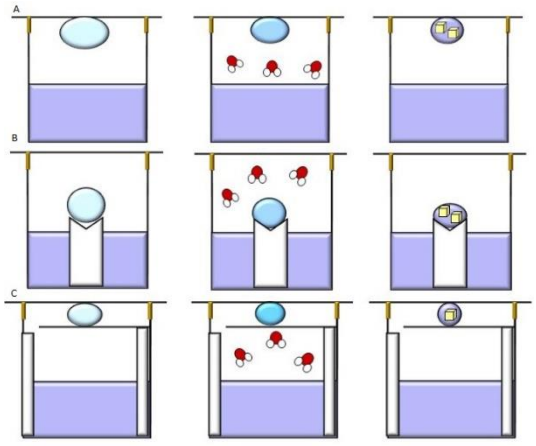


Figure 9. Vapor diffusion. There are three ways to perform this technique (A) hanging drop, (B) sitting drop and (C) sandwich drop [56]. From left to right, a drop of protein solution is mixed with a crystallization buffer (light blue) and sealed in the presence of a larger volume reservoir of the same crystallization solution (dark blue). Vapor diffusion of solvent (ball-and-stick models) slowly decreases the volume of the drop without altering the absolute number of molecules present. By increasing the supersaturation, nucleation starts, so crystals (yellow cubes) can form and grow slowly.

The kinetics of supersaturation and nucleation are determined by the kinetics of water and any other volatile species present [46]. Different parameters such as temperature, the chemical nature of crystallization agents, water pressure, distance between drop and the crystallization agent solution and initial volume of the drop, determine the equilibration rate. As a result, the time it takes for water to reach 90% equilibrium can vary from about 25 hours to more than 25 days [46]. As such, even changing the droplet size changes the kinetics of crystallization and can affect crystal size [38].

Consequently, one drawback of this technique is that it is difficult to control the level of supersaturation as a function of the time [63].

1.1.3.3.2 Dialysis

There are different available methods for crystallization using dialysis, but it generally refers to the use of microdialysis cells [38]. As illustrated in Figure 10, after adding the protein, the protein chamber is covered by a semi-permeable membrane. An elastic O-ring is then used to fix the membrane to the button. The button is immersed in one of the wells of the crystallization (Linbro) plate that has previously been filled with the desired crystallization agents which is then well sealed by a coverslip[46].

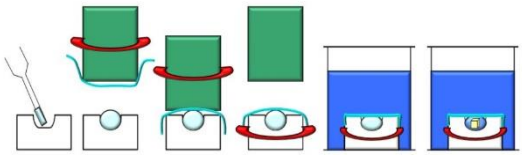


Figure 10. Microdialysis method [56]. From left to right, a drop of protein solution (light blue) is placed in a microdialysis button and covered with a semi-permeable membrane (light blue line) sealed in place with an elastic O-ring (red line). This is then submerged in a sealed well with crystallization solution (dark blue) allowing dialysis through the membrane which slowly increases precipitant concentration to the point of nucleation and crystallization (yellow cube).

The semi-permeable membrane allows small molecules like additives, buffers and ions to pass through the membrane but prevents diffusion of the macromolecule [38]. Therefore, diffusion and equilibration of crystallization agents occur through the membrane. Consequently, the crystallization agent concentration inside the dialysis chamber slowly approaches the concentration required for crystallization of the protein.

In the dialysis method, equilibration kinetics are dependent on several factors such as the molecular mass cut-off the membrane, the volume of the dialysis buttons, temperature, and protein and crystallization agent concentration [46].

One of the advantages of this method is that crystallization agents can be exchanged during the experiment [63]. Section 2.2 and 2.3 of Vahdatahar et al. (Appendix I) detail the protocol of how the crystallization bench (OptiCrys) performs this exchange in real time in automatic mode in order to grow crystals.

1.1.3.3.3 Batch method

The batch method is one of the simplest methods for crystallization [46]. In this method, shown in Figure 11, the undersaturated protein solution is directly mixed with the crystallization solution in a range of ratios [4]. Therefore, samples are in their final concentration at the beginning of the experiment [39], [60], [62]. With this method, supersaturation is obtained immediately [39], [60] so exploring the phase diagram with only one sample is not possible [56]. This method is also used to improve crystal quality [1] and size by seeding [42]. This technique is not the preferable technique for screening a range of conditions, as formation of nuclei can happen immediately after mixing protein with the crystallization solution [64].

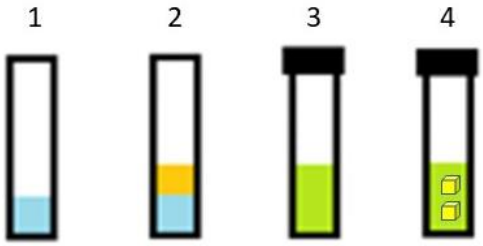


Figure 11. Batch method. In the first step, the crystallization solution (blue) and then the protein sample (orange) are added to the tube. After gently mixing the solution (green), it is kept in the desired temperature.

1.1.3.4 Comparison of kinetic pathways in the presented crystallization techniques

It is not possible to determine the phase diagram for all biological macromolecules, but understanding how solubility typically changes when using different crystallization methods (as illustrated in Figure 12) is still useful [38]. This is especially true when growing large crystals, such as those required for neutron crystallography (see section 1.1.5.2). Knowledge of the key variables when using different crystallization techniques is crucial in attempts to make crystallization of proteins more predictable.

For the dialysis method, protein concentration remains constant but the crystallization agent concentration changes over time. By increasing the concentration of the crystallization agent over time, a state of supersaturation will be reached and after that, the final concentration of the crystallization agent determines whether the protein crystallizes or precipitates [38].

For the vapor diffusion method, both the crystallization agent and protein concentration in the droplet increase with time. If equilibration occurs in the metastable zone no crystals will form. But if equilibration happens in the nucleation zone, nucleation will occur [38].

In the batch method, three different events can happen. If the final solution is undersaturated, no crystals will appear. If the solution is located in the metastable zone, crystallization can occur. If the solution is located in a high supersaturation region the protein will precipitate immediately [38].

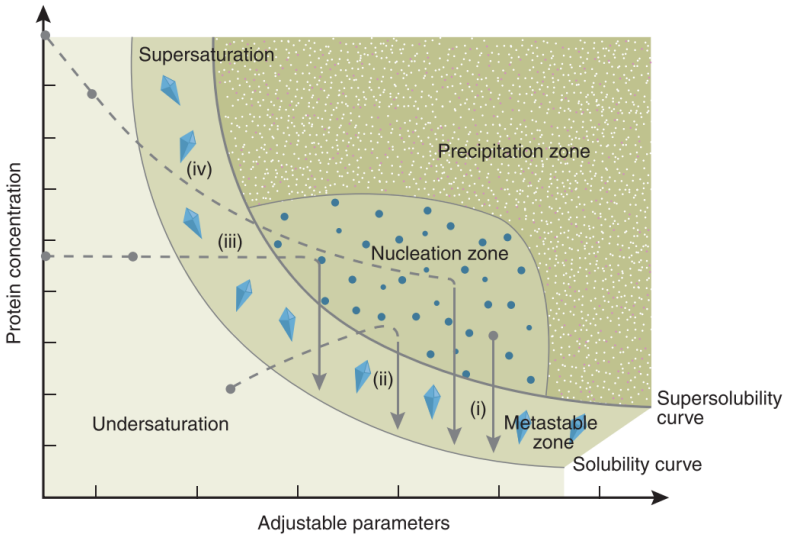


Figure 12. Different routes for each crystallization technique- (i) batch method, (ii) vapor diffusion method, (iii) dialysis method, and (iv) free-interface diffusion (FID) method- to reach nucleation and metastable zones in protein crystallization phase diagram [65].

1.1.4 Role of detergent for solubilization and crystallization of membrane proteins

As mentioned previously, membrane proteins have both hydrophobic and hydrophilic regions [7]. To preserve their native state and prevent aggregation [66] during solubilization, they need an environment that mimics the membrane. As such, surfactants, which can mimic the membrane, are typically used for a given protein [7]. Detergents are the most widely used surfactant for the extraction and stabilization of membrane proteins in aqueous solution [67], [68]. Understanding detergent behavior, micelle structure and protein detergent complexes is crucial for membrane protein studies [69].

Depending on the head group, detergents can be categorized into three major categories: ionic, non-ionic and zwitterionic [7], [67], [69]. Each category has unique features briefly discussed as follows. In ionic detergents the head group is positively (cationic) or negatively (anionic) charged.

One example of an ionic detergent is sodium dodecyl sulfate (SDS). These disrupt protein-lipid, lipid-lipid and protein-protein interactions and often denature the protein [67]. Conversely, non-ionic detergents do not disrupt protein-protein interactions and because of this, are often considered as mild. Non-ionic detergents are therefore more commonly used in protein purification and crystallization [67]. Zwitterionic detergents are electro-neutral because both charges are present in one molecule. These detergents are harsher than non-ionic but milder than ionic detergents [67], [68].

Two phase transitions can occur for detergent monomers, crystallization in aqueous solution or micelle formation [69]. There is a broad range of detergent concentrations referred to as the critical micelle concentration or CMC, where monomeric forms self-associate and create micelles (below CMC detergents are in their monomeric forms).

The size and saturation of the alkyl chains of the detergent affect CMC [70]. For example, addition of the double bonds increases the CMC value and the length of the alkyl chain decreases it [70]. The CMC value determines the size of the micelle formed by a detergent, lower CMC detergents form large micelles, making it difficult to exchange these with other detergents [70]. However, for solubilization and purification, if detergents with high CMC values are used higher concentrations are required [70]. For solubilization and purification, detergents with concentrations ranging from 0.5 to 50 mM are typically used [70].

Detergents interact with the phospholipids of membranes and after fragmentation of membranes, they can act to solubilize proteins [7]. For this to occur, it is important that detergent can create micelles because detergent monomers are not able to disrupt the membrane [7]. Membrane proteins can then form part of a detergent-protein complex. These complexes are dynamic and detergent molecules can be exchanged with monomer detergents in the solution [52]. During purification of proteins, detergent concentration should be kept above CMC (between 1.5 to 3 times more than CMC) [7] as lower or much higher detergent concentration than the value mentioned can cause protein aggregation and inactivation respectively [7]. Higher concentrations of detergent also lead to phase separation which can cause protein denaturation [67], [70].

Another important factor that affects micelle shape in detergents is the aggregation number, which is the average number of monomers in the micelle [7]. Aggregation number depends on the length of the carbon acyl group. For example in detergents with a higher aggregation number the carbon acyl chain is longer [7]. Detergents with lower aggregation numbers create spherical shaped micelles and those with higher aggregation numbers create ellipsoid shaped micelles [7]. For solubilizing membrane proteins, detergents with aggregation numbers between 50 to 100 are used [7]. The pH, temperature, detergent purity and ionic strength all affect CMC and the aggregation number of the detergent [7].

It should be noted that micelles are not static structures. Type, size and chemical properties of the monomers and the environment that detergent is in, determine the size and shape of the

micelle. These features are described by the average molecular weight, hydrodynamic radius, and aggregation number [69]. Due to the packing defects of monomers, some detergents are not able to create well-ordered micelles and because of that, they are not a good candidate for solubilization of proteins [69].

Also, in a mixed system containing a membrane protein, a detergent and lipids, the system will not behave as a pure solution for each of them [69], so changes in CMC, micelle shape and size can occur [69]. These changes are not predictable and can affect protein solubility, stability, aggregation and crystallization [69]. Apart from these, solution composition can also affect detergent behavior [69].

Different detergents are used in membrane protein purification and crystallization [67]. The following section details the detergents used in the purification, crystallization and kinetic studies described later in this thesis.

1.1.4.1 Detergent phase diagram

The phase transition of the detergent used impacts the crystallization of membrane proteins. An understanding of detergent behavior is essential for preventing protein denaturation and aggregation during the crystallization process [69].

In Figure 13, the phase diagram of non-ionic detergent as a function of detergent concentration and temperature is presented. As mentioned above, critical micelle concentration or CMC is the minimal detergent concentration at a given temperature which determines the formation of micelles [71]. Below CMC, only detergent monomers are in solution, whereas above CMC, both monomer and micelles are in solution. If detergent concentration increases much higher than CMC, nonmicellar phases are created that can be hexagonal, reverse hexagonal, or lamellar in structure (Figure 13, liquid crystalline) [71]. Apart from concentration, temperature, pH and ionic strength influence phase separation and micelle formation [70]. One example of phase separation at high temperature is the disordered aggregation of micelles [71]. The upper consolute or Krafft point is the temperature at which micelles are formed from monomers, thus below this temperature (Krafft point), no micelles are observed [70]. The lower consolute or cloud point is the temperature that micelles start to aggregate and phase separation occurs [70], [71]. These interactions between micelles can play an important role in binding detergent parts from protein detergent complexes that can initiate nucleation and lead to crystal packing [72].

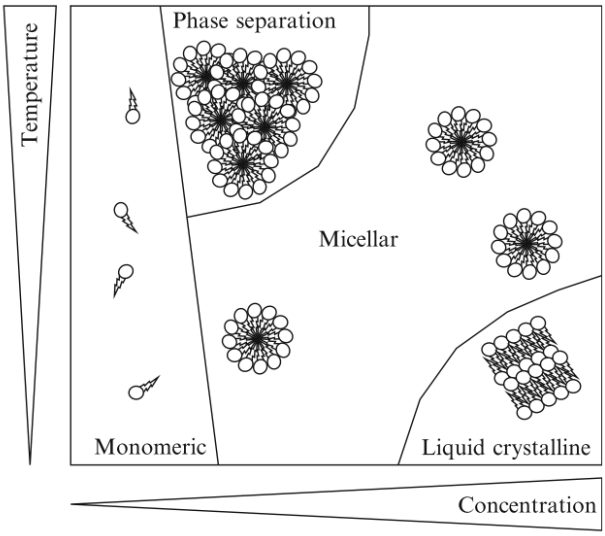


Figure 13. Phase diagram for detergent within the lower consolute boundary (most nonionic detergents) as a function of temperature and detergent concentration [71], [73].

1.1.4.2 Detergents in purification and crystallization of membrane proteins

For membrane protein purification and crystallization, detergents with different structures and properties are commercially available. Finding a suitable detergent for extracting membrane proteins and preserving their native structure is assisted by an understanding of the physiochemical properties of detergents and their interactions with the protein.

Solubilization of the membrane is the first step in membrane protein purification [71]. To achieve this, detergent concentration should be high enough to create micelles that can accommodate all of the protein. In this step, the choice of detergent highly depends on the purpose of the protein purification. If the protein is for crystallization, excess detergent should usually be removed or existing detergent should be exchanged with another detergent more suited to crystallization. There are different methods for this such as chromatography, dialysis, BioBeads or the addition of cyclodextrins that can be used to decrease or exchange the detergent in solution [71]. The following section focuses on the detergents used in this thesis for purification and crystallization of model membrane proteins.

1.1.4.2.1 n-Dodecyl-β-D-Maltopyranoside

The most frequently used detergent in membrane protein studies is n-Dodecyl-β-D-Maltopyranoside, DDM or Lauryl maltoside. DDM is a non-ionic detergent with a low CMC (~ 0.0087 (w/v %) or 0.17 mM) but relatively large micelle size (~ 65-70kDa). The low CMC value is

an advantage of this detergent because less detergent is needed to reach the CMC [67]. DDM, as it is illustrated in Figure 14 has a maltose headgroup.

1.1.4.2.2 n-Octyl- β -D-Glucopyranoside

n-Octyl- β -D-Glucopyranoside, Octyl glycoside or OG (Figure 14) is another non-ionic detergent that is widely used for crystallization of membrane proteins [74]. CMC of OG is very high (0.53 (w/v %) or 20 mM) and it has a compact micelle (~ 25kDa) [67], [74]. OG and other detergents with short alkyl chains (C7 – C10) often cannot sufficiently pack the hydrophobic region of a protein, leaving hydrophobic surfaces exposed to water which can lead to protein aggregation [70].

1.1.4.2.3 CYMAL-6

6-Cyclohexyl-hexyl-beta-D-maltoside also known as CYMAL-6 (Figure 14) is a non-ionic detergent with the maltose head group. This detergent has a micelle size of ~ 32 kDa with a CMC value ~ 0.028 (w/v %) or 0.56 mM [67].

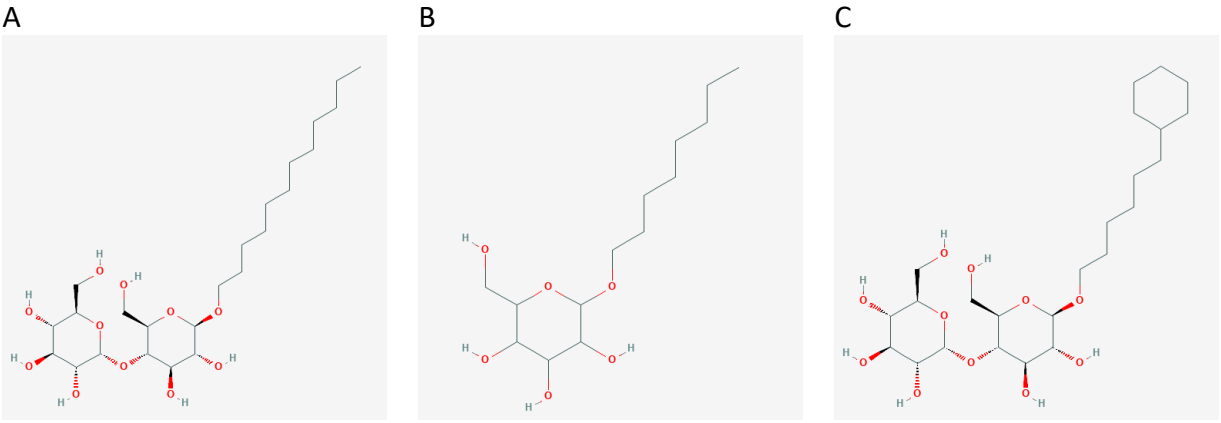


Figure 14. Structure of A. DDM (PubChem CID: 114880), B. OG (PubChem CID: 548230) and C. CYMAL6 (PubChem CID: 447688) [75].

Key features of these detergents and two other detergents (n-Octyl-oligo-oxyethylene (OPOE) and Triton X-100) that were used in this study are summarized in Table 2.

Table 2. Some properties of DDM,OG, CYMAL-6 [67], OPOE [76] and Triton X-100 [77].

Detergent	CMC: (w/v%) (mM)	M_w (Da)	Micelle size M_w (kDa)	Aggregation number, N
DDM	0.0087 (0.17)	510.6	65-70	80-150

OG	0.53 (20)	292.4	25	30-100
CYMAL-6	0.028 (0.56)	508.6	32	91
OPOE	(6.6)	330	-	-
Triton X-100	0.016 (0.01)	630 (average)	80	75–165

1.1.4.3 Membrane protein crystallization

One of the main goals in the crystallization of membrane proteins, as with soluble proteins, is achieving supersaturation and subsequently obtaining well-ordered packing of protein molecules [70]. Rather than protein macromolecules alone, a membrane protein-detergent complex (PDC) [70] is the species that crystallizes. In addition to successful crystallization, membrane proteins should be in their native structure and the protein detergent complex should be homogeneous [70].

In Figure 15, two types of crystal packing are illustrated that can occur during membrane protein crystallization [7]. Contacts between molecules are desirable in order to form crystals [7]. In type I crystals, hydrophobic and polar interactions are the main force for creating protein contacts [7]. This type of crystal lattice is found in the lipidic cubic phase [7] and bicelles [70]. In type II crystal packing, hydrophobic parts of the molecules are covered by detergent and crystals are formed by polar interaction between proteins [7], [78]. These types of crystals are obtained by using *in surfo* method.

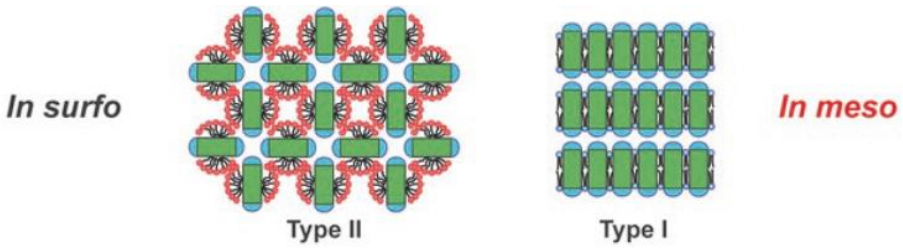


Figure 15. Crystal packing in membrane protein crystallization [79]. Hydrophobic surfaces of proteins are shown in green, hydrophilic surfaces in blue.

As shown in Figure 16, for the *in surfo* method, membrane proteins are incorporated into detergent micelles and then crystallization can occur using standard crystallization techniques like vapor diffusion, batch and dialysis [80]. In this method, crystal contacts are between hydrophilic parts of the PDC. Although the presence of detergent preserves the native structure of the protein, it may prevent protein contacts within the crystals [80]. Therefore choosing the appropriate detergent plays an important role in crystal packing and contacts. To date, detergents with alkyl chains, like DDM, DM, OG and LDAO, are the most successful detergents

for crystallization of membrane proteins [70]. Micelles within the crystals can also have polar interactions that can stabilize the lattice [70]. Although in this method, the protein is complexed with detergent that can hinder crystallization, the protein detergent solution can be used directly for crystallization [81].

Crystallization of membrane proteins also can occur within the lipidic mesophase. This is referred to as *in meso* crystallization. The lipid that is generally used for this technique is monoolein [81]. After absorbing the detergent around the protein into lipids of the cubic phase, contact with the hydrophobic parts of the protein leads to the formation of crystals [70]. Similar to the *in surfo* method, different components like buffer, salts, PEG, protein and detergent that are present in the crystallization solution can affect the mesophase and therefore crystallization [61]. In the next two sections, the HiLiDe and LCP methods for crystallization of membrane proteins are discussed.

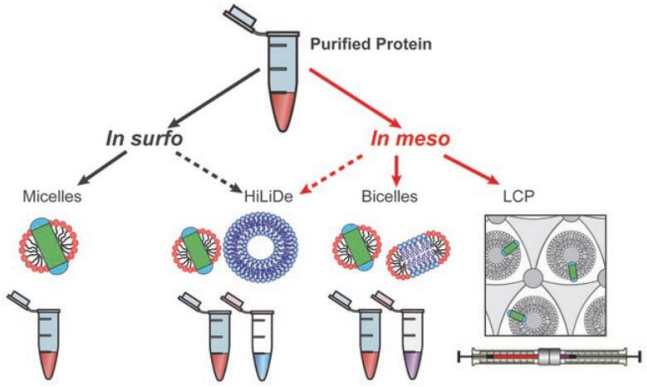


Figure 16. Membrane protein crystallization [79].

1.1.4.3.1 HiLiDe method

The High concentration of Lipid and Detergent (HiLiDe) method is a re-lipidation process used prior to crystallization for systematically screening high concentrations of lipid and detergents [78]. As shown in Figure 16, this technique takes advantage of both *in surfo* and *in meso* crystallization for membrane proteins [79] but the crystals obtained from this technique are type I [78]. The phospholipid that is used for this technique is 1,2-dioleoyl-sn-glycero-3-phosphatidylcholine (DOPC) [78]. After the re-lipidation step, crystallization is set up by conventional crystallization techniques like vapor diffusion or the batch method. By adding detergents for solubilization of the protein during the purification step, lipids surrounding the protein are removed which can adversely affect the stability of the protein. Adding this extra lipid to the PDC as part of the HiLiDe method, can improve the quality of crystals [79] by preserving native and stable protein conformation.

1.1.4.3.2 Lipidic Cubic Phase (LCP)

For this method, protein is first isolated from the membrane by standard purification techniques and then a solution containing the protein and detergent is concentrated [81]. This solution is then mixed with lipid to form a cubic mesophase [81]. The typical protein to lipid ratio at 20 °C is two to three [82]. After adding the crystallization solution, components in the solution diffuse through the mesophase so that finally crystal nucleation and growth can occur directly from the lipidic mesophase [81]. During this growth, membrane proteins diffuse through a network of interconnected lipid bilayers and aqueous channels to reach and join the growing crystal [79]. One of the advantages of LCP is that it acts as a filter so impurities and aggregates cannot diffuse through it. However, it can be problematic for large membrane proteins (larger than 100 kDa), as these may not diffuse easily in a uniform manner. To solve this problem, increasing the precipitant concentration leads LCP to transform to LSP (Lipidic Sponge Phase) that can support larger proteins. Examples of such precipitants are PEG 400, 2-methyl-2,4-pendanediol (MPD), pentaerythritol propoxylate, and 1,4-butanediol [79]. A class of lipids that is commonly used for LCP crystallization is monoacylglycerols (MAGs) and one example is monoolein or 9.9 MAG [79]. Lipids in this class have a glycerol headgroup with a monounsaturated fatty acid chain [79].

One of the challenges of using this method is that the cubic phase is extremely viscous and sticky [81]. Therefore, visualization and crystal harvesting require special consideration.

1.1.5 Protein Crystallography

Crystallographic applications are divided into two parts. The first concerns determination of the atomic structure of naturally occurring crystals, such as calcium carbonate ($CaCO_3$). This is mainly relevant in mineralogy, earth science and metallurgy. The second part instead involves the determination of the molecular structure of artificially formed crystals of a molecule [83]. Protein and biomacromolecule structure determination are almost always examples of the second type of crystallography [83]. The atomic structure of these crystals is obtained by exploiting their ability to diffract different particles, typically photons of X-ray wavelengths, neutrons or electrons.

1.1.5.1 X-ray crystallography

In 1943, John D. Bernal and Dorothy Crowfoot published the first X-ray diffraction of protein crystals [83], [84]. Monochromatic X-rays are typically generated either using an in-house or laboratory source with a sealed tube or a rotating anode or more commonly from synchrotron radiation [85]. In the early years of using in house X-ray sources, experiments were limited by a need to grow large enough crystals to compensate for these relatively weak X-ray sources. In 1976, the first X-ray diffraction of a protein crystal using synchrotron radiation was published. Since then crystallographers have generally used more intense synchrotron sources instead of in

house sources, reducing the need for long exposure times and large, well-diffracting crystals. Over the years, synchrotron beam sizes have decreased and flux density increased. However, increased flux density can lead to sample damage induced by X-rays [79]. To reduce such damage data collection at cryogenic temperatures has become routine. Often a cryoprotectant solution is necessary to protect protein crystals from damage or adventitious ice formation at these temperatures.

Crystals are formed from periodically repeated three-dimensional arrangements of molecules termed unit cells. To determine crystal structure, knowledge of both the direction and intensity of scattered radiation from an incident X-ray beam passing through the crystal is needed. The directions of the diffracted beams (reflections) are defined by the shape and symmetry of the unit cell. The intensities of the diffracted beams are defined by the location of the atoms in the unit cell [84].

To obtain the electron distribution in the unit cell, the Fourier transform of structure factors (F values) is calculated which displays the amplitudes and phases of the reflections [84]. However, from diffraction data only the amplitude of the structure factor and not phases are recorded making it impossible to generate an electron density map without more data. Solving this "phase problem" can be achieved by direct methods, experimental methods and molecular replacement [86]. In the direct method, phase is estimated from the measured intensities by probability theory and other mathematical tools. The calculations required for this method become more challenging the more atoms are present. Therefore, it is only used to solve the structure of small molecules or small proteins with X-ray diffraction better than 1.2 Å. Experimental phasing is divided into three categories:

1. Multiple or single isomorphous replacement (MIR or SIR). By co-crystallizing or soaking crystals in a buffer containing heavy atoms, the resulting crystals contain heavy atoms in their structure. As native and heavy atom derivative crystals should be isomorphous the X-ray scattering of heavy atoms can be used to calculate phase information.

2. Multi-wavelength anomalous diffraction (MAD) is where multiple sets of diffraction data are collected at X-ray wavelengths above and below the absorption edge of specific heavy atoms within the unit cell. Typical heavy atoms are metals included in protein cofactors such as iron or molybdenum or selenium which can be incorporated into most proteins in the form of selenomethionine. The differences in the diffraction patterns arising from the anomalous scattering of these heavy elements can be used to calculate their position and therefore gain phase information.

3. Single-wavelength anomalous diffraction (SAD) is a related technique where a single dataset is collected on or close to the absorbance edge peak for a heavy element within a crystal.

Molecular replacement uses a similar known structure to position a model in the new unit cell and deduce a set of phases [85], [87]. By using one of these techniques, the phase problem can

be solved allowing an electron density map to be generated from the experimental X-ray diffraction data [84].

Data collection is the last experimental step in structural analysis. Following this, electron density maps can be calculated by computer and a structural model can be fitted to these maps [79]. Different software packages can be used for data processing like HKL2000, MOSFLM and XDS (X-ray detector software) [85]. The next step is refinement of the atomic model by varying the model parameters. The goal of this process is to obtain the best agreement between observed (F_{obs}) and calculated structure factor (F_{calc}). R-factor is a value that reports on the level of this agreement [84]. Figure 17 summarizes the steps following data collection to obtain the final model and the corresponding appropriate software used at different stages.

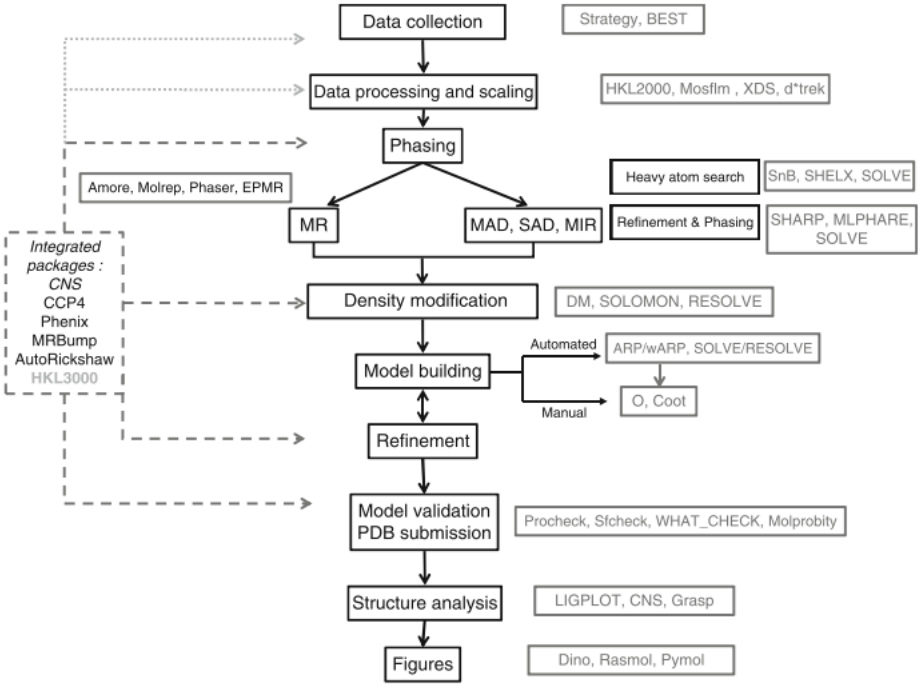


Figure 17. Workflow to determine a crystallographic structure [85].

1.1.5.2 Neutron crystallography

Approximately 50% of atoms in proteins are H-atoms [39]. In order to understand biological pathways, structure-function relationships and reaction mechanisms, it is often necessary to know the positions of H-atoms and how they transfer between biomacromolecules [88], [89]. H-atoms with only one electron scatter X-rays weakly so their positions cannot be determined with X-rays unless very high resolution data are available (higher than 1 Å) [31], [39], [84]. Neutron crystallography is a technique that can be used to obtain an accurate position of H-atoms in biological macromolecules even with medium resolution [31], [39]. Neutrons interact with the nuclei of atoms so localization of light atoms like H is possible [90]. Neutron diffraction studies

can be especially useful when studying enzymatic mechanisms or even for determining the distribution of detergent molecules in crystals of membrane proteins [91] as detergents that are tightly bound to proteins are often not visible to X-rays [91].

Biomacromolecular structural studies using neutron diffraction are similar to X-ray diffraction structural studies. The neutron beam passes through a protein crystal and the diffraction pattern is recorded by a detector. Then data processing software can be used to calculate a nuclear density map in which a structural model can be built. As neutrons are diffracted by nuclei and neutron scattering length of elements does not have a significant difference, methods like MIR, SIR and MAD cannot be used to solve the phase problem [87]. Therefore most neutron structures published to date were solved using X-ray model structures for molecular replacement [92]. However, because of negligible radiation damage during neutron crystallography experiments, data collection can mostly be conducted at room temperature. By collecting data from the same crystal, at the same temperature for both techniques, both data sets can be used simultaneously for refinement [92].

Nuclear reactors and spallation sources are two neutron sources used for neutron diffraction studies [91], [93]. In spallation sources, high energy protons strike a target (for example liquid metal) to produce pulsed neutron beams [90], [91]. At nuclear reactors, nuclear fission provides a continuous beam [90]. In both sources when neutrons are initially produced, they have high velocities and short wavelengths so they cannot be used for biological samples [91]. To moderate the velocity of the neutrons, they are passed through a cold medium like hydrogen/methane at 26 K [91]. Wavelengths of these 'cold neutrons' are approximately 2-20 Å and are suitable for biological studies [91].

Each type of neutron source uses a different detector for macromolecular crystallography. For example, a cylindrical neutron–sensitive image-plate (NIP), which is used at reactor neutron sources [94], covers a sample completely and thus it has large coverage of the reciprocal space [94], [95]. However, in spallation sources large position-sensitive detectors (PSDs) are used [93]. In reactor sources, quasi-Laue diffractometers are used, which are able to collect high-resolution data between 1.5 and 2.5 Å [39]. Spallation sources instead use time-of-flight (TOF) Laue diffractometers to increase the data collection efficiency and speed [96].

Two practical obstacles for using neutron crystallography are the weak flux (in comparison to typical x-ray beams) and weak scattering of neutrons [39], [89]. To be able to get experimental results from neutron diffraction studies, it is important to minimize the noise and maximize the diffraction signal [89]. The factors influencing neutron diffraction intensity (I) are given by the following equation:

Equation 1: $\mathbf{I = I_0 F^2 V A / \upsilon^2_0}$

In this equation, I is the diffraction intensity that is directly related to incident intensity (I_0), the size of the crystals (V), structure factor (F) and the detector area (A). In addition, it is indirectly

proportional to the square of the unit cell volume [89]. Therefore, by increasing intensity, diffraction intensity will be increased. Although with development of next generation neutron spallation sources like European Spallation Source (ESS) higher neutron intensity will be obtained, still more instrumental improvement is needed before neutron diffraction becomes commonplace. From Equation 1 it is clear that size of the crystal affects the diffraction intensity, and increasing the size of the crystal causes an increase in the intensity [89]. However, growing large volume crystals is challenging because of the dynamic and flexible nature of macromolecules. So even a partially denatured protein can interfere in this process, because the denatured protein molecules can poison the surface of the crystal and therefore prevent crystal growth [39]. For smaller crystals the exposure time can be increased to obtain a measurable diffraction signal [97]. Increasing the detector area can also be used to collect more diffracted neutrons.

Signal to noise ratio is another factor to be considered when employing neutron crystallography. H atoms have large incoherent scattering cross-section ($8.027 \times 10^{-23} cm^2$) [98], so they produce considerable background noise and limit the resolution of diffraction data. Conversely, deuterium atoms have around 40 times lower incoherent scattering contribution (2.05×10^{-24} cm^2) [89], [98] and a positive coherent scattering length, which is approximately twice that of H atoms. Consequently, D atoms are more easily located in neutron maps [93]. For solving the signal to noise ratio problem, deuteration by H/D exchange or perdeuteration techniques is used [89] which can also decrease the necessary exposure time [99].

There are two ways to deuterate a sample. The first involves treating a pre-existing crystal either by vapor exchange or soaking in solution containing D_2O. In this method, H atoms bound to nitrogen and oxygen that are in contact with the solvent will exchange with deuterium atoms. However, H atoms bonded to carbons will not exchange [90]. Therefore, approximately 66% of H atoms in unit cells and 26% of the H atoms in the protein can be exchanged using this method [39].

The second technique is the perdeuteration of samples. This can be achieved by growing bacteria expressing the protein on deuterated media. In this method, all H atoms in the protein are replaced with D atoms. The data from these samples have better signal to noise ratios, which decrease data collection time and improve (increase) diffraction. Perdeuteration can also be used to compensate for smaller crystals or larger unit cells [93], [98]. Perdeuterated crystals with crystal volume at least 0.1 mm^3 can be used for data collection at LADI III, ILL [95], although larger crystals are routinely required for successful data collection and analysis [39], [95].

Other than a necessity for large crystals and potentially the incorporation of deuterium rather than hydrogen, growing crystals for neutron diffraction studies is identical to the crystal-growing techniques that are suitable for X-ray diffraction studies.

1.1.5.3 Electron diffraction

As previously mentioned, the majority of structures deposited in the protein data bank are obtained by X-ray diffraction. Most of the remaining structures are obtained from single particle cryo-electron microscopy, neutron diffraction and spectroscopy methods. Use of electron diffraction is still limited but is growing rapidly [100] especially with the development of methods to study 3D crystals. The limiting factor and bottleneck for X-ray and neutron crystallography is the crystal size. In electron diffraction (ED), crystal volumes that are two or three orders of magnitude smaller than adequate samples for X-ray diffraction can be used and high-resolution diffraction data can be collected from these very small protein crystals [100], [101]. For many years, only 2D crystals were used for data collection for electron diffraction because of the worries of dynamical scattering [102], but Gonen and coworkers showed that by rotating the crystals they could collect data and solve structures from 3D protein crystals [103]. Crystals should still be thin enough (few hundred nanometers [104]) but by using this technique the completeness of data collection with ED is maximized and dynamical scattering at zone axis is reduced. Large crystals can be broken by vortexing or sonication to prepare crystals with desire size for ED data collection [105]. In 3D ED, which uses 3D crystals, data collection at different angles is achieved using extremely low electron doses [106]. The technique also gives the possibility of data collection in both the diffraction and imaging from the same sample giving both reciprocal and real space information. Finally, because of strong Coulomb interactions between electrons and matter, signal to noise ratio is higher in comparison to X-ray diffraction. So, it is easier to identify light atoms like hydrogen [100]. Diffraction patterns obtained in ED are related to the Coulomb potential distribution. So, instead of an electron density map obtained from X-ray diffraction a Coulomb potential map is obtained from ED. According to the environment and charge of the residues, protein structures may look different in this map in comparison to the electron density map from X-ray diffraction [103]. This can be used to determine the charge state of protein residues and metals [107], which is useful for understanding the function of proteins [103]. Overall ED can be considered a complementary technique for X-ray diffraction and preferable when the size of crystals is too small to use currently available synchrotron X-ray sources [101].

Another possibility for collecting diffraction data from extremely small crystals is by X-ray serial crystallography using synchrotron or X-ray free-electron laser (XFEL). High-energy ‘free’ electrons passed through periodic alternating magnetic fields can effectively create brief X-ray laser pulses that are intensely bright. These can be used for protein structural studies as pulses are short enough in duration (in the order of femtoseconds) that diffraction can be recorded before radiation damages the exposed protein crystal. Due to the highly-focused nature of these pulses, it is possible to use small crystals, although hundreds of micrograms to hundreds of milligrams of crystals are often required to solve a complete structure [108]. 3D ED has several advantages over XFEL, when collecting a full data set with 3D ED, one to three crystals are often enough and for samples with low symmetries, several data sets from different crystals can be merged. After collecting data, statistics like indexing, merging and scaling can be done with standard X-ray

crystallography software and are relatively straightforward in comparison to XFEL. XFEL facilities are limited and finding beam time is not easy whereas transmission electron microscopes are more accessible [109].

After growing crystals to a size sufficient for ED (but less than 500 nm as crystals that are greater than 500 nm in thickness must be thinned), microcrystals are applied to an EM grid. For electron diffraction studies a cryo-transmission electron cryo-microscope (cryo-EM) is used [106], [109]. Then the extra liquid is removed using filter paper (blotting) and after vitrification, the sample is loaded into the microscope. Test diffraction is used to check the quality of the crystals on the grid. After finding high-quality crystals, a diffraction data set can be collected. Data collection is done at a low electron dose (~ 0.01 $e^-/Å^2/s$) with continuous rotation of the crystal on the grid [106]. Standard crystallography software can then be used for data analysis, phasing and refinement [106]. Electron scattering factors are tabulated. Nevertheless, R factors are relatively high for ED-resolved protein structures in large part because of dynamical scattering [109].

1.1.6 Rationalizing protein crystallization

In the next two sections, two crystallization benches, previously developed in our laboratory, are discussed. In this study, for the first time, these instruments have been employed in rational crystal growth strategies of model membrane proteins. In these instruments, temperature is controlled automatically. Crystallization solution exchange is also possible in a manual (first generation) and automated (second-generation) mode. Controlled crystal growth with these benches is possible by using knowledge of the phase diagram and the ability to vary the temperature and chemical composition of crystallization solution (see section 1.1.3.1).

1.1.6.1 First generation crystallization bench

The first-generation crystallization bench is a semi-automated instrument using combined temperature control and seeding for investigation of the phase diagram, which can be used to induce nucleation, control crystal growth and determine the solubility curve. This instrument uses both dialysis and batch methods (see section 1.1.3.3.2 and 1.1.3.3.3). The batch method is initially used in combination with crystal seeding to investigate the phase diagram and measure the solubility curve. Data from these studies are then used to grow large crystals using precise control of the temperature to keep crystals in the metastable zone during the course of a crystallisation experiment. In this case, no new nucleation will occur and no new crystals will appear and only seeded crystals will grow. This way seeding and real time temperature control are used to select between nucleation and crystal growth [42], [98].

The study of solubility before crystal growth is of particular importance when large crystals are desired for neutron crystallography (see section 1.1.5.2 and 1.1.3.1). H/D exchange, which can improve the signal to noise ratio in such experiments, also alters physico-chemical features of

the protein sample in solution that can lead to changes in the crystal-solution equilibrium. Therefore, systematic studies to investigate the phase diagram and measure the solubility under deuterated conditions are crucial [42], [98].

After the experimental investigation has been carried out to determine the location of the metastable zone and the solubility curve, this information can be used to grow large crystals (see section 1.1.3.1). If the protein sample has direct solubility, i.e. solubility of the protein increases with increasing temperature, four different behaviours can be observed in seeding experiments [42].

1. Seeded crystal dissolves. This means that a solution is in the undersaturated zone of the phase diagram. By reducing the temperature, supersaturation will increase.

2. No nucleation occurs and the seeded crystal grows. This represents the metastable zone where due to too low supersaturation, the nucleation rate is extremely low and no nucleation occurs in a reasonable time.

3. Seeded crystal grows and nucleation occurs. This means that a solution is in the nucleation zone. To decrease the supersaturation, the temperature should be increased.

4. Aggregation and precipitate form. This represents the precipitation zone of the phase diagram, where supersaturation is high. In this zone, aggregate and precipitate formation is faster than crystallization.

These principles can be used in order to grow large crystals as illustrated in Figure 18. For this, a crystallization solution is seeded with a crystal at some point in the supersaturation zone in the vicinity of or in the metastable zone. This seed will grow until equilibration between protein solution and crystal is achieved. Then the temperature can be changed to keep the crystal in the metastable zone. As a result, the crystal will start to grow again. As long as temperature changes can be used to keep the crystal in the metastable zone, the crystal will continue to grow until the crystal reaches the desired size (e.g. for neutron crystallography).

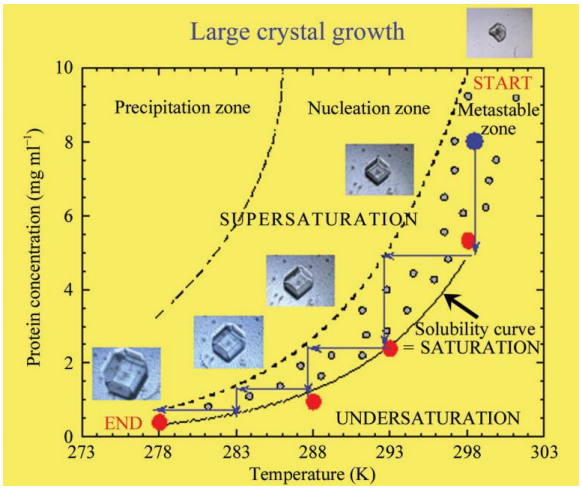


Figure 18. Optimization of crystal growth using temperature variations for proteins with direct solubility. By changing the temperature to 298, 293, 288, 283 and 278 K, seeded crystal is kept in the Metastable zone of the phase diagram [42].

1.1.6.2 Second generation crystallization bench

To improve the size and diffraction quality of crystals, in addition to the temperature control already offered by the first-generation crystallization bench, controlling variables such as the concentration of protein and crystallization agents and equilibration rate are also important. This led to the development of the second crystallization instrument (OptiCrys). OptiCrys contains a fluidic system for exchanging crystallization agents in an automated mode during crystallization experiments. As shown in Figure 19, the dialysis flowing set up has three main parts: the crystallization (dialysis) chamber, reservoir chamber and an airtight cap. The dialysis set up is then inserted inside a titanium support. The support is in contact with Peltier elements that are connected to a chiller. With this design, precise control of temperature is possible. This automated system also allows exchange of crystallization agents during the experiment using a pumping system. With these features, it is possible to screen different conditions with a small amount of protein sample to find the best one for nucleation and crystal growth. Therefore, the second generation instrument is a powerful system to control and separate nucleation and crystal growth. Another advantage of using this instrument is that it allows the optimization of the crystal growth based on the temperature-precipitant concentration phase diagram. This method is one of the best methods to grow not only a small number of large crystals for neutron crystallography studies but also a homogeneous population of a large number of small crystals that can be of interest of microED or serial synchrotron crystallography. There are several examples of successfully grown crystals with this method [63].

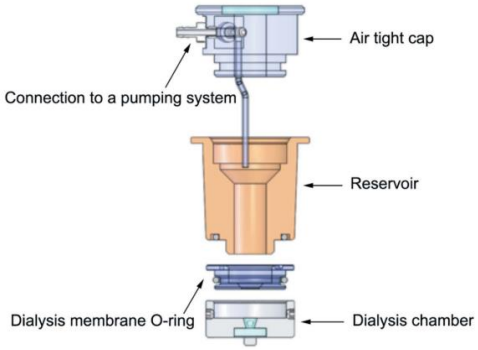


Figure 19. Schematic of the temperature-controlled flowing reservoir dialysis setup of the second generation crystallization bench [63].

1.1.7 Aim of this thesis

The objective of this thesis project is to find rational approaches to understand macromolecular (especially membrane protein) crystallization in a systematic manner with conventional and

modern crystallization techniques. These kinds of studies can be used to help obtaining high-quality crystals in a predictable manner instead of the traditional trial-and-error approach.

Accordingly, the first aim of the present study is on optimizing the purification of two model membrane proteins introduced early in this chapter. Different chromatography techniques (affinity, size exclusion and anion exchange chromatography) were applied for the purification of these proteins.

Following this, there is a particular focus on optimization of crystal growth using different crystallization techniques such as vapor diffusion, dialysis and batch methods. Next, will be to demonstrate how physico-chemical variables, such as temperature, supersaturation rate, and biochemical variables like purity of the protein can be controlled to improve crystal quality and crystal size in combination with knowledge of phase diagrams.

This work will show how crystallization benches can be used to control the size and quality of the crystals for different crystallography approaches in an automated and reversible manner.

Quality of the obtained crystals from different crystallization techniques (vapor diffusion, dialysis and batch method) will be compared.

Finally, mass transport of the detergent through a semipermeable membrane is systematically studied for further characterization of the membrane protein crystallization using the dialysis method.

Chapter 2: Protein purification

2.1 Introduction

To characterize the protein structures and thus understand their functions and related biological mechanisms, highly pure proteins are needed. Although the principle of purification for soluble and membrane proteins is the same, the purification of membrane proteins is typically challenging in practice. Such proteins in their native environment are embedded in a cell membrane and have hydrophobic and hydrophilic parts corresponding to the membrane embedded and exposed areas of the protein respectively. Therefore, to keep these proteins in their native structure during the purification process, and to replace the membrane phospholipids, detergents are used to stabilize the hydrophobic surfaces of these proteins. Finding an appropriate detergent, as well as the pH, buffer and salts to keep the protein stable in solution is therefore key to successful purification.

As protein has different characteristics, it is impossible to find a single method for purification of all membrane proteins [110]. However, the general steps for protein purification are as follows. In the first step of the purification, the membrane is isolated from the cells. There are two main categories of techniques used for cell lysis: mechanical and non-mechanical methods. In mechanical methods, high-pressure homogenizers, sonicators and bead mills are used. In particular, high-pressure homogenizers like microfluidizers, homogenizers or French presses are widely used for the disruption of the bacterial cells [111]. Non-mechanical methods are divided into three different categories: physical (for example heating and osmotic shock), chemical (such as Alkalis and detergents) and biological (such as enzymes). Regardless of the specific technique used for cell lysis, the unbroken cells are then removed by low speed centrifugation followed by ultracentrifugation to isolate the membrane. After this step, detergents are used to solubilize the membrane. Finally, chromatographic techniques like affinity, size exclusion or ion exchange (cation or anion) chromatography can be used to purify the desired protein from others present in the solubilized cell membrane extract [111].

Although there are different methods available to study the structure of proteins, this work focuses on crystallography. For that, crystals with high quality are required. As mentioned previously, the purity and homogeneity of the protein are the biochemical variables that affect the crystal quality. The aim of this chapter is to describe the purification of two model membrane proteins, ShuA of *Shigella dysenteriae* and AcrB of *Escherichia coli*. In the first section of this chapter, the methods used for the purification of these two selected membrane targets/transporters are summarized, including the modifications made to improve the purification process. The next section presents the results of these experiments, which are then discussed in the last section of the chapter.

2.2 Methods

General reagents were purchased from Sigma-Aldrich, Fluka, Merck, Bio-Rad and Roth. All solutions, unless otherwise stated, were prepared with deionized 'Milli-Q' water (resistivity = 18.2 MΩ·cm) using a Millipore water purification system.

The ShuA overexpression vector was kindly provided by David Cobessi. The AcrB overexpression vector was kindly provided by Eva Pebay-Peyroula. For the production of both proteins, *E. coli* BL21 (DE3) cells were used.

2.2.1 Plasmid DNA isolation

For plasmid DNA isolation from the bacteria cells by alkaline lysis method [112], a plasmid DNA isolation Kit from NucleoSpin was used. *E.coli* cells containing a plasmid of interest were grown overnight at 37 °C on antibiotic-infused LB agar plates (100 µg/mL ampicillin for ShuA, 50 µg/mL kanamycin for AcrB). A single bacterial colony from this agar plate was then used to inoculate 5 mL of LB containing the same antibiotic and incubated at 37 °C with 200 rpm shaking overnight. The bacterial cells were pelleted by centrifugation for 30 seconds at 13,000 rpm. The supernatant was completely removed and 250 µL of the Resuspension Buffer (Buffer A1) was added to the pelleted bacteria. By vortexing, the cell pellets were completely suspended with buffer A1. 250 µL of blue Lysis Buffer (buffer A2) was then added to the sample and mixed gently by inverting the tube 6 to 8 times. This was incubated for 5 min at room temperature to release the plasmid DNA from the cells. When the lysate was apparent, 300 µL of Neutralization Buffer (buffer A3) was added and the tube was inverted again 6 to 8 times, mixing the sample with the buffer until it was colorless. The sample was then centrifuged for 5 min at 13,000 rpm leaving a clear supernatant. A NucleoSpin Plasmid Column was placed in a 2 mL Eppendorf tube and the supernatant was pipetted into the top of the column. After centrifugation for 1 min at 13,000 rpm, 500 µL of Wash Buffer (Buffer AW) preheated (at 50°C) after reaching room temperature was added to wash the silica membrane of the column. To remove it, the column was centrifuged for 1 min at 13,000 rpm. Then 600 µL of the second Wash Buffer (buffer A4) was added to the column and centrifuged again for 1 min. After this step, the column was centrifuged for an extra 2 min at 13,000 rpm to dry. The column was then transferred to a new 2 mL tube where 50 µL of Elution Buffer (AE buffer that contains 5 mM Tris-HCl, pH 8.5) was added. After 3 min incubation at room temperature, the column was centrifuged for 1 min at 13,000 rpm to elute the plasmid.

The absorbance of the sample at 260 nm was recorded (Spectrophotometer UV5Nano) and used to estimate nucleotide concentration (using ε_{260} = 0.020 $(\mu g/mL)^{-1}$ cm^{-1} as an average value for double-stranded DNA). The sample was then stored at -20 °C and thawed as needed.

2.2.2 Restriction enzyme digestion

The volumes of restriction enzyme and buffer required for the DNA digestion were used according to the manufacturer's instructions (ThermoFisher Scientific). First, the following reaction components were added to Eppendorf tubes: 15 µL of Nuclease-free water, 2 µL of 10X fast digest, 2 µL of DNA (up to 1 µg) and the fast digest enzyme.

After mixing and spinning down the reaction mixture by centrifuging the sample, it was incubated at 37 °C for 10 min. The sample was then mixed with a loading buffer and loaded onto an agarose gel.

2.2.2.1 Agarose gel electrophoresis

The size of the DNA determines the concentration of Agarose used for DNA gel electrophoresis. For ShuA and AcrB samples, 300 mg of Agarose was melted in 30 mL of 0.5X TBE (45 mM Tris-borate and 1mM EDTA, pH 8.0) in a microwave oven. The solution was then cooled to around 50 °C and 3 µL of gel red were added to the solution. This solution was then poured into the gel cast and the comb was fixed in position. After 30 minutes, the gel had polymerized. 5 µL of smart ladder was added to the first well. Next, 10 µL of the sample was mixed with 2 µL of the gel loading buffer Blue (DNA 6 X, ThermoFisher Scientific) and 10 µL of the final solution was added to the well. After loading all samples, TBE was added to the tank and electrophoresis was started by applying an appropriate potential difference (100 V). The experiment was then left for 45 min to allow maximum separation of the DNA fragments for the gel size. The DNA was visualized by a UV transilluminator at a wavelength of 302 nm.

2.2.2.2 Transformation of *E.coli* cells

3 µL of plasmid was added to 100 µL of bacterial cells (BL21(DE3)). The mixture was incubated for 45 min on ice in a Biological Safety Cabinet (Supcris). The cells were then placed in a water bath at 42 °C for 45 seconds for heat shock, then the tube was kept on ice for 3 min. This sample was divided into two tubes (each containing 50 µL). The sample from the first tube was transferred to an agar plate containing an antibiotic (100 µg/mL ampicillin for ShuA and 50 µg/mL kanamycin for AcrB) and incubated at 37 °C overnight. 250 µL of fresh LB was added to the second sample and incubated for 1 - 2 hours at 37 °C. This was then centrifuged for 10 min at 13,000 rpm. The supernatant was discarded and 100 µL of LB was added to the cell pellet which was then resuspended with a micropipette. 50 µL of this sample was spread on an LB agar plate containing an antibiotic, and the second portion was pipetted into a tube containing 5 mL of LB and the same concentration of antibiotic. All samples were then incubated at 37 °C overnight.

2.2.3 Culture media

2.2.3.1 Luria Broth (LB)

20 g Luria Broth powder (from Sigma-Aldrich) was dissolved in 1 L of Milli Q water and the medium was sterilized by autoclaving.

2.2.3.2 LB Agar

Autoclave sterilized LB agar was cooled to approximately 45 °C before adding an appropriate antibiotic with a final concentration of 100 µg/mL for ampicillin or 50 µg/mL for kanamycin. This solution was then poured into sterile plastic Petri dishes in a Biological Safety Cabinet. After being allowed to cool and polymerize, they were stored at 4 °C for future use.

2.2.4 Protein expression in *E. coli*

A single colony of *E. coli* cells transformed with plasmids (see 2.2.2.2) containing the DNA sequence of AcrB was used to inoculate 5 mL of LB containing 50 µg/mL of kanamycin. This was then incubated at 37 °C with shaking at 220 rpm. After 6 hours, 50 µL of this solution was transferred to a flask containing 50 mL of LB and kanamycin (50mg/mL). The cells were then grown overnight with constant shaking at 200 rpm at 37 °C to obtain a primary culture.

For *E.coli* transformed with plasmids containing the ShuA DNA sequence, a single colony was used to inoculate 70 mL of LB containing 100 µg/mL ampicillin and incubated at 37 °C with 200 rpm agitation overnight.

The next day, the optical density at 600 nm wavelength (OD_{600}), proportional to the number of bacterial cells present in the pre-culture solution, was recorded and an appropriate amount (16 mL of AcrB or 25 mL of ShuA) was transferred to a flask containing 1 L of fresh LB medium and the appropriate antibiotic and grown at 37 °C, 220 rpm to an OD_{600} of 0.6. The temperature was reduced to 30 °C and after 30 min, 0.5 mM isopropyl-β-D-1-thiogalacto pyranoside (IPTG) was added to each flask to induce protein overexpression. Cells were incubated (in the case of AcrB for 4 hours and in the case of the ShuA overnight) at 30 °C, 220 rpm. Cells were harvested by centrifugation for 25 min at 5500 rpm at 4 °C (SCL-6000 rotor). The cell pellets were washed with buffer (for ShuA 200 mL of 50 mM Tris-HCl pH 8.0 and for AcrB 150 mL of 20 mM Tris-HCl pH 8.0, 500 mM NaCl) to remove IPTG. These were then centrifuged for 1 hour at 3600 rpm at 4 °C (Thermo Scientific multifuge X1R, TX-400 rotor) to pellet the cells. These cell pellets were stored at -20 °C and then used for purification.

2.2.5 AcrB purification

2.2.5.1 Ni-NTA affinity chromatography under native conditions

Each cell pellet was resuspended with 25 mL of buffer B (see Table 3 below for details of buffers). Potter-Elvehjem was used to homogenise the solution to prevent clogging when it was subsequently transferred to a microfluidizer (LM20 Microfluidizer®) for cell lysis. Ice and ethanol were added to the microfluidizer tray to keep the suspension as cold as possible. The microfluidizer tubes were washed three times with Milli Q water to remove all traces of ethanol. Passing buffer B through them then equilibrated the tubes. The homogenized sample was passed 3 - 5 times through the microfluidizer then through buffer B. The tubes were washed with water and the machine was stored in 50% ethanol. To break up the cells, a pressure of 15,000 psi was applied. The resulting lysate was then centrifuged at 10,000 rpm for 15 min at 4 °C to remove unbroken cells. The supernatant was removed and ultra-centrifuged at 46,000 rpm for 1 hour at 4 °C (Beckman, rotor 70 Ti). The resulting membrane pellets were resuspended in 6 mL of buffer A and the solution was homogenized using the Potter. Next, 40 mL of solubilisation buffer was added and the final solution was kept at a rotating wheel for 1 hour at 4 °C. The solution was transferred to ultracentrifuge tubes and centrifuged at 40,000 rpm for 1 hour at 4 °C (Beckman, rotor 70 Ti). The resulting pellet was discarded and the supernatant containing the solubilized membrane extract was recovered and kept on ice for the next step.

Ni-nitrilotriacetic acid (Ni-NTA) agarose beads (Qiagen) were resuspended by gently shaking the vial and 6 mL of the solution was transferred to a 15 mL falcon tube. 3 mL of Milli Q water was added to the falcon and the solution was centrifuged for one minute at maximum speed (Thermo Scientific multifuge X1R). The supernatant was discarded. This washing process was then repeated three more times. In the following steps, the Ni-NTA resin was added to a gravity flow column (BioRad) and equilibrated with buffer C. The supernatant containing the solubilised membrane extract was added to the resin and incubated overnight on a rotating wheel at 4 °C. For this experiment, the sample and resin were mixed in batches (batch method) instead of on-column binding [113]. The solution was deposited on the column and the flow through was recovered. Then, 15 mL of buffer C followed by 15 mL of buffer D was added to rinse the column. All rinses were collected and kept on ice for further investigation. Passing 9 mL of buffer E in six fractions of 1.5 mL eluted protein. Fractions 2 to 4 were incubated for 5 min and then collected.

To calculate the protein concentration of each sample, absorbance at 280 nm (A_{280}) was measured by nanodrop spectrophotometer (Spectrophotometer UV5Nano), ($\varepsilon_{280} = 89855\ M^{-1} cm^{-1}$).

After measuring the concentration, 20 µL of each sample was added to an Eppendorf tube for SDS-PAGE analysis (see 2.2.7)

Fractions containing pure AcrB protein, as determined by SDS-PAGE (fractions 2-6), were pooled and the protein was concentrated with a 15 mL Amicon Ultra Centrifugal Filter with a cut-off of 100 kDa. The protein was concentrated to approximately 500 µL by three rounds of

centrifugation for 15 min at 3000 rpm. After each round, the sample was re-suspended with a pipette to avoid aggregate formation. Buffer F was added to the remaining solution and the concentration was repeated for another four rounds to dilute the buffer E from the previous step to a negligible amount. The absorbance of the concentrate at 280 nm (A_{280}) was measured again to calculate the sample concentration and the concentrated sample was stored at 4°C.

Table 3. Buffer for AcrB purification by affinity chromatography

Buffer A:

Component	Final concentration	Stock
Tris-HCl pH 8.0	20 mM	1 M
NaCl	500 mM	5 M

Buffer B:

Component	Final concentration	Stock
Tris-HCl pH 8.0	20 mM	1 M
NaCl	500 mM	5 M
$MgCl_2$	2 mM	1 M
Protease inhibitors	¼tablet*	

* Protease inhibitor cocktail tablets (complete, EDTA free) (Roche)

Solubilisation buffer:

Component	Final concentration	Stock
Glycerol	10 %	≥99 %
Tris-HCl pH 7.0	50 mM	1 M
Imidazole pH 7.0	10 mM	1 M
DDM	2 %	20 %
Protease inhibitors	½ tablet*	

* Protease inhibitor cocktail tablets (complete, EDTA free) (Roche)

Buffer C:

Component	Final concentration	Stock
Tris-HCl pH 7.0	20 mM	1 M
DDM	0.2 %	20 %
Glycerol	10 %	≥99 %
Imidazole pH 7.0	10 mM	1 M

Buffer D:

Component	Final concentration	Stock
Tris-HCl pH 7.0	20 mM	1 M

DDM	0.02 %	20 %
Glycerol	10 %	≥99 %
Imidazole pH 7.0	50 mM	1 M

Buffer E:

Component	Final concentration	Stock
Tris-HCl pH 7.0	20 mM	1 M
DDM	0.02 %	20 %
Glycerol	10 %	≥99 %
Imidazole pH 7.0	200 mM	1 M

Buffer F:

Component	Final concentration	Stock
Hepes pH 7.0	10 mM	1 M
DDM	0.02 %	20 %

2.2.5.2 Size exclusion chromatography

Samples from Ni-NTA method (see 2.2.5.1) were further purified either by size exclusion chromatography with HiLoad 16/60 Superdex 200 prep grade (GE Healthcare) or by anion exchange chromatography (see 2.2.5.3). In size exclusion chromatography macromolecules are separated according to their size. After washing the column with Milli Q water, the column was equilibrated by using a buffer described in Table 4.

The sample obtained from Ni-NTA method was injected onto the Superdex column after 10 min centrifugation at 13,200 rpm, 4 °C. Then one and half column volumes of buffer were passed through the column and the eluate was collected in 2 mL fractionation tubes. To determine which fractions contain pure AcrB, SDS-PAGE of 20 µL samples from each fraction was performed (see 2.2.7). From the results of the resolved SDS-PAGE gel, fractions containing pure AcrB were pooled together. To concentrate these pooled fractions, an Amicon Ultra Centrifugal Filter with a 100 kDa cut-off membrane was used. After determining the protein concentration with the nanodrop spectrophotometer, the purified AcrB was stored at 4 °C for the next step of the study.

Table 4. Buffer for AcrB purification by size exclusion chromatography

Component	Final concentration	Stock
HEPES pH7.0	10 mM	1 M
DDM	0.02 %	20 %

2.2.5.3 Mono Q anion exchange chromatography

Samples from the Ni-NTA method were also purified alternatively using anion exchange chromatography (rather than size exclusion chromatography, see 2.2.5.2). After the first step of purification with affinity chromatography, the AcrB-containing fractions, determined by SDS-PAGE, were pooled. The Mono Q column (Mono Q 5/50 GL) was first washed with Milli Q water and then equilibrated with mono Q buffer A (as detailed in Table 5) before the pooled proteins were injected onto the column. After the protein was passed through, unbound proteins were desorbed by washing with the same buffer (Mono Q A). By increasing the percentage of the elution buffer (Mono Q B buffer), bound proteins including AcrB were eluted. The gradient volume was 20 mL at a flow rate of 0.5 mL/min.

Table 5. Buffer for AcrB purification by ion exchange chromatography

Mono Q A: 200 mL

Component	Final concentration	Stock
Tris-HCl pH 8.8	50 mM	1 M
DDM	0.02 %	20 %

Mono Q B: 50 mL

Component	Final concentration	Stock
Tris-HCl pH 8.0	50 mM	1 M
DDM	0.02 %	20 %
NaCl	1 M	5 M

2.2.6 ShuA purification

2.2.6.1 Ni-NTA affinity chromatography under native conditions

ShuA was purified using Ni-NTA affinity chromatography using the batch method in a similar way to AcrB (see2.2.5.1). The protocol followed is mentioned in brief.

Two bacterial cell pellets from 2 L culture were resuspended by adding 25 mL of buffer A (lysis buffer) to each. The potter homogenized the solution before the cells were lysed by microfluidizer (15000 psi pressure). The lysed cell mixture was centrifuged at 10,000 rpm for 15 min at 4 °C to separate the soluble fraction from cell debris and the insoluble fraction.

To keep the solution cold, a mixture of ice and ethanol were added to the tray. Milli Q water was used to clean the ethanol from the microfluidizer tubes. Then by passing buffer A followed by the sample, the tubes were equilibrated. The cells were passed through the microfluidizer three to five times to ensure thorough cell lysis. In the following steps, buffer A was first used to remove all cells from the tubes. Then the tubes were washed with Milli Q water and stored in 50 % ethanol. To remove unbroken cells and cell debris, the lysate was centrifuged at 10,000 rpm for

15 minutes at 4 °C. The supernatant was then added to a falcon tube and 1 % (w/v) sodium lauroyl sarcosinate (Sigma-Aldrich) was added to this solution to remove the inner membrane. The solution was gently shaken to dissolve any remaining powder. The solution was transferred to ultracentrifuge tubes and after calibrating the weight of the centrifuge tubes, centrifugation was performed at 40,000 rpm for 1 hour at 4 °C (Beckman, rotor 70 Ti, or 45 Ti).

Pellets were resuspended in 25 mL of buffer A and the resulting solution was homogenised using the potter. To solubilize the membrane, 2 % n-dodecyl-β-D-maltoside (DDM)(Anatrace) was added to the solution and the solution was incubated for 1 hour at 4 °C on the rotating wheel. The solution was transferred to ultracentrifuge tubes and centrifugation was performed at 40,000 rpm for 45 min at 4 °C. A gravity flow column containing 3 mL of Ni-NTA agarose beads was washed three times with Milli Q water to remove the ethanol storage solution. 20 mL of Buffer A containing 1 % n-Octylpolyoxyethylene (OPOE) (Bachem) (Table 6) was added to the column to equilibrate it. The supernatant was mixed with the resin and they were incubated overnight at 4 °C on the rotating wheel. After packing the column, it was washed with buffer A containing 1 % OPOE. Protein was eluted by buffer B containing 1% OPOE.

SDS-PAGE was performed and samples containing the protein were pooled. This sample was then concentrated with a 15 mL Amicon concentrator with 50 kDa cut-off. The concentration process was achieved using three rounds of centrifugation for 15 min at 3000 rpm to reach a volume of 2 mL. The sample was pipetted after each round to avoid aggregation.

Thus ShuA purified first by affinity chromatography, was then purified by gel filtration chromatography using ÄKTAprime plus size exclusion chromatography, as detailed in the following section.

Table 6. Buffers for ShuA purification with affinity chromatography

Buffer A:

Component	Final concentration	Stock
Tris-HCl pH 8.0	50 mM	1 M
Imidazole	10 mM	1 M
NaCl	250 mM	5 M

Buffer B:

Component	Final concentration	Stock
Tris-HCl pH 8.0	50 mM	1 M
Imidazole	250 mM	1 M
NaCl	100 mM	5 M
OPOE	1 %	100 %

2.2.6.2 Size exclusion chromatography

For further purification of the affinity chromatography samples, size exclusion chromatography was carried out using an ÄKTAprime plus. As mentioned earlier in the case of AcrB purification (see 2.2.5.2), macromolecules are separated according to their size. As a first step, the column was washed with Milli Q water to remove ethanol and then equilibrated with one and half volumes of column buffer (Table 7). The protein sample was centrifuged for 10 min at 13,200 rpm at 4 °C and then injected into the column. The sample was passed through the column using one and half volumes of column buffer (180 mL). The eluted samples were collected in fractionation tubes. SDS-PAGE was performed and fractions containing ShuA were pooled and concentrated using an Amicon 50 kDa cut-off concentrator. This pure protein sample was then stored at 4 °C for further crystallization experiments.

Table 7. Buffers for ShuA purification with size exclusion chromatography

Component	Final concentration	Stock
Tris-HCl pH 8.0	10 mM	1 M
NaCl	50 mM	5 M
OPOE	1 %	100 %

2.2.6.3 Mono Q anion exchange chromatography

Similar to the AcrB purification process (section 2.2.5.3), Mono Q column anion exchange chromatography was used for further purification of ShuA samples obtained by affinity chromatography. However, the detergent used in the buffer for the purification of ShuA was OPOE (Table 8). After collecting the fractions and running an SDS PAGE gel, samples containing ShuA were pooled together. An Amicon tube concentrator with 50 kDa cut off was used for concentrating the samples and then the buffer was exchanged for 10 mM Tris-HCl pH 8.0 with 1.4 % OG (Anatrace).

Table 8. Buffers for ShuA purification with Ion exchange chromatography

Mono Q A: 200 mL

Component	Final concentration	Stock
Tris-HCl pH8	50 mM	1 M
OPOE	1 %	100 %

Mono Q B: 50 mM

Component	Final concentration	Stock
Tris-HCl pH8	50 mM	1 M
NaCl	1 M	5 M
OPOE	1 %	100 %

2.2.7 SDS-PAGE Analysis of proteins

SDS-PAGE was carried out under reducing conditions [114]. For ShuA and AcrB, 10 % and 12 % acrylamide gels were used respectively, prepared according to Table 9. Electrophoresis was achieved by setting a potential difference of 150 V between the anode and cathode of a Gel Electrophoresis System (Biorad) for approximately 1 hour or until the sample-loaded Bromophenol Blue had reached the base of the gel. Once this was completed, the gel was removed from the Electrophoresis System and stained using Coomassie Brilliant Blue solution for 30 min. As for the gel destaining solution, it contains 45 % (v/v) methanol and 10 % (v/v) acetic acid.

Table 9. Components for gel preparation for SDS-PAGE [115].

Component	12 % separation gel	10 % separation gel	Stacking gel
Water	1.6 mL	1.9 mL	1.15 mL
Tris-HCl 1.5M pH 8.8	1.3 mL	1.3 mL	0.5 mL 0.5M Tris-HCl pH 6.8
Acrylamide 30%	2 mL	1.7 mL	0.33 mL
SDS 20%	25 µL	25 µL	10 µL
APS 10%	50 µL	50 µL	20 µL
TEMED	5 µL	5 µL	2 µL

2.2.8 Western blot analysis

After SDS-PAGE analysis, an Immobilion-P Transfer PVDF Membrane (Millipore, Prod. No. IPV H00010, pore size 0.45 µm) was activated by pre-wetting it for 1 min in pure ethanol and then rinsed with TBS-T (20 mM Tris pH 7.5, 150 mM NaCl, 0.05 % Tween 20). To start protein transfers from the gel to the membrane, a Western blot sandwich was assembled; the potential difference was set at 100 V for 1 hour. In the next steps, the membrane was removed from the assembly and soaked in blocking buffer (TBS-T with 5 % low fat milk powder) for one hour at room temperature with continuous gentle agitation. The membrane was then i) washed three times (each time for 5 min) by using TBS-T, ii) incubated for 1 hour at room temperature with antibody (α-His HRP) diluted in a blocking buffer (1:2500), iii) washed again by TBS-T three times. Finally the membrane was soaked in 3,3-diaminobenzidine (DAB) staining solution for approximately 1 min and then this solution was removed using Milli Q water [116].

2.3 Results

2.3.1 Plasmid purification

There are several methods for isolating plasmid DNA from bacterial cells. Here the alkaline lysis method was used where bacterial cells are in contact with an anionic detergent at high pH. This condition opens the cell wall and denatures proteins and chromosomal, but not plasmid DNA. By

centrifuging the sample, the cell wall, denatured protein and chromosomal DNA are removed by sedimentation and soluble plasmid DNA can be recovered from the supernatant [112]. The purified plasmid DNA concentration was then obtained by measuring the absorbance at 260 nm by nanodrop spectrophotometer.

In the first experiment, plasmid DNA (45 µg/mL) from two colonies of *E.coli* bacteria modified to express the ShuA gene were extracted using a Nucleospin kit and sequenced using Sanger sequencing (GENEWIZ, Leipzig). Both plasmids had a complete DNA sequence for ShuA. Multiple sequence alignment and map of the pET20b are presented in appendix II.

In a second experiment, AcrB plasmids were purified by the same method for subsequent restriction enzyme digestion to determine that the lengths of the inserted DNA fragments were consistent with the genes to be expressed. From 10 mL cultures, the final concentration of extracted plasmid DNA for AcrB was 240 µg/mL as estimated from absorbance at 260 nm (see 2.2.1).

The restriction enzymes NdeI and XhoI were used for isolation/cutting of the AcrB sequence from the pET24 (5.31 kbp). For AcrB, a fragment with 3.15 kbp was expected [117]. Consistent with this a 1 % agarose gel (Figure 20) of digested plasmid samples features a clear band with a mass of ~ 3 kbp where AcrB DNA sequence is present.

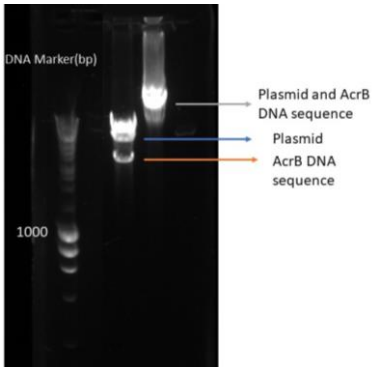


Figure 20. Agarose gel electrophoresis of plasmids and restriction digestion products. In Line 1, 5µL of the smart ladder was loaded. DNA migration is in the range: 10,000 bp to 200 bp (from top to bottom). Line 2 is plasmid with AcrB after double digestion. Line 3 represents the plasmid DNA containing the AcrB DNA sequence. 10 µL of the DNA sample was mixed with 2 µL of the gel loading buffer Blue (DNA 6x) and 10 µL of the final solution was added to the gel.

2.3.2 AcrB purification

Single colony of the *E.coli* BL-21(DE3) host strain containing the vector encoding for the AcrB was grown in LB supplemented with kanamycin (see 2.2.4). AcrB has His-tagging sequence so Ni-NTA affinity chromatography was used to purify the protein from the cell membrane extract. Ni-NTA bound protein was eluted with a buffer containing 200 mM Imidazole (buffer E). Figure 21 shows

the SDS PAGE analysis (see 2.2.7) of the eluted fractions after staining with Coomassie Blue. A dark band with a mass of ~ 100 kDa is visible in the E2-6 fractions with or without thermal denaturation, which corresponds to the predicted mass of 114 kDa of the AcrB protomer [22].

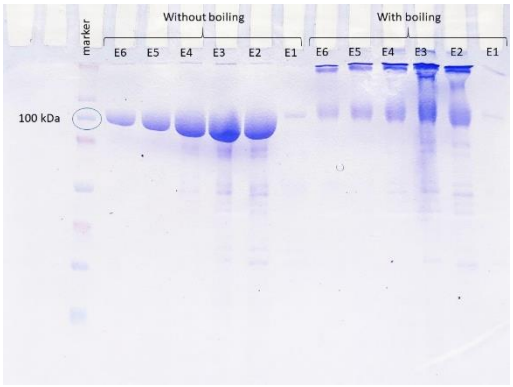


Figure 21. Purification of the AcrB by using Ni-NTA affinity chromatography. SDS PAGE gel of the sample after elution by buffer E containing 20 mM Tris-HCl pH7, 0.02 % DDM, 10 % Glycerol and 200 mM imidazole. Two sets of samples were loaded into the gel to check the effect of boiling the sample for SDS-PAGE electrophoresis.

On the basis of SDS-PAGE analysis, fractions containing AcrB were pooled and further purified using a Mono Q anion exchange column. A sodium chloride gradient was used to elute protein bound to the column. For this experiment, a flow rate of 0.5 mL/min was used (Figure 22) and fractions were collected by 1 mL. A peak absorbance at 280 nm was visible in the chromatogram from a concentration of 400 mM sodium chloride (Figure 22 A). Samples from fractions 6 to 18 corresponding to this peak (Figure 22 B) were analyzed by SDS PAGE.

Based on the chromatogram and SDS PAGE gel, fractions 6-18 were divided into three groups: 7 – 9 (group 1), 10 - 12 (group 2) and 13 – 18 (group 3) according to their level of purity. Each group was concentrated and used for separate crystallization experiments. Based on absorbance at 280 nm, 200 µL of 11.5 mg/mL AcrB was present in group 1, 100 µL of 7.5 mg/mL for group 2 and 100 µL of 11.5 mg/mL for group 3.

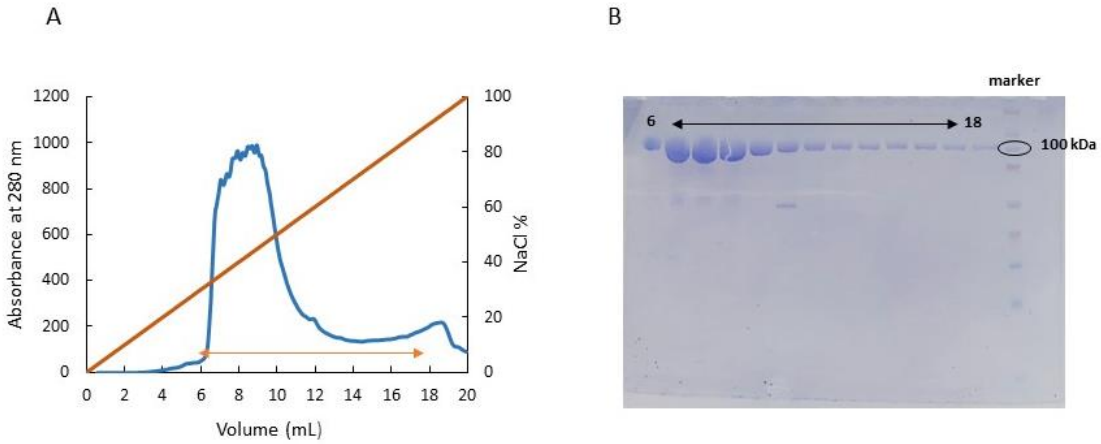


Figure 22. Purification of the AcrB by anion exchange chromatography. (A) Chromatogram (B) SDS PAGE gel. Orange arrow shows the fractions 6 to 18. For this experiment 10% SDS-PAGE gel was used. 20 µL of the sample was mixed with 5 µL of Coomassie Blue and 6 µL of the final solution was loaded into the gel.

A second AcrB purification was also performed for which size exclusion chromatography was added after affinity chromatography. Cell membrane extract was purified by Ni-NTA and analyzed by the SDS-PAGE as before. On the basis of the ~ 100 kDa bands present, fractions containing proteins were pooled and concentrated to 2 mL before injection onto a Superdex 200 size exclusion column. Second SDS-PAGE analysis was also performed and this gel was used for Western blot analysis (see2.3.2.1).

After collecting the AcrB fractions from size exclusion chromatography (Figure 23), SDS PAGE was used to analyze fractions 41 to 51, which give rise to a peak absorbance at 280 nm in the resulting chromatogram shown in Figure 23 A. On the other hand, the Figure 23 B shows the SDS-PAGE gel for AcrB samples after size exclusion chromatography. On the basis of the presence of a band corresponding to a species with a mass of ~ 100 kDa consistent with the 114 kDa mass of the AcrB protomer, fractions 42 to 51 were pooled and concentrated.

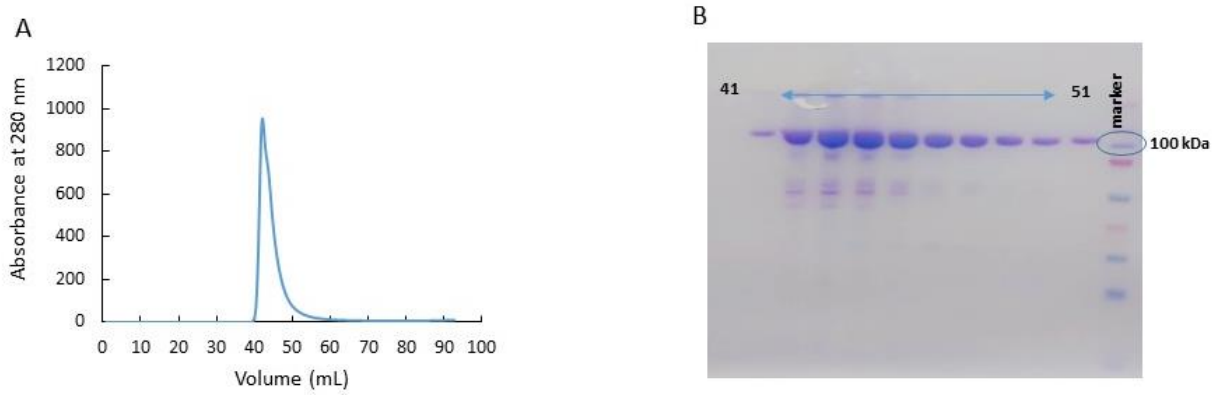


Figure 23. AcrB purification by size exclusion chromatography. (A) Size exclusion chromatogram and (B) SDS PAGE gel of the samples from fractions 41 to 51.

Concentrated protein (15.8 mg/mL) was kept at 4 °C for the next step of the experiment. Even after size exclusion chromatography, other bands are visible in the SDS-PAGE gel (Figure 23 B). In order to check if these bands correspond to AcrB molecules in a protein mixture, Western blot analysis was carried out.

2.3.2.1 Western blot analysis

Western blot analysis of the samples eluted after Ni-NTA chromatography is presented in Figure 24. Recombinant AcrB that is used here is a His-tagged protein, so for the Western blot α-His HRP (horseradish peroxidase) was used. DAB (3,3'-Diaminobenzidine) was used as a chromogenic

substrate for the α-His HRP giving rise to the brown colored precipitate visible on the membrane. As seen in Figure 24, all bands in the SDS-PAGE gel after AcrB purification have developed a brown color after Western blotting, meaning that all are likely to have a poly-His affinity tag. This analysis allowed us to clearly identify the presence of molecules belonging to the AcrB protein in the studied protein mixture.

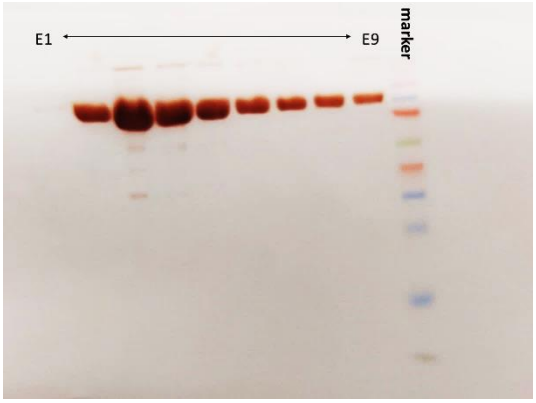


Figure 24. Western blot analysis of the AcrB samples after purification by affinity chromatography by Ni-NTA column.

2.3.3 Optimization of ShuA purification

In order to optimize ShuA purification two other experiments were performed where different detergents were used to solubilize the protein and where different conditions were used for Ni NTA elution. The results of these two experiments are presented in the following sections.

2.3.3.1 Using different detergents for ShuA purification

The homogenized cell membrane extract was divided into three equal aliquots of 20 mL. For solubilization in each aliquot, 2 % DDM, 10 % OPOE or 1 % Triton X-100 was added. All samples were stored at 4 °C for one hour on a rotating wheel. After ultracentrifugation, 20 µL of supernatant from each sample were mixed with 5 µL of Coomassie Blue and 15 µL of the solution was analyzed by SDS-PAGE (Figure 25). Fractions were selected based on the presence of a ~ 70 kDa band corresponding to the 70503.57 Da mass of the ShuA. A darker band is observed in lines 1 and 4 compared to the others, suggesting that DDM is the best detergent for solubilization of ShuA. Also, better crystallization results were obtained from this sample compared to the other protein samples.

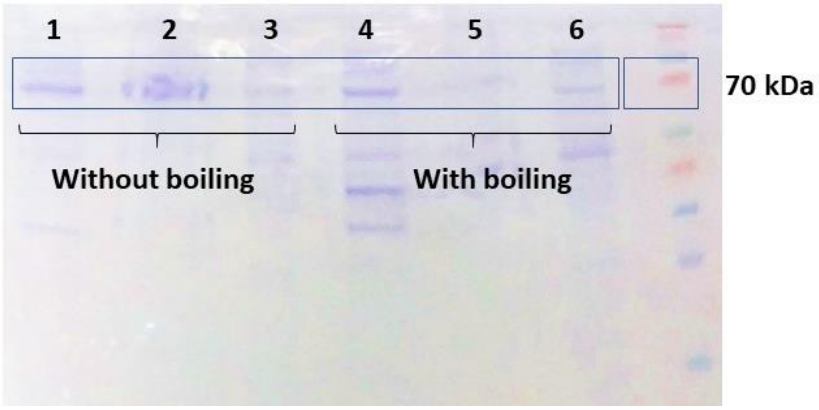


Figure 25. SDS-PAGE gel after solubilization of the membrane by three different detergents for purification of ShuA. 5 µL of the Coomassie Blue was added to 20 µL of the supernatant solutions containing the protein and 15 µL of each was loaded to the gel. Samples 1 and 4 contain DDM, 2 and 5 contain OPOE and samples 3 and 6 contain Triton X-100.

After this SDS PAGE, purification was continued using Ni-NTA column affinity chromatography.

2.3.3.2 Finding the best concentration of imidazole for elution buffer in affinity chromatography

The cell membrane extract containing ShuA was loaded onto a Ni-NTA column, and elution buffers with different concentrations of imidazole were compared. Imidazole competes with the histidine side chain in his-tagged proteins for binding to the Ni-NTA column. By increasing the concentration of imidazole, proteins that bound to the column were eluted and recovered in fractionation tubes. For this experiment, elution buffers with six different concentrations of imidazole were used sequentially. All buffers contained 100 mM NaCl and 50 mM Tris-HCl pH 8.0 and a different concentration of imidazole as summarized in Table 10.

Table 10. Imidazole concentration in elution buffer used in ShuA purification by affinity chromatography.

Buffer name	B1	B2	B3	B4	B5	B6
Imidazole (mM)	20	50	100	200	300	500

After washing a protein-loaded column with buffer A (washing buffer), 15 mL of each elution buffer was used. The fractions were collected and 25 µL of each was mixed with 5 µL of the Coomassie Blue and 15 µL of each sample was loaded onto a SDS-PAGE gel. The resolved electrophoresis gel is shown in Figure 26. A band of ~70 kDa corresponding to the mass of ShuA is visible in the gel lines corresponding to elution with B1-4 buffers, so 200 mM imidazole is sufficient to elute the protein.

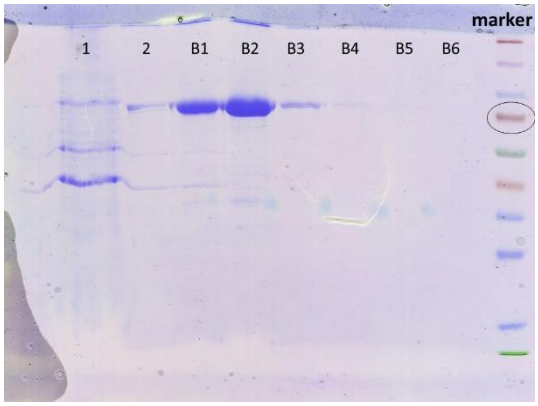


Figure 26. SDS page gel after using elution buffers with different concentration of imidazole in affinity chromatography for the purification of ShuA. Line 1 after packing the column, line 2 after washing the column with buffer A and line B1 to B6 after elution with buffers containing 20 to 500 mM imidazole.

2.3.4 ShuA purification

The *E.coli* BL-21(DE3) host strain containing the vector encoding ShuA was grown overnight in LB containing ampicillin. Cells were harvested by centrifugation and the pellets were washed in buffer containing 50 mM Tris-HCl pH 8.0. Two cell pellets were used for the purification step. After lysis of the cells by the microfluidizer, the sample was centrifuged to remove any unbroken cells. Solubilization of the inner membrane was achieved using 1 % (w/v) sodium lauroyl sarcosinate. For solubilization of the outer membrane containing ShuA, 2 % (w/v) DDM was used. After the second ultracentrifugation, the supernatant was mixed with the Ni-NTA column resin, previously washed with water and equilibrated with buffer A + 1 % OPOE, and incubated overnight at 4 °C on a rotating wheel. After packing the column and washing it with buffer A + 1 % OPOE, the fractions were collected by using the elution buffer. 20 µL of each fraction was mixed with 5 µL Coomassie Blue and 5 µL of the sample was loaded on a SDS-PAGE gel, shown in Figure 27.

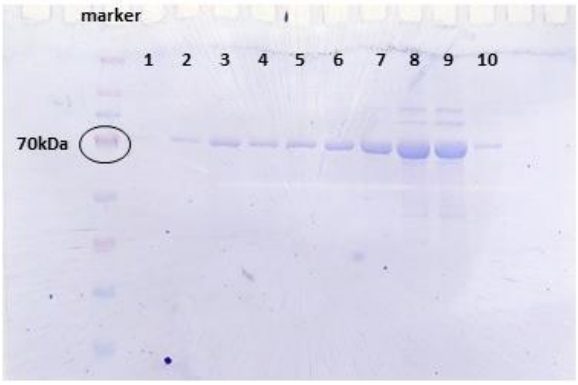


Figure 27. SDS-PAGE gels of ShuA samples after purification by Affinity chromatography with Ni-NTA column. The gels used for this experiment are 10 % acrylamide and 5 µL of each sample was loaded. Samples were eluted by buffer B containing 250 mM of imidazole.

Fractions 2 to 10 show a band with a mass of approximately 70 kDa in agreement with the expected mass for ShuA, 70 kDa [118]. Fractions 6 to 9 (sample 2) and 2 to 5 (sample 1) were concentrated separately. After concentration, around 100 µL of 4.9 mg/mL protein was present in sample 1 and 350 µL of 5.5 mg/mL in sample 2. To purify the protein further, either anion exchange or size exclusion chromatography were used.

In order to study the effect of sodium lauroyl sarcosinate on purification of ShuA, one of the purifications was performed by using affinity and anion exchange chromatography. In this example, after breaking the cells by microfluidizer and centrifugation, the sample was divided into two groups. In Group one, 1 % sodium lauroyl sarcosinate was added and in Group two, purification continued without this step. Both samples were purified by Ni-NTA affinity chromatography and then imidazole was removed by dialysis overnight in 50 mM Tris-HCl pH 8.0, 0.5 M NaCl with 1 % OPOE. A Mono Q anion exchange column was then used to purify both samples. Fractions of 1 mL were collected and a flow rate of 0.5 mL/min was used. For Group 1 a peak absorbance at 280 nm with a slight shoulder was observed once the salt gradient reached a concentration of 200 mM sodium chloride (Figure 28 A). The elution profile of Group 2 was identical (data not shown). Fractions were collected corresponding to the peak of the chromatogram and SDS-PAGE was used to analyze these fractions (Figure 28). The absorbance intensity at 280 nm correlates with the presence of a ~70 kDa band visible on the SDS-PAGE gel, which corresponds to the presence of ShuA.

Fractions 4 and 5 from Group one and 5 and 6 from Group two were concentrated using Amicon Ultra Centrifugal Filters 50 kDa. The final sample volume was 200 µL with the concentration ~ 6.3 mg/mL. Here only the chromatogram of Group one is shown.

Pure protein was obtained from both groups, although the bands in the gel are weaker when sodium lauroyl sarcosinate was used in the purification, suggesting that this step decreases the yield in protein production.

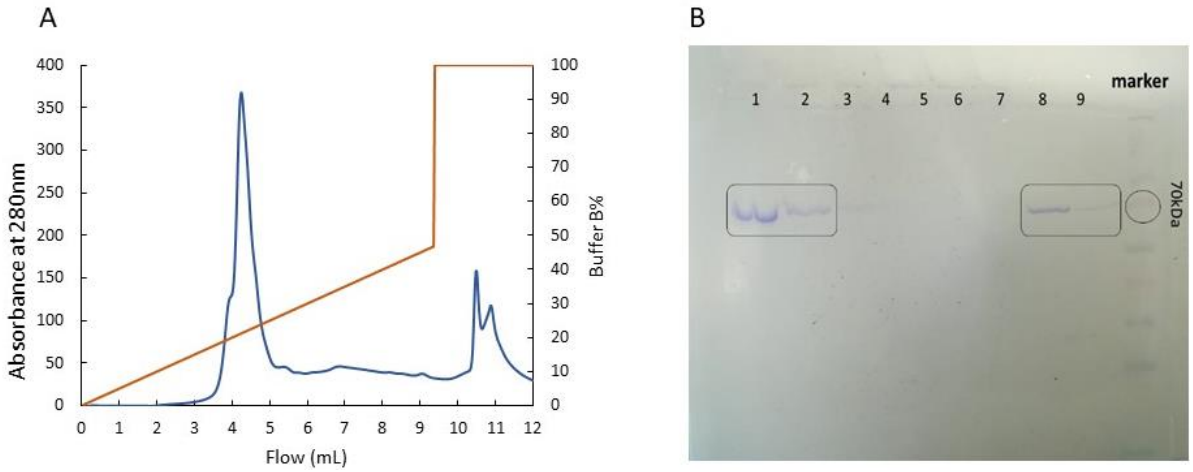


Figure 28. Purification of the ShuA by anion exchange chromatography. Increasing the sodium chloride concentration from 0 to 100 % on a 5 ml Mono Q column eluted protein fractions. (A) Chromatogram of Group one. (B) SDS-PAGE of fractions from Group one and two. Numbers one to three are fractions 6, 5, 4 of Group two and four to nine are fraction 10, 8,7, 6, 5 and 4 of Group one.

In the last purification presented here, size exclusion chromatography was used after affinity chromatography. The fractions from affinity chromatography were analyzed using SDS-PAGE and showed the typical presence of a ~70 kDa band indicating the presence of ShuA (Figure 29). Fractions 2 to 9 were pooled and concentrated to 2 mL. This was centrifuged at 13,200 rpm for 10 minutes at 4 °C to remove precipitated protein and then was injected onto the S200 column where it was eluted with 10 mM Tris-HCl pH 8.0, 0.5 M NaCl, with 1% OPOE.

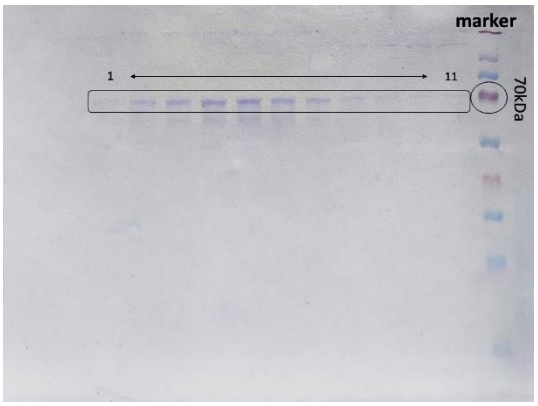


Figure 29. Purification of ShuA by affinity chromatography using a NiNTA column. A 10 % SDS PAGE gel was used and 5 µL of the solution containing 10 µL of the fraction and 2.5 µL of Coomassie Blue was loaded into each well of the gel. The samples were not boiled before loading onto the gel.

Fractions of 1 mL were collected at a flow rate of 1 mL/min. Fractions 67 to 74 corresponding to the largest absorbance peak of 280 nm in the chromatogram (Figure 30 A) were analyzed using SDS-PAGE (Figure 30 B). Each of these fractions showed a band of ~70 kDa corresponding to ShuA and no other species.

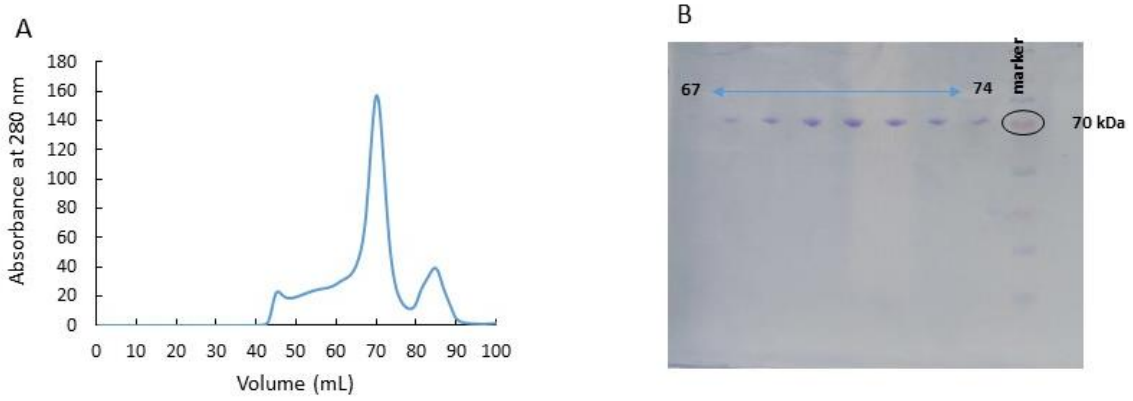


Figure 30. Purification of ShuA by size exclusion chromatography. 5 µL of the Coomassie Blue was added to the 20 µL of sample and 5 µL of the final solution after boiling was loaded onto the gel.

Samples 68 to 74 were therefore pooled together and concentrated to ~4 mg/mL in 10 mM Tris-HCl pH 8.0, 1.4 % OG. Then, the buffer containing OPOE was exchanged with the new buffer containing OG. For that, 5 mL of the new buffer was added to the Amicon tube at each round of centrifugation. To remove the buffer containing OPOE and avoid mixed micelles this process was repeated four times. This protein sample was then stored at 4 °C for crystallization.

2.4 Discussion

2.4.1 Protein purification

Protein purity is one of the biochemical variables [4] that can highly affect protein crystallization [3]. Although some proteins can be crystallized from a heterogeneous mixture, the likelihood of reproducible crystallization is significantly increased when pure and homogenous samples are used [4]. Longenecker et al. even suggest that when screening for the best crystallization conditions, the protein itself can be considered the most important variable, even more than the crystallization solution [119].

When purifying proteins for structural biology, one major goal is to preserve the native structure of those proteins. In the case of membrane proteins, structure preservation is a challenge, as they have both hydrophobic and hydrophilic regions. To stabilize membrane protein structure, detergents are generally used throughout purification. Detergents replace the lipids surrounding the protein which can lead to loss of native protein activity. As a result, screening different detergents is crucial to find the optimal conditions to both solubilize the membrane and keep the protein in its native and active state [120]. Apart from detergent, some other methods are also used for solubilizing these proteins. These methods aim to preserve the lipid surrounding the protein, often when preservation of native protein structure and activity is impossible when detergents are used. One such method is using styrene maleic acid (SMA) lipid particles (SMALPs). In this method, the protein and surrounding lipid are surrounded and encapsulated by SMA. SMALPs contain a central lipid bilayer supported by an outer SMA polymer. One limitation of this method is that the size of the protein cannot be more than ~400 kDa to fit easily inside the bilayer. Also pH values lower than 6.5 or divalent cations like Mg^{2+} and Ca^{2+} make the SMA polymer insoluble, disrupting the SMALPs [120]. Nanodiscs, formed of apolipoprotein A-I (ApoA-I) associated with a fragment of the phospholipid bilayer, are another species used to solubilize membrane proteins. Nanodiscs can accommodate a larger size range of membrane proteins as they have 9.8 to 17 nm diameter [121]. Using this method, proteins are solubilized with the lipid surrounding them retained and such samples will be monodisperse. However, this method often cannot be used to maintain the functionality of membrane proteins which require a different environment [122].

In this study, two model membrane proteins (ShuA from *S. dysenteriae* and AcrB from *E.coli*) were used. Although the purification of these proteins was achieved previously [117], [123], this process has been further optimized using detergent both for purification and to stabilize the protein in its native state for structural studies by crystallography. Thus, several optimization procedures are presented here which have been used over the course of the development of existing purification protocols.

2.4.1.1 AcrB purification

From a survey of the current AcrB literature, it is apparent that most research groups used a single-step purification by affinity chromatography [25], [117], [124]–[126] for the purification of AcrB. Some other groups used a two-step purification combining affinity and size exclusion chromatography [21], [127]–[129]. Some studies have also successfully employed SMALPs to solubilize AcrB with some of its native membrane for cryo-EM studies. In these reports, no aggregation or polydispersity was observed in protein samples [130], [131]. In this thesis, one- and two- step purification by affinity, size exclusion chromatography and anion exchange chromatography have been used to purify AcrB in the presence of detergent.

The recombinant AcrB used by most groups to date incorporates a poly histidine tag (His-tag) with a high affinity to nickel (II). Thus, in most cases, one step affinity chromatography using an immobilized nickel resin is sufficient to obtain a high purity protein [113]. Here a nickel-charged affinity resin (Ni-NTA agarose) was used for the purification of AcrB. Although SDS-PAGE indicated that a large amount of intact AcrB protein was present, other lower molecular weight species were also observed, suggesting that further purification steps would be appropriate to obtain a homogenous sample, such as gel filtration (size exclusion chromatography) or anion exchange chromatography.

Even after these steps, some bands were still observable on SDS-PAGE gels especially where a high concentration of AcrB is also present (Figure 22, and Figure 23). Western blot analysis was performed to determine if these bands are related to AcrB or not. As shown in Figure 24, all of these bands have a His-tag present suggesting that they are truncated forms of AcrB, rather than other co-purified proteins. One of the reasons for this phenomenon is related to the structure of the protein. AcrB is a homotrimeric protein and each protomer has a molecular weight of approximately 114 kDa [22], [125], [132]. This relatively large size and tendency towards oligomerization may increase the risk of aggregation during the purification process. Moreover, partially aggregated proteins can act as nucleation sites for other proteins leading to the formation of further aggregates. Such aggregates are likely to dissociate when samples are heated to denature them before SDS-PAGE. From Figure 21, there is a clear difference between heat-denatured samples and those denatured in SDS only. The SDS added to the protein breaks ionic bonds and also gives additional negative charge to the protein sample. Thus, different segments of a protein aggregate or complex are likely to be separated under denaturing SDS-PAGE conditions when they would be associated in solution.

It is worth noting that in crystallization experiments (examples of which are shown in the next chapter), protein samples prepared using one- or two- step purifications have clearly visible differences. Protein precipitate is significantly higher in experiments using protein from one-step purification compared to protein from two-step purification. Through a systematic and quantitative approach, the quality of the crystals generated from these purifications is compared. Such comparison is complicated because crystallographic results are affected by many variables which may or may not be a direct consequence of protein purity. For example, the different crystallization conditions from which the crystals are obtained, the size of the crystals, their robustness to freezing, the cryoprotection conditions and the characteristics of the beamline on which the data is collected are all important considerations.

2.4.1.2 ShuA purification

Previous purifications of ShuA protein used OPOE and Triton X-100 for solubilization [29], [123]. Here three different detergents (DDM, OPOE and Triton X-100) were tested to see if solubilization of the outer membrane could be improved, thereby improving the yield and stability of ShuA. SDS-PAGE analysis suggests that DDM is better at solubilizing ShuA in comparison to the other detergents. In addition, six different concentrations of imidazole were compared to optimize purification by affinity chromatography. Almost all the protein bound to a Ni-NTA agarose column is eluted using 250 mM of imidazole in the elution buffer. As with AcrB, additional gel filtration (size exclusion) and anion exchange chromatography steps were used in parallel experiments for comparison.

Using size exclusion chromatography, the final amount of pure protein obtained was 0.3 mg per liter of cell culture. When anion exchange chromatography was used as a second step for purification of ShuA, the final amount of protein was 0.6 mg per liter of cell culture. As such, twice as much protein was obtained using anion exchange chromatography compared to size exclusion. It seems using anion exchange chromatography in the a second step, is the better choice when higher amount of pure protein is needed.

However, for membrane proteins the concentration of the detergent and the conditions of the buffer (such as salt concentration or the pH of the buffer) are critical for successful purification with a good yield of protein. This can also have an effect on subsequent crystallization experiments and structural studies. These conditions vary for each membrane protein and are likely to be something that needs to be optimized in the purification process.

Using such optimization, the existing purification protocols for ShuA and AcrB have been adapted to significantly improve the purity of protein obtained and there is evidence that this may also improve the diffraction quality of crystals of these proteins.

Generally, two-step purification yields pure, high-quality proteins suitable for structural studies. However, the second step can be omitted to increase yield and save time when sample purity is less critical, such as for initial screening of suitable crystallization conditions.

Chapter 3: Crystallization and optimization of crystal growth

3.1 Introduction

From the thermodynamic and kinetics point of view, the (protein) crystallization process consists of at least two steps that must be considered and treated separately: nucleation and crystal growth [133]. In the first step of the crystallization process protein molecules associate and form a “critical nucleus” that will develop in the next step to a crystal [133]. The crystal growth stage consists of the addition of new protein molecules or its oligomers into the characteristic arrangement of the crystal lattice. Any factor that prevents the regular arrangement of molecules within the crystal lattice, such as low protein purity or homogeneity, affects the quality of the crystals [4]. Many physical, chemical and biochemical parameters must therefore be taken into account during the crystallization process [32]. These parameters will have a direct impact on the solubility of the protein being studied [52]. A protein crystallization phase diagram can therefore be used to study the crystallization process as a function of each variable considered. (see 1.1.3.1).

In order to generate high quality crystals, it is important to control the entire crystallization process. The best strategy is to provoke the nucleation in the vicinity of the metastable zone to avoid too much nucleation [133]. Different methods can be applied to bring the solution into the nucleation zone to enable crystallization, each using a different kinetic pathway in the phase diagram [65] (see 1.1.3.4). In addition, different variables exist that control the supersaturation state of sample. As crystal growth is a multi-parametric process [133], knowing each parameter and how it affects the crystallization process is crucial. Unfortunately, for the majority of proteins this information is currently poorly characterized and crystallization phase diagrams are only available for a small number of model protein systems [53]. As such, empirical optimization is instead necessary in order to obtain well diffracting crystals. During the optimization process, the effect of various experimental variables on the protein is examined to find the best condition for nucleation and crystal growth.

This chapter is devoted to the crystallization of model soluble and membrane proteins and the optimization of crystal growth for different crystallographic techniques (X-ray, neutron and electron diffraction). First, the crystallization results of different conventional and non-conventional crystallization methods are compared. Secondly, the effect of different variables such as temperature, pH, salt and protein concentrations on protein crystallization is demonstrated. Finally, the effect of varying these parameters is illustrated via the kinetic path taken in the qualitative phase diagram to reach the nucleation zone for each protein studied.

Soluble proteins used for crystallization experiments are from Sigma-Aldrich and Roche. Model membrane proteins were purified according to the protocols in 2.2.6 for ShuA and 2.2.5 for AcrB. ShuA crystallization was carried out after exchanging the buffer with 10 mM Tris-HCl pH 8.0 and 1.4% OG. AcrB crystallization was carried out after exchanging the buffer with 10 mM Hepes pH 7.0 and 0.02% DDM. Reservoir volume for the vapor diffusion technique was 500 µL and for the dialysis method was 2 mL. All chemicals are purest grade (>99%) and a Millipore ‘Milli Q’ water

purification system was used to deionize water (resistivity = 18.2 MΩ·cm) for sample preparation. Protein solutions were centrifuged at 13,200 rpm for > 10 min before use to remove any precipitate and chemical solutions were filtered using 0.2 µm filters. It is worth noting that all phase diagrams presented in this work are qualitative and the solubility curves are representative and have not been measured experimentally.

3.2 Methods and Results

3.2.1 Protein crystallization

Conventional and unconventional crystallization techniques were employed for crystallization of model soluble (Lysozyme and Thaumatin) and membrane proteins (AcrB and ShuA). LCP method and preliminary results obtained from crystallization of lysozyme and ShuA with this method are in appendix III.

3.2.1.1 Vapor diffusion

The vapor diffusion method (hanging drop and sitting drop) (see 1.1.3.3.1) was used for crystallization of both model soluble and membrane proteins. 500 µL of crystallization solutions were added to the wells of a 24 well crystallization plate. A droplet of crystallization solution was added to each clean coverslip (hanging drop) or the bridge inserted inside the well of the crystallization tray (sitting drop) and then a droplet of the protein sample was added to droplets of crystallization solution. In the hanging drop method, the coverslips with droplets were then flipped and fixed on top of the well containing the appropriate crystallization solution. For the sitting drop, after adding the protein the well was sealed by a clean coverslip to ensure a tight seal. Volume of the drop and protein: crystallization solution ratio is different for each protein. These ratios for each experiment are presented in the relevant section.

3.2.1.1.1 Lysozyme

Both hanging drop and sitting drop vapor diffusion were used for crystallization of lysozyme with 30 mg/mL concentration. Crystallization solutions contained a range of concentrations of sodium chloride (0.7 to 1.2 M) and a constant concentration of sodium acetate (0.4 M). The final volume of the crystallization solution in each well was 500 µL. Lysozyme crystals obtained with vapor diffusion are shown in Figure 31. Here, increasing the sodium chloride concentration leads to higher supersaturation that moves the solution further in the nucleation zone, as a result the number of crystals in the drops increases and their size decreases. The same trend is visible using both hanging drop (row A) and sitting drop (row B) methods. In both experiments the same protein: crystallization solution ratios were used with different overall volumes (5µl:5µl for hanging drop and 10µl:10µl for sitting drop).

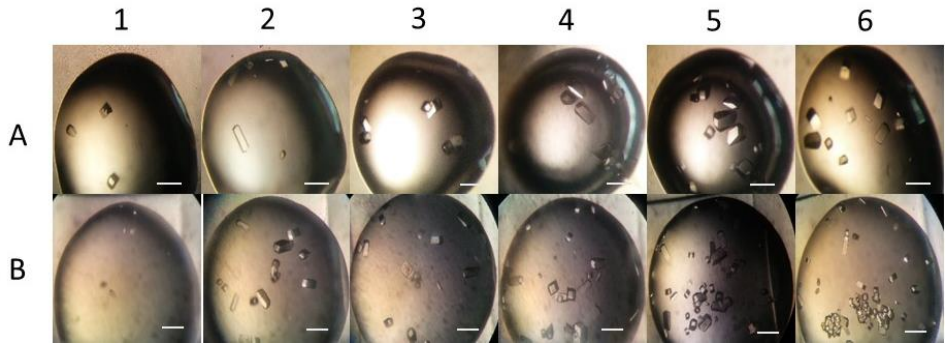


Figure 31. Crystallization of lysozyme by vapor diffusion. Salt concentration increases from 0.7 to 1.2 M from column 1 to 6. The experiment was performed at 20 °C with 0.4 M sodium acetate pH 4.0. Row A shows results from hanging drop while row B from sitting drop crystallization experiments. Scale bar represents 300 µm.

3.2.1.1.2 Thaumatin

The hanging drop vapor diffusion method was used to crystallize 15 mg/mL thaumatin in 10 mM ADA pH 6.5. Crystallization solutions consisted of a range of concentrations of sodium/potassium tartrate (0.85 to 1.1 M) in the ADA buffer (50 mM ADA pH 6.5). In this experiment (Figure 32), two sets of drops were prepared. In the first set (row A), protein volume to crystallization solution volume ratio is 2µl:2µl and for the second set (row B) this ratio is 4µl:2µl. In both experiments, prismatic crystals are observed. The drops show the same tendency as observed for lysozyme crystals, i.e. with increasing supersaturation the number of nuclei and consequently the number of crystals increases in the drops.

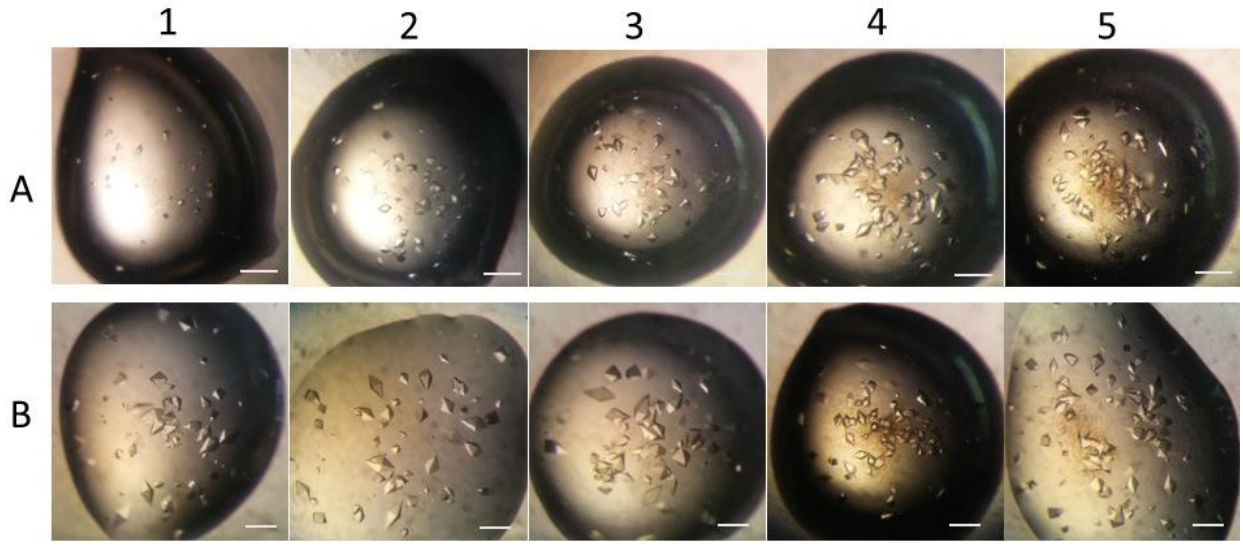


Figure 32. Crystals of Thaumatin formed using the hanging drop vapor diffusion method. In row A the protein volume to crystallization solution volume ratio is 2:2 and in row B it is 4:2. From A (B) 1 to A (B) 5 salt concentration was increased from 0.85 to 1.1 M. Images are taken after three days of incubation at 20 °C. Scale bar represents 300 µm.

Figure 33 (A) demonstrates how to screen different conditions in a single drop using the hanging drop vapor diffusion method by changing the ratio of protein to crystallizing agent concentration in the crystallizing solution mixture until equilibrium is reached. Also, the kinetics trajectories of two samples with the same salt concentration but different protein volume: salt volume ratios are illustrated.

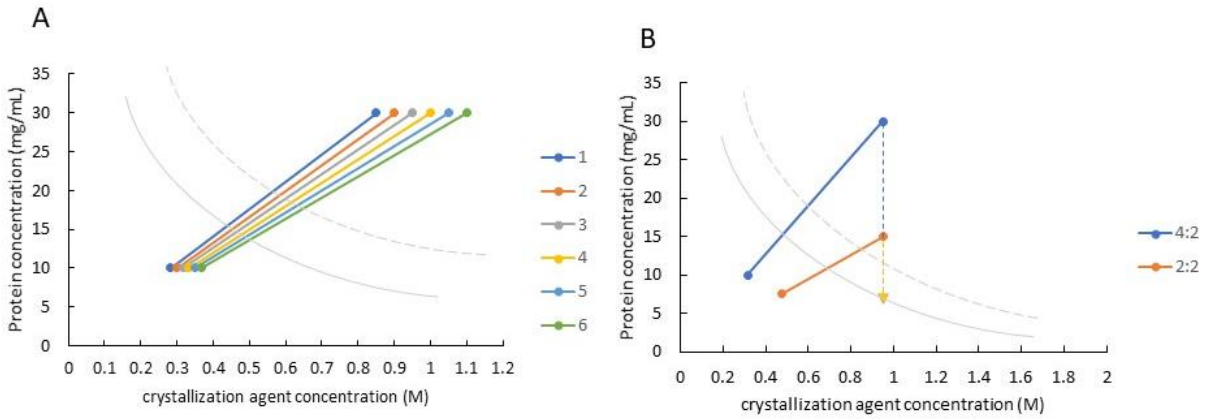


Figure 33. Schematic crystallization phase diagram in case of hanging drop vapor diffusion method. A. Kinetic pathways for samples with protein volume to the salt volume ratio 4:2. From 1 to 6 final concentration of the Na/K tartrate increased from 0.85 to 1.1 M with 0.05 intervals. B. Comparison between kinetic pathways for samples with protein to the salt volume ratio 4:2 and 2:2, final concentration of the Na/K tartrate is 0.95 M. Solubility curve is qualitative.

Because of the different protein to crystallization solution volume ratios, the initial salt concentrations in row B are less than row A but protein concentrations are higher (Figure 33, B). The final protein concentrations in drops in row B are higher than row A, while the final salt concentrations are the same in both groups (0.85 to 1.1 M). Here increasing the protein concentration leads the solution in the final state to go further into the nucleation zone of the phase diagram (to higher supersaturation). In agreement with this, more crystals can be observed in line B compared to line A for both experiments. Furthermore, as there is a higher amount of protein in the drop B and the equilibration rate is altered (related to the different ratios of protein volume: precipitant volume in the drop), an increase in crystal size can also be observed.

3.2.1.1.3 AcrB

The vapor diffusion hanging drop method was also used for crystallization of AcrB. During the optimization process different salts, PEGs, buffers and additives were screened (see appendix IV for some of the crystallization conditions used for the crystallization optimization). Figure 34 shows examples of crystals obtained from protein purified using affinity chromatography and size exclusion chromatography (A,B) compared to those crystallized from protein purified using affinity chromatography alone (C,D). In the crystallization drops shown in Figure 34C and 34D, more precipitate is observed, compared to Figure 34A and 34B, which show less precipitate,

possibly reflecting a different level of protein purity between the batches of protein used in these crystallization trials. Furthermore, in sample B, the crystals generated are needle-shaped, whereas in sample D, they are cubic. Similar needle-like crystals were also observed using the same crystallization conditions as shown in Figure 34D but with a twofold lower protein concentration (6.25 mg/mL, image is not shown). It seems that the packing and/or the geometry of the crystals are affected by the concentration of the protein, which can therefore be considered as a variable in the crystallization. Interaction between hydrophilic regions of protein molecules is important in crystal packing when detergent is used for growing crystals (*in surfo* method (see 1.1.4.3)) [134]. Increasing the protein concentration by increasing the supersaturation [52] may increase the chance of this interaction and different crystal packing. However, because crystallization is multiparametric, different variables affect this process simultaneously for example purity of the protein, sample ageing and crystallization sample preparation.

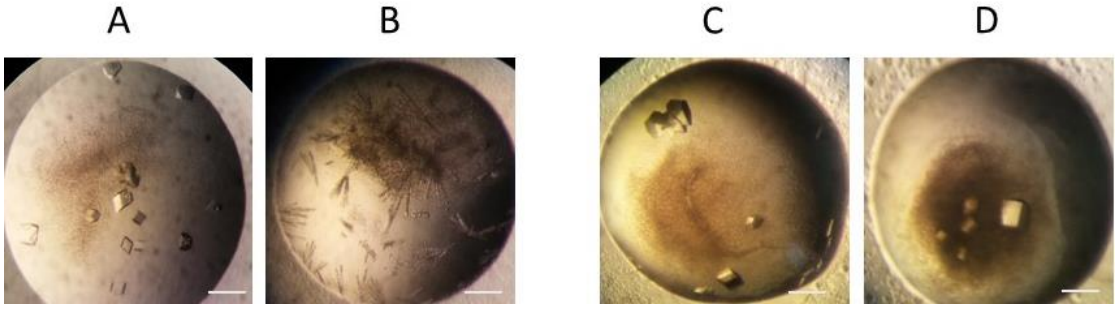


Figure 34. Crystals of AcrB obtained by hanging drop vapor diffusion method. Protein for (A) and (B) were obtained from affinity and size exclusion chromatography purification protocol whereas sample (C) and (D) are from affinity chromatography alone. Crystallization solution contains 9% PEG 3350 and 0.1 M Li_2SO_4 for (A) and (C) or 0.1 M $MgSO_4$ for (B) and (D) respectively. Protein concentration is 10 mg/mL for samples (A) and (B) and 12.5 mg/mL for (C) and (D) respectively. All volume ratios between the protein and the crystallization solution are 1:1. Scale bar represents 500 µm.

3.2.1.1.4 ShuA

After ShuA was purified in a buffer containing OPOE, it was exchanged into 10 mM Tris-HCl pH 8.0 and 1.4% OG for crystallization. The crystallization solution consisted of 0.1 M MES pH 6.5, 0.1 M sodium chloride or sodium acetate with 15 % PEG 1000 [123]. The samples for the crystallization experiment in Figure 35 were obtained after affinity chromatography followed by anion exchange chromatography using a mono Q column (A), size exclusion chromatography (B) and no further purification (C). According to the previous work [123] crystals represented in Figure 35 B and 35C were obtained by co-crystallization of the protein with Europium (III) chloride (see 3.2.2.1.1). Crystals have almost the similar shape despite being obtained from different purification procedures. The quality of these crystals will be discussed in chapter 4.

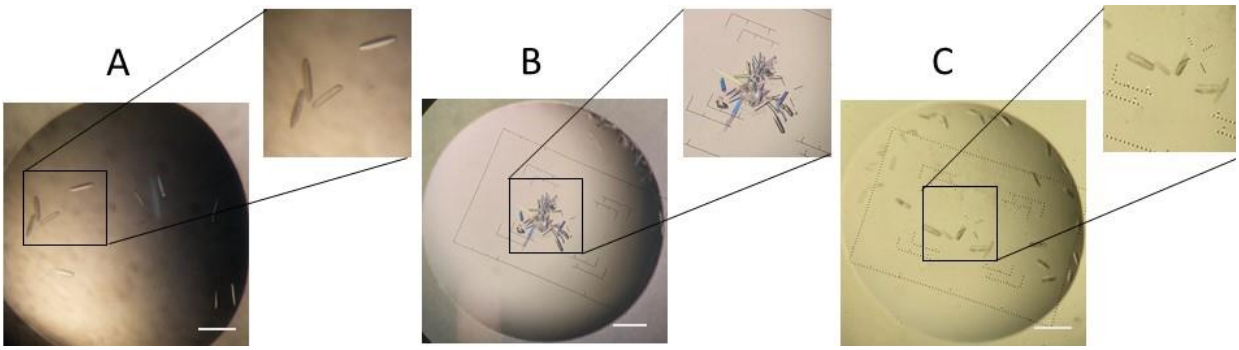


Figure 35. Crystals of ShuA obtained using the hanging drop vapor diffusion technique. Crystallization solutions contain 0.1 M MES pH 6.5, 0.1 M salt and 15% PEG 1000. A) 6.25 mg/mL ShuA purified using affinity and anion exchange chromatography. Drop ratio (protein volume:reservoir volume) is 2:2 with 0.1 M sodium chloride. B) 3.6 mg/mL ShuA purified using affinity and size exclusion chromatography. Drop ratio (protein volume:reservoir volume) is 1:1 with 0.1 M sodium acetate. C) 5.5 mg/mL ShuA purified using affinity chromatography only. Drop ratio (protein volume:reservoir volume) is 1:1 with 0.1 M sodium chloride. Scale bar represents 200 µm.

3.2.1.2 Dialysis

Model soluble and membrane proteins were pipetted to the chamber of the dialysis buttons with the desired volume. Dialysis membrane was used to cover the chamber and then was fixed using an elastic O-ring. Buttons were then transferred to the wells of a 24 well crystallization plate previously filled with 2 mL of crystallization solutions. Each well was sealed using vacuum grease and coverslip. A detailed protocol for the crystallization and optimization of crystal growth in dialysis buttons and using the crystallization bench (OptiCrys) in our laboratory has been recently published [133] (Appendix I).

3.2.1.2.1 Lysozyme

The whole protocol used for the crystallization of lysozyme by dialysis as well as the presentation of the results with crystals obtained are presented in appendix I.

3.2.1.2.2 Thaumatin

Thaumatin crystallization conditions identical to those used in vapor diffusion were used with the dialysis method, 15mg/mL protein in 0.5 M ADA pH 6.5. The crystallization agent is sodium/potassium tartrate, in a range of concentration from 0.4 to 0.6 M with 0.05 M interval. Thaumatin crystals obtained from dialysis are shown in Figure 36. Although the dialysis method was used under microgravity for crystallization of Thaumatin [135], it is not a common and routine method for crystallization of this protein. Thaumatin crystals obtained under microgravity are larger in comparison to the crystals obtained on the Earth under similar/same crystallization conditions [135], which could be due to the lack of convection and absence of sedimentation of impurity on surfaces of the crystals under microgravity [136], [137].

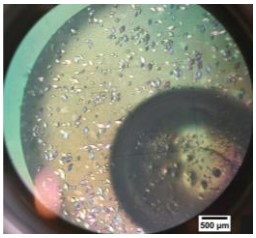


Figure 36. Crystals of Thaumatin obtained using the dialysis crystallization method. Protein concentration is 15 mg/mL and the crystallization solution is 0.45 M sodium/potassium tartrate, 0.05 M ADA pH 6.5. Dialysis button volume is 30 µL.

3.2.1.2.3 AcrB

Building on previous conditions used successfully in hanging drop vapor diffusion (see 3.2.1.1.3), dialysis method was also used for crystallization of AcrB. Two different concentrations of the protein (6 and 12 mg/mL), and different salts and buffers were used for this experiment (detailed fully in appendix IV). In Figure 37, one example of the crystals obtained from this method is presented. In the series of experiments performed with AcrB, no needle shape crystals were observed when the dialysis method was employed and generated crystals were much thicker and robust in comparison to the crystals obtained with other techniques that were used in this study. AcrB crystal quality is discussed further in chapter 4.

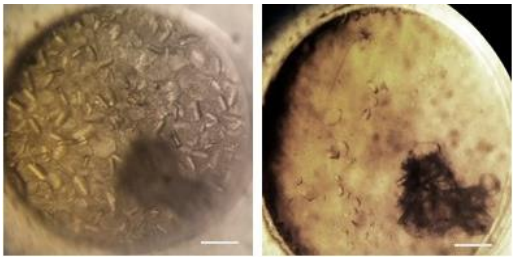


Figure 37. Crystals of AcrB obtained using the Dialysis method. Protein concentration is 6 mg/mL and the crystallization solution is 10% PEG 4000, 0.5 M ADA pH 6.5, 5% Glycerol, 0.2 M Ammonium sulfate, and 0.02% DDM. Dialysis button volume was 15 µL. Scale bar represents 500 µm.

From these experiments, it appears that the size of PEG affects the crystallization of AcrB. Larger crystals are formed when PEG 4000 is used in the crystallization solution, compared to the low molecular weight PEGs (400 and 300) previously used for AcrB crystallization by the vapor diffusion method. Taking into consideration the size of our protein, the higher molecular weight PEGs appear to have a proportionally greater efficiency to force protein out of solution [32]. In addition, diffusion speed of the different solutes is different, in particular the one of the different PEGs [138], [139].

3.2.1.2.4 ShuA

ShuA was purified with size exclusion chromatography as an additional step to increase purity (see 2.2.6.2) and the buffer was exchanged for 10 mM Tris-HCl pH 8.0 and 1.4% OG. Crystallization of ShuA in the presence or absence of heavy atoms was performed using the dialysis method. Crystallization solution was added to each well of a crystallization tray. ShuA was added to the dialysis button and the protein sample was covered by a dialysis membrane with a 12-14 kDa cut off. After this, the button was transferred to a well containing the crystallization agents, detergent and heavy atom salt. As detergents are crucial for maintaining native protein structure, they should be added to the crystallization solution. If the equilibrium between the compartments is not maintained, the concentration of detergent could decrease with respect to its CMC, therefore the protein-detergent complex will be destabilized and the protein will precipitate. Adding heavy atom salt to the crystallization solution in order to maintain its interaction with the protein is crucial too. This well was then sealed using a coverslip. The crystallization tray was incubated at 20 °C. In Figure 38 crystals of ShuA obtained by co-crystallization with Europium (III) chloride ($EuCl_3$) are presented. These crystals are similar to the crystals obtained with hanging drop vapor diffusion (See Figure 35) but are smaller in size.

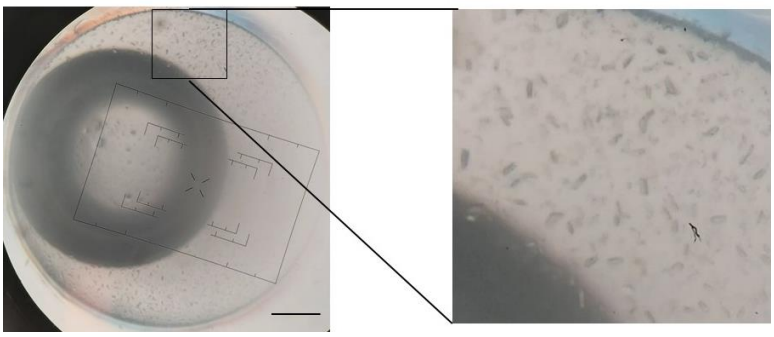


Figure 38. Crystals of ShuA obtained from the dialysis method in presence of Europium (III) chloride. Protein concentration was approximately 4 mg/mL with a protein to heavy atom molar ratio of 1:5. The crystallization experiment was performed in presence of 2 mL of crystallization solution containing 0.1 M MES pH 6.5, 0.1 M sodium acetate, 12% PEG 1000, 1.4% OG and 0.2 mM $EuCl_3$. Dialysis button with 15 µL volume was used. Scale bar represents 500 µm.

In this experiment, after mixing the heavy atom salt solution with the protein solution in a 5:1 or 10:1 (data not shown) molar ratio, a white precipitate forms that is not observed in parallel experiments without heavy atoms. This was removed by centrifuging the sample for 10 min at the 13,200 rpm at 4 °C and retaining just the supernatant for use in the subsequent crystallization experiment. ShuA crystals formed in the absence of heavy atoms were generally larger but diffracted at lower resolution (chapter 4). Co-crystallization is fully detailed in section 3.2.2.1.1.

3.2.1.3 Crystallization batch

For the batch method protein solutions were simply mixed gently with the crystallization solution in the appropriate ratio targeting spontaneous nucleation or crystal seeding experiment.

The batch crystallization method was used to grow ShuA and AcrB crystals in the presence of various additives detailed fully in section 3.2.2.1 where the entire sample preparation is described.

3.2.1.3.1 AcrB

Protein and rifampicin with a molar ratio of 1:10 were incubated overnight at 4 °C (0.11 mM AcrB, 1.1 mM rifampicin). The following day this sample was centrifuged for 10 min at 13,200 rpm and 4 °C to remove any possible dust or aggregate before starting the crystallization experiment. 50 µL of each crystallization solution was then added to 1.5 mL Eppendorf tubes to reach final concentration of 5 to 10% PEG 4000, 0.05 M ADA pH 6.5, 5% glycerol, 0.2 M ammonium sulfate, 0.5 mM rifampicin and 0.02% DDM in the final solution. Then 50 µL of protein-rifampicin solution was added slowly to the tube and mixed by pipetting. Samples were incubated at 20 °C.

For this experiment, four different concentrations (5 mg/mL, 7 mg/mL, 10 mg/mL, 13 mg/mL) of protein in 0.01 M HEPES pH 7.0 and 0.02% DDM with rifampicin were prepared.

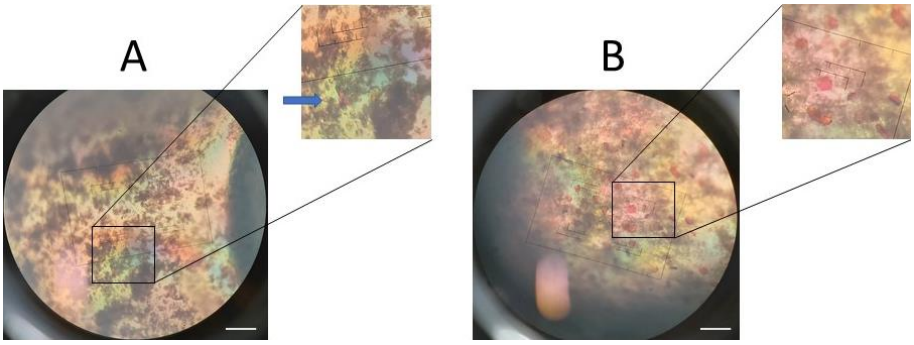


Figure 39. Co-crystallization of AcrB with rifampicin with batch crystallization method. Crystallization solution is 10% PEG 4000, 0.5 M ADA pH 6.5, 5% glycerol, 0.2 M ammonium sulfate, 1 mM rifampicin and 0.02% DDM. Protein concentrations are 6.5 mg/mL (A) and 2.5 mg/mL (B). The final volume of the mixture is 100 µL. Scale bar represents 500 µm.

As shown in Figure 39, the crystals obtained with a higher concentration of the protein (A) are numerous and smaller. Although the kinetic trajectory for both samples is similar, as the batch method is used in both cases, the starting points are different (see Figure 40) due to the different protein concentrations. In the batch method, the final supersaturation is achieved immediately after mixing the protein and crystallization solution and the sample is in its final state after mixing. According to the starting point, various numbers or morphologies of crystals can be achieved.

The crystals obtained using the batch method appeared more fragile in comparison to the crystals obtained using the vapor diffusion or dialysis methods. This could be due to the different kinetic pathways and supersaturation rates achieved by these methods during the crystallization process (see Figure 12 and Figure 40) allowing the features of the generated crystals to be controlled (e.g. size, morphology and number of the crystals) [11,12].

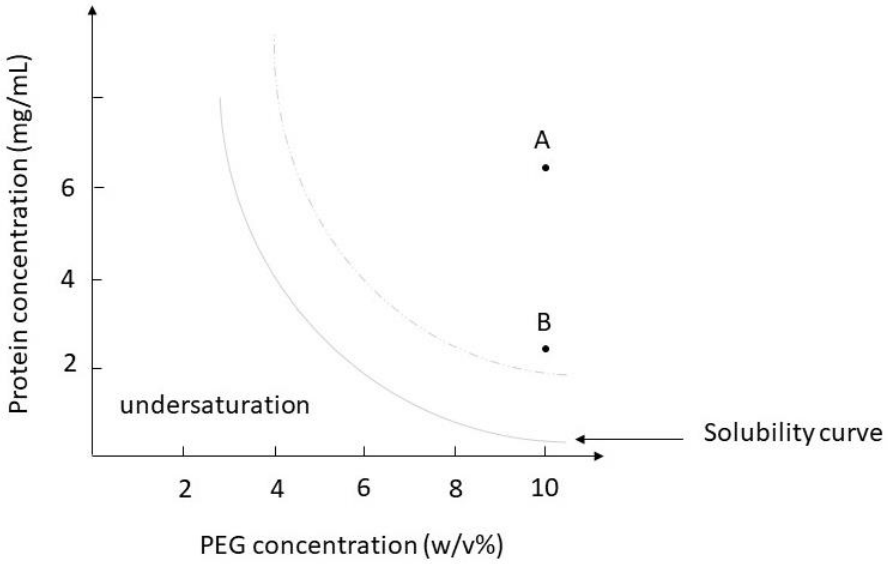


Figure 40. Kinetic pathway of protein crystallization by crystallization batch method. AcrB-Rifampicin concentration for sample A is 6.5 mg/mL and B is 2.5 mg/mL. Crystallization solutions contain 0.5 M ADA pH 6.5, 5% glycerol, 0.2 M ammonium sulfate, 1 mM rifampicin, 0.02% DDM and 10% PEG 4000. The solubility curve is qualitative.

3.2.1.3.2 ShuA

For crystallization of ShuA using the batch method, first ShuA was mixed with 8 µL of europium chloride (stock concentration 10 mM) to reach a 10:1 molar ratio (europium chloride: ShuA). Crystallization solutions were prepared in order to obtain 0.1 M MES pH 6.5, 15% PEG 1000, 0.1 M sodium acetate, 0.8 mM europium chloride, and 1.4% OG. After centrifuging the protein, 50 µL of it was added to an Eppendorf tube containing 50 µL of crystallization solution. After mixing the sample gently, tubes were stored at 20 °C.

In contrast to AcrB, in all ShuA crystallization solutions, no crystals or precipitates were observed once a similar white precipitate as the one observed for ShuA dialysis experiments had been removed by centrifugation. This observation inferentially indicates that the final mixture solution was either in the undersaturated or metastable region of the phase diagram and therefore nucleation could not take place. This is likely to have occurred if the removed precipitate from the solution contained protein. Therefore, protein concentration in the final crystallization solutions is significantly lower than the concentration needed for the nucleation.

3.2.1.4 HiLiDe method

Two samples with protein to lipid mass ratios of 3:1 and 2:1 were prepared for re-lipidation of AcrB and ShuA (see Figure 41 for a schematic representation). For AcrB a 9.8 mg/mL protein stock and a 40 mg/mL stock solution of 1,2-dioleoyl-sn-glycero-3-phosphatidylcholine (DOPC) in chloroform ($CHCl_3$) was used. In order to obtain the desired ratios 3:1 and 2:1, ~9 µL and ~13.5 µL of the DOPC stock solution were used with 110 µL stock protein, respectively. For ShuA a 10.46 mg/mL stock was used. In samples with 100 µL protein, ~8.7 µL (for 3:1 ratio) and ~13.1 µL (for 2:1 ratio) of the DOPC stock solution was added.

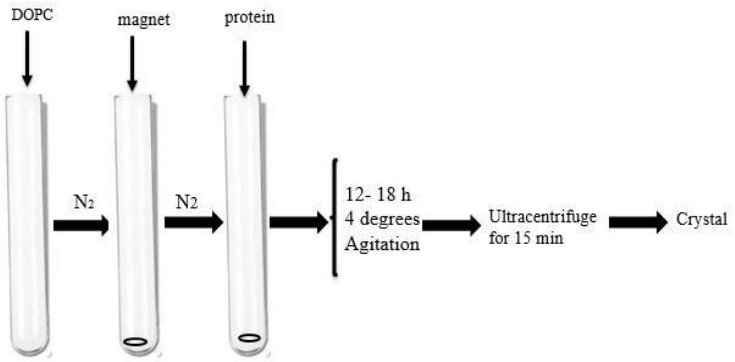


Figure 41. Re-lipidation of membrane proteins using the HiLiDe method. (From left to right) 1,2-dioleoyl-sn-glycero-3-phosphatidylcholine (DOPC) in chloroform is added to a tube which is dried by nitrogen gas. A magnetic stirrer is added and the tube dried again. Then the protein solution is added to the tube. The tube is cooled for overnight incubation with stirring. Ultracentrifugation is used to sediment any solid matter so the supernatant is suitable for crystallization experiments.

First, a glass tube was dried using nitrogen gas at room temperature and then lipid was added to create a thin layer at the bottom of the tube. Then a magnetic stirrer and afterwards the protein sample were added to the tube. After closing the tube with a cap and sealing it with a piece of parafilm, the sample was then stirred at low speed (50 rpm) overnight (12 to 18 hours) at 4 °C. The following day, the sample was ultracentrifuged at 70,000 rpm for 15 min at 4 °C [78]. Then the supernatant was removed and used for crystallization by hanging drop vapor diffusion.

For the crystallization of AcrB, polyethylene glycol (PEG) and polyethylene glycol monomethyl ether (MME) with average chain lengths 400 and 550 MME at 4 different concentrations (10 %, 20 %, 30 %, 40 %) and 6 different salts at two different concentrations (100 mM and 200 mM) were used. In hanging drop vapor diffusion experiments, each drop contained 0.9 µL of the protein sample and 0.9 µL of the crystallization solution.

For the crystallization of ShuA, PEG 2000 at four concentrations (8 %, 10 %, 12 % and 14 %) and six different salts at 0.1 M or 0.2 M concentrations were used. (Appendix V summarizes the crystallization conditions tested).

3.2.1.4.1 AcrB

After sample preparation and overnight re-lipidation of AcrB, white precipitate was observed (Figure 42). Before starting the crystallization process, samples were ultracentrifuged and just the precipitate-free supernatant was used for crystallization with the hanging drop vapor diffusion method using 0.9 µL protein and 0.9 µL reservoir solution for each drop.

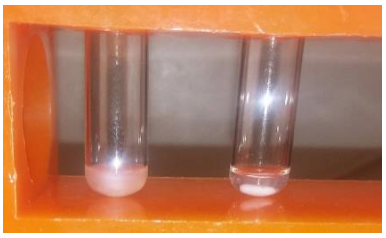


Figure 42. Re-lipidation of AcrB with a protein:lipid ratio of 3:1 (right sample) and 2:1 (left sample). White precipitate, likely to be precipitated lipid.

Four plates were prepared (with conditions that are shown in appendix V), but no crystals were observed. AcrB was purified again and concentrated using an Amicon centrifugal filter with a 100 kDa cut off instead of 50 kDa, (used during the first purification) so that the detergent with the 72 kDa micelle size would be removed prior to re-lipidation. After re-lipidation, a greater amount of white aggregate was observed than the previous preparation. This established that the white aggregate is likely to be lipid precipitation. In order to test this hypothesis two experiments were designed: 1) adding extra detergent to account for the need to solubilize the lipid and then using this sample for crystallization, 2) trialing re-lipidation with different protein-lipid ratios, 6:1 and 12:1.

For the first experiment, excess detergent (~12 µL of 0.02% DDM) was added to the 3:1 and 2:1 samples and they were incubated with stirring at 4 °C overnight. The next day both samples were completely clear without any white aggregate. After ultracentrifugation, these were used for hanging drop vapor diffusion crystallization experiments where crystals were formed (Figure 43).

The second experiment was designed to check the effect of different protein lipid ratios on the crystallization process. For this experiment, re-lipidation was performed as before but with variant protein-lipid ratios, 6:1 and 12:1. For these samples, 90 µL 13 mg/mL AcrB protein was employed. After overnight re-lipidation at 4 °C these samples were ultra-centrifuged and crystallization plates were prepared using the supernatant. No crystals were obtained from samples containing 12:1, 6:1 and 3:1 ratios of protein:lipid, suggesting that detergent concentration and lipid to protein ratio are critical factors in crystallization of AcrB with the HiLiDe method.

In Figure 43 the crystals obtained from two different crystallization conditions by hanging drop vapor diffusion using the HiLiDe method are shown. Other images of the crystals obtained from 2:1 ratio are presented in appendix V. This is the first time that this method has been applied to this protein. And because of the crystal size, these crystals could become good candidates for MicroED studies. In chapter 4, X-ray diffraction quality of these crystals is discussed.

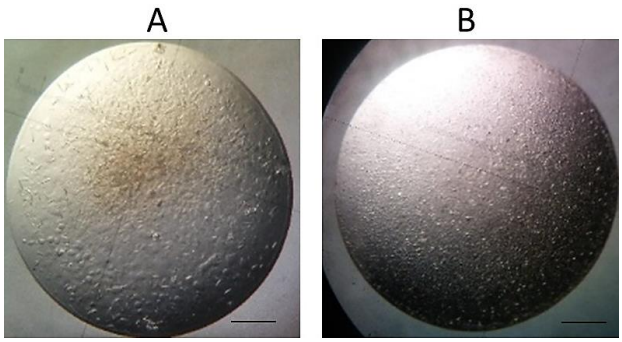


Figure 43. AcrB crystals obtained from HiLiDe method with 2:1 protein to lipid ratio. The crystallization condition for panel (A) is 0.1 M Lithium chloride, 0.1 M MES pH 6.5 and 12% PEG 400 and for panel (B) is 0.2 M Sodium chloride, 0.1 M MES pH 6.5 and 12% PEG 550 MME. Scale bar represents 500 µm.

3.2.1.4.2 ShuA

In contrast to AcrB, no aggregate/precipitate was observed during the re-lipidation process of ShuA (Figure 44). This sample was used for hanging drop vapor diffusion experiments after centrifugation.

Figure 44. Re-lipidation of ShuA with protein:lipid ratios of 3:1(right) and 2:1 (left). In contrast to AcrB, no lipid precipitation is observed.

In Figure 45, two images of the ShuA crystals obtained with the HiLiDe method are presented. More images are in appendix V. From Figure 34, Figure 35, Figure 43 and Figure 45 it is apparent that crystals obtained from HiLiDe method by vapor diffusion are smaller than otherwise comparable crystals obtained from vapor diffusion without the re-lipidation process. Similar to AcrB, this is the first time that this method has been applied to this protein. ShuA crystals

obtained from this method could also be considered as good candidates for structural studies using MicroED.

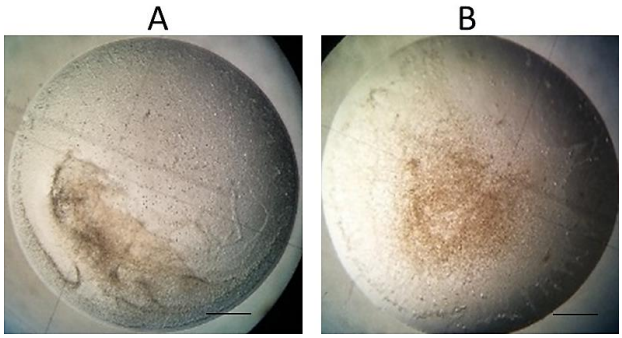


Figure 45. ShuA crystals obtained using the HiLiDe method with 2:1 (panel A) and 3:1 (panel B) protein to lipid ratio. A) Crystallization with 0.1 M Magnesium sulfate, 0.1 M MES pH 6.5 and 12% PEG 2000. B) Crystallization with 0.1 M Sodium citrate, 0.1 M MES pH 6.5 and 8% PEG 2000. Scale bar represents 500 µm.

Using a lower protein:lipid mass ratio (2:1) and an Amicon centrifugal filter with a 100 kDa cut-off resulted in the appearance of more lipid aggregation in AcrB samples. However, during the re-lipidation process of ShuA, no aggregation was observed. The mass ratio of AcrB and lipid was initially calculated without considering that this protein has both membrane (lipophilic) and periplasmic (hydrophilic) domains (see Figure 3). It is clear that treating the entire protein as lipophilic is not accurate and therefore overestimates the amount of added lipid. The initial mass of lipid added was excessive, which leads to aggregation of this extra lipid. Adding detergent compensated for this in solubilizing the excess lipid resulting in a clear solution. For ShuA samples, the calculation of mass ratios was appropriate and no aggregate was observed. This highlights the importance of finding the optimum protein: detergent: lipid mass ratio for this technique.

3.2.2 Optimization of crystal growth

The following section presents the results of crystallization trials, obtained mainly with model membrane proteins using different methods, and identifies the physical and chemical (such as temperature, and protein and PEG concentration) variables useful for crystallization optimization. The outcomes from co-crystallization and seeding experiments are presented first, followed by the experiments carried out using a crystallization equipment developed in our laboratory in order to control the variables mentioned above. The effects of these variables on the crystallization of model proteins are then discussed in relation to the phase diagram.

3.2.2.1 Co-crystallization

Co-crystallization of AcrB was carried out using two antibiotics, rifampicin and erythromycin [140]. Stock solutions at the concentration of 0.05 M were prepared in dimethyl sulfoxide (DMSO).

Co-crystallization of ShuA with heavy atom salts was carried out using Europium (III) chloride and Lead (II) nitrate dissolved in 10 mM Tris-HCl pH 8.0 with 1.4% OG [123] .

The molar ratio of the protein additive depends on the binding affinity of the additive to the protein and it should not be less than three times the K_d (i.e. dissociation constant of ligand protein K_d = [P][L]/[PL]) [141]. Here, an appropriate amount of the additives was added to the purified protein to give a 10:1 molar ratio of additive to protein. After gentle mixing, the final solution was incubated at 4 °C overnight.

In this section, results of the co-crystallization of the model membrane proteins ShuA and AcrB are presented.

3.2.2.1.1 Co-crystallization of ShuA with heavy atoms

Co-crystallization of ShuA with Lead (II) nitrate ($Pb(NO_3)_2$) and Europium (III) chloride ($EuCl_3$) generated crystals with slightly different shapes to those formed in the absence of heavy atoms. Both vapor diffusion and dialysis techniques were employed. ShuA without additives forms needles and urchins-shaped crystals whereas crystals with heavy atoms are robust and single columns (Figure 46). As such, co-crystallization with heavy atoms is likely to result in crystals suitable for X-ray (or neutron) diffraction experiments where larger crystals are required. Section 4.3.1.2.1 describes the X-ray diffraction quality of these crystals. The process of sample preparation in the presence of Lead (II) nitrate or Europium (III) chloride is however more challenging. Even when incubation is performed at 4 °C and the additive to protein molar ratio is optimized, aggregated protein forms leading to some sample loss and ambiguity in the protein concentration value.

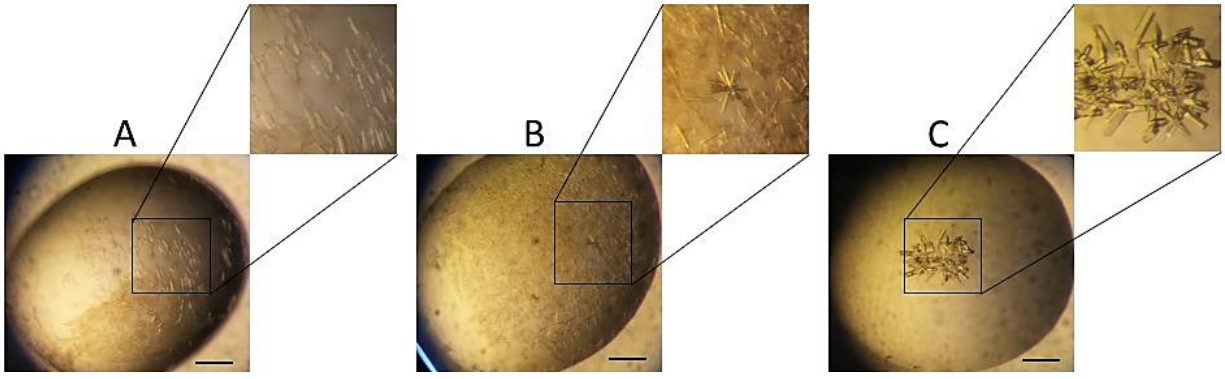


Figure 46. Crystals of ShuA obtained using hanging drop vapor diffusion with and without heavy atom based co-crystallization method. Crystallization solution is 0.1 M MES pH 6.5, 15% PEG 1000 and 0.1 M sodium chloride

(acetate for C). A) ShuA co-crystallized with Lead (II) nitrate. B) ShuA crystallized without heavy atoms. C) ShuA co-crystallized with Europium (III) chloride. Scale bar represents 300 µm.

3.2.2.1.2 Co-crystallization of AcrB with drugs

AcrB was co-crystalized in the presence of antibiotics (erythromycin or rifampicin) using different crystallization techniques (Figure 39, Figure 48 and Figure 50). Crystallization experiments were performed in either 0.01 M HEPES pH 7.0 or 0.02 M sodium phosphate pH 6.4. Crystallization conditions were employed by optimizing those reported in previous works [126], [140].

The results of our experiments showed that crystallization is more reproducible when AcrB-Rifampicin solution is in HEPES pH 7.0. In addition, less aggregate was observed in these samples in comparison to the protein in the phosphate buffer. The reason can be due to a higher rate of degradation of the drug in the phosphate buffer [142].

3.2.2.2 Seeding

A seeding method was used where seed crystals were transferred to clear drops or batches, at equilibrium (macro seeding) [143].

To control the size of the AcrB crystals grown with aim to obtain sample exploitable for neutron crystallography experiments, a batch method with the first generation crystallization bench [42] (see 1.1.6.1) was used at the initial temperature of 20 °C. For this experiment 50 µL protein was added to 50 µL of the crystallization solution to obtain the final mixture containing 7.5 mg/mL protein, 10% PEG 4000, 0.05 M ADA pH 6.5, 5% glycerol, 0.2 M ammonium sulfate and 0.02% DDM. Large amounts of protein precipitate were observed after gentle mixing of the solution. To remove it, the sample was centrifuged for 5 min at 13,200 rpm. Supernatant then transferred to a quartz tube and the solution was seeded with a crystal of AcrB (macroseeding) at a higher temperature (25°C). In the crystallization solution, the formation of crystal nuclei was observed which indicates the solution was supersaturated enough to provoke spontaneous nucleation. This experiment was repeated with the protein at the concentration of 5 mg/mL and identical results were observed (see Figure 47, A).

Macroseeding was also attempted in case of three clear drops, obtained by vapor diffusion, containing ShuA-Europium (III) chloride in crystallization mixture with 0.1 M MES pH, 0.1 M sodium acetate and 12 to 14% PEG 1000. Crystal macroseeds from a drop at higher PEG concentration (16%) were introduced into those pre-equilibrated drops. The crystallization plate was incubated at 20 °C and the following day many small crystals were observed in the drops (Figure 47, B) although the seeded crystals were still clearly visible, even growing, in these drops.

Despite the fact that seeding is a powerful technique for separating nucleation and crystal growth [42], the results of our few trials (microseeding and macroseeding) with membrane proteins did not convince/encourage us to proceed further and we focused on optimizing crystallization using

alternative strategies that we master in the laboratory, such as control of the chemical composition of the crystallization solution as well as the temperature control. It is also important to note that due to the significant amounts of protein samples that we had to use for our different crystallization experiments, at some point a careful choice of strategy had to be made.

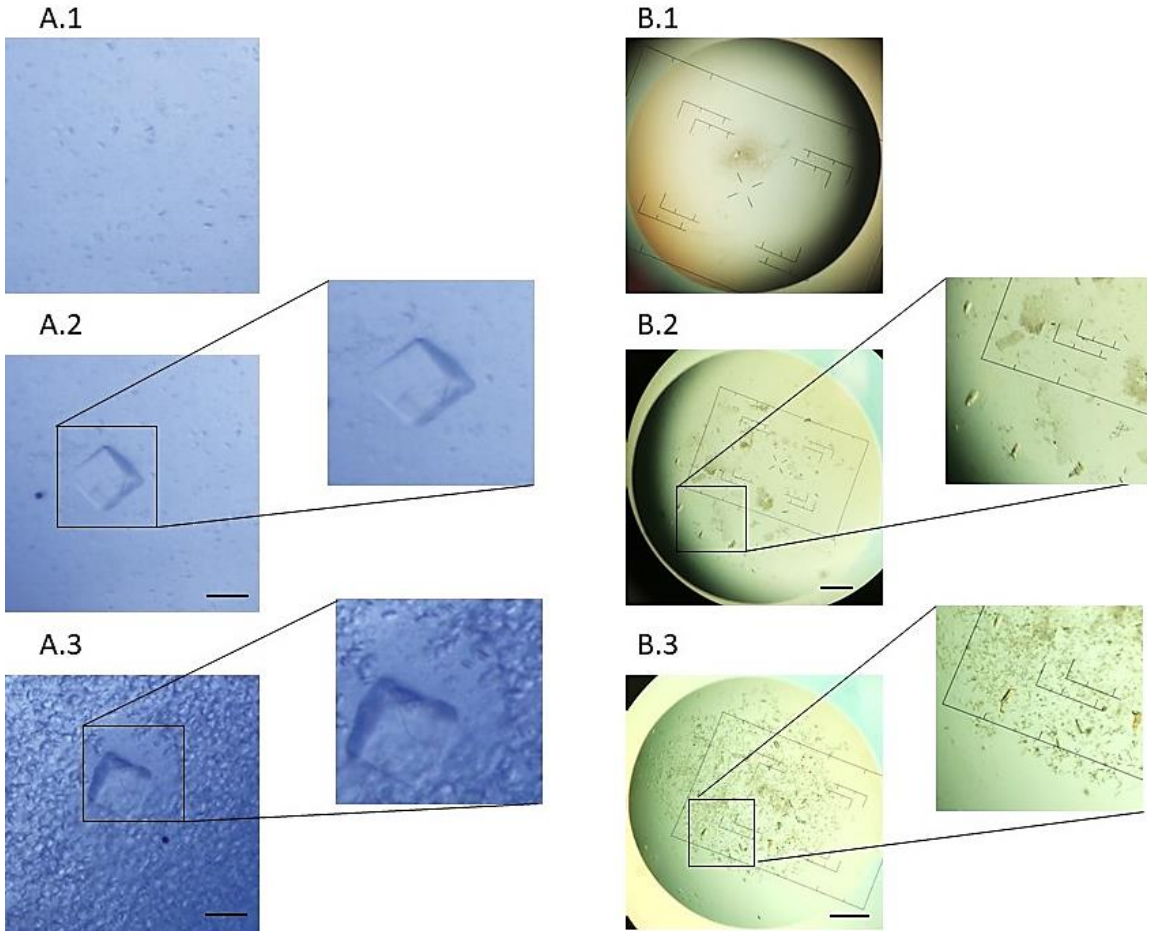


Figure 47. Macroseeding of AcrB (A) and ShuA (B). Initial crystallization batch (A.1) and vapor diffusion hanging drop (B.1). Seeded batch (A.2) and hanging drop (B.2) incubated at 20 °C, initial state. After seeding, nucleation in both samples is observed (A.3, B.3), final state. Scale bar represents 300 µm.

3.2.2.3 Effect of the temperature and protein concentration on crystal growth

AcrB crystallization experiments using the dialysis method were performed at three different temperatures (4, 20 and 25 °C) and protein concentrations (2.5, 5 and 10 mg/mL) to investigate the effect of these parameters on the crystallization process and crystal growth. The crystallization solution condition was otherwise similar for all experiments. Crystallization solutions contained 10% PEG 4000, 0.05 M ADA pH 6.5, 5% Glycerol, 0.2 M ammonium sulfate,

0.5 mM rifampicin, 0.02% DDM. 2 mL crystallization solution was used with a dialysis button volume of 50 µL. Rifampicin was used as an additive with an additive to protein molar ratio of 10:1.

At 4 °C, non-amorphous protein precipitates were observed for all samples. As shown in Figure 48, increasing the protein concentration at 20 °C results in an increase in the number of optically visible crystals formed inside the dialysis button. For all samples, the same chemical composition (crystallization liquor) is used, so the only variable is the protein concentration. Figure 49 illustrates schematically the kinetic path for each sample to reach the nucleation zone of the schematic crystallization phase diagram. Although the quantitative phase diagram of AcrB has not been studied, it seems from the obtained results, that in sample A2 the solution should be inside the nucleation zone as shown in Figure 48,49, while in the B2, the number of crystals is less than A2 which is illustrated by the fact that this sample is in the nucleation zone but near the metastable zone. Accordingly, in C2 (Figure 48,49) a lower number of crystals are observed which indicates that the solution is in the nucleation zone but in the vicinity of the metastable zone. However, because of the low protein concentration in this sample and still present significant protein precipitate, crystals did not grow to larger dimensions. It should be noted that in all the samples shown, an AcrB-rifampicin precipitate is observed at the bottom of the dialysis button. This precipitation seems to decrease with an increase in temperature, notably visible with the variation from 20 to 25 °C (Figure 48: A3, B3 and C3). Therefore, if these conditions are used for crystal growth optimization experiments, the sample will have to be driven to be kept in the metastable region of the phase diagram and the crystals will grow if the physico-chemical parameter, such as temperature, is changed.

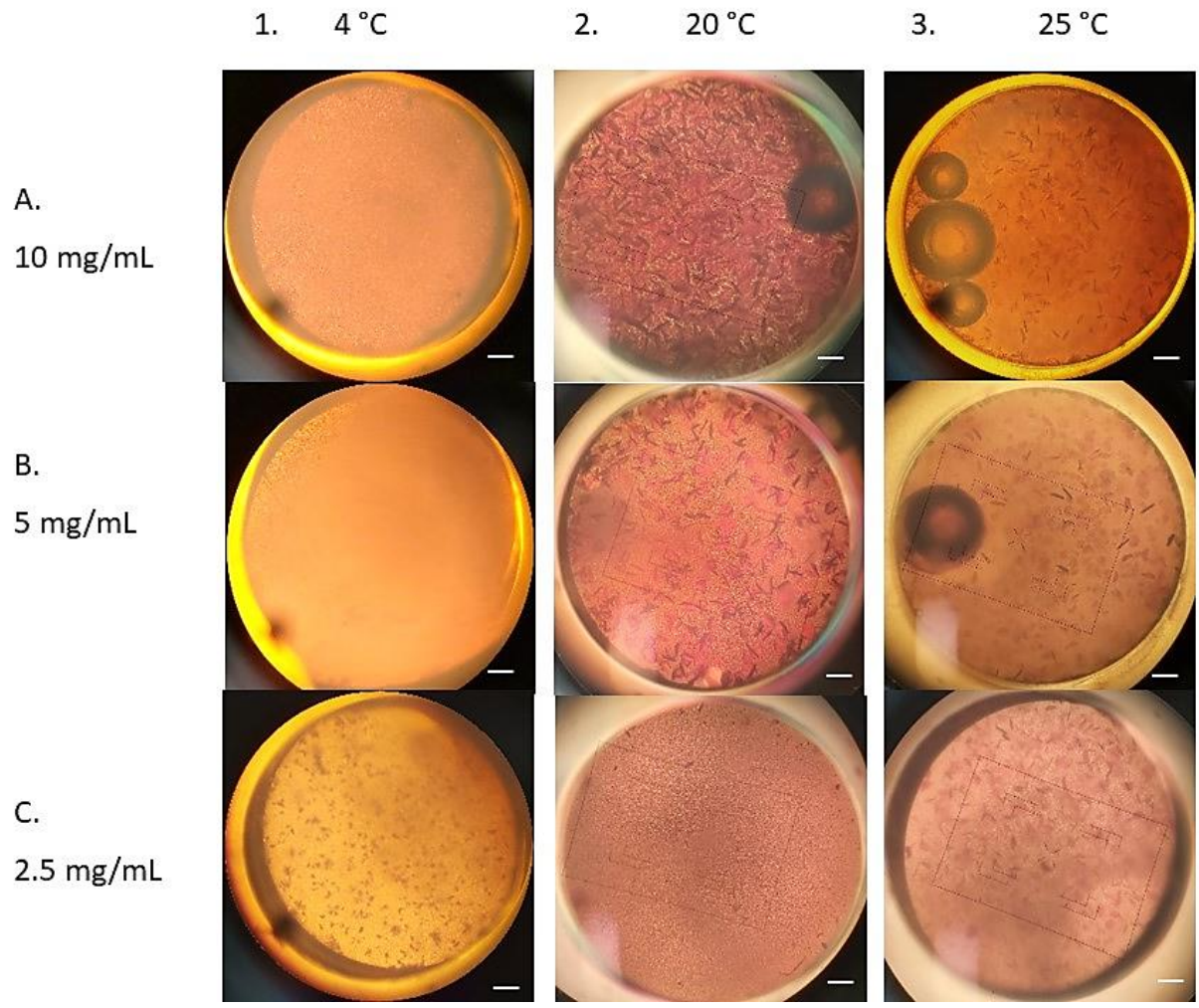


Figure 48. Investigating the effect of temperature and protein concentration on crystallization of AcrB-Rifampicin using the dialysis technique. Three temperatures (4, 20 and 25 °C) and protein concentrations (2.5, 5 and 10 mg/mL) were used for this experiment. Scale bar represents 500 µm.

All the buttons at 25 °C contain crystals. Comparing these results to the results at 20 °C it would be expected that less number of crystals would form (case of direct solubility). This appears to be generally the case (A3, B3), although in the case of the lowest concentrated sample (2.5 mg / ml) the number of crystals is higher than that observed at 20 ° C. But this can be explained by the fact that at 25 °C, the strong precipitate previously formed at 20 °C, was dissolved in favor of the formation of crystals of the AcrB-rifampicin complex. From this series of experiments, temperature effect could be demonstrated. It seems that the temperature range 20-25 °C is interesting for the growth of AcrB-rifampicin crystals. These crystals were then characterized by X-ray diffraction (Chapter 4).

As shown in the following paragraph, at a chosen constant temperature (20 °C) with a PEG concentration of 10% in the crystallization solution, a lower concentration of the protein is

needed to reduce the number of crystals inside the dialysis button. To obtain larger crystals, decreasing the PEG or salt concentrations without reducing the protein concentration will lead to fewer crystals and the excess protein still present will feed the growth of the already formed crystals. Therefore, the precipitant concentration can systematically be decreasing to find the optimum precipitant concentration of the crystallization solution.

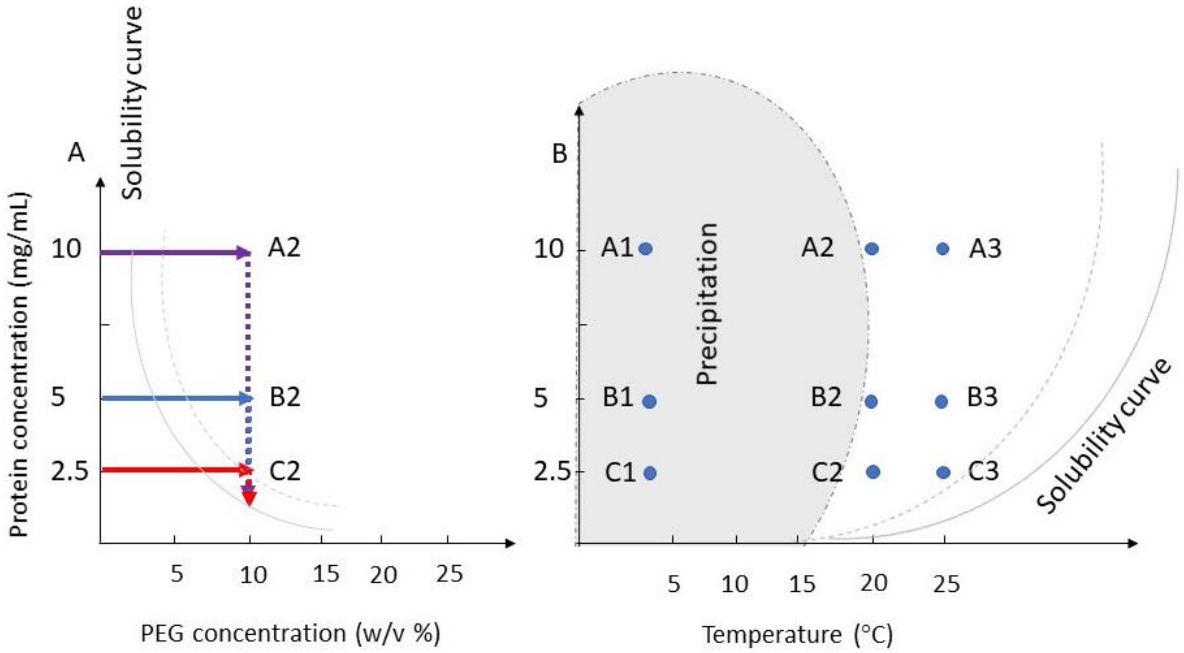


Figure 49. Schematic representation of crystallization phase diagram. Representation of kinetic pathway of protein crystallization by dialysis method with different protein concentrations as a function of (A) crystallization solution concentration and (B) temperature. Three different protein concentrations were used for this experiment. By increasing the crystallization solution concentration inside the button, crystals appeared inside the button. Each sample uses a different pathway through the phase diagram to reach the nucleation zone. Therefore, supersaturation is different in each sample, which leads to differences in nucleation and crystal growth. The solubility curve is qualitative.

3.2.2.4 Effect of the PEG concentration on crystal growth

Hanging drop vapor diffusion experiments were performed in the presence of a range of PEG concentrations (PEG 4000 was used from 6 to 16 % in 2 % intervals). Two concentrations of AcrB were employed (12 mg/mL and 6 mg/mL). Crystals appeared in hanging drops with a PEG concentration higher than 10 %, however increasing the concentration beyond 12% leads to crystalline precipitate as illustrated in Figure 50 (A5). Further increase in the PEG concentration beyond 14 % results in protein precipitation.

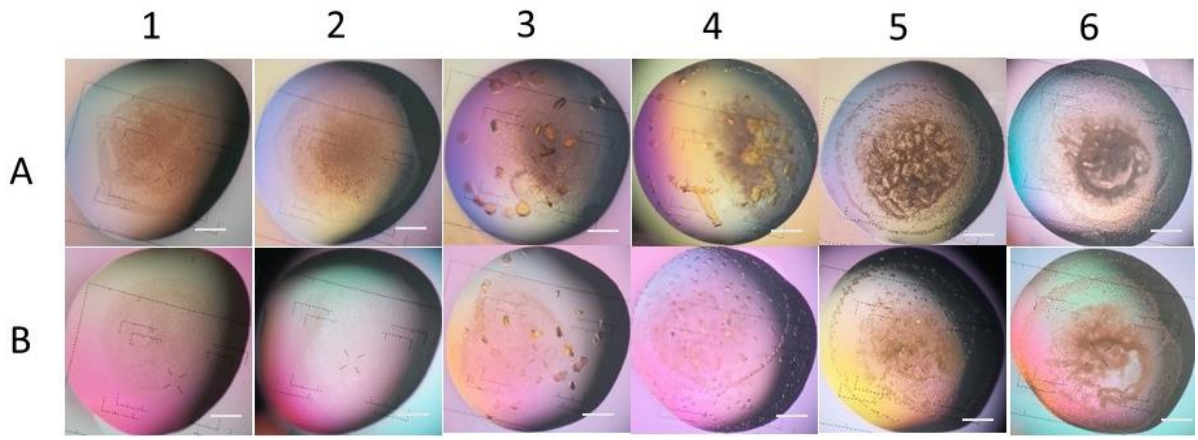


Figure 50. Co-crystallization of AcrB with Rifampicin. Hanging drop vapor diffusion was used with different PEG 4000 concentrations. From 1 to 6 PEG concentration increases from 6% to 16%. In row A protein concentration is 12 mg/mL and in row B it is 6 mg/mL. Drop ratio (protein volume:reservoir volume) is 1:1 and crystallization was performed at 20 °C. Scale bar represents 500 µm.

In order to determine the kinetic pathway in the schematic crystallization phase diagram for hanging drop vapor diffusion experiments, protein concentration was plotted against PEG concentration (Figure 51). According to the experimental data (see Figure 50) in case of AcrB-rifampicin complex at concentration of 12 mg/mL, increasing the PEG concentration beyond 14% leads the crystallization solution to precipitate. Although only schematic phase diagrams can be presented (more data are needed to obtain a quantitative phase diagram), invaluable qualitative information can still be obtained from this experiment.

According to these data, in a crystallization drop in which a volume of 12 mg/mL AcrB-rifampicin is initially mixed with the same volume of 10% of PEG 4000 (thus the initial protein concentration in drop is 6 mg/mL and PEG concentration is 5%), nucleation occurs in the vicinity of the metastable zone of the phase diagram at the same time as protein precipitation is kinetically sustained (Figure 50, A3/B3 and Figure 51, A3). By increasing the PEG concentration, the solution goes further into the nucleation zone, resulting in an increase in the number of crystals formed, still in the presence of the precipitate (Figure 50, A4/B4 and B5 and Figure 51, A4). By further increasing the PEG concentration, protein precipitation is kinetically favored over protein nucleation (Figure 50, A6/B6 and Figure 51, A6).

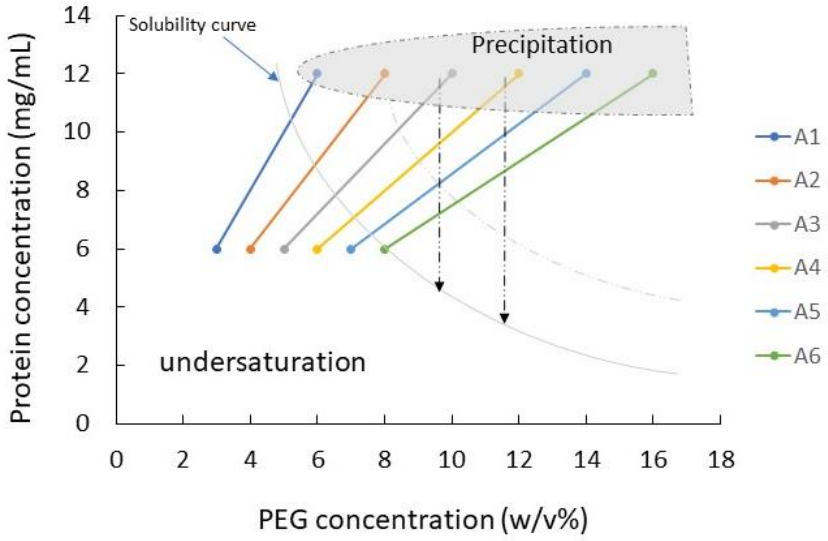


Figure 51. Schematic illustration of kinetic pathways for hanging drop vapor diffusion crystallization method. Samples with a 1:1 AcrB to PEG concentration ratio are represented. From 1 to 6, final concentration of PEG 4000 is increased from 6 to 16% with 2% intervals. Solubility curve is qualitative.

3.2.2.4.1 Large crystals of AcrB

The largest crystals of AcrB obtained from hanging drop vapor diffusion and batch methods as well as the standard size crystals obtained by dialysis are illustrated in Figure 52. It should be noted that the largest crystal ever obtained by dialysis, of the same morphology, approached 750 micrometres in its longest dimension. These crystals are obtained in presence (dialysis and batch) and absence (hanging drop) of rifampicin.

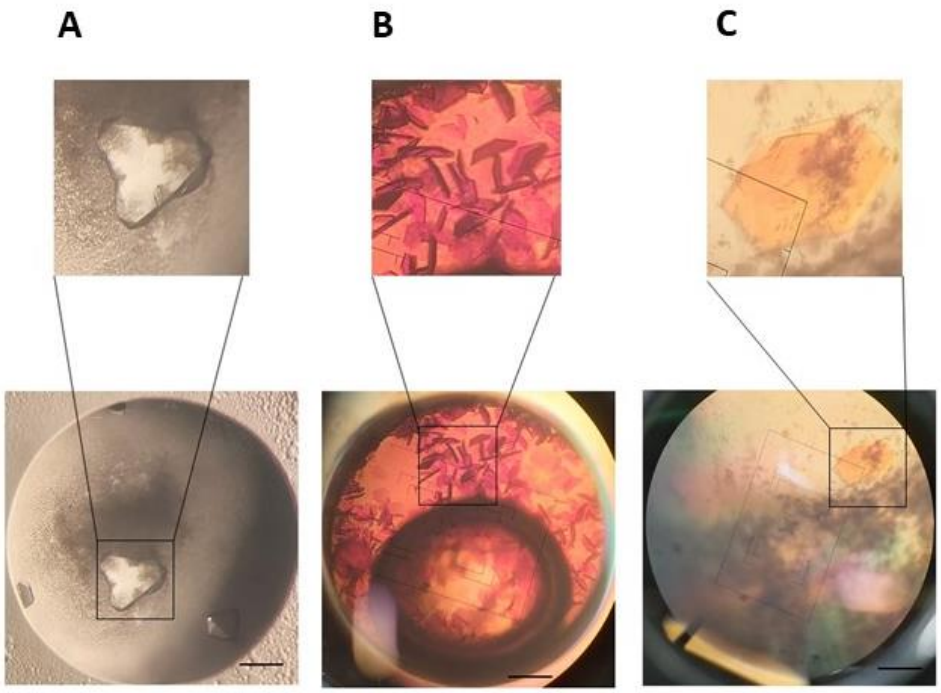


Figure 52. Crystals of AcrB in presence and absence of rifampicin by (A) hanging drop vapor diffusion (B) dialysis and (C) batch method. (A) Protein concentration is ~7.5 and the crystallization solution contains 13% PEG 400, 0.1 M MES pH 6.5, 0.1 M NaCl, 0.1 M Li_2SO_4. Drop ratio (protein volume:reservoir volume) is 2:2. (B) Protein concentration is 10 mg/ml and the crystallization solution contains 10% PEG 4000 , 0.05 M ADA pH 6.5, 5% glycerol, 0.2 M ammonium sulfate, 1 mM rifampicin and 0.02% DDM. Crystallization button volume is 50 µL and (C) Protein concentration is 6.4 mg/mL. Crystallization solution is similar to (B) only 8% PEG 4000 is in the crystallization solution. The final mixture volume is 100 µL. Scale bar represents 500 µm.

3.2.2.5 Crystallization bench (OptiCrys)

A semi-automated Crystallization Bench (OptiCrys) was recently developed for *in situ* monitoring and controlling of the crystallization processes of soluble proteins (see 1.1.6.2) [63]. This second-generation crystallization apparatus in addition to the temperature control uses dialysis as a crystallization method. By using OptiCrys, crystals with specific sizes and shapes can be obtained that are suitable for different crystallographic approaches [53], [63], [133]. In this section, crystals of model soluble and membrane proteins (lysozyme and ShuA) obtained using this instrument are shown. Examples of optimization of the crystallization process are presented to demonstrate the reversibility of the protein crystallization process as well as the ability to control the size of the crystals generated. Modifications to the flow cell design of this instrument to facilitate the crystallization of membrane proteins are also discussed. Finally, the discussion of the use of detergents in relation to dialysis implemented as a crystallization method in the developed instrument is included.

3.2.2.5.1 Protein crystallization using OptiCrys

3.2.2.5.1.1 Lysozyme, screening initial conditions

For the crystallization of lysozyme with the crystallization bench, two different temperatures (18 and 22 °C), protein with 30 mg/mL and a dialysis membrane with 6-8 kDa cut off were tested. The crystallization solution contains 0.1 M sodium acetate buffer pH 4.0 and two different concentrations of sodium chloride (0.75 M and 0.9 M). Illustrative results from these experiments are shown in Figure 53.

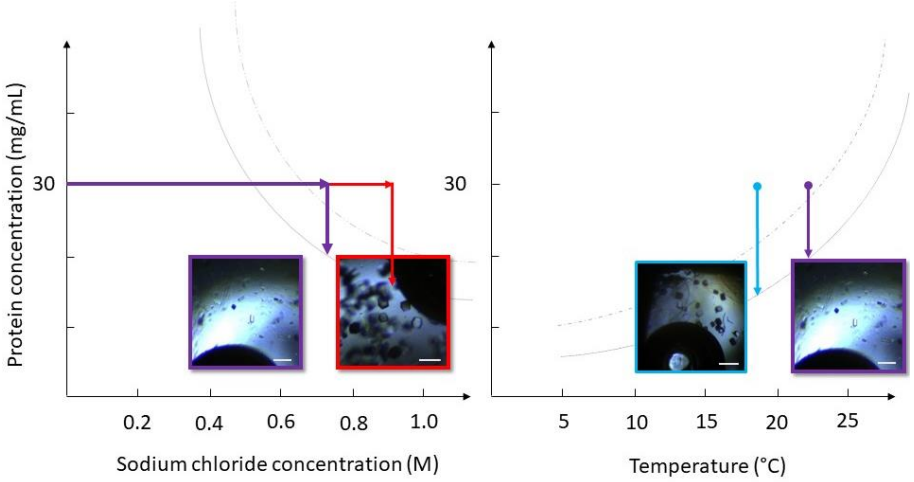


Figure 53. Schematic crystallization phase diagram representing protein concentration as a function of sodium chloride concentration (A) and temperature (B). Crystallization solution contains 0.75 and 0.9 M sodium chloride. By increasing salt concentration, the crystallization solution goes further in the nucleation zone of the phase diagram; as a result, the number of crystals increases. Crystallization is carried out at the constant temperature of 22 °C (A). At constant chemical composition of crystallization solution (NaCl 0.75 M, 0.1 M sodium acetate buffer pH 4.0) only variable in this experiment is the temperature. At lower temperature a higher number of nuclei appeared (B). Different zoom was used for this experiment. Scale bar represents 100 µm. Solubility curve is qualitative.

3.2.2.5.1.2 Optimization of the crystallization process: Growing large number of small crystals of Lysozyme

In this section, an example of the crystallization process optimization with lysozyme to generate a large number of small crystals of uniform size, using the crystallization bench, is presented. Detailed information on the crystallization bench, sample preparation and other examples of the optimization crystallization process are described in a recently published article included (see appendix I).

In order to grow a large number of small lysozyme crystals, the temperature of the crystallization experiment was controlled at a constant chemical composition (0.75 M sodium chloride in 0.1 M sodium acetate pH 4.0). First, protein (30 mg/mL) was added to the crystallization chamber (see appendix I [133]). Then, after adding the crystallization solution to the reservoir chamber, temperature was initially set at 18 °C (Figure 54, frame 1). The first crystals appeared after approximately 120 min. After several days, these crystals had a size of around 50 µm (Figure 54,

blue frame 2). Temperature was then increased to 35 °C gradually (over 20 min) to dissolve the crystals (Figure 54, red frame 3). When all the crystals were dissolved (Figure 54, red frame 4), around 24h later, temperature decreased quickly (in one step taking a few minutes) to 18 °C again (Figure 54, point 5). After several hours, a large number of smaller crystals (~ 10 µm) appeared inside the crystallization chamber (Figure 54, green frame 6).

In this experiment, rapid decreasing the temperature from 35 to 18 °C causes an increase in the rate of supersaturation by changing the width of the metastable zone [53], which leads to the fast simultaneous nucleation of smaller crystals. Thus, many small crystals with a narrow range of sizes were grown. It is likely these would be well suited for serial synchrotron crystallography or XFEL and further optimization to reduce crystal size could result in crystals suitable for electron diffraction studies. Alternatively, vortexing, sonication or pipetting can be used to break larger crystals to a size suitable for microED studies [144]. In serial crystallography, the optimal size and volume of the crystals depends on the sample delivery system used. For example, the liquid jet delivery method requires a higher volume of crystals in comparison to other delivery methods (i.e. LCP jet, fixed target chip and drop-on-demand) [145]. In contrast, the drop-on-demand delivery method requires a 20 µL to 100 µL sample volume only. Batch, LCP and vapor diffusion are the most frequently used techniques for serial crystallography sample preparation [146]. Because of the amount of protein needed for optimization and scaling up sample volumes, obtaining crystals for these methods can be costly. Use of the crystallization bench can reduce both the time and amount of protein required for sample preparation as it is possible to screen many different crystallization conditions with the same protein sample. Although the sample volume required for this crystallography method is high, the new flow cell designed for this instrument (volume ~ 160 µL) is ideally suited for this purpose. In the next section, a new design of dialysis flow cells is introduced.

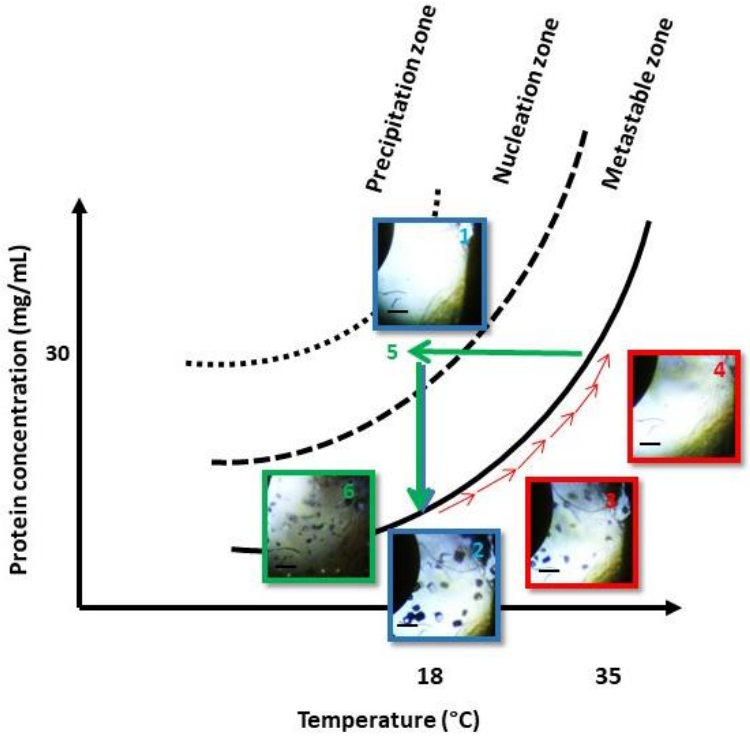


Figure 54. Optimization of lysozyme crystallization: growth of a large number of small crystals using the crystallization bench. Figure adapted from Junius et al [53]. Scale bar represents 100 µm. Solubility curve is qualitative.

3.2.2.5.2 Modification of flowing reservoir dialysis set-up of OptiCrys for protein crystallization

The design of the temperature-controlled flowing reservoir dialysis set-up was modified to make it easier for manipulation and avoid air bubble formation during assembly. This new design is similar to the older version [63] with a dialysis chamber, overchamber, reservoir and plug (see Figure 55). Two dialysis chambers were designed for this instrument. One of them has volume of around 160 µL. The new chamber has a lower height, circular shape and much larger diameter in comparison to the first prototype. With this new design assembly of the crystallization chamber and overchamber is facilitated (the bottom of the overchamber is spherical while the inside of the chamber is conical), which avoids formation of the bubbles during sample loading. The use of this new dialysis chamber is designed to promote the generation of protein crystals. Particular attention is given to the growth of large protein crystals for neutron crystallography and/or large volumes of small, uniformly sized crystals for serial synchrotron (and XFEL) or MicroED crystallography. The second dialysis chamber has a volume of approximately 40 µL. Modifications in the design of this chamber increase the contact surface between the dialysis chamber and overchamber to make it easier in applying the smaller sample volumes (Figure 55,B).

Another modification is decreasing the height of the dialysis chamber from 5 to 2.5 mm. This is intended to decrease the convection in the dialysis chamber; flow that occurs when one part of

the solution volume is denser than the other part of the solution. The Rayleigh number can be calculated to determine if convection can occur in the dialysis chamber [138]. It is directly proportional to the height cubed. By decreasing the height, convection will be slower but still present as this number is high (order of millions for 5 mm [138]). Decreasing the height also affects the mixing time, which is a result of the balance between diffusion, convection and flux through the semipermeable membrane. Diffusion is directly proportional to height squared and flux through the semipermeable membrane varies linearly with the height [138]. Therefore, the mixing time for the new chamber is quicker than that of the previous chamber.

Finally, Titanium was used for manufacturing the new flow cell to avoid any possible harmful reactions (oxidation) between the chemical solutions used for the crystallization and the materials used for making the flow cell. As a reminder, the old flow cell was made of stainless steel, which did not meet these requirements and suffered from insufficient purity.

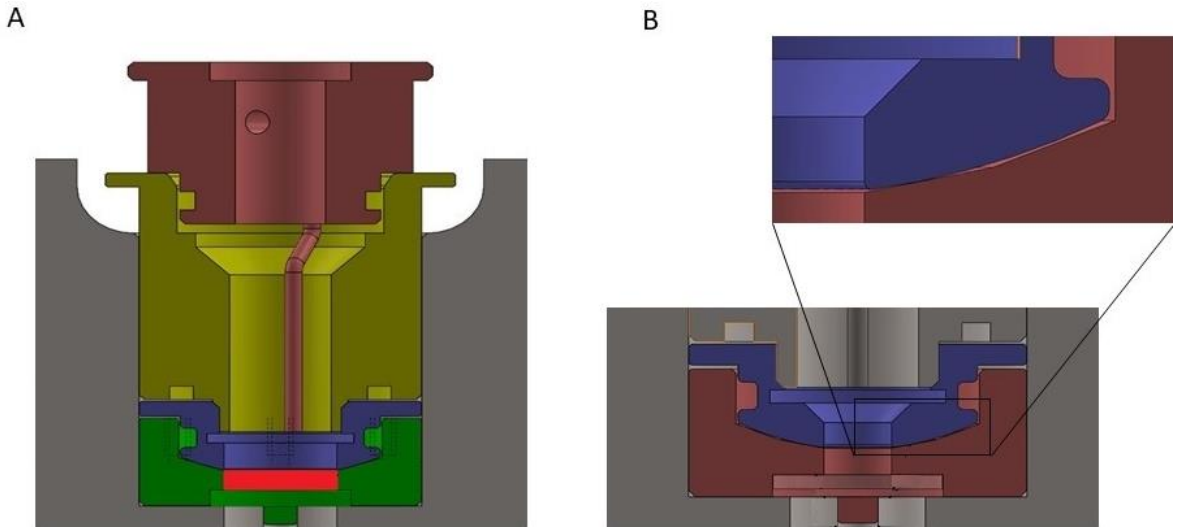


Figure 55. Schematic view of the new temperature-controlled flowing reservoir dialysis setup for OptiCrys with A) ~160 µL and B) ~40 µL dialysis chamber volumes. Four parts of the flow cell (dialysis chamber (red), overchamber (blue), reservoir (yellow) and plug (brown) are shown in 'A' with different colors. In 'B' points of contact between dialysis chamber (with 40 µL volume) and overchamber are shown.

3.2.2.5.2.1 Optimization of the crystallization of ShuA

Crystallization of ShuA was carried out at a constant temperature (20 °C) while the chemical composition of the crystallization solution during the experiment was modified manually using the dialysis set-up of the first generation crystallization bench to minimize the amount of protein used in this study. It is noteworthy that the set-up used mimics perfectly the second generation instrument (OptiCrys) with the option of using smaller volume dialysis buttons in the experiments. Therefore, less than 30 µL of 10 mg/mL ShuA was needed and added to the dialysis chamber of the first generation crystallization bench. A crystallization solution with 0.1 M MES pH 6.5, 0.1 M sodium acetate, 12% PEG 1000 and 1.4% OG was prepared and added to the reservoir chamber. Nucleation started in the corresponding area of the schematic crystallization

phase diagram shown in Figure 56, point 2. Crystals then grew until they reached their maximum size (Figure 56, blue frame 3). The PEG concentration was then decreased until 8% and finally PEG has been exchanged with protein buffer containing 10 mM Tris-HCl pH 8 and 1.4% OG, dissolving the crystals until no precipitate and crystals were observed in the chamber (Figure 56, red frame 5). After three days once the ShuA crystals have been completely dissolved (complete dissolution took a few days), new crystallization solution with a lower PEG 1000 concentration has been added to the reservoir chamber in two steps: i) (8%) during first 5 days which no crystals were optically visible and ii) then increased to 10% that eventually led to the re-nucleation of ShuA crystals at point 6 (Figure 56), in vicinity of the metastable zone of the schematic phase diagram. Then crystals grew until they reached the solubility limit of the phase diagram.

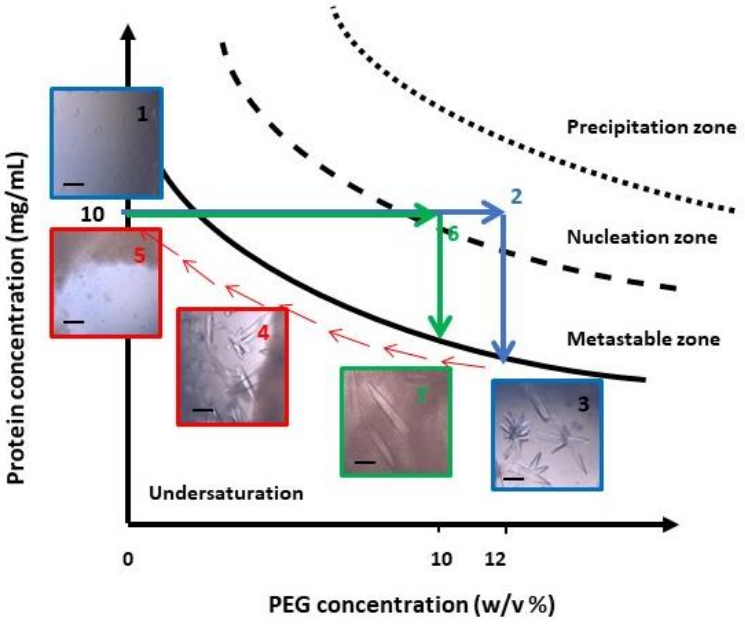


Figure 56. Crystallization optimization of ShuA. Reversibility of the crystallization process using the dialysis set-up of the first generation crystallization bench. In the first crystallization experiment (first nucleation event), a crystallization solution with a higher PEG concentration was used. Decreasing the PEG concentration then resulted in dissolving crystals. In the second crystallization experiment (second nucleation event), a crystallization solution with a lower concentration of PEG was used which leads to the formation of a lower number of larger crystals. Scale bar represents 100 µm. Solubility curve is qualitative.

In the first crystallization growth experiment presented in Figure 56 (blue frames and arrows), ShuA crystals are numerous with rosette shapes. By decreasing PEG concentration during the second crystal growth experiment with otherwise identical crystallization solution conditions, fewer and larger crystals were observed (see Figure 56, green frame and arrows). This experiment demonstrates the potential of using the crystallization bench for the crystallization of a model membrane protein in the presence of detergent.

In the next experiment, temperature was varied with constant crystallization solution concentration (0.1 M MES pH 6.5, 0.1 M sodium acetate, 12% PEG 1000 and 1.4% OG) to grow large crystals of ShuA. After crystals reached their maximum size at 20 °C, temperature was increased to dissolve crystals resulting from the first nucleation. No change in the size of the crystals was observed after 7 days at 35 °C. The temperature was then increased incrementally to 40 °C after which still no change was observed. Finally, the crystallization solution was changed (decreasing the PEG concentration) to dissolve the crystals. It is obvious that in optimization of membrane protein crystallization, changing temperature affects both protein and detergent. Therefore, the optimization process by variation of temperature is more complicated than variation of the crystallization solution concentration. Phase diagrams of both, detergent and protein, should be carefully explored in order to find the optimal conditions for crystal nucleation and/or growth.

Using these instruments, the reversibility of the crystallization process has been validated allowing suitable conditions to be found for the growth of (membrane) protein crystals of appropriate size to meet the requirements of structural studies. There are also other high-throughput instruments and robots used for setting up crystallization experiments. Some examples include robotic tools for *in meso* crystallization [147], or automation of microseed matrix-screening [148], the semi-automated Microcapillary Protein Crystallization System (MPCS) [149] and the Crystal Former using microfluidic technology [150]. These instruments facilitate the crystallization process, screening and image processing [32]. In addition, miniaturized devices such as microfluidic chips are using ultra-low volume of samples for screening and crystal growth. Crystals grow on chips and can be used directly for *in situ* X-ray crystallography, eliminating potential damage due to manual sample handling [4]. Most of these tools require low sample volume for screening crystallization conditions in a short time. However, there are other instruments compatible for both screening and optimization of the crystallization process using the batch method with a variation of temperature [50], using on-chip crystallization with temperature and precipitant solution variations [151] or using dialysis and variation of protein volume by syringes to control the size and quality of the formed crystals [152]. The crystallization bench combines dialysis with an automated control over temperature and precipitant concentration variables, keeping up with the state-of-the-art instrumental developments in applications related to protein crystallography.

3.2.2.6 Modeling of mass transport of detergent for optimization of membrane protein crystallization

The OptiCrys crystallization bench uses the dialysis method to crystallize proteins [63]. In this method, diffusion of the crystallization solution occurs through a semipermeable membrane (see 1.1.3.3.2 and 1.1.6.2). Using this instrument, mass transport has been observed and well characterized using a systematic study [139]. As part of this, the diffusion of salts and PEGs across

the semipermeable membrane was measured and computationally modeled [138]. The objective of this study was to decipher the mixing mechanisms (a solution of the protein with a solution of a precipitant) that occur in the OptiCrys crystallization chamber. This can then be used to establish effective strategies for the optimization of soluble protein crystallization. Here, this study was extended so that these concepts could be applied to membrane protein crystallization.

In order to preserve the native structure of membrane proteins during solubilization, purification and crystallization, the presence of detergents is crucial. Therefore, an experiment was designed to study and model the diffusion of detergents across dialysis membranes (with different molecular weight cutoffs) and calculate the mixing time for them. For this experiment, three detergents (OG, DDM and CYMAL-6) and four dialysis membranes (12-14, 20, 50 and 100 kDa) were used. These detergents are among the most frequently used detergents in the crystallization of membrane proteins [67], for example in this study OG and DDM were used in the crystallization process of ShuA and AcrB, respectively. These detergents have a range of different CMC values (0.17, 20 and 0.56 for DDM, OG and CYMAL6) and micelle sizes and so are good candidates for modelling the mass transport through semipermeable membranes (see Table 2). To ensure that detergent micelles, as well as monomers, can pass through the membrane, dialysis membranes with four different molecular weight cut-offs were used. Although lower molecular weight cut-off membranes are typically used for crystallization, higher cut-off membranes are routinely used for concentrating protein samples after purification. Since the colorimetric method is used in this study for calculating detergent concentration, after obtaining calibration curves for these detergents, it seems that 0.5 mM is the best concentration that can be used in this study.

For the first step, only Milli Q water was added to the chambers of 200 µL dialysis buttons. Semipermeable dialysis membrane was then used to cover each chamber. These buttons were then transferred into 50 mL tubes containing 5 mL detergent with 0.5 mM concentration and incubated at 20 °C. After appropriate time intervals (every 2h), the liquids inside the chamber of the dialysis buttons were transferred to an Eppendorf tube. Colorimetric method was then used to evaluate the amount of detergent inside the dialysis buttons [153]. Under a fume hood, 250 µL of 5 % phenol (Sigma-Aldrich) was added to Eppendorf tubes containing 50 µL of the solution from the dialysis buttons, then 600 µL concentrated (95-97%) sulfuric acid (Fluka) was added to the solution. After the addition of sulfuric acid, hexoses are dehydrated to hydroxymethyl furfural that after reaction with phenol produces a yellow colored product [154]. Absorbance of this product was measured at 490 nm [153], [154]. The first membrane used for this experiment was a cellulose ester membrane Spectra/Por (Spectrum™ Labs 100 kD Biotech CE Dialysis) with a 100 kDa molecular weight cut-off. DDM with 0.5 mM concentration was used for this experiment. Samples were incubated at 20 °C, and four samples were taken every two hours to measure the detergent concentration. After performing the reaction, the Beer–Lambert equation was used to calculate the detergent concentration. Data then were fitted using **Equation 2:**

Equation 2 : y = a*(1-exp (-b*x))

Where x is time and y is the concentration of the product in the chamber, normalized to the reservoir concentration. The results from one set of these experiments are presented in Figure 57. As can be seen from this figure and other complementary experiments that were performed (appendix VI), equilibrium was not reached and as a result mixing time could not be accurately calculated.

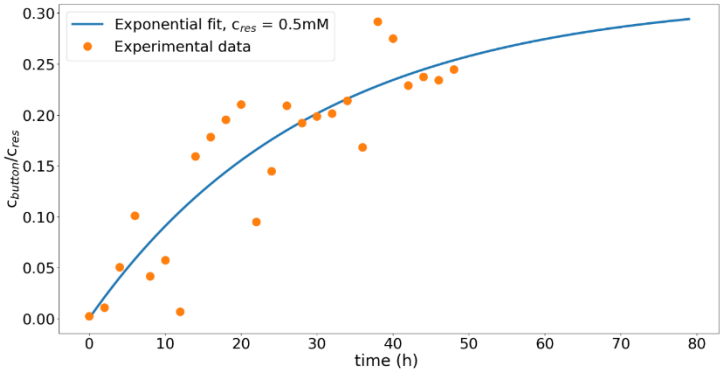


Figure 57. Measured detergent concentration inside the dialysis button as a function of the time. Dots are experimental data and the curve is the exponential fit. Each dot is the mean value of three measurements from one dialysis button.

In order to understand why equilibrium was not achieved, another study was designed. In this study, two dialysis membranes were used. Membrane 1. was a standard RC membrane Spectra/Por with 12–14 kDa molecular weight cut-offs and 2. was a cellulose ester membrane Spectra/Por with 100 kDa molecular weight cut-offs. DDM, OG and CYMAL-6 with 0.5 mM were used as model detergents for the experiment. Absorbance was measured at time zero and after 72h for all detergents (Figure 58 A and B). To check if the membrane adsorbs the detergent, a piece of membrane (around 1 cm×1 cm) was transferred to a second 50 mL tube containing 5 mL detergent (sample C). Milli Q water was added to the chamber of a 200 µL dialysis button and the dialysis membrane was fixed to the top of the chamber using an elastic O-ring. This was then transferred to a third 50 mL tube previously filled with 5 mL detergent. After 72 h the dialysis button was removed and the solution inside the button was analyzed using the same colorimetric experiment described above (sample D). Samples of the detergent taken from the inside of the same 50 mL tube (third tube) were also used for this experiment (sample E). Figure 58 shows a summary of the results for this preliminary experiment.

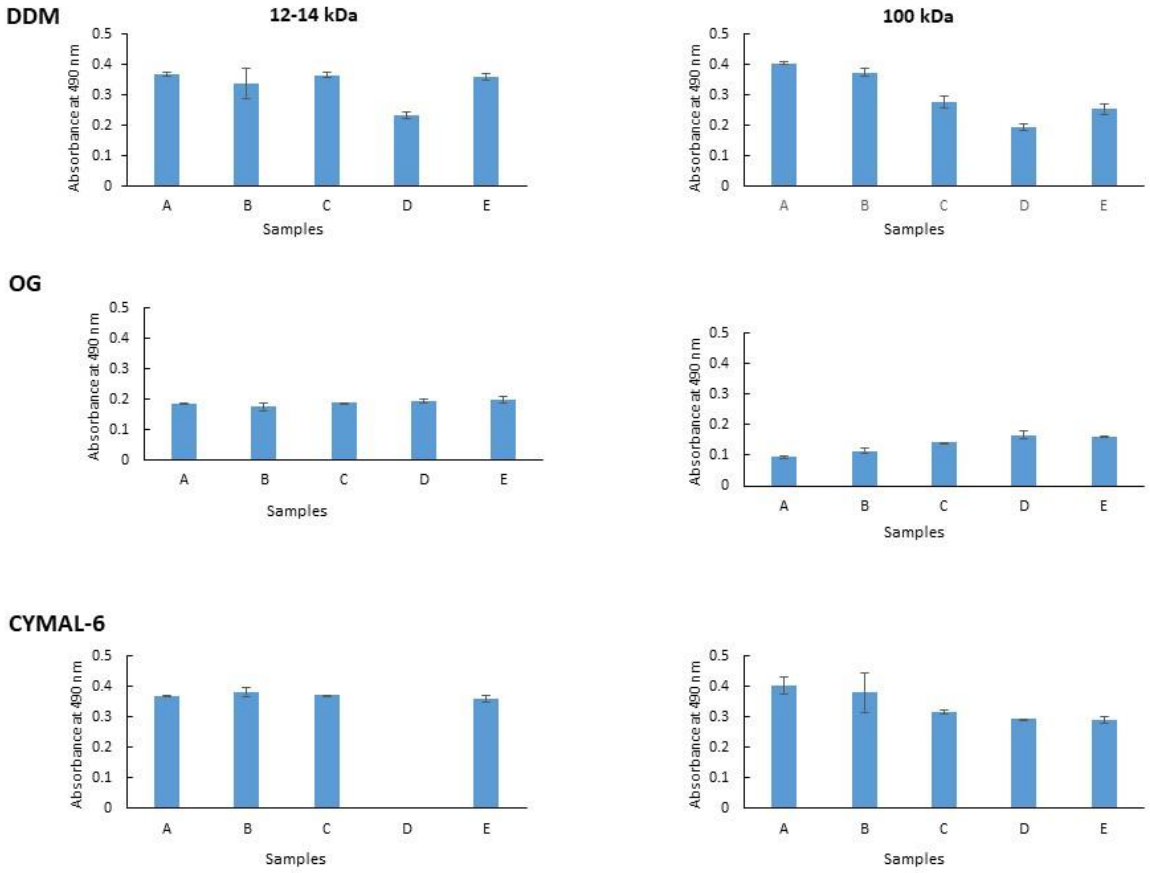


Figure 58. Measured absorbance for DDM, OG and CYMAL-6 at time zero (A), after 72h (B), after 72 h for the sample of 5 mL detergent in presence of a membrane (C), after 72 h for the solution from the inside of the dialysis button, (D) and finally after 72 h for the detergent inside the tube with the dialysis button inside it. Each column corresponds to three measurements from one sample.

It is worth mentioning that the absorbance for the samples with OG is approximately half of the absorbance for samples with DDM and CYMAL-6. The reason is that DDM and CYMAL-6 have a maltoside headgroup but OG has a glycosidic head group. So using the colorimetric method less colored product will be produced for OG samples and as a result, absorbance at 490 nm will also be less.

As clearly shown in Figure 58, in samples with DDM and dialysis membrane with 100 kDa molecular weight cut-offs, absorbance is less than the similar experiment with 12-14 kDa molecular weight cut-offs. The 100 kDa dialysis membrane is cellulose acetate (CE membrane) while the 12-14 kDa dialysis membrane is a regenerated cellulose membrane (RC membrane). Both dialysis membranes are hydrophilic, however RC membranes are more hydrophilic than the CE membranes (less hydrophilic). Therefore detergent flux reduction at concentrations higher than the CMC is more probable when a CE membrane is used.

Each detergent has a different CMC and micellar size (see Table 2 and 1.1.4.2). Above the CMC micelles are formed. The 100 kDa cut-off membrane was chosen to be sure that even DDM detergent micelles (65-70 kDa, see Table 2) are able to pass through the membrane. Despite this equilibrium was not achieved after 72h for DDM Samples. The term used for decrease of flux is fouling [155]. Fouling in this study could be due to hydrophobic interactions between the membrane and the detergents [155]. Alternatively, it could be due to retention of the micelles or even adsorption of the detergent molecules to the pores of the membrane [155]. Detergents have a polar head group and a hydrophobic tail [69] and hydrophobic solutes may adsorb onto the surface of a membrane more readily than hydrophilic solutes [155]. Using a nonionic detergent with a cellulose acetate membrane has been reported to affect flux through the membrane when the detergent concentration is higher than CMC [155]. This is in agreement with the findings from this experiment for at least the case of DDM. The CMC of DDM is ~0.17 mM (see Table 2) and the detergent concentration used for this study is 0.5 mM. OG has CMC ~ 20 mM but the smallest micellar size of the detergents tested (Table 2) and no changes in the flux were observed. However for CYMAL-6 (CMC is ~0.56 mM, micellar size 32 kDa) a decrease in the absorbance at 490 nm was also observed in this preliminary test, which may be due to the use of detergent concentrations approaching CMC. Adsorption was not observed when a 12-14 kDa RC membrane was used (comparing columns B and C). However, in case of DDM equilibrium was not achieved after 72 h, most likely due to the low CMC and high molecular weight of this detergent. During the dialysis, free detergent monomers pass through the membrane [70], but for DDM with concentration higher than CMC micelles are present in the solution which may decrease this rate. It seems that 72 h is not enough time to reach equilibrium for this detergent.

Mass transport through the dialysis membrane can be controlled and adjusted during the crystallization process using the crystallization bench OptiCrys. To control the size and quality of crystals, the gradient created due to the mass transport of crystallization agents can be controlled. Knowing the mechanism of detergent mass transport and mixing time are important to maintain detergent concentration during the crystallization process and ensure there is always sufficient detergent to preserve the membrane protein native structure. However, as this information is currently only partially available, further experiments with RC membranes, which are commonly used in dialysis crystallization, should be performed. Until then, detergents with a similar concentration in the protein solution should be added to the reservoir crystallization solution to maintain the required concentration on both sides of the membrane. The final chapter features a discussion about possible follow-up experiments that could be performed to build on these preliminary experiments.

3.3 Discussion

When determining the structure of a biomacromolecule using crystallography, the first and often most arduous step is finding conditions that generate crystals. Typically it will then be necessary

to improve these initial conditions to generate the appropriate volume, size or diffraction quality of these crystals. To benefit from current generation diffraction techniques more specific crystal volumes are required. For example, crystals of at least 0.1 mm^3 are required [39] for neutron protein crystallography, and in microED submicron crystals are used [156]. The experiments performed in this chapter demonstrate how different conventional and non-conventional crystallization methods and physico-chemical variables affect the crystallization process in a way that can often be rationalized by phase diagrams. This information once studied can then be used to obtain suitable crystals reproducibly. This chapter demonstrates how this methodology can also be applied to model membrane proteins.

Using synchrotron radiation X-ray sources, single protein crystals with sizes from 10 to 100 µm can be used. During this work, crystals of soluble and membrane proteins with sizes within this range were obtained which were subsequently analyzed by using X-ray crystallography (see chapter 4). For X-ray crystallography, the quality of the crystals also plays an important role in obtaining accurate structural information. Therefore, the strategy for optimal crystallization can only be fully established and validated when the excellent X-ray diffraction characteristics from these crystals are recorded (e.g. low mosaicity, high diffraction power).

Two membrane proteins were studied in depth with regard to their crystallization: ShuA and AcrB. Here we will discuss and compare the crystal growth aspects related to various crystallization techniques employed; in chapter 4 the diffraction properties of these are discussed.

Foremost, the best technique identified by this study for obtaining large crystals of a size theoretically suitable for neutron diffraction experiments, is dialysis. Crystals of AcrB with size ~ 750 µm in their largest dimension were obtained using this method. The dialysis technique is also practically convenient as the crystallization conditions can be precisely varied by exchanging the crystallization solution during the experiment in a way that would be impossible using the other methods studied. As a result, there is more control over the crystallization process in comparison to other techniques, which is especially useful when attempting the rational formation of crystals in the framework of the crystallization phase diagram. Furthermore, as illustrated in the case of Crystallization Bench/OptiCrys (see 1.1.6.2); this process is amenable to being fully automated. In a conventional crystallization process, exchange of the crystallization solution would necessitate removal of the coverslip. However, the crystallization bench has a pumping system that can automate this process via an user-friendly software interface which reduces the chance of error and improves the reproducibility of the experiment [53], [63].

ShuA was crystallized using the dialysis set-up of the first generation crystallization bench (3.2.2.5.2.1). Optimization was achieved by controlling the crystallization solution to first dissolve and then to regrow crystals in a controlled and rational manner. Crystals obtained from the second nucleation were larger in comparison to those from the first nucleation. However, this is not the case in experiments using temperature control of the same protein system. One rationalization for this is that temperature has an effect on both the protein and the detergent

phase and as such, the phase diagram description of this system becomes more complex. For example, high temperatures used to dissolve the crystals (35 °C to 40 °C) can cause detergent-based phase separation which leads to micelle aggregation [71] (see 1.1.4.1).

In crystallization of AcrB by dialysis method, three temperatures (4, 20 and 25 °C) at the constant PEG concentration were used (see 3.2.2.3 and Figure 48) the precipitate was observed at the bottom of all the dialysis buttons. However, at 25 °C the precipitate was dissolved partially in favor of the crystals. Therefore, although the higher temperature cannot be used in a system that contains protein/detergent complexs, careful controlling of the temperature could be beneficial in the optimization of crystal growth. This optimization process highly depends on the conditions of crystallization and the protein studied.

From the schematic qualitative phase diagrams presented in this chapter and knowledge of the reversibility of the crystallization process, (see Figure 56, and [53]) the dialysis technique used by the crystallization bench allowed for many crystallization conditions to be screened without consuming large amounts of the protein. These optimization data then can be used in successive experiments to obtain crystals of appropriate size for different crystallographic techniques reproducibly, without the need for extensive empirical screening experiments.

Large crystals were also obtained using the hanging drop vapor diffusion but these appeared fragile in comparison to the crystals obtained from the dialysis method (more robust and thicker). The diffraction quality of these crystals will be compared in the next chapter. Although batch seeding is another powerful technique for crystal growth, in this case it has not yet yielded spectacular results. Further study and experimentation may be required to achieve this goal with our model membrane proteins, but for the time being other methods such as dialysis are more promising.

In addition to being able to grow large crystals, the dialysis (see 3.2.1.2.4, Figure 38), LCP (appendix III) and HiLiDe (see 3.2.1.4, Figure 43, Figure 45) methods show strong potential for growing a large uniform population of small crystals potentially suitable for other diffraction techniques, such as MicroED. This finding is in agreement with a previous study that emphasizes the small size of the crystals generated by LCP [7]. Optimization of crystal size and morphology with HiLiDe method however clearly depends on a finely balanced protein/lipid/detergent ratio [78] as illustrated by experiments with AcrB. From these experiments, 2:1 (protein: lipid) for both ShuA and AcrB and 3:1 for ShuA can be used in order to obtain small enough crystals for MicroED.

Mass transport of detergent through dialysis membranes was also studied in order to optimize membrane protein crystallization. It is clear that this technique is promising for crystallizing membrane proteins and amenable to automation, so having a good model of what is going on at a molecular level within the dialysis cell will certainly be useful. This is especially important to identify any minor effects that might accumulate over the long reversible crystallization optimization experiments that are possible using the crystallization bench. However, it is clear that more experiments are likely to be necessary to improve theoretical models in this regard as

using dialysis membranes with 100 kDa cut-offs and DDM with 0.5 mM concentration, mass transport could not be simply modeled. One factor influencing this could be the level of hydrophilicity of the cellulose acetate membrane. Additionally, the concentration of DDM used in the experiment was higher than CMC, which may reduce the flux through the membrane. In agreement with similar experiments performed with more hydrophilic RC membranes, the crystallization of membrane proteins by the dialysis method should only be performed using this type of membrane, minimizing possible detergent adsorption. Alternatively, a more simple method of stabilizing the protein-detergent complex in the dialysis chamber during crystallization is to ensure that the crystallization solution of the reservoir contains the same amount of detergent so passage through the membrane is not so critical.

Chapter 4: Crystallography

4.1 Introduction

X-ray diffraction studies, the most widely used, are now increasingly supplemented by neutron and electron diffraction for structural studies of biomacromolecules [157]. Figure 59 is a comparison between the numbers of deposited structures in the Protein Data Bank (PDB) according to the method used for structure determination. Although the above-mentioned techniques are slowly becoming popular, X-ray diffraction has been used to determine the majority of protein structures (78%). Indeed, according to the protein data bank [17] from 1976 until the end of 2020, 152,800 structures were deposited to PDB using X-ray crystallography.

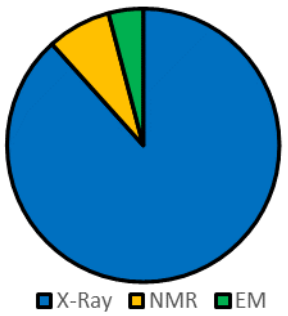


Figure 59. Comparison between the methods used for the determination of structures in the PDB from 1976 - 2021. X-ray diffraction (contains X-ray, fiber and powder diffraction) is the most frequently used technique for structure determination among the structures deposited in the PDB. This image is adapted from the statistics section on the PDB website. NMR data from both solution and solid-state NMR and EM data from electron microscopy, electron crystallography and electron tomography are shown in this graph. For multiple techniques the number of deposited structures is low (184 structures, 71 of which are related to neutron crystallography) that cannot be visualized in the chart. Structures that are solved by both X-ray diffraction and neutron diffraction are in this category.

In X-ray diffraction, X-rays are scattered by the electron clouds of the atoms in protein crystals [157] and the position to which these X-rays are scattered is recorded by a detector. Such two-dimensional diffraction patterns can be recorded across a range of positions in single-crystal X-ray diffraction, typically by rotating (oscillating) the crystal in a stationary X-ray beam. Combined diffraction patterns from a range of angles can then be used to calculate a three-dimensional electron density map that contains the location of electrons in the average unit cell of a crystal [84]. The analysis of electron density is then used to build a three-dimensional model of the molecular structure of the crystallized protein. Although this method for studying the structure of biomacromolecules provides invaluable information about their three dimensional structure, for example the positions of amino acids, cofactors, or substrates to discuss catalytic mechanism, protein function, mechanism of the disease, designing new drugs [83], information about the position of H atoms is extremely difficult to interpret from the electron density map without high resolution diffraction [95]. The reason for this problem is that H atoms have only one electron,

so they scatter X-rays weakly. Neutrons, on the other hand, interact with nuclei and H atoms have a scattering length similar to other atoms like C, N and O. Therefore neutron diffraction can be used to localize H atoms more easily than X-ray diffraction [158] and provide complementary information in combination with X-ray diffraction studies [157].

As detailed in chapter 3, crystal growth can be rationally optimized to obtain high quality crystals for crystallography. In the case of membrane proteins and macromolecular complexes especially [103], this optimization process is not always sufficient to obtain large enough crystals for X-ray crystallography. Using electron crystallography, even microcrystals can be used to solve the structure of a protein. By obtaining Coulomb potential maps from electron diffraction experiments, we can also gain information about the charge states of the amino acids [103].

This chapter presents the results from X-ray crystallography and MicroED experiments using the crystals obtained from two model membrane proteins (ShuA from *Shigella dysenteriae* and AcrB from *Escherichia coli*) used for this study. First, the results from X-ray crystallography are presented which compare the qualities of AcrB crystals generated from different crystallization techniques employed. Following, the 3D structure of ShuA obtained using molecular replacement is presented. Comparative preliminary results from electron diffraction studies of ShuA and AcrB are then presented. Finally, at the end of the chapter, these results are discussed.

4.2 Methods

4.2.1 X-ray diffraction

4.2.1.1 Sample preparation

Individual crystals were fished using a cryo-loop (polymer loops from MiTeGen and Nylon loops from HAMPTON RESEARCH) and immediately immersed in appropriate cryoprotectant solution(s) as detailed and then flash frozen in liquid nitrogen.

4.2.1.2 Data collection, model building and refinement

Diffraction data were collected on several beamlines: BM30A, ID30A-1 at the European Synchrotron Radiation Facility (ESRF), BL14.1 at the Berlin Electron Storage Ring Society for Synchrotron Radiation (BESSY) and BL13-XALOC at ALBA.

Data were collected at cryogenic temperatures (100 K). Crystals were transferred from storage in liquid nitrogen and mounted on a goniometer automatically using a 6-axis robotic arm. Prior to full data collection, 4 to 10 frames were collected to characterize crystals and optimize the data collection strategy.

Diffraction data for ShuA and AcrB were processed and scaled using XDS [159]. ShuA structure was determined using molecular replacement with PHENIX [160] and the published 2.6 Å structure of ShuA *Shigella dysentriae* (PDB entry: 3FHH [30]) was used as a search model. Then Coot was used for model building [161]. After each step, structure refinement was performed using PHENIX. During the last cycles of refinement, water molecules were added to the model automatically in PHENIX. Manual corrections using Coot were made to the refined structure.

4.2.2 Micro ED

4.2.2.1 Sample preparation

Around 4 µL of crystals in mother liquor were pipetted onto a lacey carbon transmission electron microscopy (TEM) grid. Excess solution was removed by applying a piece of filter paper (Whatman filter paper no.4) to the backside of the grid. The grid was then plunged into liquid ethane cooled by liquid nitrogen in a manual plunger. This frozen grid was then cryo-transferred into the TEM for observation and data collection.

4.2.2.2 Data collection

Diffraction data were collected on a ThermoFisher Tecnai F20 microscope operating at 200 kV (FEG). Images were recorded using a 4kx4k Ceta CMOS camera and diffraction patterns were recorded on a hybrid pixel detector Cheetah (Amsterdam Scientific Instruments). For the screening of the entire grid, low magnification (~ 1000×) was used. After a suitable crystal was located, the microscope was switched to diffraction mode and data collection was carried out with <1 e $Å^{-2}$ s^{-1} dose rate.

4.3 Results

4.3.1 X-ray diffraction

In this section, data collection statistics for AcrB crystals obtained from different crystallization techniques and conditions are compared. Refinement of the model required time beyond the scope of this thesis. As such, preparation of a structure was deferred for future work.

Then the crystal quality of ShuA obtained in presence and absence of heavy atoms is compared. Finally, the new structure obtained for ShuA is compared to the structure previously deposited in the PDB (3FHH).

4.3.1.1 AcrB

4.3.1.1.1 Comparison on the basis of purification techniques

To compare the quality of protein crystals obtained from different purification methods (affinity, anion exchange and size exclusion chromatography) more than 10 crystals of each prepared using the hanging drop vapor diffusion method were studied. The crystallization conditions for tested crystals are shown in Table 11. Crystals of protein obtained from affinity chromatography alone did not diffract in a reasonable range so no full data set is presented for these. However, when AcrB was co-crystallized with rifampicin a better diffraction was obtained.

Table 11. Crystallization condition for AcrB crystals prepared using different purification methods.

Sample name	A	B	C
Additive	Yes / rifampicin	No	No
PEG 4000	14%	10%	10%
Buffer	0.1 M sodium phosphate pH 6.2	0.05 M ADA pH6.5	0.05 M ADA pH6.5
Salt	0.1 M sodium chloride	0.2 M ammonium sulfate	0.2 M ammonium sulfate
Glycerol	No	5%	5%
Cryoprotectant	Yes	Yes	no

For cryoprotection, glycerol concentration was increased from 10 to 30% in a solution similar to crystallization solution but a with 5% higher PEG concentration.

In Table 12, the statistics for the best diffraction prepared using each of the three purification methods is presented. These crystals belong to the space group R32 with unit-cell parameters for crystal A: a= b= 144.519, c = 522.466 Å, crystal B: a= b= 143.306, c= 529.907, and crystal C: a =b = 145.074, c= 526.206. Resolution cut-off was based on the I/σ (I) ≧ 1. For comparing the data quality, completeness of data, signal to noise ratio (I/σ), resolution, R-meas and CC (1/2) were compared.

Table 12. Best data collection statistics for AcrB crystals prepared using different purification methods. (Statistics in parentheses are the highest resolution subshell).

	AcrB-Rifampicin Crystal A	AcrB Crystal B	AcrB Crystal C
Purification technique	Affinity	Anion exchange	Size exclusion
Crystallization method	Hanging drop	Hanging drop	Hanging drop
Beamline	BL13-XALOC	BL13-XALOC	BL14.1
Wavelength (Å)	0.979	0.979	0.9184
Space group	R32	R32	R32

a, b, c (Å)	144.519,522.466	143.306,529.907	145.074,526.206
Angle (˚)	90, 90, 120	90, 90, 120	90, 90, 120
Resolution limit	47.97-4.29 (4.55-4.29)	48.07-4.24 (4.50-4.24)	48.47-4.19 (4.45-4.19)
Number of reflections			
Observed	96205 (15651)	139850 (18991)	152457 (23831)
Unique	13942 (2185)	14716 (2076)	16013 (2490)
Completeness (%)	93.9 (93.3)	97.8 (87.1)	99.7% (98.7)
I/sigma	8.89 (1.05)	4.49 (1.08)	6.39 (1.03)
R-meas (%)	11.6 (168.4)	21.3 (115.5)	23.8% (216.6)
CC(1/2) (%)	100 (52.1)	99.8 (70.6)	99.8 (44.8)

R-meas (%) [162]

From these data, it seems that, although having pure protein is important for obtaining better resolution, the right additive can clearly improve crystal quality compared to the purer samples. From the deposited structure on the PDB, best diffraction for the AcrB crystals obtained from hanging drop vapor diffusion in absence of the additive is at 2.8 Å resolution (PDB 2DHH). Crystals for this experiment were prepared using affinity chromatography purified protein [25], [132]. Crystals with additives were also studied by these authors and showed lower resolution in comparison to the first structure, PDB 2DR6 with 3.30 Å resolution for AcrB-doxorubicin and PDB 2DRD with 3.10 Å resolution for AcrB-minocycline [132]. These crystals belong to the monoclinic crystals system in comparison to the results obtained in the current study with the space group R32 (trigonal crystal system). PDB 1IWG is another structure of AcrB (no additive) solved by these authors in space group R32 at 3.5 Å resolutions.

In this study, none of the AcrB crystals from co-crystallization with erythromycin diffracted sufficiently, so no full data set is available from these crystals. One reason for the poor diffraction quality of these crystals could be due to the physical characteristics of them. The AcrB-erythromycin crystals that were frozen for the data collection were small (10 µm diameter loops were used) and fragile in comparison to the AcrB-rifampicin crystals, so it is possible that poor diffraction was caused by damage sustained by these during sample handling. Another reason could be that only a limited number of crystals formed under these conditions reducing the probability of the serendipitous formation of well diffracting crystals. There are four AcrB-erythromycin structures deposited in the PDB (PDB 4ZJL, 4ZJO, 4ZJQ [140] and 3AOC [126] with resolution 3.47, 3.6, 3.59 and 3.33 Å, respectively). It is therefore still possible that if more crystals

were prepared under these conditions for diffraction studies, optimization of the crystallization and sample preparation could improve the diffraction quality of the crystals.

From this initial optimization experiment it is clear that AcrB co-crystallized with rifampicin has the highest signal to noise (I/sigma = 8.89 compared to <6.39) and the lowest R-measured values (11.6 compared to >20), which is promising. However only crystals obtained from hanging drop vapor diffusion were compared. So for the next step of optimization, crystals from other crystallization techniques were prepared and tested.

AcrB protein obtained from affinity chromatography was co-crystallized with rifampicin. Crystallization conditions are summarized in Table 13. Cryoprotection of these crystals was achieved by increasing the glycerol concentration from 10 to 30% in a solution similar to the crystallization solution but with a 5% higher PEG concentration. A single crystal was transferred to a droplet of each cryoprotectant before flash freezing it with liquid nitrogen.

Table 13. Crystallization condition for AcrB crystals using different crystallization methods.

	Vapor diffusion	Dialysis	Batch
Protein concentration (mg/mL)	5, 7, 10 and 13	5, 7, 10 and 13	2.5, 3.5, 5 and 6.5
Crystallization solution:			
PEG 4000 (%)	5, 7 or 10	5, 7 or 10%	5, 7 or 10%
ADA pH 6.5 (M)	0.05	0.05	0.05
Glycerol (%)	5	5	5
Ammonium sulfate (M)	0.2	0.2	0.2
DDM(%)	no	0.02%	0.02%
Rifampicin (mM)	no	0.5	0.5

For hanging drop vapor diffusion, the quality of the crystals obtained from three different protein concentrations (5, 7 and 10 mg/mL) using a crystallization solution containing 10% PEG 4000 was studied. For dialysis, crystals, obtained from 5 and 13 mg/mL using a crystallization solution containing 10% PEG 4000, were frozen. For the batch method, crystals, with 6.5 and 5 mg/mL of protein in presence of 7% PEG 4000, were frozen.

Table 14, compares the best diffraction statistics obtained for crystals prepared using each technique. In this experiment, more than 10 crystals of AcrB-Rifampicin from each technique were cryo-cooled under identical conditions. Data collection was performed at BL13-XALOC, ALBA at cryogenic temperature (100 K).

Table 14. Best data collection statistics for AcrB-Rifampicin crystals prepared using different crystallization techniques. Protein concentration and PEG 4000 concentration for hanging drop are 10 mg/mL and 10%, for dialysis are 5 mg/mL and 10%, and for batch are 6.5 mg/mL and 7%, respectively. (Statistics in parentheses are the highest resolution subshell).

	AcrB-Rifampicin	AcrB-Rifampicin	AcrB-Rifampicin
Protein purification technique	Affinity	Affinity	Affinity
Crystallization method	Hanging drop	Dialysis	Batch
Space group	R32	R32	P1
a, b, c (Å)	143.773,522.040	142.721,522.105	142.924,143.243,196.790
Angle (˚)	90, 90, 120	90, 90, 120	69.321, 89.225, 60.412
Resolution limit	48.29-3.90 (4.14-3.90)	47.66-3.76 (3.99-3.76)	48.14-7.77(8.23-7.77)
Number of reflections			
Observed	191238 (31071)	201539(32621)	21345 (1955)
Unique	19456 (3020)	20717(3141)	2469(311)
Completeness (%)	99.7 (98.4)	99.2 (95.2)	96.2 (79.5)
I/sigma	10.41 (1.06)	6.95 (1.02)	10.73 (1.11)
R-meas (%)	13.9 (227.5)	20.1 (162.3)	12.2 (121.7)
CC(1/2) (%)	99.9 (61.8)	100.0 (74.0)	100.0 (87.1)

R-meas (%) [162]

I/σ (I) ≧ 1 was used for resolution cut-off. Resolution, completeness of data, R-meas and CC (1/2) were compared using Table 14. It is clear that higher quality crystals were not obtained using the batch method compared to either dialysis or hanging drop. The resolution for batch method crystals is much less (7.77 Å compared to 3.76 or 3.90 Å) suggesting that clear information on the positions of amino acids could not be achieved by this method. It is also possible either that batch generated crystals have different, low symmetry (P1 compared to R32) or that the rather poor quality of the data is insufficient to accurately determine the unit cell. In the batch method, as the protein solution is simply mixed with the crystallization solution in one ratio that does not

change, equilibrium is more-or-less instant. There is no vapor diffusion or mass transport, and this rapid approach could explain the formation of less ordered crystals. As for dialysis, some crystals were retained in cryoprotectants some time before freezing, potentially increasing the risk of damage at this stage of sample preparation. Although no AcrB structures have currently been deposited to the PDB which were crystallized using the batch method, there is a structure with P1 symmetry (PDB 2HRT, a=127.33, b=134.87, c=140.84, α=103.9, β=94.64, γ=09.11) from hanging drop vapor diffusion. Using the same crystallization conditions another crystal with C2 symmetry (PDB 2GIF) was obtained. These structures feature one and two nonsymmetric trimers in the unit cell, respectively [163]. The crystallization solution for this experiment contained PEG 300 and 400, glycerol and sodium citrate pH 4.6 [163]. It could be possible that this low pH causes different crystal packing in comparison to the crystals obtained in the current study. The pH affects the protonation state of the amino acids side chains on the protein surface as well as the carboxyl (C-terminal) and amino (N-terminal) groups [164]. This in turn affects protein-protein interactions, potentially creating salt bridges or changing hydrogen bonds which can be crucial in crystal packing [164]. Studies of selected point mutations can show how changing the protonation state of key interacting residues can affect crystal packing. In another study, point mutagenesis was used to substitute aspartic acid (D407A, PDB 2HQC and D408A, PDB 2HQD), lysine (K940A, PDB 2HQF) and threonine (T978A, PDB 2HQG) with non-charged alanine [124]. In crystals obtained from wild-type AcrB, these residues are immobilized by the strong interactions between them; however, in crystals from mutant proteins this network of interaction is interrupted. Although this study did not use identical crystallization conditions, sodium citrate pH 5.6 was used [124]. With these modifications in charged residues, all of the crystals in this study belong to the R32 space group [124]. It therefore seems that substitution of these selected charged amino acids with alanine does indeed significantly affect protein-protein interactions resulting in modified crystal packing.

Generally, AcrB crystals obtained with hanging drop vapor diffusion and batch method, in the present study, were more fragile in comparison to the crystals obtained from dialysis. Some of these crystals could even be quickly dissolved in the cryoprotectant. Crystals prepared using dialysis or vapor diffusion diffract to a resolution suitable for structural experiments whereas those prepared by the batch method do not. Before discarding this method, it is worth exploring if preparation of crystals by the batch method can be improved. Other experiments were therefore designed to optimize the quality of crystals obtained by the batch method. The best diffraction data from these experiments are shown in Table 15. Although the optimization of the crystallization conditions resulted in better quality crystals, the diversity of crystals prepared by the batch method is high. Some crystals did not diffract and those that did had disparate maximum resolutions ranging from about 10 to 4 Å, making the best crystals comparable in resolution to those prepared by dialysis (3.76 Å) or by hanging drop (3.90 Å).

Table 15. Summary of the diffraction characterization of AcrB-Rifampicin crystal obtained with batch method. Final protein concentration: 6 mg/mL, with 10% PEG 4000 in the final solution: Statistics for the highest resolution shell are in parentheses. Data collection was performed at ID30A1, ESRF, at cryogenic temperature (100 K).

Space group	R32
a, b, c (Å)	142.133,524.766
Angle (˚)	90, 90, 120
Resolution limit	48.17-4.08 (4.33-4.08)
Number of reflections	
Observed	59923(8878)
Unique	16315(2516)
Completeness (%)	98.1 (96.4)
I/sigma	7.70 (1.24)
R-meas (%)	9.1 (95.2)
CC(1/2) (%)	99.5 (49.8)

R-meas (%) [162]

The best diffraction statistics of crystals formed by hanging drop vapor diffusion is the one presented in Table 14. We have prepared and tested around 100 crystals using this method and analyzed them using X-ray diffraction. Significant diversity in shape and diffraction quality of the crystals was observed. For the similar number of crystals obtained with the dialysis method, diversity in crystal quality observed is less than for the other two crystallization methods employed. When more than one crystal is taken from the same dialysis button, however, the quality of crystals gets worse proportional to the time they remain in the cryoprotectant. We have thus accumulated and generated information that is rarely reported in the literature.

One of the factors affecting the diffraction resolution of AcrB crystals in X-ray studies could be the amount of the solvent in the unit cell. Less solvent causes more contacts between molecules in the crystals which leads to well-ordered packed crystals with better diffraction [84], [165]. The Matthews coefficient (V_M) can be determined using Equation 3 which can in turn be used to calculate the volume of solvent (V_S).

Equation 3: $\mathbf{V_M = V_A/M}$

V_A is the crystal (asymmetric unit) volume (Å^3) and M is the molecular weight of the unit of the protein (Da). Then Equation 4 is used to calculate the solvent volume. 1.230 is a constant conversion factor (Da/Å^3) [165].

Equation 4: $\mathbf{V_S = 1 - (1.230/V_M)}$

The solvent content of AcrB crystals with R32 symmetry is approximately 72% if one molecule per asymmetric unit is assumed [117]. This relatively high crystal solvent content may therefore physically limit the resolution to which AcrB crystals will diffract X-rays, despite any other attempts to optimize crystal quality. Consistent with this, AcrB crystals with P1 symmetry (PDB 2HRT) have a lower solvent concentration (64.78 %) in comparison to crystals with R32 symmetry [163]. Structure of this crystal was solved at 3 Å resolution compared to 3.38 – 3.65 Å in the case of the R32 crystals. This could also explain why AcrB-rifampicin crystals obtained using hanging drop, dialysis (Table 14), and batch method (Table 15) have lower resolutions in comparison to the AcrB-rifampicin structure with PDB ID: 3AOB [140]. This crystal was obtained using sitting drop vapor diffusion, has C2 symmetry and slightly higher resolution (3.35 Å) in comparison to the crystals presented in this work. The solvent content for this crystal is lower (65%) than that calculated for the crystals with R32 symmetry. This is also consistent with the idea of the solvent content of a crystal physically limiting the diffraction resolution of a given crystal.

It is noteworthy that there are several studies using Cryo-EM. To date no structure with resolution better than 3 Å has been obtained [130]. Among published cryo-EM structures, there is only one structure of wild-type AcrB (PDB ID 6BAJ) with 3.2 Å resolution. For the AcrB D407A mutant (PDB ID 6CSX) a 3.0 Å resolution was obtained [130], which is comparable to X-ray structures solved for the same AcrB mutant (PDB 2HQC with 3.56 Å and PDB 2HQD with 3.65 Å). The complexes of AcrB with AcrZ and minocycline (PDB ID 6SGS with 3.2 Å) and AcrB-DARPin (PDB ID 6SGU with 3.27 Å) were also solved using cryo-EM [166]. However, the best resolution (1.90 Å) AcrB-DARPin complex was solved using X-ray crystallography (PDB ID 4DX5) [125].

4.3.1.2 ShuA

4.3.1.2.1 Diffraction data

In order to compare the quality of ShuA crystals obtained in presence and absence of heavy atoms, diffraction data were collected at BL13-XALOC, ALBA (for ShuA crystals with heavy atoms) and BM30A, ESRF (for ShuA crystals without). Before data collection, crystals were frozen in liquid nitrogen in the absence and presence of cryoprotectants. For ShuA crystals (without heavy atoms), characterization prior to data collection did not identify any crystals with diffraction better than 4 Å. As such, no data were collected from these crystals.

Several data sets were collected from single crystals of ShuA-Eu and data were processed and scaled using XDS. Data collection statistics are summarized in Table 16. The resolution cut-off was selected based on the I/σ(I) ≧ 2 and the best diffraction was at 2.49 Å resolution. These crystals belong to the space group $P2_12_12_1$ with unit-cell parameters a = 79.09, b = 114.24, c = 117.95 Å. Co-crystallization of ShuA with Europium (III) chloride significantly improved the resolution consistent with previous reports [123]. ShuA-Pb crystals only diffracted to low resolution and no data were collected from them.

Table 16. Summary of the diffraction characterization of ShuA-Eu crystal: Statistics for the highest resolution shell are in parentheses.

Space group	$P2_12_12_1$
a, b, c (Å)	79.09, 114.24, 117.95
Angle (˚)	90, 90, 90
Resolution limit	47.24 -2.49 (2.63-2.49)
Number of reflections	
Observed	228347 (39453)
Unique	37089 (5958)
Completeness (%)	96.6 (97.6)
I/sigma	14.82 (2.07)
R-meas (%)	13.3 (89.9)
CC(1/2)%	99.7 (83.9)

R-meas (%) [162]

Solvent content was also calculated for this crystal (Equation 3), using unit cell parameters and molecular weight, if one molecule is present in the asymmetric unit crystals contain approximately 67% solvent.

4.3.1.2.2 Model building and refinement

Molecular replacement was used for phase determination (see 1.1.5.1) using the previously determined structure of ShuA as a search model. PHENIX was used for refinement and Coot was used for model building. The refinement statistics for the model are presented in Table 17. Around 5% of the reflections were used for calculation of R-free.

Table 17. Refinement statistics for ShuA-Eu.

Resolution range	47.24 - 2.49 (2.579 - 2.49)
Unique reflections	37349 (3653)
Completeness (%)	98.11 (98.38)
Wilson B-factor	43.53
Reflections used in refinement	37288 (3650)
Reflections used for R-free	1810 (145)
R-work	0.2369 (0.2509)
R-free	0.2747 (0.3380)
Number of non-hydrogen atoms	5024
macromolecules	4808
ligands	81
Protein residues	621
RMS(bonds) (Å)	0.007
RMS(angles) (°)	0.93
Ramachandran favored (%)	96.91
Ramachandran allowed (%)	2.61
Ramachandran outliers (%)	0.49
Average B-factor	46.38
macromolecules	46.28
ligands	61.69
solvent	40.7

The refined model includes residues 1-640 and has R-work = 0.2369 and R-free = 0.2747 (R-work and R-free for the PDB 3FHH are 0.237 and 0.286, respectively). The overall structure has an rmsd of 0.461 Å when compared to PDB 3FHH, Figure 60. PDBeFold was used for 3D alignment (https://www.ebi.ac.uk/msd-srv/ssm/cgi-bin/ssmserver) of the two structures for comparative purposes.

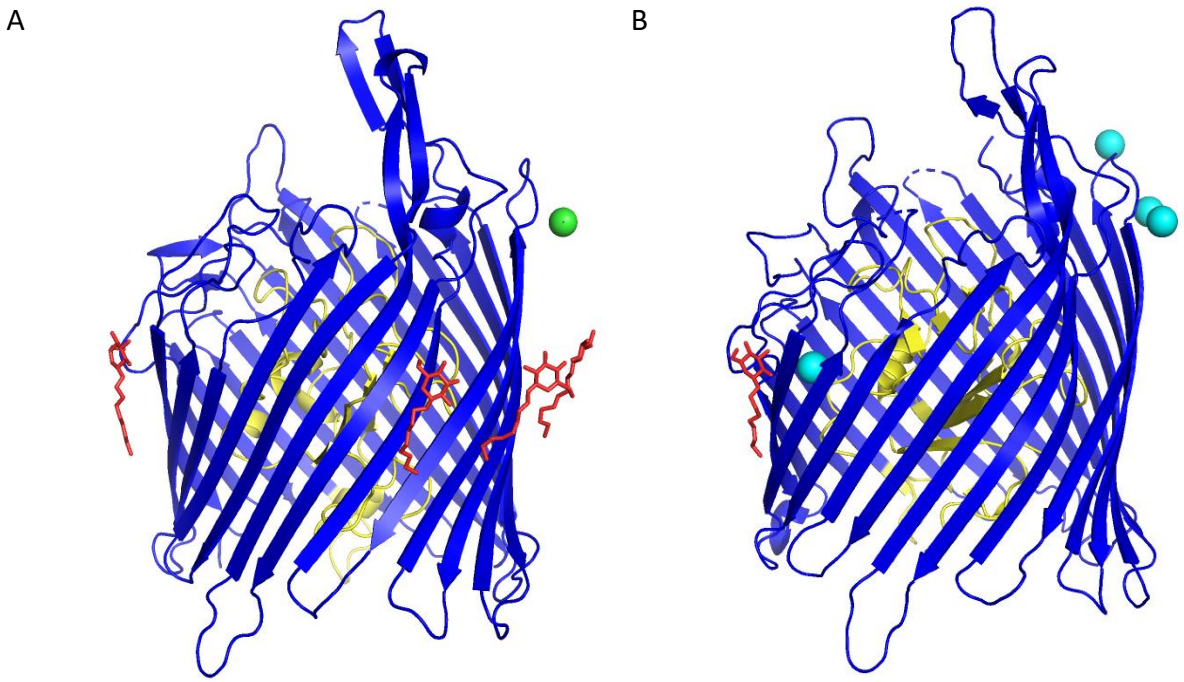


Figure 60. ShuA structures. A) Is a model built using B) (PDB F3HH) as a search model. Four detergent molecules and Eu^{3+} are present in A. However only one molecule of detergent with four Pb^{2+} ions are presented in B. Figure was rendered in PyMol. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

Differences between the two ShuA models can be visualized by a comparison of the distances between the amino acid Cα in each structure. These distances were calculated and a plot of these is shown in Figure 61. For better visualization, the square of distance was used in this plot. This identifies ten regions with significant deviation, marked in the plot (Figure 61, A-J). These regions are located in solvent-exposed loops and these differences are consistent with the expected flexibility of such regions (see appendix VII).

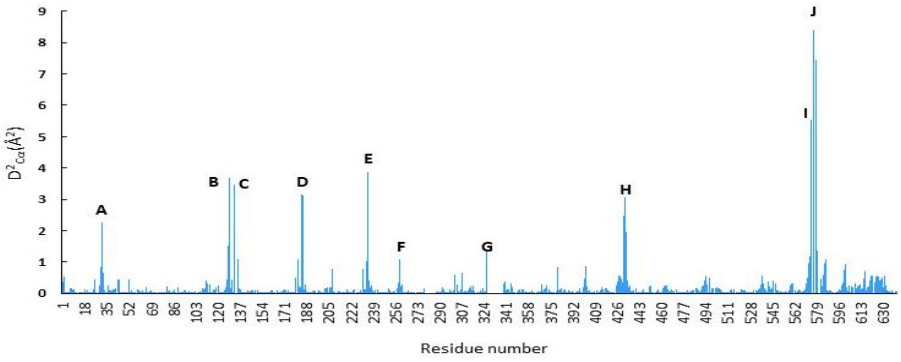


Figure 61. Square of distance between Cα of the identical residues in both refined and the search model (3FHH).

The search model was obtained from co-crystallizing ShuA with $Pb(NO_3)_2$ [30]. In this model, there are four Pb^{2+} ions, however, in ShuA-Eu only one Eu^{3+} is present (see Figure 60) [30]. This ion is located 2.9 Å from the carboxylic acid group of Asp 374 (see Figure 62) which is close to the upper limit of previously observed Eu-O bonds [167] suggesting that this residue is involved in interactions with heavy metals. It also binds lead in the model structure (see Figure 60/B). Additionally, four detergent molecules (OG) were included and modeled in the electron density map (see Figure 60) compared to the search model which only models one detergent molecule.

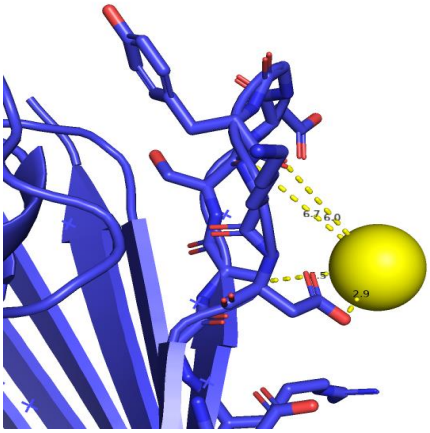


Figure 62. View of the Eu^{3+} binding site to the surface of the ShuA. Distances between residues and the Eu^{3+} are shown with the dashed lines. PyMol was used for preparing the figure. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

In the ShuA-Eu^{3+} structure, electron density sufficient to confidently model 620 amino acid residues was observed. In this model, residues 279-286, 327-338 are missing compared to residues 280-285, 328-337, 579-581 in the search model (see Figure 63). Although these regions are all within loops, it was possible to confidently build one of these loops from the new electron density map.

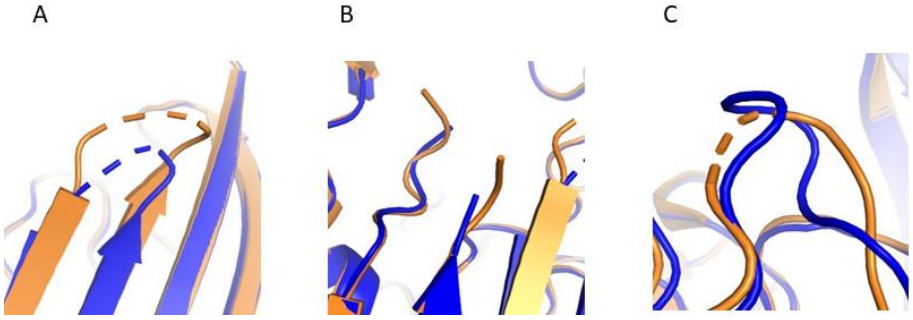


Figure 63. Missing residues in the refined model (blue) and search model (orange). A) Residues 279-286 (in search model 280_285) that are located in loop 4, B) Residues 327-338 (in search model 328-337) that are located in loop 5 and C) Residues 579-581 of the search model. In the refined model, electron density corresponding to these residues is present and loop 10 was built using COOT. PyMol was used for preparing the figure. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

Unfortunately, current data resolution does not allow a more accurate description of ShuA geometry and would require higher-resolution diffraction data for understanding the mechanism of hemoglobin recognition and extraction of heme from it.

4.3.1.3 Diffraction results for AcrB and ShuA using HiLiDe method

Characterization of ShuA and AcrB crystals by X-ray diffraction was performed but diffraction was only at low resolution (best diffraction ~10 Å) precluding full structural determination. Apart from the optimization for the crystallization process to obtain higher quality crystals, one reason for poor diffraction could relate to the cryoprotectant used for these samples as crystals were damaged and in some cases obscured by ice crystals. 20% Glycerol in a protein buffer was used as a cryoprotectant, which is different from the optimized cryoprotectant protocol for both proteins.

4.3.2 Microcrystal electron diffraction (MicroED)

Crystals of AcrB-Rifampicin obtained from batch and dialysis methods were analyzed using electron diffraction, as shown in Figure 64. Data collection was performed at cryogenic temperatures (~100 K) at the Institut de Biologie Strucutrale (IBS) in Grenoble, France using a hybrid pixel detector (Cheetah, Amsterdam Scientific Instruments). These crystals did not diffract well enough (not enough reflections) to solve the structure of this protein (Figure 64). However, crystal lattices can be visualized using TEM [168].

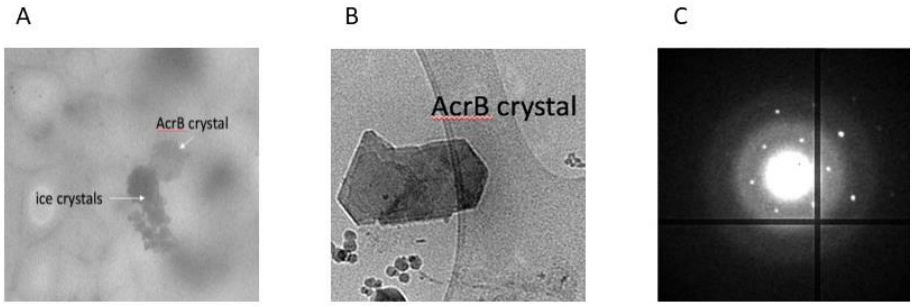


Figure 64. A, B) AcrB Crystals and C) electron diffraction pattern of AcrB from (B). Crystallization condition was 10% PEG 4000, 0.05 M ADA pH6.5, 5% glycerol, 0.2 M ammonium sulfate, 50 mM Rifampicin and 0.02% DDM. The batch method was used for crystallization of this sample.

In Figure 65, cryo-images of ShuA-Eu crystals are illustrated. These crystals were transferred from the dialysis button (see 3.2.1.2.4) to a 3-well glass plate. Samples were then transferred to grids and prepared as explained in 4.2.2.1. Cryo images of some of these crystals are shown Figure 65. From Figure 65, it is clear that the crystals were dissolving or degrading before full data collection

could be performed. In particular, the edges of the crystal A1.1 (see Figure 65,) and the whole crystal B1.1, are not straight. There are several hypotheses for this observation. The irregular edges may indicate that these crystals are not yet mature to have the Wulff shape but are closer to the nucleation state. However, we cannot exclude the possibility that they might have been damaged during sample preparation. Although crystals were transferred from the dialysis button to the glass plate at the same temperature for the crystallization (20 °C), the crystallization conditions were changed after the buttons were removed from the well containing crystallization solution. Another reason, especially for the second sample, could be the addition of a crystallization solution to the button to dilute the sample. The time between the removing samples from the button and grid preparation could also be another source of problem. Finally, cryo-plunging during grid preparation may have damaged delicate samples like membrane proteins [144].

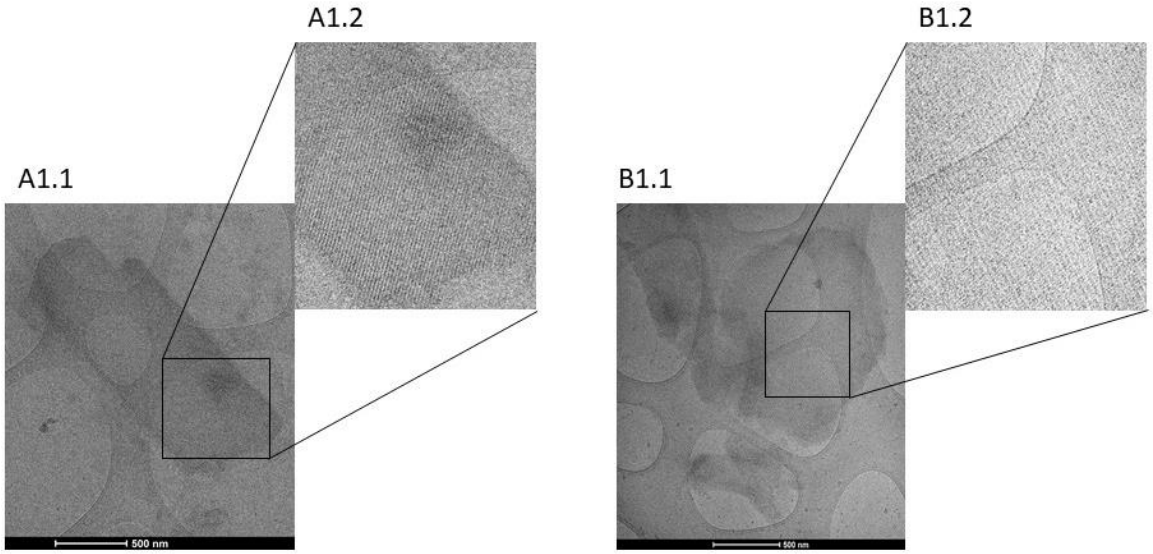


Figure 65. A/B 1.1 are cryo-images and A/B 1.2 are lattices of the ShuA crystals. Both crystals were obtained using the dialysis method. See 3.2.1.2.4 for the crystallization conditions.

4.3.3 Neutron diffraction: preliminary considerations

Crystals of AcrB-rifampicin were obtained which from general observation would seem to be large enough for data collection (~ 750 µm) by neutron diffraction as the typical crystal size used is ~ 1 mm^3. However, unit cell size is also an important consideration as it is the effective volume, $V_{crystal}/(V_{cell})^2$, which determines achievable crystallographic diffraction and resolution [169]. Some studies have successfully collected using small isometric crystals (0.1- 0.2 mm^3) [93]. The unit cell parameters of AcrB-rifampicin crystals are unfortunately quite large (142.721, 522.105 Å) and the X-ray diffraction resolution (~ 3.5 Å) is quite low suggesting that data collection in current-generation beamlines such as LADI III is not realistic. Typically smaller unit cell parameters (<100-150 Å for each edge) and higher diffraction resolutions (1-2.5 Å) are recommended for biological crystals for neutron diffraction [170]. When the unit cell is large like

this, there is even more emphasis on physical crystal size to determine the success of data collection. Therefore, even larger AcrB-rifampicin crystals suitable for neutron diffraction could still be prepared using an extension of the same methodology used in this work (see Equation 1, chapter 1). In addition, perdeuteration (see 1.1.5.2) of the protein is likely to improve the chances of successful data collection [171]. Alternatively, as reflection intensity is proportional to the beam flux [169] the high flux European Spallation Source (ESS), currently under construction may be suitable for data collection with these crystals when it comes into operation in the near future. This source allows data to be collected in much shorter time from much smaller crystals (0.01 mm^3) than is possible to be done today [172].

In contrast to AcrB, ShuA crystals have smaller unit cell parameters that make them suitable candidates for data collection with current neutron beamlines (79.09, 114.24, 117.95 Å). Unfortunately, crystals > 300 µm have not been obtained to date so data collection is still not feasible. For ShuA, the rational strategies for growing large crystals presented in this thesis such as crystallization bench and knowledge of the phase diagram could be used to grow larger crystals (see 3.2.2.5.2.1). As such, data collection with current neutron sources should be achievable. But having a sufficient amount of protein sample for further optimization would be required to obtain crystals of the size exploitable for data collection with the current neutron sources. Alternatively, the current crystals may also be suitable candidates for data collection using a high flux spallation source, when these become available for use.

4.4 Discussion

One of the predominant methods used for the structure determination of biomacromolecules is crystal diffraction [157]. Successful use of the various diffraction techniques can give considerable information on (membrane) protein structure in a single experiment. Crystallographic methods are complementary [173] and combining the information obtained from each of them can lead to a better understanding of the structure and function of target proteins. Earlier in this chapter, crystals of two model membrane proteins were prepared and different diffraction techniques were used in order to assess the quality of these crystals. With X-ray crystallography, the quality of the AcrB-Rifampicin crystals obtained from different crystallization techniques was also compared. This showed that for all methods, better quality crystals can be obtained, but some modifications and optimizations are needed. In particular, the vapor diffusion and batch methods generated highly variable crystals. Therefore, preparing and testing numerous crystals was necessary in order to find the crystal with the best diffraction.

All data presented in this thesis from X-ray diffraction was obtained at cryogenic temperatures (~100 K) using synchrotron sources. It should be noted that preliminary room temperature measurement of AcrB crystals prepared without additives was also attempted using an in-house X-ray laboratory source, however, diffraction was not observed (data not shown). This

preliminary experiment was performed before the majority of optimization experiments were carried out.

In X-ray crystallography, using synchrotron radiation sources increases the chance of radiation damage to crystals during data collection. Flash-cooling crystals using liquid nitrogen and data collection at cryogenic temperatures prolongs crystal lifetime as the reactive products of photochemical reactions such as radicals should be immobilized at low temperature preventing damage to non-exposed parts of the crystal [174].

In crystals of AcrB and ShuA, the solvent contents are around 72% and 67% respectively. Flash freezing the crystals is key to avoid ordered ice formation which could otherwise damage them. As an additional precaution, cryoprotectant was applied before flash freezing the crystals in most cases. For AcrB, several cryoprotectants were successfully used by other researchers previously [117], [126]. Pos et al. used two different cryoprotectants for AcrB crystals. In the first, crystals formed in PEG 4000 and 5% glycerol were protected in the same solution with 30% glycerol. In the second, 15% PEG 400 was used for crystallization and 30 % for cryoprotection. Although crystals cryocooled using these cryoprotectants diffracted to approximately 3.5 Å, the size of crystals handled in this way seemed to affect the quality of data and larger crystals (larger than 400 µm in each edge) diffracted poorly [117]. Increasing the glycerol concentration stepwise from 10% to 30% in 5% steps [124], [129], [132], [175] and using 5% higher concentration of PEG in comparison to the crystallization solution seemed to be optimal for the crystals detailed earlier in this chapter. One example of successful use of this cryoprotection method for AcrB crystals without additives is demonstrated by PDB ID 2DHH, a structure that diffracted to 2.8 Å [132]. Use of a suitable cryoprotectant when fishing crystals is therefore an important post-crystallization process, especially when the size of the crystal is large. Otherwise, it is easy to damage such crystals leading to poor diffraction quality [176].

Large crystals from hanging drop vapor diffusion and batch crystallization were more sensitive to the 3 to 6% step increase in cryoprotectant glycerol concentration in comparison to the crystals from dialysis. This is in agreement with previous work that showed large crystals obtained from vapor diffusion being sensitive to glycerol concentration during cryoprotection. For these crystals, increasing the glycerol concentration in the drops lead to cracks forming in them before flash freezing [175]. In this work, the soaking method was used to apply cryoprotectant, but it is also possible to use dialysis when soaking is not successful [174]. Where crystals can be grown in a crystallization solution that can also be used as a cryoprotectant, less crystal handling is necessary which decreases the risk of damage to the crystals [174]. In the case of AcrB, some crystals were grown and cryoprotected in the presence of PEG 400 (data not shown) but these did not diffract to an expected resolution. This could be due to the crystallization process generating crystals of lower diffraction quality or this crystallization solution may not have been sufficient to act as a cryoprotectant to prevent damage to these crystals during freezing. High pressure (210 MPa) also can be used for vitrification of crystals before flash freezing. AcrB crystals flash frozen using this method without cryoprotectant diffracted to 4.2 Å [176].

For ShuA, crystals were flash-frozen both in the presence and absence of cryoprotectant and crystals diffracted to the same resolution (2.6 [123]-2.49 Å) and had comparable diffraction statistics (4.3.1.2.1). This highlights how finding suitable conditions for sample handling prior to data collection varies depending on protein and is still very much an empirical process. Where cryoprotectant was used, the heavy atom additives necessary for high quality crystal formation were not added to the cryoprotectant. Similarly for the previously published ShuA structure (3FHH) the cryoprotectant used was the crystallization solution plus 15% (v/v) glycerol and no additional heavy atom solution [123]. Cryoprotectants can sometimes compete with the additives positioned within a crystal [174] so even for short soaking times it may be better to add the same additive to the cryoprotectant solution as well. Otherwise, this is another benefit of not using cryoprotectants when such is not necessary.

The quality of diffraction data determines the quality of the electron density map and therefore the possibility of accurately modeling atoms within a crystal structure. Parameters generally used to check the quality of diffraction data are resolution, data completeness, $CC_{1/2}$, I/σ(I), and R_{meas}. For crystals analyzed in this work, these parameters were checked both in total and for the highest resolution shell determined by XDS [84]. Opinions vary about the most important parameter to be used to determine the true resolution cutoff. For example Karplus et al. suggest that $CC_{1/2}$ is the best parameter for this [162] whilst I/σ(I) is favored by others [177]. For ShuA data I/σ(I) $\geqq$ 2 were used as this value is traditionally used for resolution cutoff and deposition of the structure to the PDB. However, for comparing crystals obtained from AcrB, I/σ(I) $\geqq$ 1 was used instead as these data were not used for model building and refinement, rather just for comparison of the diffraction data.

After choosing the resolution cutoff, data obtained from AcrB crystals were compared. From Table 12 it seems generally that further purification of protein prior to crystallization improves the quality of AcrB crystals. Additional steps in the purification of the proteins allows for the removal of many types of aggregates and impurities. As a result, a homogeneous solution of protein oligomers will then generate better quality crystals with a regular lattice. However, research groups that were using size exclusion and affinity chromatography for protein purification did not overall obtain higher resolution crystals in comparison to the other groups [127]–[129]. On the other hand, the highest resolution of AcrB crystal without additives (PDB 2DHH) was obtained from protein purified with affinity chromatography.

Co-crystallization with additives can improve crystal quality, although optimization of this is necessary as illustrated by co-crystallization of AcrB with erythromycin, which resulted in poorly diffracting crystals. Well-chosen additives can also aid crystal lattice formation by increasing the probability of interaction between macromolecules or stabilizing otherwise flexible areas of proteins by bond formation [4], [178]. Stabilization of these areas by additives has allowed for determination of their electron density. In contrast, when these areas were flexible before, they had no consistent position in the crystal unit cell and therefore did not contribute to the electron density resulting in missing areas in the electron density map [179].

It is noteworthy that all AcrB crystal structures deposited in the PDB were obtained using hanging and sitting drop vapor diffusion, and no structures were found which used the batch or dialysis crystallization methods.

In electron diffraction studies of biological samples like protein crystals, samples are vitrified before data collection to preserve the crystals in their hydrated state and to protect them from radiation damage [100], [180], [19]. Therefore, sample preparation and freezing the samples are even more important steps for ED. Initial structural studies of our two model membrane proteins with MicroED look promising, even though exploitable diffraction patterns were not achieved using the crystals detailed in this work. The next chapter discusses what can be done next to increase the chance of solving membrane protein structures with this technique.

Despite all efforts to obtain large enough crystals of ShuA for neutron crystallography, crystals were still not large enough for data collection at LADI III, where perdeuterated crystals with 0.1-0.2 mm^3 are typically used for data collection [95]. Nevertheless, the unit cell volume of these crystals is in a good range. Crystallization benches could be used for the further size optimization of crystals of this protein without consuming a large amount of sample, which is another bottleneck in crystal growth for a protein with a low yield. With the current beamline (LADI III, Grenoble France), it is not possible to study AcrB crystals for the opposite reason. Although the crystal size after optimization appeared initially acceptable for study, the large volume of the unit cell of AcrB crystals makes this protein a poor candidate for data collection. On the other hand, neutron diffraction data collection on the crystals of this protein can be attempted in the future at the ESS where a high flux of neutrons will make it possible to collect data from smaller crystals or crystals with a large unit cell volume in a shorter time.

Chapter 5: Conclusion and future work

5.1 Summary and conclusions

Protein crystallization and optimization of this process is a bottleneck for structural studies using crystallographic techniques [181]. To begin with, conditions that lead to crystallization should be identified through screening and then different variables should be tuned to optimize the crystal quality [4]. The crystallization process, which includes nucleation and growth, can be investigated using a phase diagram in order to rationalize and guide this optimization [182]. The phase diagram describes different phases of the protein (liquid or solid) as a function of different variables such as temperature (physical variable) or crystallization solution composition (chemical variables) [39].

The goal of this PhD study was to optimize crystal growth of two model membrane proteins (ShuA from *Shigella dysenteriae* and AcrB from *Escherichia coli*) for structural studies using X-ray and neutron crystallography and MicroED. Each crystallography technique requires high-quality crystals within a specific range of sizes. For example for MicroED crystals with micro- and nanometer dimensions are used [183] whereas for neutron crystallography large crystals (> 0.1 mm^3) are required [184]. X-ray crystallography using synchrotron radiation sources is less stringent and can use crystals of 20 µm or even smaller [185]. This thesis describes an optimization process to obtain the best quality crystals for crystallographic technique.

5.1.1 Optimizing crystallization of AcrB

One of the most important biochemical variables affecting the quality of protein crystals is the purity of the protein sample [4]. Purification is reported in this work either as a one-step purification using affinity chromatography with nickel-charged affinity resin or a two-step process using affinity and size exclusion chromatography or affinity and anion exchange chromatography.

One-step purification by affinity chromatography has generally been used for purifying this protein previously [25], [124]–[126], although some research groups used two-step purification [21], [127]–[129]. Different methods were used to optimize the purification process in this work. Different methods were used to disrupt the cell membrane, including lysis buffers, ultrasound and high-pressure lysis with a microfluidizer. The microfluidizer was identified as the preferable technique for this protein because it was more efficient than other methods used in this study for cell membrane lysis and consequently higher yield of protein was obtained. Optimization of the incubation time and other modifications lead to more than 6 mg of protein per 1 liter of cell culture being obtained after affinity chromatography purification. Even after two-step purification, some proteins co-purify with AcrB as visualized by SDS-PAGE. Western blot analysis revealed that these bands are His-tagged species and therefore likely to originate from AcrB.

Then, the purified AcrB samples were used for crystallization experiments applying different techniques (vapor diffusion (hanging drop), dialysis, batch method and HiLiDe). AcrB samples

obtained from different purification methods were also used in parallel crystallization trials. Crystals were formed using protein samples from all purification methodologies mentioned above, although samples prepared using only affinity chromatography led to significant visible protein precipitation during crystallization trials. This suggests that the two-step purification protocol gives better results, in agreement with previous studies of protein crystallization [1]. Finally, AcrB crystals were obtained using the HiLiDe method with 2:1 protein to lipid ratio and changing the ratio did not improve crystallization.

Once crystals were obtained, crystallization variables like temperature, protein and precipitant concentration and pH were further studied in order to optimize crystallization of AcrB. Different additives were also used for the optimization of crystal growth. AcrB was co-crystallized with Erythromycin and Rifampicin. More optimization experiments would be needed to find the optimal conditions for using Erythromycin with AcrB, as the current experiments resulted in crystals unusable for X-ray diffraction studies.

In chapter 4, the diffraction quality of the AcrB crystals was studied. Crystals obtained by hanging drop vapor diffusion of protein purified using different purification protocols were compared by X-ray diffraction. Generally, AcrB crystals showed better diffraction quality when the protein had undergone more purification steps and was assessed to be pure from SDS-PAGE. The homogeneity of a crystallizing protein sample is likely to affect both how quickly a crystal lattice forms and how regular the lattice is. Impurities in the proteins may hinder crystal growth or introduce defects in the crystals, resulting in the observed poor diffraction quality [186].

The diffraction quality of AcrB crystals prepared using various crystallization techniques was also assessed. X-ray diffraction experiments were performed to compare AcrB crystals that were grown using the vapor diffusion hanging drop, dialysis or batch methods. After the crystallization conditions were optimized, crystals with comparable diffraction quality were obtained from all methods. It is noteworthy though, that the crystal with the best observed diffraction was generated by the dialysis method. Moreover, the diversity in the quality of crystals grown with the batch and the hanging drop vapor diffusion methods was higher than that of crystals obtained with the dialysis method. Therefore, the dialysis method is preferable if crystals of consistent quality are desired. Although different crystallization techniques may be more suitable depending on the goal of a study.

Although the purity of the protein sample is one of the crucial factors to obtain high-quality crystals, the appropriate additive can also improve the crystal quality significantly. In this study, co-crystallization of AcrB with rifampicin significantly improved the quality of the crystals. Using $I/\sigma(I) \geqq 1$ for resolution cut-off, similar results were observed from crystals of proteins obtained from one-step purification by affinity chromatography plus rifampicin. This suggests that Rifampicin affects the interactions and facilitates the contact between protein molecules in the crystal lattice [4]. Erythromycin was also used in this study for co-crystallization with AcrB. However, the AcrB-erythromycin crystals did not diffract well enough to collect full data sets, despite some research groups obtaining high-resolution data by co-crystallization of AcrB with

erythromycin [126], [140]. Only a limited number of small size (~ 10 µm) AcrB-erythromycin crystals were used for diffraction studies in comparison to the AcrB-rifampicin crystals. Therefore, further optimization may still result in improved crystal size or diffraction properties for AcrB-erythromycin crystals as has been observed in previous work.

During the optimization of the crystallization conditions of the AcrB-rifampicin complex, crystals of approximately 750 µm in the largest dimension were obtained. The unit cell parameters of these crystals were identified using X-ray diffraction and the respective values were a = b =142.721 and c = 522.105 Å. As such, despite being a suitable physical size for neutron diffraction studies, the large unit cell of these crystals precludes data collection using current neutron sources. As diffraction intensity is related to the effective crystal volume, $V_{crystal}/(V_{cell})^2$. It is possible that using the crystallization bench and the qualitative phase diagram now obtained for Acrb-rifampicin, even larger crystals could be grown with relative ease that would be suitable for neutron diffraction. Despite all the effort applied during this study, the highest resolution achieved for AcrB crystals was 3.76 Å. It is likely that the physical properties of AcrB make it difficult to obtain higher resolution data. Although a resolution of 2.8 Å (PDB ID: 2DHH) was reported for AcrB and 1.90 Å (PDB ID: 4DX5) for the AcrB/DARPin in complex with minocycline [132], the most common resolution range is between 3 to 4 Å. Therefore using this protein for advanced crystallographic techniques is still highly challenging.

The 2.8 Å AcrB structure (2DHH) along with another at 3.5 Å (1IWG) were reported by Murakami *et al*. For this report, two different detergents, dodecanoyl sucrose (0.3 mM CMC) and DDM (0.17 mM CMC) were used for purification and crystallization of the protein. The crystallization solutions contained PEG 4000 and PEG 2000, respectively. The symmetry and solvent content of these two structures is different. 1IWG has the space group R32 and 73.07 % solvent content, however, 2DHH has 65.65 % solvent content with C2 symmetry [25], [132]. Although, crystallization is a multi-parametric process, using different detergent used for solubilization and crystallization of the protein in this case also affects the packing of molecules in crystals. Lower CMC detergents form large micelles [70], which can hinder molecular contacts and as a consequence change the packing of the molecules in the crystals. In addition, a tenfold higher DDM concentration (0.2%) was used for the preparation of 1IWG than that used in the crystallization solution of AcrB-rifampicin in this study (0.02%). Both structures have comparable resolutions, 3.5 Å and 3.76 Å and the same space groups of R32. This illustrates how the detergent CMC/micellar size (not concentration) of detergent can clearly influence molecular contacts and crystal packing to the extent that symmetry can be altered. The higher resolution of 2DHH was not achieved in this study that could be related to the different symmetry for this structure compared to the symmetry obtained for the crystals in this study. It shows the simultaneous effect of different parameters on the final resolution.

5.1.2 Optimizing crystallization of ShuA

Similar to AcrB, one-step (affinity chromatography using nickel-charged affinity resin) and two-step (affinity and size exclusion or anion exchange chromatography) purification protocols were used.

Also in the ShuA purification, three different detergents were compared for the solubilization of the membrane. The results from these experiments suggested that DDM solubilizes ShuA better in comparison to the other two detergents. After initial solubilization, the purification was then continued using OPOE. Different concentrations of imidazole were then used in parallel purifications to find the optimum concentration for elution from the Ni-NTA column during affinity chromatography purification. These experiments identified that a buffer containing 250 mM of imidazole could elute almost all of the protein bound to the column. Similar to AcrB, one- and two- step purifications were compared. Overall, 0.3 mg of protein per 1 liter of culture was obtained after size exclusion chromatography, whereas 0.6 mg of protein per 1 liter of culture was obtained after anion exchange chromatography. On the basis of the experiments performed, crystals obtained from purified protein using anion exchange chromatography compared to size exclusion, diffracted at a better resolution by X-ray crystallography.

Crystals of ShuA were obtained using vapor diffusion hanging drop, dialysis, HiLiDe and LCP methods. However, the batch method did not yield any ShuA crystals. From comparison to the results from other methods, the final protein concentration in all batch experiments is likely to be below supersaturation. Consistent with this, no nucleation or aggregation was observed in these batches, suggesting that the supersaturation of the mixture is not sufficient to cause nucleation. Further experiments using higher concentrations of protein solution in the presence and absence of additives could be explored to optimize crystallization with this method. Although the large number of unsuccessful experiments and the fact that other techniques have more easily identified the supersaturation point highlights one of the weaknesses of batch crystallization for initial crystallization trials in case of this protein. Alternatively macroseeding could be used (see 3.2.2.2) to determine if the solution can sustain the crystal growth of crystal seeds in the metastable zone of the phase diagram.

The HiLiDe method with 2:1 and 3:1 protein to lipid ratios was used and both samples gave small crystals suitable for electron diffraction studies. LCP was also attempted for crystallization of lysozyme and ShuA. Lysozyme crystals grown inside the LCP were smaller compared to the crystals from the same drop outside the LCP (see appendix III). On the basis of this observation, which is in agreement with previous works [7], LCP was used to crystallize ShuA specifically to obtain small crystals for MicroED experiments. Using the conditions suggested by Franck Brachet (IBPC Paris, personal communication), small ShuA crystals were generated using this method.

For optimization of crystal growth, the crystallization bench was used to control some physico-chemical variables affecting the protein crystallization phase diagram in an automated way, starting with lysozyme as a standard soluble protein. The first nucleation yielded a large number

of small lysozyme crystals (approximately 50 µm) by maintaining the temperature at 18 °C. Then the temperature was increased incrementally to 35 °C, whereupon all crystals dissolved. The temperature was then rapidly decreased to 18 °C inducing rapid supersaturation. This effectively changes the width of the metastable zone and causes a fast second nucleation at lower supersaturation. Because of this, lysozyme crystals from the second nucleation were smaller (approximately 10 µm) compared to the crystals formed during the first nucleation [53].

A complementary experiment was performed using the first generation crystallization bench to show the crystallization optimization of ShuA at constant temperature, this time by controlling the chemical composition of crystallization solution during the experiment. Higher precipitant concentration was used for the first nucleation and after equilibrium was achieved, the precipitant concentration was decreased to dissolve the crystals. For the second nucleation a lower concentration of precipitant was used causing nucleation to occur near the metastable zone, leading to a smaller number of nuclei in comparison to the first nucleation. Due to this, fewer, larger crystals formed (approximately 300 µm).

ShuA crystal growth experiments were also performed by controlling the temperature with the crystallization bench. Once crystals were obtained at 20 °C, temperature was incrementally increased similarly to the study conducted using lysozyme. However, high temperature values could lead to protein denaturation so a sensible limit to the temperature range was used for ShuA experiments. At the maximum point of this temperature range (40 °C), ShuA crystals still did not dissolve. To avoid the risk of heat denaturation, ShuA crystals were instead dissolved by changing the concentration of the crystallization solution. This highlights one of the complexities inherent in the crystallization of membrane proteins, where a complex of protein and detergent molecules is the species being crystallized. Detergent micelles can be formed in a temperature-dependent manner complicating the crystallization process [71] and as such, the effect of temperature on both protein and detergent should be considered. Therefore, optimizing the crystallization of membrane proteins by controlling temperature cannot be simply described by the protein crystallization phase diagram. Instead, a detergent phase diagram should also be taken into account, which complicates the control and optimization of this process by temperature but does not detract from its importance or potential utility.

Lead(II)nitrate and europium(III)chloride were used for co-crystallization with ShuA. More optimization would be needed to find the best conditions for using both additives (especially Lead(II)nitrate). When ShuA was mixed with heavy atom salts, even after extensive optimization, some protein precipitate forms. This makes the quantification of ShuA in these experiments difficult even though medium quality crystals were obtained from co-crystallization of ShuA with europium(III)chloride. In this study, ShuA-Lead(II) nitrate crystals did not diffract well enough that data collection was practical, although other research groups have reported high-resolution structures (2.6 Å) from this species [30].

The quality of ShuA-Eu crystals obtained from hanging drop vapor diffusion crystallization experiments was studied using X-ray crystallography. Diffraction data were used to build a 3D

structural model using the molecular replacement approach with the search model PDB ID: 3FHH. This model has an 87% similarity in secondary structure to the search model despite the different additives used for co-crystallization. Therefore, the addition of Europium(III) chloride salt has improved the diffraction quality of ShuA crystals without altering the structure of the protein. Therefore, disorder in these crystals is less than the native crystals, which is due to higher intermolecular interaction and bridging the molecules by these positively charged heavy ions [30]. This enhanced diffraction quality has allowed for the modelling of three extra detergent molecules in this structure in comparison with the PDB ID: 3FHH model. In addition, there is electron density in the new map sufficient to confidently model loop 10 which was missing in 3FHH.

5.2 Recommendations for future work

This PhD thesis investigates the optimization process that can be used to improve crystal quality and control the size of protein crystals required for different advanced diffraction techniques. This work aims to give a better understanding of the use of protein crystallization phase diagrams for the rational optimization of this process, especially for membrane proteins. A qualitative phase diagram for the crystallization of AcrB has been established, in support of the experiments in this study, however, supplementary experiments exploring a wider range of crystallization conditions are suggested in order to obtain a more accurate quantitative phase diagram. One approach would be by using the crystallization bench to control the concentration of precipitating agents present in the crystallization solution. This way, the position of the solubility curve in the phase diagram could be determined without any need to consume a large amount of protein. In the case of the model membrane proteins, incremental variation of temperature can also be used to find the precise optimum temperature for the crystallization, similar to the result obtained by using different temperatures (20 and 25 °C) for crystallization of AcrB-rifampicin. Unlike this experiment, the crystallization bench requires only one sample, so the effect of variation of temperature on crystallization can be studied with less protein and in more detail.

Despite optimization of the purification process and crystallization conditions, high-resolution X-ray diffraction data were not obtained for AcrB samples. Focusing on a different model membrane protein (ShuA is a more suitable candidate) for optimization of crystallization might be more practical to give better information about this process. However, some other modifications can still be attempted to improve the purification and crystallization of AcrB. Using different detergents like dodecanoyl sucrose [132] or another purification method like using SMALPs, which preserves the lipid around the protein, may still improve crystal quality. Also, complexes of AcrB with designed ankyrin repeat proteins (DARPins) have been used to prepare crystals that diffract to high resolution (1.90Å [125]), although screening and optimization of crystallization conditions would be required.

In this study, control of precipitant concentration using the crystallization bench was used to grow crystals of ShuA. However, experiments where temperature was adjusted as a crystallization variable, did not proceed as expected. Therefore, the effect of temperature on the crystallization of a detergent-protein complex should be further investigated. Mass transport of detergent through semi-permeable membranes at different temperatures could also be studied using the crystallization bench. Detergent can be added to the protein chamber and Milli Q water on the reservoir chamber. In the time intervals, liquid from the reservoir can be taken and used for colorimetric studies. As the volume of detergent is less than Milli Q water less time is needed to reach the equilibrium. In addition, some complementary experiments could be performed to investigate the effect of salts and PEGs on detergent mass transport at different temperatures, as each of these variables can affect the CMC of the detergent and thus the mixing time.

Using the crystallization bench to screen different precipitant concentrations and temperatures would also be worthwhile for ShuA crystal growth in order to explore the phase diagrams for this protein. Using different crystallization techniques and comparing crystals grown from them can give insights for the optimization of crystal size for neutron and electron diffraction studies. Although the crystallization bench was used to grow crystals with sufficient effective volume (especially in the case of ShuA) for neutron crystallography, sample mounting and H/D exchange is a challenging process that would require its own optimization. Generally, data collection for neutron crystallography is performed at room temperature, however, for crystals that are not stable at room temperature neutron cryo-crystallography is also possible [93].

HiLiDe and LCP can be used to produce small uniformly sized crystals of AcrB and ShuA for MicroED. However, there are still challenges for sample preparation, especially for the viscous samples obtained using the LCP method. Using some additives such as MPD (2-Methyl-2,4-pentanediol), the cubic phase can be converted to the less viscous sponge phase. Lipase can also be used to convert the cubic phase to liquid phases containing water/glycerol and oleic acid [187]. By applying these modifications, membrane crystals could be attempted to grow in the LCP with the objective to be used for MicroED studies.

Another challenge of using MicroED is grid preparation, especially for delicate samples like membrane proteins [144]. This step can be optimized by using several blotting times or cryoprotectants. It is also possible to use TEM to have a direct image with near atomic resolution of protein structure in liquid water [188]. In this method, a sample is loaded into a microfabricated liquid cell with silicon nitride windows. This way the sample is isolated from the vacuum necessary for TEM and data collection can be done at the room temperature [189]. There are also some significant challenges when using this method such as poor resolution due to the liquid layer and silicon nitride windows and heat damage caused by the electron beam, electrostatic charging and radiolysis. Although this method has some limitations there is high potential for improvement in the future [190].

References

[1] J. A. Gavira, “Current trends in protein crystallization,” *Arch. Biochem. Biophys.*, vol. 602, pp. 3–11, Jul. 2016, doi: 10.1016/j.abb.2015.12.010.

[2] A. Govil and G. Govil, “Physical Methods for Determination of Biomolecular Structures,” *Proceedings of the National Academy of Sciences India Section A - Physical Sciences*, vol. 88, no. 2. Springer India, pp. 163–169, 2018, doi: 10.1007/s40010-018-0512-8.

[3] G. E. Dale, C. Oefner, and A. D’Arcy, “The protein as a variable in protein crystallization,” *J. Struct. Biol.*, vol. 142, no. 1, pp. 88–97, Apr. 2003, doi: 10.1016/S1047-8477(03)00041-8.

[4] A. McPherson and J. A. Gavira, “Introduction to protein crystallization,” *Acta Crystallogr. Sect. FStructural Biol. Commun.*, vol. 70, no. 1, pp. 2–20, 2014, doi: 10.1107/S2053230X13033141.

[5] A. Krogh, B. Larsson, G. Von Heijne, and E. L. L. Sonnhammer, “Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes,” *J. Mol. Biol.*, vol. 305, no. 3, pp. 567–580, Jan. 2001, doi: 10.1006/jmbi.2000.4315.

[6] J. G. Almeida, A. J. Preto, P. I. Koukos, A. M. J. J. Bonvin, and I. S. Moreira, “Membrane proteins structures: A review on computational modeling tools,” *Biochimica et Biophysica Acta - Biomembranes*, vol. 1859, no. 10. Elsevier B.V., pp. 2021–2039, Oct. 01, 2017, doi: 10.1016/j.bbamem.2017.07.008.

[7] J. Birch, D. Axfordc, J. Foadid, A. Meyere, A. Eckhardte, Y. Thielmannf, I. Moraes, “The fine art of integral membrane protein crystallisation,” *Methods*, vol. 147, pp. 150–162, Sep. 2018, doi: 10.1016/j.ymeth.2018.05.014.

[8] M. C. Deller, L. Kong, and B. Rupp, “Protein stability: A crystallographer’s perspective,” *Acta Crystallogr. Sect. Struct. Biol. Commun.*, vol. 72, pp. 72–95, 2016, doi: 10.1107/S2053230X15024619.

[9] J. M. Finkelstein, “Structures of membrane proteins,” *Nature*, vol. 511, no. S7509, pp. 21–21, Aug. 2014, doi: 10.1038/nature13372.

[10] N. Asheric, C. Ginsberg, A. Greenbaum, S. Blass, and S. Knafo, “Effects of protein purity and precipitant stereochemistry on the crystallization of thaumatin,” *Cryst. Growth Des.*, vol. 8, no. 12, pp. 4200–4207, 2008, doi: 10.1021/cg800616q.

[11] A. Shahmohammadi, “Lysozyme separation from chicken egg white: a review,” *Eur. Food Res. Technol.*, vol. 244, no. 4, pp. 577–593, 2018, doi: 10.1007/s00217-017-2993-0.

[12] N. Arnheim, M. Inouye, L. Law, and A. Laudin, “Chemical studies on the enzymatic specificity of goose egg white lysozyme.,” *J. Biol. Chem.*, vol. 248, no. 1, pp. 233–236, 1973, doi: 10.1016/S0021-9258(19)44466-9.

[13] L. Callewaert and C. W. Michiels, “Lysozymes in the animal kingdom,” *Journal of*

Biosciences, vol. 35, no. 1. pp. 127–160, Mar. 2010, doi: 10.1007/s12038-010-0015-5.

[14] Blake, D. Koenig, G. Mair, A. North, D. Phillips, and V. Sarma, "STRUCTURE OF HEN EGG-WHITE LYSOZYME A Three-dimensional Fourier Synthesis at 2 A Resolution," *Nature*, vol. 206, no. 49, pp. 757–761, 1965.

[15] S. Strazdaite, E. Navakauskas, J. Kirschner, T. Sneideris, and G. Niaura, "Structure Determination of Hen Egg-White Lysozyme Aggregates Adsorbed to Lipid/Water and Air/Water Interfaces," *Langmuir*, vol. 36, no. 17, pp. 4766–4775, 2020, doi: 10.1021/acs.langmuir.9b03826.

[16] L. Dufossé, F. Almeida, J. F. M. Van Impe, J. A. Joseph, S. Akkermans, and P. Nimmegeers, "Bioproduction of the Recombinant Sweet Protein Thaumatin: Current State of the Art and Perspectives," *Front. Microbiol,* vol.10. pp. 695, 2019, doi: 10.3389/fmicb.2019.00695.

[17] H. M. Berman et al., "The Protein Data Bank," *Acta Cryst. D*, vol. 58, no. 6, pp. 899-907, 2002.

[18] S. Tan, T. T. Hwee, and M. C. M. Chung, "Membrane proteins and membrane proteomics," *Proteomics*, vol. 8, no. 19. pp. 3924–3932, Oct. 2008, doi: 10.1002/pmic.200800597.

[19] M. C. Wiener, "A pedestrian guide to membrane protein crystallization," *Methods*, vol. 34, no. 3, pp. 364–372, 2004, doi: 10.1016/j.ymeth.2004.03.025.

[20] H. Michel, "Crystallization of membrane proteins," pp. 56–59, 1983.

[21] S. Törnroth-Horsefield, P. Gourdon, R. Horsefield, L. Brive, N. Yamamoto, H. Mori, A. Snijder, and R. Neutze, "Crystal Structure of AcrB in Complex with a Single Transmembrane Subunit Reveals Another Twist," *Structure*, vol. 15, no. 12, pp. 1663–1673, Dec. 2007, doi: 10.1016/j.str.2007.09.023.

[22] D. Das, Q. S. Xu, J. Y. Lee, I. Ankoudinova, C. Huang, Y. Lou, A. DeGiovanni, R. Kim, and S-H Kim, "Crystal structure of the multidrug efflux transporter AcrB at 3.1 Å resolution reveals the N-terminal region with conserved amino acids," *J. Struct. Biol*., vol. 158, no. 3, pp. 494–502, Jun. 2007, doi: 10.1016/j.jsb.2006.12.004.

[23] C. A. Elkins and H. Nikaido, "3D structure of AcrB: The archetypal multidrug efflux transporter of Escherichia coli likely captures substrates from periplasm," *Drug Resist. Updat*., vol. 6, no. 1, pp. 9–13, 2003, doi: 10.1016/S1368-7646(03)00004-9.

[24] E. W. Yu, G. McDermott, H. I. Zgurskaya, H. Nikaido, and D. E. Koshland, "Structural Basis of Multiple Drug-Binding Capacity of the AcrB Multidrug Efflux Pump," *Science*, vol. 300, no. 5621, pp. 976-980.2003. doi: 10.1126/science.1083137.

[25] S. Murakami, R. Nakashima, E. Yamashita, and A. Yamaguchi, "Crystal structure of bacterial multidrug efflux transporter AcrB," *Nature*, vol. 419, no. 6907, pp. 587–593, 2002, doi: 10.1038/nature01050.

[26] S. A. Hare, "Diverse structural approaches to haem appropriation by pathogenic bacteria," Biochimica et Biophysica Acta - Proteins and Proteomics, vol. 1865, no. 4. *Elsevier B.V*., pp.

422–433, Apr. 01, 2017, doi: 10.1016/j.bbapap.2017.01.006.

[27] N. Noinaj, M. Guillier, T. J. Barnard, and S. K. Buchanan, "TonB-Dependent Transporters: Regulation, Structure, and Function," *Annu. Rev. Microbiol.*, vol. 64, no. 1, pp. 43–60, Oct. 2010, doi: 10.1146/annurev.micro.112408.134247.

[28] L. M. Meneghini, S. Tripathi, M. A. Woodworth, S. Majumdar, T. L. Poulos, and G. A. Weiss, "Dissecting binding of a β-barrel membrane protein by phage display," *Mol. Biosyst.*, vol. 13, no. 8, pp. 1438–1447, 2017, doi: 10.1039/c7mb00163k.

[29] K. A. Burkhard and A. Wilks, "Characterization of the outer membrane receptor ShuA from the heme uptake system of Shigella dysenteriae: Substrate specificity and identification of the heme protein ligands," *J. Biol. Chem.*, vol. 282, no. 20, pp. 15126–15136, May 2007, doi: 10.1074/jbc.M611121200.

[30] D. Cobessi, A. Meksem, and K. Brillet, "Structure of the heme/hemoglobin outer membrane receptor ShuA from shigella dysenteriae: heme binding by an induced fit mechanism," *Proteins Struct. Funct. Bioinforma.*, vol. 78, no. 2, pp. 286–294, 2010, doi: 10.1002/prot.22539.

[31] W. B. O. Dell, A. M. Bodenheimer, F. Meilleur, W. B. O'Dell, A. M. Bodenheimer, and F. Meilleur, "Neutron Protein Crystallography: A Complementary Tool for Locating Hydrogens in Proteins," *Arch. Biochem. Biophys.*, vol. 602, pp. 48–60, Jul. 2015, doi: 10.1016/j.abb.2015.11.033.

[32] A. McPherson, "Introduction to protein crystallization," *Methods*, vol. 34, no. 3, pp. 254–265, 2004, doi: 10.1016/j.ymeth.2004.03.019.

[33] J. Gulbis, "Protein crystallography: methods and protocols," *Crystallogr. Rev.*, vol. 24, no. 2, pp. 136–143, 2018, doi: 10.1080/0889311x.2017.1401613.

[34] R. Giegé, "A historical perspective on protein crystallization from 1840 to the present day," *FEBS Journal*, vol. 280, no. 24. pp. 6456–6497, Dec. 2013, doi: 10.1111/febs.12580.

[35] A. McPherson, "A brief history of protein crystal growth," *J. Cryst. Growth*, vol. 110, no. 1–2, pp. 1–10, Mar. 1991, doi: 10.1016/0022-0248(91)90859-4.

[36] J. B. Sumner, "The Isolation and Crystallization of the Enzyme Urease," *J. Biol. Chem.*, vol. 69, no. 2, pp. 435–441, 1926, doi: 10.1016/s0021-9258(18)84560-4.

[37] W. M. Stanley, "Isolation of a crystalline protein possessing the properties of tobacco-mosaic virus," *Science* (80)., vol. 81, no. 2113, pp. 644–645, 1935, doi: 10.1126/science.81.2113.644.

[38] A. Ducruix and R. Giege, Crystallization of nucleic acids and proteins. A practical approach, Second edi. 1999.

[39] J. D. Ng, J. K. Baird, L. Coates, J. M. Garcia-Ruiz, T. A. Hodge, and S. Huang, "Large-volume protein crystal growth for neutron macromolecular crystallography," *Acta Crystallogr. Sect. FStructural Biol. Commun.*, vol. 71, no. 4, pp. 358–370, Apr. 2015, doi:

10.1107/S2053230X15005348.

[40] J. Narayanan and X. Y. Liu, “Protein interactions in undersaturated and supersaturated solutions: A study using light and X-ray scattering,” *Biophys. J.*, vol. 84, no. 1, pp. 523–532, Jan. 2003, doi: 10.1016/S0006-3495(03)74871-1.

[41] N. Asherie, “Protein crystallization and phase diagrams,” *Methods*, vol. 34, no. 3, pp. 266–272, Nov. 2004, doi: 10.1016/j.ymeth.2004.03.028.

[42] M. Budayova-Spano, F. Dauvergne, M. Audiffren, T. Bactivelane, and S. Cusack, “A methodology and an instrument for the temperature-controlled optimization of crystal growth,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 63, no. 3, pp. 339–347, Feb. 2007, doi: 10.1107/S0907444906054230.

[43] R. J. Gray, W. B. Hou, A. B. Kudryavtsev, and L. J. DeLucas, “A new approach to the measurement of protein solubility by Michaelson interferometry,” *J. Cryst. Growth*, vol. 232, no. 1–4, pp. 10–16, Nov. 2001, doi: 10.1016/S0022-0248(01)01142-3.

[44] N. E. Chayen, T. J. Boggon, A. Cassetta, A. Deacon, T. Gleichmann, J. Habash, S. J. Harrop, J. R. Helliwell, Y. P. Nieh, M. R. Peterson, J. Raftery, E. H. Snell, A. Hädener, A. C. Niemann, D. P. Siddons, V. Stojanoff, A. W. Thompson, T. Ursby, and M. Wulff, “Trends and Challenges in Experimental Macromolecular Crystallography,” *Q. Rev. Biophys.*, vol. 29, no. 3, pp. 227–278, Aug. 1996, doi: 10.1017/s0033583500005837.

[45] M. Benvenuti and S. Mangani, “Crystallization of soluble proteins in vapor diffusion for x-ray crystallography,” *Nat. Protoc.*, vol. 2, no. 7, pp. 1633–1651, Jun. 2007, doi: 10.1038/nprot.2007.198.

[46] C. Sauter, B. Lorber, and R. Giegé, “Crystallography of Biological Macromolecules,” in *International Tables of Crystallography*, 2n edition., M. G. R. E. Arnold, D.M. Himmel, Ed. John Wiley and Sons, 2012, pp. 99–121.

[47] X. K. Wang, D. C. Yin, C. Y. Zhang, Q. Q. Lu, Y. Z. Guo, and W. H. Guo, “Effect of temperature programmes on protein crystallisation,” *Cryst. Res. Technol.*, vol. 45, no. 5, pp. 479–489, 2010, doi: 10.1002/crat.201000097.

[48] R. Q. Chen, Q.-Q. Lu, Q.-D. Cheng, L.-B. Ao, C.-Y. Zhang, H. Hou, Y.-M. Liu, D.-W. Li and D.-C. Yin, “An ignored variable: Solution preparation temperature in protein crystallization,” *Sci. Rep.*, vol. 5, pp. 1–6, 2015, doi: 10.1038/srep07797.

[49] J. R. Luft, J. R. Wolfley, M. I. Said, R. M. Nagel, A. M. Lauricella, J. L. Smith, M. H. Thayer, C.K. Veatch, Snell, M. G. Malkowski, G. T DeTitta, “Efficient optimization of crystallization conditions by manipulation of drop volume ratio and temperature,” *Protein Sci.*, vol. 16, no. 4, pp. 715–722, Feb. 2007, doi: 10.1110/ps.062699707.

[50] J. P. Astier and S. Veesler, “Using temperature to crystallize proteins: A mini-review,” *Cryst. Growth Des.*, vol. 8, no. 12, pp. 4215–4219, 2008, doi: 10.1021/cg800665b.

[51] Y. Suzuki, E. Konda, H. Hondoh, and K. Tamura, “Effects of temperature, pressure, and pH

on the solubility of triclinic lysozyme crystals," in *Journal of Crystal Growth*, Mar. 2011, vol. 318, no. 1, pp. 1085–1088, doi: 10.1016/j.jcrysgro.2010.11.107.

[52] A. McPherson and B. Cudney, "Optimization of crystallization conditions for biological macromolecules," *Acta Crystallogr. Sect. F Structural Biol. Commun.*, vol. 70, pp. 1445–1467, Nov. 2014, doi: 10.1107/S2053230X14019670.

[53] N. Junius, E. Vahdatahar, E. Oksanen, J.-L. Ferrer, and M. Budayova-Spano, "Optimization of crystallization of biological macromolecules using dialysis combined with temperature control," *J. Appl. Crystallogr.*, vol. 53, no. 3, Jun. 2020, doi: 10.1107/s1600576720003209.

[54] S. Veesler, N. Fertè, M. S. Costes, M. Czjzek, and J. P. Astier, "Temperature and pH effect on the polymorphism of aprotinin (BPTI) in sodium bromide solutions," *Cryst. Growth Des.*, vol. 4, no. 6, pp. 1137–1141, 2004, doi: 10.1021/cg0498195.

[55] E. L. Forsythe, D. L. Maxwell, and M. Pusey, "Vapor diffusion, nucleation rates and the reservoir to crystallization volume ratio," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 58, no. 10 II, pp. 1601–1605, 2002, doi: 10.1107/S0907444902014208.

[56] I. R. Krauss, A. Merlino, A. Vergara, and F. Sica, "An overview of biological macromolecule crystallization," *International Journal of Molecular Sciences*, vol. 14, no. 6. pp. 11643–11691, 2013, doi: 10.3390/ijms140611643.

[57] J. D. Schmit and K. Dill, "Growth rates of protein crystals," *J. Am. Chem. Soc.*, vol. 134, no. 9, pp. 3934–3937, 2012, doi: 10.1021/ja207336r.

[58] M. M. Ries-Kautt and A. F. Ducruix, "Relative effectiveness of various ions on the solubility and crystal growth of lysozyme.," *J. Biol. Chem.*, vol. 264, no. 2, pp. 745–748, 1989.

[59] D. Friedmann, T. Messick, and R. Marmorstein, "Crystallization of macromolecules," *Curr. Protoc. Protein Sci.*, vol. 1, no. SUPPL.66, p. Unit17.4, Nov. 2011, doi: 10.1002/0471140864.ps1704s66.

[60] N. E. Chayen, "Comparative Studies of Protein Crystallization by Vapour-Diffusion and Microbatch Techniques," *Rev. Acta Cryst*, 1998.

[61] P. M. Martins, F. Rocha, and A. M. Damas, "Understanding water equilibration fundamentals as a step for rational protein crystallization," *PLoS One*, vol. 3, no. 4, Apr. 2008, doi: 10.1371/journal.pone.0001998.

[62] J. D. Ng, J. A. Gavira, and J. M. García-Ruíz, "Protein crystallization by capillary counterdiffusion for applied crystallographic structure determination," *J. Struct. Biol.*, vol. 142, no. 1, pp. 218–231, 2003, doi: 10.1016/S1047-8477(03)00052-2.

[63] N. Junius, E. Oksanen, M. Terrien, C. Berzin, J. L. Ferrer, and M. Budayova-Spano, "A crystallization apparatus for temperaturecontrolled flow-cell dialysis with real-time visualization," *J. Appl. Crystallogr.*, vol. 49, pp. 806–813, 2016, doi: 10.1107/S1600576716004635.

[64] A. E. Gutiérrez-Quezada, R. Arreguín-Espinosa, and A. Moreno, "Protein Crystal Growth

Methods," in *Springer Handbook of Crystal Growth*, Springer Berlin Heidelberg, 2010, pp. 1583–1605.

[65] N. E. Chayen and E. Saridakis, "Protein crystallization: From purified protein to diffraction-quality crystal," *Nat. Methods*, vol. 5, no. 2, pp. 147–153, Feb. 2008, doi: 10.1038/nmeth.f.203.

[66] L. R. Eriks, J. A. Mayor, and R. S. Kaplan, "A strategy for identification and quantification of detergents frequently used in the purification of membrane proteins," *Anal. Biochem.*, vol. 323, no. 2, pp. 234–241, Dec. 2003, doi: 10.1016/j.ab.2003.09.002.

[67] A. Stetsenko and A. Guskov, "An overview of the top ten detergents used for membrane protein crystallization," *Crystals*, vol. 7, no. 7, Jul. 2017, doi: 10.3390/cryst7070197.

[68] M. Le Maire, P. Champeil, and J. V. Møller, "Interaction of membrane proteins and lipids with solubilizing detergents," *Biochimica et Biophysica Acta - Biomembranes*, vol. 1508, no. 1–2. pp. 86–111, 2000, doi: 10.1016/S0304-4157(00)00010-1.

[69] R. M. Garavito and S. Ferguson-Miller, "Detergents as Tools in Membrane Biochemistry," *Journal of Biological Chemistry*, vol. 276, no. 35. pp. 32403–32406, Aug. 31, 2001, doi: 10.1074/jbc.R100031200.

[70] A. Anandan and A. Vrielink, "Detergents in membrane protein purification and crystallisation," in *Advances in Experimental Medicine and Biology*, vol. 922, 2016, pp. 13–28.

[71] D. Linke, "Detergents. An Overview," in *Methods in Enzymology*, 1st ed., vol. 463, no. C, Elsevier Inc., 2009, pp. 603–617.

[72] C. Hitscherich, V. Aseyev, J. Wiencek, and P. J. Loll, "Effects of PEG on detergent micelles: Implications for the crystallization of integral membrane proteins," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 57, no. 7, pp. 1020–1029, Jul. 2001, doi: 10.1107/S0907444901006242.

[73] T. Arnold and D. Linke, "Phase separation in the isolation and purification of membrane proteins," *Biotechniques*, vol. 43, no. 4, pp. 427–440, 2007, doi: 10.2144/000112566.

[74] G. G. Privé, "Detergents for the stabilization and crystallization of membrane proteins," *Methods*, vol. 41, no. 4, pp. 388–397, Apr. 2007, doi: 10.1016/j.ymeth.2007.01.007.

[75] S. Kim, P. A. Thiessen, E. E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B. A. Shoemaker, J. Wang, B. Yu, J. Zhang, and S. H. Bryant, "PubChem substance and compound databases," *Nucleic Acids Res.*, vol. 44, no. D1, pp. D1202–D1213, 2016, doi: 10.1093/nar/gkv951.

[76] J. F. van Dyck, A. Konijnenberg, and F. Sobott, "Native mass spectrometry for the characterization of structure and interactions of membrane proteins," in *Methods in Molecular Biology*, vol. 1635, Humana Press, New York, NY, 2017, pp. 205–232.

[77] T. W. Allen and F. Separovic, *Membrane protein structure and function*, vol. 1818, no. 2.

2012.

[78] P. Gourdon, J. L. Andersen, K. L. Hein, M. Bublitz, B. P. Pedersen, X-Y. Liu, L. Yatime, M. Nyblom, T. T. Nielsen, C. Olesen, J. V. Møller, P. Nissen, and J. P. Morth, "HiLiDe-systematic approach to membrane protein crystallization in lipid and detergent," *Cryst. Growth Des.*, vol. 11, no. 6, pp. 2098–2106, Jun. 2011, doi: 10.1021/cg101360d.

[79] A. Wlodawer, Z. Dauter, and M. Jaskolski, *Protein Crystallography Methods and Protocols*, 1st ed. Humana Press Inc., 2017.

[80] A. M. Seddon, P. Curnow, and P. J. Booth, "Membrane proteins, lipids and detergents: Not just a soap opera," *Biochimica et Biophysica Acta - Biomembranes*, vol. 1666, no. 1–2. Elsevier B.V., pp. 105–117, Nov. 03, 2004, doi: 10.1016/j.bbamem.2004.04.011.

[81] M. Caffrey and V. Cherezov, "Crystallizing membrane proteins using lipidic mesophases," *Nat. Protoc.*, vol. 4, no. 5, pp. 706–731, 2009, doi: 10.1038/nprot.2009.31.

[82] M. Caffrey, "A comprehensive review of the lipid cubic phase or in meso method for crystallizing membrane and soluble proteins and complexes," *Acta Crystallogr. Sect. FStructural Biol. Commun.*, vol. 71, pp. 3–18, Jan. 2015, doi: 10.1107/S2053230X14026843.

[83] X. D. Su, H. Zhang, T. C. Terwilliger, A. Liljas, J. Xiao, and Y. Dong, "Protein crystallography from the perspective of technology developments," *Crystallogr. Rev.*, vol. 21, no. 1–2, pp. 122–153, 2015, doi: 10.1080/0889311X.2014.973868.

[84] A. Wlodawer, W. Minor, Z. Dauter, and M. Jaskolski, "Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures," *FEBS J.*, vol. 275, no. 1, pp. 1–21, Jan. 2008, doi: 10.1111/j.1742-4658.2007.06178.x.

[85] D. Jain and V. Lamour, "Computational Tools in Protein Crystallography," in *Computational Biology, Methods in Molecular Biology*, no. 10, Springer Science Business Media, 2010, pp. 129-137.

[86] I. Usón and G. M. Sheldrick, "Advances in direct methods for protein crystallography," *Current Opinion in Structural Biology*, vol. 9, no. 5. Current Biology Ltd, pp. 643–648, Oct. 01, 1999, doi: 10.1016/S0959-440X(99)00020-2.

[87] gale rhodes, "Crystallography made crystal clear," doi: 10.1016/0307-4412(94)90199-6.

[88] D. A. Myles, "Neutron protein crystallography: current status and a brighter future," *Current Opinion in Structural Biology*, vol. 16, no. 5. Elsevier Current Trends, pp. 630–637, Oct. 01, 2006, doi: 10.1016/j.sbi.2006.08.010.

[89] E. H. Snell, M. J. Van Der Woerd, M. Damon, R. A. Judge, D. A. A. Myles, and F. Meilleur, "Optimizing crystal volume for neutron diffraction: D-xylose isomerase," *Eur. Biophys. J.*, vol. 35, no. 7, pp. 621–632, Sep. 2006, doi: 10.1007/s00249-006-0068-4.

[90] G. C. Schröder and F. Meilleur, "Neutron crystallography data collection and processing for modelling hydrogen atoms in protein structures," *J. Vis. Exp.*, vol. 2020, no. 166, pp. 1–38,

Dec. 2020, doi: 10.3791/61903.

[91] J. H. Lakey, “Neutrons for biologists: A beginner’s guide, or why you should consider using neutrons,” in *Journal of the Royal Society Interface*, 2009, vol. 6, no. SUPPL. 5, doi: 10.1098/rsif.2009.0156.focus.

[92] P. V Afonine, M. Mustyakimov, R. W. Grosse-Kunstleve, N. W. Moriarty, P. Langan, and P. D. Adams, “Joint X-ray and neutron refinement with phenix.refine,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 66, no. 11, pp. 1153–1163, 2010, doi: 10.1107/S0907444910026582.

[93] M. P. Blakeley, S. S. Hasnain, and S. V. Antonyuk, “Sub-atomic resolution X-ray crystallography and neutron crystallography: Promise, challenges and potential,” *IUCrJ*, vol. 2, no. 4. International Union of Crystallography, pp. 464–474, Jul. 01, 2015, doi: 10.1107/S2052252515011239.

[94] N. Niimura and R. Bau, “Neutron protein crystallography: Beyond the folding structure of biological macromolecules,” *Acta Crystallographica Section A: Foundations of Crystallography*, vol. 64, no. 1. International Union of Crystallography, pp. 12–22, Jan. 01, 2008, doi: 10.1107/S0108767307043498.

[95] M. P. Blakeley, S. C. M. Teixeira, I. Petit-Haertlein, I. Hazemann, A. Mitschler, M. Haertlein, E. Howard, and A. D. Podjarny, “Neutron macromolecular crystallography with LADI-III,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 66, no. 11, pp. 1198–1205, 2010, doi: 10.1107/S0907444910019797.

[96] S. C. M. Teixeira *et al.*, “New sources and instrumentation for neutrons in biology,” *Chem. Phys.*, vol. 345, no. 2–3, pp. 133–151, 2008, doi: 10.1016/j.chemphys.2008.02.030.

[97] M. P. Blakeley, P. Langan, N. Niimura, and A. Podjarny, “Neutron crystallography: opportunities, challenges, and limitations,” *Current Opinion in Structural Biology*, vol. 18, no. 5. pp. 593–600, Oct. 2008, doi: 10.1016/j.sbi.2008.06.009.

[98] E. Oksanen, M. P. Blakeley, F. Bonneté, M. T. Dauvergne, F. Dauvergne, and M. Budayova-Spano, “Large crystal growth by thermal control allows combined X-ray and neutron crystallographic studies to elucidate the protonation states in Aspergillus flavus urate oxidase,” in *Journal of the Royal Society Interface*, Oct. 2009, vol. 6, no. SUPPL. 5, doi: 10.1098/rsif.2009.0162.focus.

[99] R. Ashkar *et al.*, “Neutron scattering in the biological sciences: progress and prospects,” *Acta Crystallographica Section D: Structural Biology*, vol. 74, no. 12. Wiley-Blackwell, pp. 1129–1168, Dec. 01, 2018, doi: 10.1107/S2059798318017503.

[100] M. Gemmi, E. Mugnaioli, T. E. Gorelik, U. Kolb, L. Palatinus, P. Boullay, S. Hovmöller, and J. P. Abraham, “3D electron diffraction: The nanocrystallography revolution,” *ACS Cent. Sci.*, vol. 5, no. 8, pp. 1315–1329, 2019, doi: 10.1021/acscentsci.9b00394.

[101] T. Gruene *et al.*, “Rapid Structure Determination of Microcrystalline Molecular Compounds Using Electron Diffraction,” *Angew. Chemie - Int. Ed.*, vol. 57, no. 50, pp. 16313–16317,

2018, doi: 10.1002/anie.201811318.

[102] B. L. Nannenga, M. G. Iadanza, B. S. Vollmar, and T. Gonen, “Overview of Electron Crystallography of Membrane Proteins: Crystallization and Screening Strategies Using Negative Stain Electron Microscopy,” *Curr. Protoc. Protein Sci.*, no. SUPPL.72, 2013, doi: 10.1002/0471140864.ps1715s72.

[103] K. Yonekura, K. Kato, M. Ogasawara, M. Tomita, and C. Toyoshima, “Electron crystallography of ultrathin 3D protein crystals: Atomic model with charges,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 112, no. 11, pp. 3368–3373, 2015, doi: 10.1073/pnas.1500724112.

[104] E. V. Beale, D. G. Waterman, C. Hecksel, J. van Rooyen, J. B. Gilchrist, J. M. Parkhurst, F. de Haas, B. Buijsse, G. Evans, and P. Zhang, “A Workflow for Protein Structure Determination From Thin Crystal Lamella by Micro-Electron Diffraction,” *Front. Mol. Biosci.*, vol. 7, Aug. 2020, doi: 10.3389/fmolb.2020.00179.

[105] T. G. M. Jason de la Cruz, Johan Hattne, Dan Shi, Paul Seidler, Jose Rodriguez, Francis E. Reyes, Michael R. Sawaya, Duilio Cascio, Simon C. Weiss, Sun Kyung Kim, Cynthia S. Hinck, Andrew P. Hinck, Guillermo Calero, David Eisenberg, “Atomic resolution structures from fragmented protein crystals by the cryoEM method MicroED,” *Physiol. Behav.*, vol. 176, no. 12, pp. 139–148, 2017, doi: 10.1038/nmeth.4178.Atomic.

[106] B. L. Nannenga and T. Gonen, “Protein structure determination by MicroED,” *Current Opinion in Structural Biology*, vol. 27, no. 1. pp. 24–31, 2014, doi: 10.1016/j.sbi.2014.03.004.

[107] M. A. Marques, M. D. Purdy, and M. Yeager, “CryoEM maps are full of potential,” *Curr. Opin. Struct. Biol.*, vol. 58, no. August, pp. 214–223, 2019, doi: 10.1016/j.sbi.2019.04.006.

[108] P. F. Jose M. Martin-garcia, Chelsie E.Conrad, Jesse Coe, Shatabdi Roy-Chowdhury, “Review: Serial Femtosecond Crystallography: A Revolution in Structural Biology,” vol. 25, no. 3, pp. 289–313, 2016, doi: 10.1016/j.abb.2016.03.036.Review.

[109] D. Shi, B. L. Nannenga, M. J. De La Cruz, J. Liu, S. Sawtelle, G. Calero, F. E. Reyes, J. Hattne, and T. Gonen, “The collection of MicroED data for macromolecular crystallography,” *Nat. Protoc.*, vol. 11, no. 5, pp. 895–904, 2016, doi: 10.1038/nprot.2016.046.

[110] S. H. Lin and G. Guidotti, “Purification of Membrane Proteins,” in *Methods in Enzymology*, 1st ed., vol. 463, no. C, Elsevier Inc., 2009, pp. 619–629.

[111] M. S. Islam, A. Aryasomayajula, and P. R. Selvaganapathy, “A review on macroscale and microscale cell lysis methods,” *Micromachines*, vol. 8, no. 3, 2017, doi: 10.3390/mi8030083.

[112] J. Sambrook and D. W. Russell, “Molecular Cloning_ A Laboratory Manual.” Cold Spring Harbor Laboratory Press.

[113] A. Spriestersbach, J. Kubicek, F. Schäfer, H. Block, and B. Maertens, “Purification of His-Tagged Proteins,” in *Methods in Enzymology*, vol. 559, 2015, pp. 1–15.

[114] U. Laemmli, “Cleavage of structural proteins during the assembly of the head of bacterophage T4,” 1970.

[115] Sino Biological Inc., “SDS PAGE-Preparation,” 2013. [Online]. Available: http://www.assay-protocol.com/molecular-biology/electrophoresis/denaturing-page.

[116] T. Mahmood and P. C. Yang, “Western blot: Technique, theory, and trouble shooting,” *N. Am. J. Med. Sci.*, vol. 4, no. 9, pp. 429–434, 2012, doi: 10.4103/1947-2714.100998.

[117] K. M. Pos and K. Diederichs, “Purification, crystallization and preliminary diffraction studies of AcrB, an inner-membrane multi-drug efflux protein,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 58, no. 10 I, pp. 1865–1867, 2002, doi: 10.1107/S0907444902013963.

[118] M. Mills and S. M. Payne, “Identification of shuA, the gene encoding the heme receptor of Shigella dysenteriae, and analysis of invasion and intracellular multiplication of a shuA mutant,” *Infect. Immun.*, vol. 65, no. 12, pp. 5358–5363, 1997, doi: 10.1128/iai.65.12.5358-5363.1997.

[119] K. L. Longenecker, S. M. Garrard, P. J. Sheffield, and Z. S. Derewenda, “Protein crystallization by rational mutagenesis of surface residues: Lys to ala mutations promote crystallization of RhoGDI,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 57, no. 5, pp. 679–688, 2001, doi: 10.1107/S0907444901003122.

[120] S. C. Lee, T. J. Knowles, V. L. G. Postis, M. Jamshad, R. A. Parslow, Y-p. Lin, A. Goldman, P. Sridhar, M. Overduin, S. P. Muench and T. R. Dafforn, “A method for detergent-free isolation of membrane proteins in their local lipid environment,” *Nat. Protoc.*, vol. 11, no. 7, pp. 1149–1162, 2016, doi: 10.1038/nprot.2016.070.

[121] T. K. Ritchie, Y. V. Grinkova, T. H. Bayburt, I. G. Denisov, J. K. Zolnerciks, W. M. Atkins, and S. G. Sligar, “Reconstitution of Membrane Proteins in Phospholipid Bilayer Nanodiscs,” *Methods Enzymol.*, vol. 464, no. C, pp. 211–231, 2009, doi: 10.1016/S0076-6879(09)64011-8.

[122] J. Borch and T. Hamann, “The nanodisc: A novel tool for membrane protein studies,” *Biol. Chem.*, vol. 390, no. 8, pp. 805–814, 2009, doi: 10.1515/BC.2009.091.

[123] K. Brillet, A. Meksem, A. Thompson, and D. Cobessi, “Expression, purification, crystallization and preliminary X-ray diffraction analysis of the tonB-dependent haem outer membrane transporter shuA from Shigella dysenteriae,” *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, vol. 65, no. 4, pp. 402–405, 2009, doi: 10.1107/S1744309109008148.

[124] C. C. Su, M. Li, R. Gu, Y. Takatsuka, G. McDermott, H. Nikaido, and E. W. Yu., “Conformation of the AcrB multidrug efflux pump in mutants of the putative proton relay pathway,” *J. Bacteriol.*, vol. 188, no. 20, pp. 7290–7296, Oct. 2006, doi: 10.1128/JB.00684-06.

[125] T. Eicher, H-j. Cha, M. A. Seeger, L. Brandstätter, J. El-Delik, J. A. Bohnert, W. V. Kern, F. Verrey, M. G. Grütter, K. Diederichs, and K. M. Pos, “Transport of drugs by the multidrug transporter AcrB involves an access and a deep binding pocket that are separated by a

switch-loop," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 109, no. 15, pp. 5687–5692, Apr. 2012, doi: 10.1073/pnas.1114944109.

[126] A. Ababou and V. Koronakis, "Structures of gate loop variants of the AcrB drug efflux pump bound by erythromycin substrate," *PLoS One*, vol. 11, no. 7, pp. 10–14, 2016, doi: 10.1371/journal.pone.0159154.

[127] J. S. Kim, H. Jeong, S. Song, H-Y. Kim, K. Lee, J. Hyun, and N-C. Ha, "Structure of the tripartite multidrug efflux pump AcrAB-TolC suggests an alternative assembly mode," *Mol. Cells*, vol. 38, no. 2, pp. 180–186, Jan. 2015, doi: 10.14348/molcells.2015.2277.

[128] D. Du, Z. Wang, N. R. James, J. E. Voss, E. Klimont, T. Ohene-Agyei, H. Venter, W. Chiu, and B. F. Luisi., "Structure of the AcrAB-TolC multidrug efflux pump," *Nature*, vol. 509, no. 7501, pp. 512–515, 2014, doi: 10.1038/nature13205.

[129] E. W. Yu, J. R. Aires, G. McDermott, and H. Nikaido, "A periplasmic drug-binding site of the AcrB multidrug efflux pump: A crystallographic and site-directed mutagenesis study," *J. Bacteriol.*, vol. 187, no. 19, pp. 6804–6815, Oct. 2005, doi: 10.1128/JB.187.19.6804-6815.2005.

[130] W. Qiu, Z. Fu, G. G. Xu, R. A. Grassucci, Y. Zhang, J. Frank, W. A. Hendrickson, and Y. Guo, "Structure and activity of lipid bilayer within a membrane-protein transporter," *PNAS,* vol. 115, no. 51, pp. 12985-12990, 2018, doi: 10.1073/pnas.1812526115.

[131] M. Parmar, S. Rawson, C. A. Scarff, A. Goldman, T. R. Dafforn, S. P. Muench, and V. LG Postis, "Using a SMALP platform to determine a sub-nm single particle cryo-EM membrane protein structure," *Biochim. Biophys. Acta - Biomembr*, vol. 1860, no. 2, pp. 378–383, 2018, doi: 10.1016/j.bbamem.2017.10.005.

[132] S. Murakami, R. Nakashima, E. Yamashita, T. Matsumoto, and A. Yamaguchi, "Crystal structures of a multidrug transporter reveal a functionally rotating mechanism," *Nature*, vol. 443, no. 7108, pp. 173–179, 2006, doi: 10.1038/nature05076.

[133] E. Vahdatahar, N. Junius, and M. Budayova - Spano, "Optimization of Crystal Growth for Neutron Macromolecular Crystallography," *JoVE*, no. 169, p. e61685, 2021, doi: doi:10.3791/61685.

[134] I. Moraes, G. Evans, J. Sanchez-Weatherby, S. Newstead, and P. D. S. Stewart, "Membrane protein structure determination - The next generation," *Biochim. Biophys. Acta - Biomembr.*, vol. 1838, no. 1 PARTA, pp. 78–87, 2014, doi: 10.1016/j.bbamem.2013.07.010.

[135] J. D. Ng, B. Lorber, R. Giegé, S. Koszelak, J. Day, A. Greenwood, A. McPherson, "Comparative analysis of thaumatin crystals grown on earth and in microgravity," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 53, no. 6, pp. 724–733, 1997, doi: 10.1107/S090744499700694X.

[136] B. Rupp, "Origin and use of crystallization phase diagrams," *Acta Crystallogr. Sect. FStructural Biol. Commun.*, vol. 71, no. 13, pp. 247–260, Mar. 2015, doi: 10.1107/S2053230X1500374X.

[137] C. L. Barnes, E. H. Snell, and C. E. Kundrot, “Thaumatin crystallization aboard the International Space Station using liquid-liquid diffusion in the Enhanced Gaseous Nitrogen Dewar (EGN),” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 58, no. 5, pp. 751–760, 2002, doi: 10.1107/S0907444902002767.

[138] V. Apostolopoulou, N. Junius, R. P. Sear, and M. Budayova-Spano, “Mixing Salts and Poly(ethylene glycol) into Protein Solutions: The Effects of Diffusion across Semipermeable Membranes and of Convection,” *Cryst. Growth Des.*, vol. 20, no. 6, pp. 3927–3936, 2020, doi: 10.1021/acs.cgd.0c00246.

[139] N. Junius, “Développements instrumentaux pour le contrôle de la cristallisation par la dialyse : approche microfluidique et analyse aux rayons X,” Université Grenoble Alpes, 2016.

[140] R. Nakashima, K. Sakurai, S. Yamasaki, K. Nishino, and A. Yamaguchi, “Structures of the multidrug exporter AcrB reveal a proximal multisite drug-binding pocket,” *Nature*, vol. 480, no. 7378, pp. 565–569, 2011, doi: 10.1038/nature10641.

[141] I. Müller, “Guidelines for the successful generation of protein-ligand complex crystals,” *Acta Crystallographica Section D: Structural Biology*, vol. 73, no. 2. pp. 79–92, 2017, doi: 10.1107/S2059798316020271.

[142] K. C. Jindal, R. S. Chaudhary, A. K. Singla, S. S. Gangwal, and S. Khanna, “Effects of buffers and pH on rifampicin stability,” *Pharm. Ind.*, vol. 57, no. 5, pp. 420–422, 1995.

[143] E. A. Stura and I. A. Wilson, “Analytical and production seeding techniques,” *Methods*, vol. 1, no. 1, pp. 38–49, 1990, doi: 10.1016/S1046-2023(05)80145-8.

[144] B. L. Nannenga and T. Gonen, “The cryo-EM method microcrystal electron diffraction (MicroED),” *Nature Methods*, vol. 16, no. 5. pp. 369–379, 2019, doi: 10.1038/s41592-019-0395-x.

[145] R. K. Cheng, “crystals Towards an Optimal Sample Delivery Method for Serial Crystallography at XFEL,” 2020, doi: 10.3390/cryst10030215.

[146] J. H. Beale, R. Bolton, S. A. Marshall, E. V. Beale, S. B. Carr, A. Ebrahim, T. Moreno-Chicano, M. A. Hough, J. A. R. Worrall, I. Tews and R. L. Owen, “Successful sample preparation for serial crystallography experiments,” *J. Appl. Crystallogr.*, vol. 52, pp. 1385–1396, 2019, doi: 10.1107/S1600576719013517.

[147] V. Cherezov, A. Peddi, L. Muthusubramaniam, Y. F. Zheng, and M. Caffrey, “A robotic system for crystallizing membrane and soluble proteins in lipidic mesophases,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 60, no. 10, pp. 1795–1807, 2004, doi: 10.1107/S0907444904019109.

[148] A. D’Arcy, F. Villard, and M. Marsh, “An automated microseed matrix-screening method for protein crystallization,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 63, no. 4, pp. 550–554, 2007, doi: 10.1107/S0907444907007652.

[149] C. J. Gerdts, M. Elliott, S. Lovell, M. B. Mixon, A. J. Napuli, B. L. Staker, P. Nollert, and L. Stewart, “The plug-based nanovolume Microcapillary Protein Crystallization System (MPCS),” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 64, no. 11, pp. 1116–1122, 2008, doi: 10.1107/S0907444908028060.

[150] V. Stojanoff, J. Jakoncic, D. A. Oren, V. Nagarajan, J-C. Navarro Poulsen, M. A. Adams-Cioaba, T. Bergfors, and M. OA Sommer, “From screen to structure with a harvestable microfluidic device,” *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, vol. 67, no. 8, pp. 971–975, 2011, doi: 10.1107/S1744309111024456.

[151] N. Junius, S. Jaho, Y. Sallaz-Damaz, F. Borel, J. B. Salmon, and M. Budayova-Spano, “A microfluidic device for both on-chip dialysis protein crystallization and: In situ X-ray diffraction,” *Lab Chip*, vol. 20, no. 2, pp. 296–310, 2020, doi: 10.1039/c9lc00651f.

[152] N. Niimura, S. Arai, K. Kurihara, T. Chatake, I. Tanaka, and R. Bau, “Recent results on hydrogen and hydration in biology studied by neutron macromolecular crystallography,” *Cellular and Molecular Life Sciences*, vol. 63, no. 3. pp. 285–300, 2006, doi: 10.1007/s00018-005-5418-3.

[153] A. Urbani and T. Warne, “A colorimetric determination for glycosidic and bile salt-based detergents: applications in membrane protein research,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 7, no. June, pp. 1–8, Jan. 2019, doi: 10.1166/jnn.2009.1577 ER.

[154] S. S. Nielsen, “Phenol-Sulfuric Acid Method for Total Carbohydrates,” pp. 47–53, 2010, doi: 10.1007/978-1-4419-1463-7_6.

[155] A. S. Jönsson and B. Jönsson, “The influence of nonionic and ionic surfactants on hydrophobic and hydrophilic ultrafiltration membranes,” *J. Memb. Sci.*, vol. 56, no. 1, pp. 49–76, 1991, doi: 10.1016/0376-7388(91)85015-W.

[156] J. A. Rodriguez, D. S. Eisenberg, and T. Gonen, “Taking the measure of MicroED,” *Current Opinion in Structural Biology*, vol. 46. Elsevier Ltd, pp. 79–86, 2017, doi: 10.1016/j.sbi.2017.06.004.

[157] D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L.-W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams and P. D. Adams, “Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix,” *Acta Crystallogr. Sect. D Struct. Biol.*, vol. 75, pp. 861–877, 2019, doi: 10.1107/S2059798319011471.

[158] L. Maveyraud and L. Mourey, “Protein X-ray Crystallography and Drug Discovery,” *Molecules*, vol. 25, no. 5, Feb. 2020, doi: 10.3390/molecules25051030.

[159] W. Kabsch, “XDS,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 66, no. 2, pp. 125–132, 2010, doi: 10.1107/S0907444909047337.

[160] P. D. Adams, R. W. Grosse-Kunstleve, L. W. Hung, T. R. Ioerger, A. J. McCoy, N. W. Moriarty,

R. J. Read, J. C. Sacchettini, N. K Sauter, T. C. Terwilliger, Thomas C., “PHENIX: Building new software for automated crystallographic structure determination,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 58, no. 11, pp. 1948–1954, 2002, doi: 10.1107/S0907444902016657.

[161] P. Emsley and K. Cowtan, “Coot: Model-building tools for molecular graphics,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 60, no. 12 I, pp. 2126–2132, Dec. 2004, doi: 10.1107/S0907444904019158.

[162] P. A. Karplus and K. Diederichs, “Linking crystallographic model and data quality,” *Science (80).*, vol. 336, no. 6084, pp. 1030–1033, 2012, doi: 10.1126/science.1218231.

[163] M. A. Seeger, A. Schiefner, T. Eicher, F. Verrey, K. Diederichs, and K. M. Pos, “Structural asymmetry of AcrB trimer suggests a peristaltic pump mechanism,” *Science (80-.).*, vol. 313, no. 5791, pp. 1295–1298, 2006, doi: 10.1126/science.1131542.

[164] A. C. Dumetz, A. M. Chockla, E. W. Kaler, and A. M. Lenhoff, “Effects of pH on protein-protein interactions and implications for protein phase behavior,” *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1784, no. 4, pp. 600–610, 2008, doi: 10.1016/j.bbapap.2007.12.016.

[165] C. X. Weichenberger, P. V Afonine, K. Kantardjieff, and B. Rupp, “The solvent component of macromolecular crystals,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 71, pp. 1023–1038, 2015, doi: 10.1107/S1399004715006045.

[166] D. Du, A. Neuberger, M. W. Orr, C. E. Newman, P-C. Hsu, F. Samsudin, A. Szewczak-Harris, L. M. Ramos, M. Debela, S. Khalid, G. Storz, and B. F. Luisi, “Interactions of a Bacterial RND Transporter with a Transmembrane Small Protein in a Lipid Environment,” *Struct. Des.*, vol. 28, no. 6, pp. 625-634.e6, 2020, doi: 10.1016/j.str.2020.03.013.

[167] O. C. Gagné, “Bond-length distributions for ions bonded to oxygen: Results for the lanthanides and actinides and discussion of the f-block contraction: Results,” *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.*, vol. 74, no. 1, pp. 49–62, 2018, doi: 10.1107/S2052520617017425.

[168] H. P. Stevenson, A. M. Makhov, M. Calero, A. L. Edwards, O. B. Zeldin, I. I. Mathews, G. Lin, C. O. Barnes, H. Santamaria, T. M. Ross, S. Michael Soltis, C. Khosla, V. Nagarajan, J. F. Conway, A. E. Cohen, and G. Calero, “Use of transmission electron microscopy to identify nanocrystals of challenging protein targets,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 23, pp. 8470–8475, 2014, doi: 10.1073/pnas.1400240111.

[169] I. Tanaka, “Necessity and Effectiveness of a Biological Diffractometer for Installation in Second Target Station as a New Neutron Pulsed Source in Japan,” vol. 011015, pp. 1–5, 2019, doi: 10.7566/jpscp.25.011015.

[170] K. Koruza, B. Lafumat, M. Nyblom, W. Knecht, and Z. Fisher, “From initial hit to crystal optimization with microseeding of human carbonic anhydrase IX—A case study for neutron protein crystallography,” *Crystals*, vol. 8, no. 11, p. 434, Nov. 2018, doi:

10.3390/cryst8110434.

[171] L. Coates and B. Sullivan, *The macromolecular neutron diffractometer at the spallation neutron source*, 1st ed., vol. 634. Elsevier Inc., 2020.

[172] M. Budayova-Spano, K. Koruza, and Z. Fisher, “Large crystal growth for neutron protein crystallography,” in *Methods in Enzymology*, 1st ed., vol. 634, Academic Press Inc., 2020, pp. 21–46.

[173] R. M. F. Leal, S. Callow, P. Callow, M. P. Blakeley, C. J. Cardin, W. A. Denny, S. CM Teixeira, E. P. Mitchell, and V. Trevor Forsyth, “Combined neutron and X-ray diffraction studies of DNA in crystals and solutions,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 66, no. 11, pp. 1244–1248, 2010, doi: 10.1107/S0907444910017713.

[174] E. Garman, “Cool data: quantity AND quality,” *Acta Cryst*, 1999.

[175] K. M. Pos, A. Schiefner, M. A. Seeger, and K. Diederichs, “Crystallographic analysis of AcrB,” in *FEBS Letters*, Apr. 2004, vol. 564, no. 3, pp. 333–339, doi: 10.1016/S0014-5793(04)00272-8.

[176] M. Kurz, B. Blattmann, A. Kaech, C. Briand, P. Reardon, U. Ziegler, and M. G. Gruetter, “High-throughput counter-diffusion capillary crystallization and in situ diffraction using high-pressure freezing in protein crystallography,” *J. Appl. Crystallogr.*, vol. 45, no. 5, pp. 999–1008, 2012, doi: 10.1107/S0021889812034061.

[177] Z. Luo, K. Rajashankar, and Z. Dauter, “Weak data do not make a free lunch, only a cheap meal,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 70, no. 2, pp. 253–260, 2014, doi: 10.1107/S1399004713026680.

[178] A. McPherson, C. Nguyen, R. Cudney, and S. B. Larson, “The role of small molecule additives and chemical modification in protein crystallization,” *Cryst. Growth Des.*, vol. 11, no. 5, pp. 1469–1474, 2011, doi: 10.1021/cg101308r.

[179] A. L. Lamb, T. J. Kappock, and N. R. Silvaggi, “You are lost without a map: Navigating the sea of protein structures,” *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1854, no. 4, pp. 258–268, 2015, doi: 10.1016/j.bbapap.2014.12.021.

[180] C. Nguyen and T. Gonen, “Beyond protein structure determination with MicroED,” *Current Opinion in Structural Biology*, vol. 64, no. January. pp. 51–58, 2020, doi: 10.1016/j.sbi.2020.05.018.

[181] E. Saridakis and N. E. Chayen, “Systematic improvement of protein crystals by determining the supersolubility curves of phase diagrams,” *Biophys. J.*, vol. 84, no. 2 I, pp. 1218–1222, Feb. 2003, doi: 10.1016/S0006-3495(03)74936-4.

[182] N. E. Chayen, “Methods for separating nucleation and growth in protein crystallisation,” *Prog. Biophys. Mol. Biol.*, vol. 88, no. 3, pp. 329–337, Jul. 2005, doi: 10.1016/j.pbiomolbio.2004.07.007.

[183] H. Xu, H. Lebrette, T. Yang, V. Srinivas, S. Hovmöller, M. Högbom, and X. Zou, “A Rare

Lysozyme Crystal Form Solved Using Highly Redundant Multiple Electron Diffraction Datasets from Micron-Sized Crystals," *Structure*, vol. 26, no. 4, pp. 667-675.e3, 2018, doi: 10.1016/j.str.2018.02.015.

[184] A. Nakamura, T. Ishida, S. Fushinobu, K. Kusaka, I. Tanaka, K. Inaka, Y. Higuchi, M. Masaki, K. Ohta, S. Kaneko, N. Niimura, K. Igarashi and M. Samajima, "Phase-diagram-guided method for growth of a large crystal of glycoside hydrolase family 45 inverting cellulase suitable for neutron structural analysis," *J. Synchrotron Radiat.*, vol. 20, no. 6, pp. 859–863, 2013, doi: 10.1107/S0909049513020943.

[185] M. P. Blakeley, M. Cianci, J. R. Helliwell, and P. J. Rizkallah, "Synchrotron and neutron techniques in biological crystallography," *Chem. Soc. Rev.*, vol. 33, no. 8, pp. 548–557, 2004, doi: 10.1039/b312779f.

[186] F. G. Müller and C. R. D. Lancaster, "Crystallization of membrane proteins," *Methods Mol. Biol.*, vol. 1033, pp. 67–83, 2013, doi: 10.1007/978-1-62703-487-6_5.

[187] L. Zhu, G. Bu, L. Jing, D. Shi, M-Y. Lee, T. Gonen, W. Liu, and B. L. Nannenga, "Structure Determination from Lipidic Cubic Phase Embedded Microcrystals by MicroED," *Structure*, vol. 28, no. 10, pp. 1149-1159, 2020, doi: 10.1016/j.str.2020.07.006.

[188] A. V. Bharda and H. S. Jung, "Liquid electron microscopy: then, now and future," *Appl. Microsc.*, vol. 49, no. 1, pp. 4–9, 2019, doi: 10.1186/s42649-019-0011-7.

[189] U. M. Mirsaidov, H. Zheng, Y. Casana, and P. Matsudaira, "Imaging protein structure in water at 2.7 nm resolution by transmission electron microscopy," *Biophys. J.*, vol. 102, no. 4, pp. L15–L17, 2012, doi: 10.1016/j.bpj.2012.01.009.

[190] S. Pu, C. Gong, and A. W. Robertson, "Liquid cell transmission electron microscopy and its applications," *R. Soc. Open Sci.*, vol. 7, no. 1, 2020, doi: 10.1098/rsos.191204.

[191] M. Caffrey and C. Porter, "Crystallizing membrane proteins for structure determination using lipidic mesophases," *J. Vis. Exp.*, no. 45, pp. 1–6, 2010, doi: 10.3791/1712.

Appendix

Appendix I:

Optimization of Crystal Growth for Neutron Macromolecular Crystallography

Elham Vahdatahar[1], **Niels Junius**[1,2], **Monika Budayova - Spano**[1]

[1] CEA, CNRS, IBS, Université Grenoble Alpes [2] ELVESYS SAS

Corresponding Author

Monika Budayova - Spano
monika.spano@ibs.fr

Citation

Vahdatahar, E., Junius, N., Budayova - Spano, M. Optimization of Crystal Growth for Neutron Macromolecular Crystallography. *J. Vis. Exp.* (169), e61685, doi:10.3791/61685 (2021).

Date Published

March 13, 2021

DOI

10.3791/61685

URL

jove.com/video/61685

Abstract

The use of neutron macromolecular crystallography (NMX) is expanding rapidly with most structures determined in the last decade thanks to new NMX beamlines having been built and increased availability of structure refinement software. However, the neutron sources currently available for NMX are significantly weaker than equivalent sources for X-ray crystallography. Despite advances in this field, significantly larger crystals will always be required for neutron diffraction studies, particularly with the tendency to study ever-larger macromolecules and complexes. Further improvements in methods and instrumentation suited to growing larger crystals are therefore necessary for the use of NMX to expand.

In this work, we introduce rational strategies and a crystal growth bench (OptiCrys) developed in our laboratory that combines real-time observation through a microscope-mounted video camera with precise automated control of crystallization solutions (e.g., precipitant concentration, pH, additive, temperature). We then demonstrate how this control of temperature and chemical composition facilitates the search for optimal crystallization conditions using model soluble proteins. Thorough knowledge of the crystallization phase diagram is crucial for selecting the starting position and the kinetic path for any crystallization experiment. We show how a rational approach can control the size and number of crystals generated based on knowledge of multidimensional phase diagrams.

Introduction

Understanding the structure-function relationship of proteins and the mechanism of physiological pathways often relies on knowing the positions of hydrogen atoms (H) and how charge is transferred within a protein[1,2]. Since hydrogen atoms scatter X-rays weakly, their positions can only be determined with very high-resolution X-ray diffraction data (>1 Å)[3,4]. Conversely, neutron crystallography can be used to obtain an accurate position of hydrogen atoms in biological macromolecules as hydrogen and deuterium (H^2, isotope of hydrogen) atoms have scattering lengths of roughly equal

magnitude as oxygen, nitrogen and carbon[5]. However, neutron flux from available neutron sources is weaker than that of X-ray beams, so this must often be compensated for[2,3]. This can be achieved by exchanging H with H^2 and/or increasing the volume of crystals to reduce the incoherent scattering of hydrogens and increase the signal-to-noise ratio of diffraction images.

There are various crystallization approaches (the corresponding schematic phase diagram is shown in **Figure 1**) for obtaining large and high-quality crystals for both X-ray and neutron bio-macromolecular crystallography[6]. In vapor diffusion, a droplet prepared from a mixture of a protein and a crystallization solution is equilibrated over time, through evaporation of water or other volatile species, against a reservoir containing a higher concentration of precipitant of the same crystallization solution. The increase in concentration of protein and precipitant in the droplet leads to the supersaturation required for spontaneous nucleation followed by crystal growth at these nuclei[6,7]. Although vapor diffusion is the most frequently used technique for growing crystals[4], the crystallization process cannot be precisely controlled[8]. In the free interface diffusion method, crystallization solution diffuses into a concentrated protein solution, very slowly directing the system towards supersaturation. This method can be considered as a batch method with a slow mixing rate[6,9,10,11,12]. In the batch method, the protein is rapidly mixed with a crystallization solution leading to rapid supersaturation and in turn uniform nucleation with many crystals[3,7]. This method accounts for approximately one-third of all structures currently deposited in the Protein Data Bank. The dialysis method is also used for growing high-quality and well-diffracting protein crystals. In the dialysis method, molecules of precipitant diffuse from a reservoir through a semi-permeable membrane into a separate chamber with the protein solution. The kinetics of equilibration is dependent on various factors, such as temperature, membrane pore size, and the volume and concentration of protein samples and crystallization agents[6].

Crystallization phase diagrams can be used to describe different states of a protein as a function of different physical or chemical variable[3]. As illustrated in **Figure 1**, each crystallization technique can be visualized as using a different kinetic trajectory to reach the nucleation and metastable zones of such a diagram[6,10,13]. This provides information about protein solubility and the protein concentration at which a thermodynamic equilibrium between crystal and solution is observed, thereby finding the optimal conditions for nucleation and growth[3,14]. In a two-dimensional phase diagram, the protein concentration is plotted as a function of one variable and the other variables are kept constant[15]. In such a phase diagram, when the protein concentration is below the solubility curve, the solution is in the undersaturated region and no nucleation or crystal growth occurs. Above this curve is the supersaturation zone where the protein concentration is higher than the solubility limit[3,14]. This is further divided into three regions: the metastable zone, the spontaneous nucleation zone, and the precipitation zone. In the metastable zone, supersaturation is not sufficient for nucleation to occur within a reasonable time but growth of seeded crystals can take place. Aggregation and precipitation are favored in the precipitation zone, where supersaturation is too high[14,15].

When sufficient supersaturation for spontaneous nucleation is achieved, the first nuclei will appear[10]. The growth of crystals leads to a reduction in the protein concentration until the limit of solubility is reached. As long as supersaturation stays in the vicinity of the solubility curve, there will be no significant

change in the size of crystals. However, it has been shown that variations in the temperature and chemical composition of crystallization solution (for example, the concentration of precipitant) will affect protein solubility and may lead to the initiation of further crystal growth[8,13,16].

As dialysis is advantageous for good quality crystal growth, the OptiCrys crystallization bench illustrated in **Figure 2**, was designed and developed in our laboratory to control crystallization in a fully automated manner[8]. For this purpose, software was written with LabVIEW that allows the control and monitoring of the temperature of a flowing reservoir dialysis setup in contact with Peltier elements, via an electronic controller and a chiller. The same software also automatically regulates the chemical composition of the crystallization solution (for example the exchange of crystallization agents) using a multichannel fluidic system. Additionally, a digital camera and an inverted microscope are used to visualize and record the crystallization process. Two crystallization chambers with 15 µL and 250 µL volumes are available for growing crystals for different purposes. As the crystallization process is reversible, screening for different conditions is possible with just a few microliters of the protein solution as long as the sample is not damaged[8]. As a result, using this method minimizes the amount of protein material used.

From previous work[8], it is apparent that during the crystal growth process, *in situ* observations need to be carried out at regular time intervals. These can range from a few seconds to several days, depending on the event under observation (precipitation, nucleation, or crystal growth).

The optimization of crystal growth with OptiCrys is based on temperature-precipitant concentration phase diagrams. In the case of proteins with solubility as a direct function of temperature, it is possible to make use of the salting-out regime[18]. This is where increasing the ionic strength of the solution, which can be visualized using protein-precipitant phase diagrams, decreases the solubility of the protein. Likewise, proteins with inverse solubility can make use of the salting-in regime[18]. Nucleation occurs in the nucleation zone, in the vicinity of the metastable zone, and crystal growth then takes place in the metastable zone of the phase diagram until the protein concentration reaches the solubility limit. As shown in **Figure 3A**, with constant chemical composition temperature can be decreased to keep the crystallization solution in the metastable zone to prevent new nucleation. Crystals grow until the second crystal/solution equilibrium is achieved and after that, no further increase in the size of crystals is observed. The temperature is decreased several times until the crystals reach the desired size. In **Figure 3B**, at constant temperature, increasing the precipitant concentration keeps the solution in the metastable zone. This process can then be repeated several times to obtain large crystals. Changing the temperature and manipulating the crystallization solution conditions, by controlling the supersaturation levels, are two powerful tools for separating nucleation and growth of crystals that are controlled precisely and automatically by OptiCrys[5,8,14].

Examples of protein crystals grown by temperature-controlled, or temperature- and precipitant concentration-controlled crystallization, as well as relative diffraction data obtained are available in the literature and PDB. Among them are human γ-crystallin E, PA-IIL lectin, yeast inorganic pyrophosphatase, urate oxidase, human carbonic anhydrase II, YchB kinase, and lactate dehydrogenase[5,14,17,18].

Although OptiCrys was commercialized by NatX-ray, there are many laboratories that do not have access to this instrument or to the serial approach it offers. The alternative

to this technique is to use commercially available plastic microdialysis buttons with various volumes. Using these, temperature and chemical composition can be adjusted and varied manually. Inspection of microdialysis buttons cannot be done *in situ* and must instead be done manually with an optical microscope. Temperature control can be achieved by keeping the sample in a vibration-free temperature-controlled incubator. It is essential to keep the temperature constant to ensure that crystallization experiments are reproducible. Significant variation in temperature may also lead to damage or destruction of crystals[5].

Here we provide a detailed protocol describing sample preparation and the use of control software for the growth of large, high-quality crystals suitable for neutron protein crystallography. This step-by-step procedure was designed to take advantage of the crystallization phase diagram in order to select a starting position and kinetic path to control the size and the quality of the crystals generated. Additionally, a detailed protocol for growing crystals with microdialysis buttons is presented which uses the same rationale to obtain large, high-quality crystals.

Protocol

1. Dialysis method with microdialysis buttons

1. **Sample preparation**

    1. Prepare protein solution by dissolving 30 mg of chicken egg-white lysozyme as a lyophilized powder in 1 mL of CH_3COONa buffer (100 mM sodium acetate, pH 4) in order to obtain a solution with a final concentration of 30 $mg \cdot mL^{-1}$.

    2. Centrifuge the sample at the 13,000 × *g* for 10 min at 277 K. This process helps to remove any aggregates before starting the crystallization process.

    3. Check the absorbance of the sample at 280 nm and calculate the protein concentration by using the Beer-Lambert equation (A= εcl).

        NOTE: According to the Beer-Lambert equation, electronic absorbance (A) is directly proportional to the concentration (c, $mg \cdot mL^{-1}$) of an absorbing species given a constant optical pathlength (l, cm). The gradient of this linear relationship is the molar extinction coefficient (ε, for lysozyme at 280 nm is 2.64 $mL \cdot mg^{-1} \cdot cm^{-1}$)[19]. The sidechains of aromatic amino acids (tyrosine, tryptophan, and phenylalanine) and disulfide bonds between cysteine residues have strong absorbance at ~280 nm arising from spectroscopically-allowed π – π* transitions. As the majority of proteins contain these residues, protein concentration can typically be calculated easily by measuring the absorbance at 280 nm, given knowledge of the extinction coefficient.

    4. Prepare crystallization solutions as shown in **Figure 4**. Filter all the stock solutions with 0.22 µm Millipore filters before preparing the crystallization solution.

2. **Crystal growth**

    1. Cut a cellulose dialysis membrane with appropriate molecular weight cutoff (6-8 kDa) and soak it in distilled water.

        NOTE: Dialysis membrane discs are commercially available for microdialysis buttons, but if the dialysis tubing is used, don't forget to cut the edges in order

to separate the two layers of membrane from the tube to have only single layer membranes.

2. Fill wells of a 24-well tray with 2 mL of crystallization solution in the same order as shown in **Figure 4**.
   NOTE: If buttons with larger volumes are used (e.g., 200 µL), fill 50 mL tubes with a minimum of 5 mL crystallization solution in order to ensure efficient exchange.

3. Add/Pipette 35 µL of lysozyme solution to the chamber of the microdialysis button as illustrated in **Figure 5A**.
   NOTE: To avoid the formation of air bubbles in a 30 µL dialysis button when it is closed, an extra volume (dead volume) of 5 µL of additional protein must be added, which means a total of 35 µL of protein sample. This additional protein sample creates a slightly domed shape on top of the chamber, as shown in **Figure 5B**, which prevents the formation of air bubbles.

4. Take an applicator of the appropriate size and place the elastic O-ring at its extremity (**Figure 5C**). Then place the membrane, previously lightly wiped/drained using a piece of fibre-free paper, on top of the chamber of the dialysis button. Be careful not to put dust in the chamber when applying the paper. Set the dialysis membrane in place by transferring the elastic O-ring from the applicator to the groove of the dialysis button (**Figure 5C**).
   NOTE: The critical moment of handling is fixing the dialysis membrane on top of the chamber by transferring the elastic O-ring from the applicator into the groove of the dialysis button. All movements must be perfectly synchronized to avoid enclosing air bubbles with the sample in the protein chamber. It is useful to practice stretching the elastic O-ring before its application to know its rigidity, use tweezers to hold part of the membrane during its application and carry out the first test experiments with a model protein.

5. Transfer the button to the well or to the 50 mL tube using tweezers (**Figure 5D**).

6. Cover the well with a coverslip, pressing it gently on the grease to seal the well (**Figure 5E**).
   NOTE: If there is no grease on top of the wells, be sure to add this before starting the experiment or use a piece of tape instead of glass coverslip. The principle of protein crystallization using microdialysis buttons is illustrated in **Figure 5**.

7. Keep the sample at 293 K in a thermoregulated incubator. Different temperatures may be necessary according to the protein and crystallization conditions used.
   NOTE: The same grid of crystallization conditions shown in **Figure 4** (allowing the concentration of precipitant to be varied) can be screened as a function of temperature. In such a case, the same crystallization plate must be reproduced and each copy must be placed in an incubator regulated at a different temperature. This requires having several vibration-free thermoregulated incubators available.

8. Check the tray or tubes for crystals (**Figure 4**) and take notes on a regular basis, typically daily, to distinguish what is in each tray. Good notes are essential to avoid false positive results and to discriminate dust from crystals.

2. Crystal growth process using OptiCrys

1. **Sample preparation**

    1. Prepare a protein solution and a dialysis membrane as described in section 1.1.

    2. Prepare stock solutions of NaCl (4 M) and of CH_3COONa pH 4 (1 M) and filter them into 50 mL tubes.

    3. Add the protein solution (15 µL of lysozyme with 30 $mg \cdot mL^{-1}$) to the dialysis chamber of the temperature-controlled flowing reservoir dialysis setup. Refer to **Figure 6** for details of the temperature-controlled flowing reservoir dialysis setup. OptiCrys has two dialysis chambers, the minimum volume is 15 µL and the maximum volume is 250 µL.

    4. Cover the overchamber with a dialysis membrane and fix the membrane with the elastic O-ring (**Figure 6B**).

        NOTE: This setup is different from microdialysis buttons where each chamber is sealed directly by a dialysis membrane. In the flow cell, the dialysis membrane is instead fixed to the overchamber allowing the mounting of crystals without its removal. For this purpose, the overchamber can simply be unscrewed from the reservoir.

    5. Flip the overchamber and place it on top of the dialysis chamber. Slowly and gently press it to remove all the air trapped between the two pieces and avoid bubbles in the chamber (**Figure 6C**). Practice with a model protein can be advantageous when training to avoid air bubbles and sample loss.

    6. Fix the reservoir in its position by gently screwing it on top of the overchamber, again being mindful to avoid trapping air bubbles. Over-tightening of the reservoir can also form bubbles in the crystallization chamber (**Figure 6D**).

    7. Add the crystallization solution and cover the reservoir chamber with the airtight cap. (**Figure 6G**). Maximum volume of the reservoir is 1 mL.

    8. Transfer this assembly and insert it into the brass support. This support is in contact with Peltier elements that are used to control the temperature.

    9. As illustrated in **Figure 6**, the airtight cap is equipped with an optical window to allow top illumination of the dialysis chamber. Put the light source (**Figure 2**) on the window allowing light to pass through the chamber.

    10. The reservoir chamber can be connected to a pump to allow it to function as a continuous flow cell. Connect the tubing on top of the 50 mL tubes that contain the stock solutions and distilled water to the rotary valve as illustrated in **Figure 7**.

        NOTE: If automatic preparation and changing of crystallization solutions are not desired during the experiment, omit step 2.1.10. In such a case, the crystallization solution must be prepared and the reservoir pre-filled manually.

2. **Software**

    1. Turn on the computer and launch the software **Croissance cristalline [Crystal growth]**. This control software is written with LabVIEW (http://www.ni.com/labview/) and offers a user-friendly graphical interface. It includes 4 different graphical interfaces (**Accueil [Welcome], Paramétrage**

[Setting], Essai [Test], and **Maintenance [Maintenance]**) (**Figure 8**).

NOTE: Translations from French are in brackets.

2. Select the **Maintenance** view by clicking the button as illustrated in **Figure 8**. This view is currently the most often used and allows users to control the majority of parameters during the experiment.

   NOTE: After clicking on the **Maintenance** view, a new window with different sections for controlling parameters such as temperature or light will appear. In **Figure 8** different parts of this view are shown with arrows and frames. In the following steps, we demonstrate how each parameter is controlled using the software.

3. **Régulateur de température [Temperature controller]** section allows the control and monitoring of temperature. Click on the button, number one **(1)** in **Figure 8**, to turn it on.

   NOTE: Temperature range in OptiCrys is 233.0 – 353.0 ± 0.1 K.

4. Set the temperature on the **consigne [setpoint]** section and press enter (**Figure 8(2)**). Below this button, there is a graph with 2 traces (red and yellow). The red trace shows the final (ordered) temperature and the yellow trace shows the current temperature. As shown in **Figure 8(2)**, the temperature is set at 20 °C.

   NOTE: In order to grow a crystal, as explained in the crystal growth section, many temperatures will need to be chosen. To change the temperature, add each new temperature in the **consigne [setpoint]** section and press the Enter button on the keyboard.

5. Turn on the light by increasing the luminosity from the **Lumières [Lights]** section in **Figure 8**. Luminosity ranges from 0 to 100, "0" indicates that the light is off and at "100" the light is set to maximum of intensity and brightness. By increasing the light, in the **Microscope** section, one can see inside the dialysis chamber. During the experiment, the brightness in the cell may vary; adjust the parameters to clearly visualize inside the dialysis chamber. Magnification can also be increased or decreased by using the + and – buttons in front of “zoom” for better viewing.

6. On the right-hand side of the **Microscope** section, there are several sections for storing relevant information for each crystallization experiment. Each user can create a folder to store information on crystallization conditions, protein names and the molecular weight cutoff of the dialysis membrane used (**Figure 8(3)**).

7. The user can define a name for the experiment by simply typing it on the **Nom Dossier [Folder name]**. Clicking on the **Dossier [Folder button]** (shown with a green frame in **Figure 8**) opens a new window. In this window, there will be a text file that contains all the information defined for the experiment. In addition, time-stamped images are saved in this folder for future processing.

8. From the **NB Images** section**,** select the number of images that should be taken during the course of the experiment. Specify the number of the images in the right-hand panel along with the desired time interval between these (e.g., min, hour, day). **Figure**

8 shows the software setup to record zero images in a minute.

NOTE: Use the **Pompe [Pump]** section for mixing stock solutions and injecting crystallization solution into the reservoir chamber. See section 2.1.10 and **Figure 7** for an explanation of the principle of the fluid mixing system.

9. Input concentrations of the stock solutions in **Etape 1 [Step 1]: solutions stocks.** For the crystallization experiment in the next section, NaCl 4 M and CH_3COONa 1 M pH 4 will be used.

   NOTE: Concentrations of stock solutions are in molar units.

10. Define the final concentration of each solution. For example 0.75 M for NaCl and 0.1 M for CH_3COONa pH 4. Input these in the final concentration section (**Figure 8**) in the **Etape 2 [Step 2]: solution à préparer [solution to prepare].** Press the **Calcul [Compute]** button, which is shown with a red frame in **Figure 8**. The final volume of each stock solution that will be used in mixing will display in the volume panel in front of each concentration panel.

11. Press the **Lancer préparation [Launch preparation]** button (**Figure 8**). As illustrated in **Figure 7**, the rotary valve takes each stock solution and injects them to the mixing tube via a switch.

12. After the crystallization solution has been prepared, click on the **Entrée solution [Solution Entry]** button in **Etape 3 [Step 3]: Flux** of the pump section (yellow frame in **Figure 8**). The switch changes to inject the new crystallization solution from mixing tube into the reservoir chamber. To stop the exchanging process, press the **Arrêt distribution [Distribution Stop]** button.

    NOTE: Observe the crystallization process during the experiment and modify parameters such as temperature, crystallization solution and zoom, in the corresponding graphical interface of the supervision software. By using the software there is no need to remove the airtight cap or the flow cell during the experiment so the only variable will be the one that user changes through the software.

3. **Large crystal growth**

   1. Add 15 µL of lysozyme with a concentration of 30 $mg \cdot mL^{-1}$ to the dialysis chamber (**Figure 6A**).

      NOTE: Prepare the protein sample as described in section 1.1.

   2. Assemble the temperature-controlled flowing reservoir dialysis setup as described in section 2.1 and **Figure 6**.

   3. Prepare the crystallization solution. Do not forget to filter all the stock solutions before sample preparation with 0.22 µm filters. For this experiment, the crystallization solution contains 0.75 M NaCl and 0.1 M CH_3COONa pH 4. This can be added manually or by using the reservoir chamber and pumping system as described in sections 2.2.10 to 2.2.12.

   4. Set the temperature to 295 K as explained in sections 2.2.3 and 2.2.4 and shown in Image 8. Under initial conditions, equilibrium between the dialysis chamber and the reservoir will be reached after approximately 90 minutes and the first visible nuclei will appear after 22 hours.

5. Allow crystals to grow until no more visible changes in the size of the crystals are observed (**Figure 9**, panel 1).

   NOTE: In order to determine the nucleation time and to measure the variation in the size of the crystals, record images every 15 or 20 min, which is respectively 4 or 3 images per hour in the **NB Images** section. For *in situ* observation of protein denaturation, aggregation and precipitation or crystal dissolution or nucleation, typically between a few seconds to a few tens of minutes are required. However, for crystal growth, this range is between a few minutes to a few hours.

6. After three days, lower the temperature to 291 K to restart crystal growth. Keep the temperature constant and let the crystal develop (**Figure 9**, panel 2). For this stage of the experiment, it will be sufficient to record images every 2 hours and check every 10 to 12 hours for any change in the size of the crystals. The experiment can be continued if no change in the size of the crystals is observed.

   NOTE: Depending on the protein and precipitant concentrations in the crystallization solution and the volumes of protein used, the time needed to reach the equilibrium for each step may vary.

7. Decrease the temperature to 288 K to restart crystal growth. In the experimental condition of the case presented here, one day is enough to reach equilibrium (**Figure 9**, panel 3).

8. Check the size of the crystals and maintain a constant temperature as long as the crystal continues to grow.

9. After 4 days, decrease the temperature to 275 K in order to restart crystal growth (**Figure 9**, panel 4).

   – In the experimental conditions of the case presented, after around 10 days a crystal that is 500 µm in one dimension will be obtained (**Figure 9**).

4. **Controlling the crystal size**

   1. Prepare a protein solution and the temperature-controlled flowing reservoir dialysis set up as described in sections 2.3.1 and 2.3.2.

   2. Prepare crystallization solutions with 0.9 M NaCl and 0.1 M CH_3COONa pH 4.

   3. Set the temperature to 291 K and allow crystals to grow. Under initial conditions, the first nucleation event will start after around one hour and numerous crystals will grow in the dialysis chamber for three hours (**Figure 10**, steps: 1 and 2). Record images every 20 minutes to check the size of the crystals during the growth process.

      NOTE: The optimization of crystallization conditions is crucial in controlling most of the final properties of generated crystals. The temperature and chemical composition of crystallization solutions can be changed to dissolve and re-grow crystals of uniform size. It should be also noted that a protein sample is not consumed in such an experiment, as conditions can be reversed to re-dissolve the sample as long as it is not denatured. When dissolving crystals by changing temperature and maintaining constant chemical composition, continue the experiment as follows:

   4. Once crystals have grown at 291 K, these can be dissolved to re-grow fewer, larger crystals. Increase the temperature gradually over 20 min to reach 313

K. It takes around one hour to dissolve all the crystals inside the dialysis chamber (**Figure 10**, steps: 3-5). Record images every 5 to 15 minutes to monitor the dissolution process.

NOTE: Many proteins are sensitive to high temperatures. Be sure to work within the temperature range where the protein is stable to avoid any damage/denaturation. In addition to the protein solubility, temperature also affects the buffer solution. For example, the pH of the buffer can change with temperature, especially in the Tris buffer. In such a case, it is crucial to set the pH according to the temperature at which the experiment is performed[18]. It should also be noted that protein dissolution takes significantly less time (from a few minutes to a few hours) compared to protein crystal growth (from a few hours to a few days). In general, during the dissolution of the crystals, the temperature increases gradually and slowly (respecting the short total dissolution time), mainly in the case of partial dissolution of the crystals to avoid the increase of the crystal mosaicity. When the crystals are growing, the temperature can decrease quickly (in less than a minute) to the set temperature (respecting the long total growth time). Regular monitoring of the crystallization chamber by recording images is advisable to prevent damage to the protein and help define the optimal time for dissolution or growth of crystals for each protein studied.

5. After all crystals have dissolved set the temperature to 295 K to initiate a second nucleation event (**Figure 10**, steps: 6,7). Record images every 5 minutes to monitor the second nucleation process. In this step, the first nuclei appear after around 18 minutes.

   NOTE: At this temperature, the solution will be in the nucleation zone, in the vicinity of the metastable zone. As a result, only a few nuclei will appear in the crystallization chamber.

6. Continue the experiment repeating the optimization workflow described in section 2.3 for growing larger crystals. The total duration of the experiment represented in **Figure 10** is only a few days.

   NOTE: If during the nucleation phase the crystals appear at different times, crystals of different sizes are obtained in the crystallization chamber. In such a case, the increase in temperature (in the case of proteins with direct solubility) will result in quicker dissolution of the smaller crystals. Depending on the kinetic ripening effect, the extra protein (gained from dissolution) can then be used for the growth of the larger crystals.

   When dissolving crystals at constant temperature by changing the chemical composition of the crystallization solution, continue the experiment as follows:

   It is also possible to dissolve crystals grown previously by changing the chemical composition of the crystallization solution during the experiment to re-grow a population of uniformly sized crystals under new conditions.

7. Prepare the protein solution (2.3.1), the temperature-controlled flowing reservoir dialysis setup (2.1) and the crystallization solution (2.4.2) as described above. Under initial conditions in the nucleation zone far from the metastable zone, numerous small crystals will appear in the

crystallization chamber and begin to grow (**Figure 11**, steps: 1,2).

8. After three hours, when many medium-sized-crystals are visible in the crystallization chamber (**Figure 11**, step: 3), decrease the NaCl concentration (0.9 M) gradually to reach zero. For this, prepare a new crystallization solution containing only buffer solution with 0.1 M CH_3COONa pH 4. Use the pumping system to exchange it with the crystallization solution in the reservoir chamber. Follow steps 2.2.10 to 2.2.12 for the preparation and injection of a new solution into the reservoir chamber. With this new solution, when solutions are exchanged the NaCl concentration decreases in the chamber until the final solution in the reservoir chamber contains no more than 0.1 M of CH_3COONa pH 4 and no NaCl. Capture images every 10 minutes to record the dissolution process.

9. Allow the crystals to dissolve completely (**Figure 11**, steps: 4,5). Dissolution time is around two hours for this experiment. As previously mentioned, dissolution time is dependent on the protein system, crystallization conditions, and dialysis chamber volume used. Regular observation of the crystallization chamber (see **Microscope** section) and recording images and notes during the experiment are essential.

10. When all crystals inside the chamber are dissolved (**Figure 11**, step: 5), use the pumping system again to prepare a new crystallization solution by injecting NaCl at a lower concentration than the previous one (0.75 M NaCl in 0.1 M CH_3COONa pH 4).

11. Inject the new solution into the reservoir chamber (**Figure 11**, steps: 6,7) and repeat the crystallization growth optimization workflow as described in section 2.3. A uniform population of larger crystals will be generated. The results shown in **Figure 11** were obtained after a few days.

Representative Results

In Sections 2.3 and 2.4, three examples of optimized crystal growth are presented, showing use of the instrument and an experimental design for growing large crystals. For this demonstration, we have used lysozyme as a model protein, although crystal growth experiments have been successfully performed with many other protein systems using this method (see above). By using and mastering the protocol presented here one can adapt it for other protein candidates.

In section 2.3 we demonstrated that established rational crystallization strategies could be beneficial in growing crystals with sufficient scattering volumes for neutron protein crystallography. Here, we demonstrate that the rational optimization strategies proposed also allow the generation of a uniform population of crystals of any specific size required for downstream structure determination approaches.

These two experiments are designed to emphasize the importance of phase diagrams in controlling crystal nucleation and growth. Here, control of the temperature and chemical composition of crystallization solutions in combination with monitoring the crystallization process in real time are used to study the qualitative phase diagram. Using this method, nucleation and crystal growth can be rationally optimized in a reversible manner. Use of such a serial approach also reduces the amount of protein and the time required to control the size and quality of the crystals.

In the dialysis method, a protein solution is separated from a crystallization solution by a semi-permeable membrane[6] (**Figure 5**). This dialysis membrane allows small molecules such as additives, buffer, and ions to pass through the membrane but not macromolecules such as proteins[6,20]. This feature allows the crystallization solution to be modified during the course of the experiment[6]. Exchange of the solution can be done manually, for example in microdialysis buttons, or in an automated manner using an instrument developed for this purpose, OptiCrys[8].

In the first set of experiments, microdialysis buttons were used for the crystallization of chicken egg-white lysozyme. Microdialysis buttons were immersed in crystallization solutions with different salt concentrations. In this simple crystallization grid experiment, the only variable is precipitant concentration whilst temperature is kept constant (293 K). As shown in **Figure 4**, slight variations in the salt concentration induce a change in the size and numbers of crystals observed, allowing investigation of the crystallization phase diagram. In **Figure 4**, panel 1, the crystallization solution contains 0.7 M NaCl and a limited number of larger crystals have appeared in the buttons. By increasing salt concentration from 0.7 to 1.2 M, supersaturation increases and the solution in the nucleation zone moves away from the metastable zone (**Figure 4**, panels 1 to 6). As a result, the number of crystals increases and their size decreases.

In the first experiment with a fully automated instrument enabling temperature-controlled dialysis crystallization, OptiCrys (**Figure 9**), the crystal growth experiment was tailored to generate large crystal growth. The experiment was launched at an initial temperature of 295 K with a crystallization solution containing 0.75 M NaCl and 0.1 M Na acetate buffer pH 4. Under these experimental conditions, the crystallization solution reached the nucleation zone in the vicinity of the metastable zone of the phase diagram (**Figure 9**, arrow 1). As a result, only a few nuclei were generated during the first stage of the experiment. In order to grow selected crystals further (shown in **Figure 9**), the crystal growth optimization workflow was driven towards the metastable zone by varying temperature as soon as the crystal-solution equilibrium was reached.

Each time equilibrium between crystal and solution was reached, the temperature was lowered, first to 291 K, then to 288 K and finally to 275 K, to keep the crystallization solution in the metastable zone. The result of this experiment is a single large crystal suitable for both macromolecular X-ray and neutron crystallography.

For most proteins, the precise quantitative phase diagram (or just a qualitative diagram) has not yet been obtained due to the lack of experimental devices capable of accurately measuring protein concentration (or just of observing/ detecting the crystallization process in real time) during crystallization experiments[18]. As a result, it is often not possible to design the experiment in such a way that crystallization begins in the optimal area of the phase diagram, in the vicinity of the metastable zone.

Therefore, a crystallization optimization study must take place before the experiment dedicated to the growth of a large volume crystal is undertaken. In this study, using temperature variations (at constant chemical composition) on the one hand and variations in chemical composition (at constant temperature) on the other hand, are necessary to identify the metastable zone and to delineate the optimal conditions for starting a large crystal growth experiment.

To this end, two other experiments are presented which were tailored to demonstrate the reversibility of the temperature-controlled dialysis crystallization experiments with OptiCrys for nucleation, crystal growth, dissolution and re-growth. The crystal growth optimization workflow was controlled so that a uniform population of fewer, larger lysozyme crystals was grown, using variation of temperature or precipitant concentration.

In the second experiment with OptiCrys, the chemical composition of the crystallization solution was kept constant throughout the experiment (0.9 M NaCl in 0.1 M CH_3COONa pH 4) with variable temperature. The initial temperature was set at 291 K. The results of this experiment are summarized in **Figure 10**. Because of high supersaturation, a large number of small crystals appeared in the crystallization chamber (**Figure 10**, panels 1 and 2). In accordance with the concept of direct protein solubility, by gradually increasing the temperature to 313 K, all of the crystals were dissolved (**Figure 10**, panels 3, 4 and 5). Finally, by lowering the temperature to 295 K, the second nucleation was initiated in the vicinity of the metastable zone and allowed controlled formation of a lower number of nuclei. Further crystal growth resulted in the uniform generation of a population of larger crystals (**Figure 10**, panel 7).

As shown in **Figure 11**, variation of the chemical composition of the crystallization solution, at a constant temperature of 291 K, can likewise be used to obtain a uniform population of larger crystals. Similar to the previous experiment, the initial condition was 0.9 M NaCl in 0.1 M CH_3COONa pH 4. The NaCl concentration was then lowered gradually from 0.9 M to zero to dissolve the crystals (**Figure 11**, panels 4 and 5). At this point, NaCl was completely replaced by a buffer solution of 0.1 M CH_3COONa pH 4. Reducing the salt concentration keeps the solution in the undersaturated zone of the phase diagram, which leads to the dissolution of the crystals. Then, a new crystallization solution with lower ionic strength, at 0.75 M NaCl in 0.1 M CH_3COONa pH 4, was injected into the reservoir chamber. At this precipitant concentration, the first nuclei appeared (**Figure 11**, panel 6) after 90 minutes. The number of generated crystals was lower and the crystals reach a larger volume (**Figure 11**, panel 7) than before.

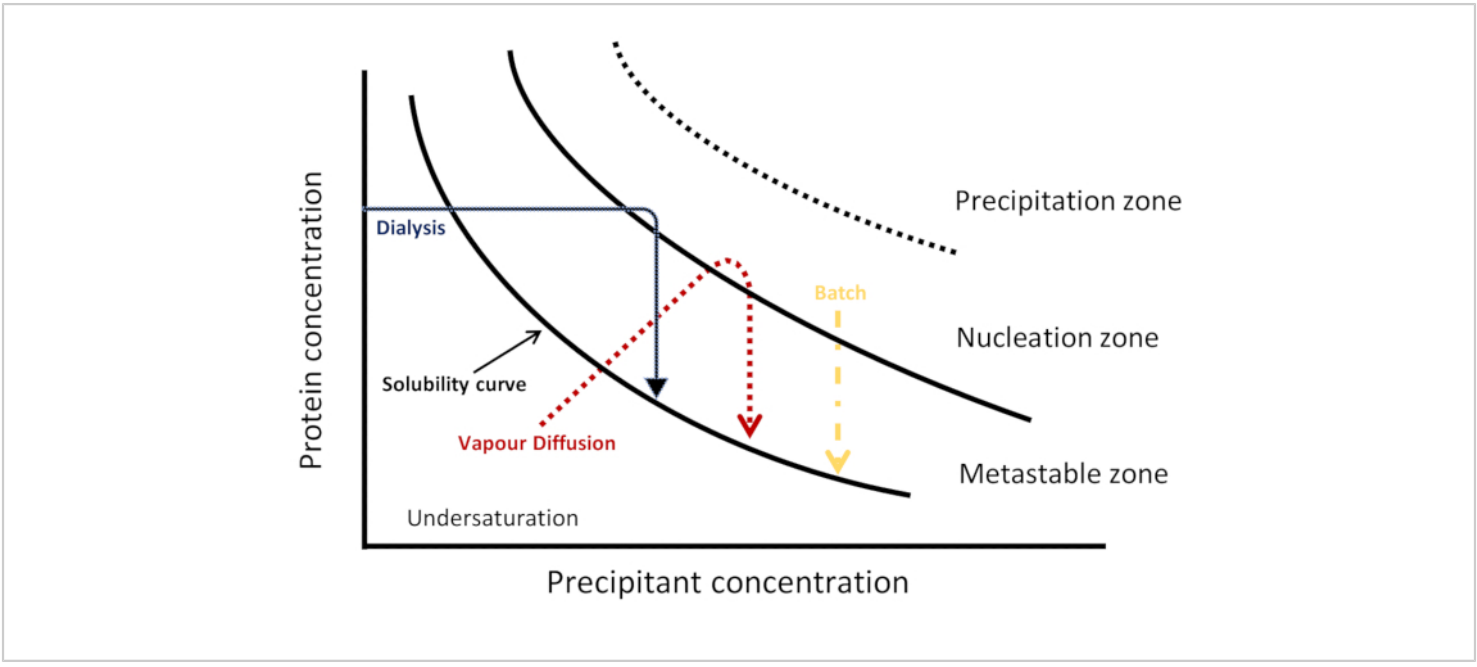


Figure 1: Schematic phase diagram. Kinetic trajectories for three crystallization techniques are represented in a salting-out regime. Each method achieves nucleation and crystallization differently, visualized by a different kinetic pathway through the phase diagram to reach the nucleation and metastable zones. The solubility curve separates undersaturation and supersaturation regions. Supersaturation is divided into three zones: metastable, nucleation and precipitation. In the nucleation zone, spontaneous nucleation occurs while in the metastable zone crystal growth takes place. This Figure is adapted from Junius *et al*.[8] Please click here to view a larger version of this figure.

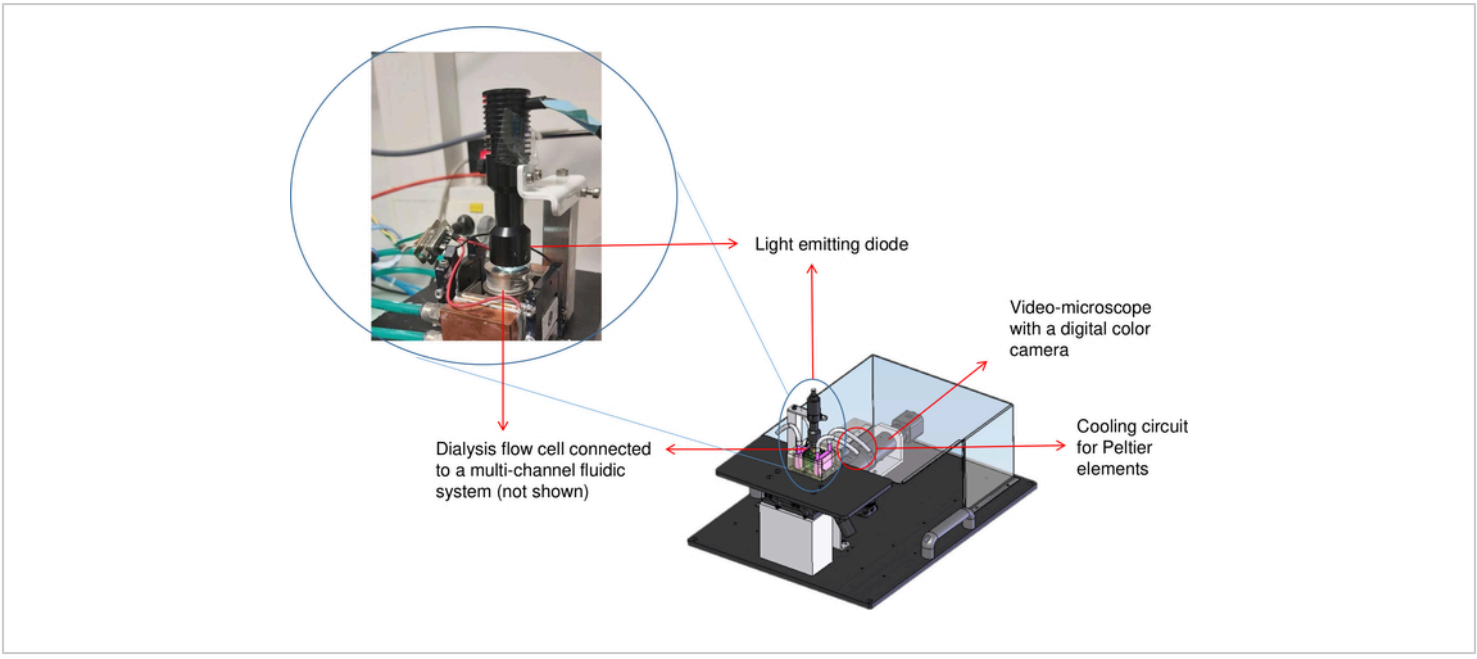


Figure 2: Schematic representation of the crystallization bench (OptiCrys). The LED light source is located on top of the temperature-controlled dialysis flow cell. An inverted microscope and the digital camera are shown at the top right of the image with the red arrow. The red circle represents the location of the chiller tubing. Please click here to view a larger version of this figure.

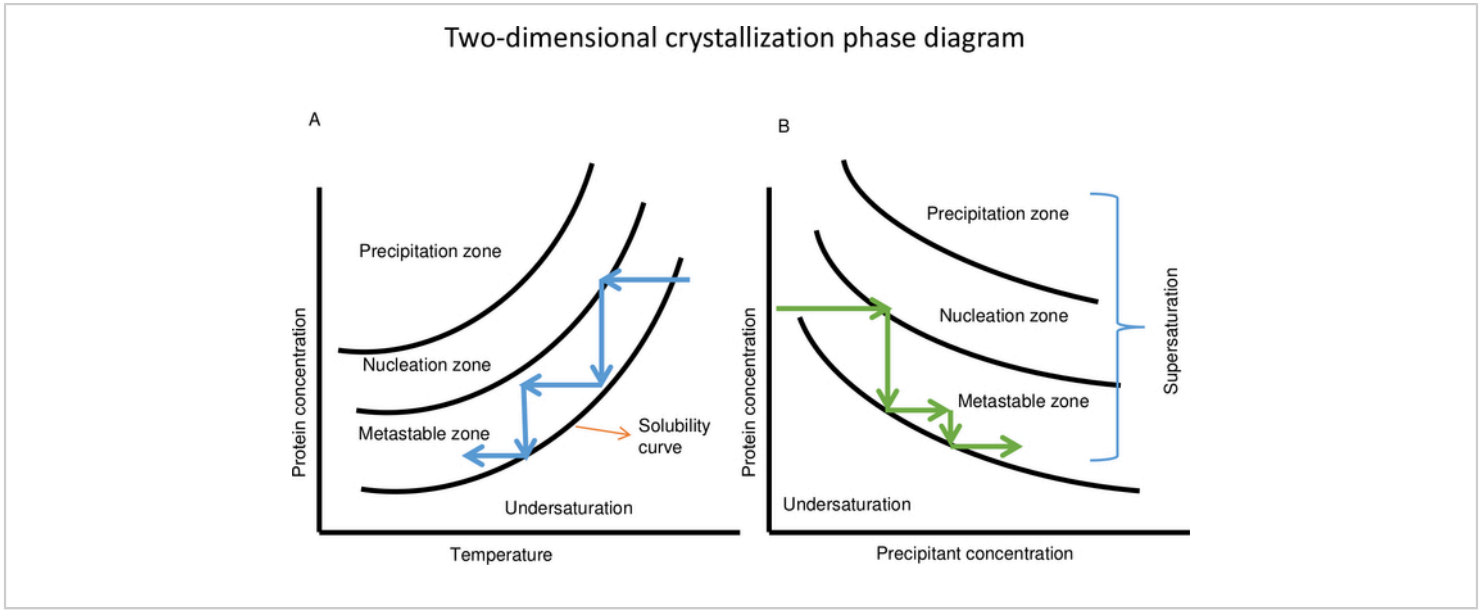
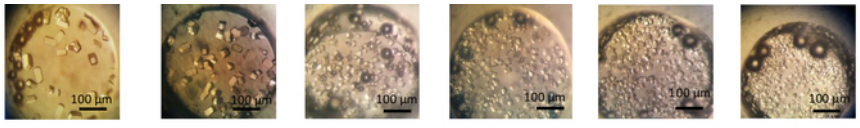


Figure 3: Schematic two-dimensional protein crystallization phase diagram as a function of temperature (A) and precipitant concentration (B). (A) In case of a protein with direct solubility, decreasing the temperature keeps the crystallization solution in the metastable zone. Temperature variation can be repeated several times to control the crystal growth process until crystals with the desired volume are obtained. (B) Changing the concentration of the precipitant solution can also be used to keep the crystallization solution in the metastable zone for growing crystals. This Figure is adapted from Junius *et al.*[8] Please click here to view a larger version of this figure.



1	2	3	4	5	6
NaCl 0.7 M 0.1 M NaOAc pH 4	NaCl 0.8 M 0.1 M NaOAc pH 4	NaCl 0.9 M 0.1 M NaOAc pH 4	NaCl 1.0 M 0.1 M NaOAc pH 4	NaCl 1.1 M 0.1 M NaOAc pH 4	NaCl 1.2 M 0.1 M NaOAc pH 4

Figure 4: Crystals of lysozyme obtained using the dialysis method. This experiment was performed at a constant temperature of 293 K in 0.1 M sodium acetate buffer pH 4. Increasing NaCl concentration from 0.7 M to 1.2 M increases the nucleation rate and results in a larger number of crystals. Please click here to view a larger version of this figure.

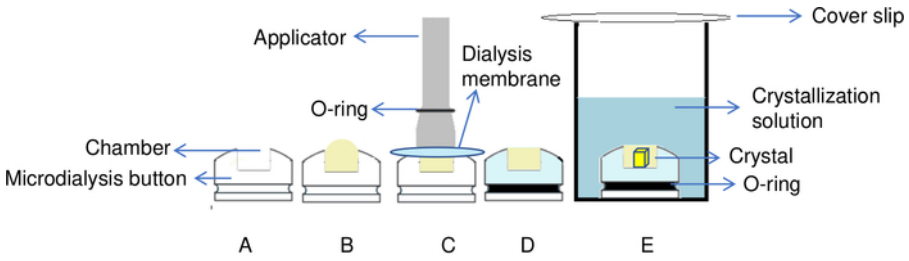


Figure 5: Overview of the protein crystallization process by the dialysis method. (A) By adding the protein to the chamber of the dialysis button, (B) a dome shape is created on the top of the chamber. (C) An applicator is used to transfer the O-ring to the groove of the dialysis button in order to fix the dialysis membrane in place. (D) The dialysis button is ready for immersion in the reservoir solution. (E) Crystallization solution passes through the semipermeable membrane and crystals start to form inside the chamber. Please click here to view a larger version of this figure.

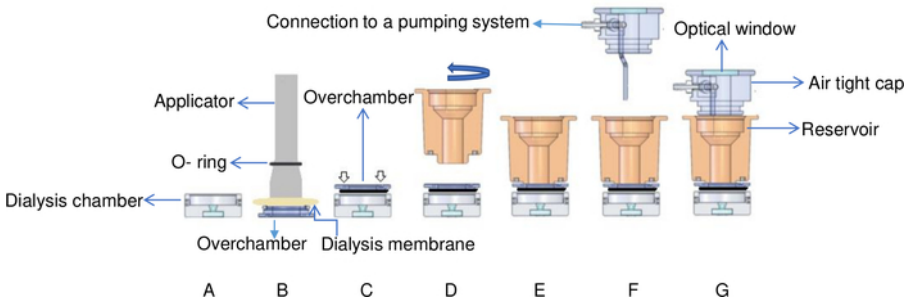


Figure 6: Schematic view of the temperature-controlled flowing dialysis setup. (A) The protein sample is added to the dialysis chamber. (B) The dialysis membrane is fixed onto the overchamber with an O-ring by using an applicator. (C) The overchamber is turned and fixed onto the top of the dialysis chamber. White arrows indicate where screws are placed on the overchamber. (D) The reservoir chamber is turned clockwise (E) and fixed on top of the overchamber. (F) The reservoir chamber is covered by an airtight cap with connectors to a pumping system and (G) the flow cell is placed in the brass support. This Figure is adapted from Junius *et al.*[8] Please click here to view a larger version of this figure.

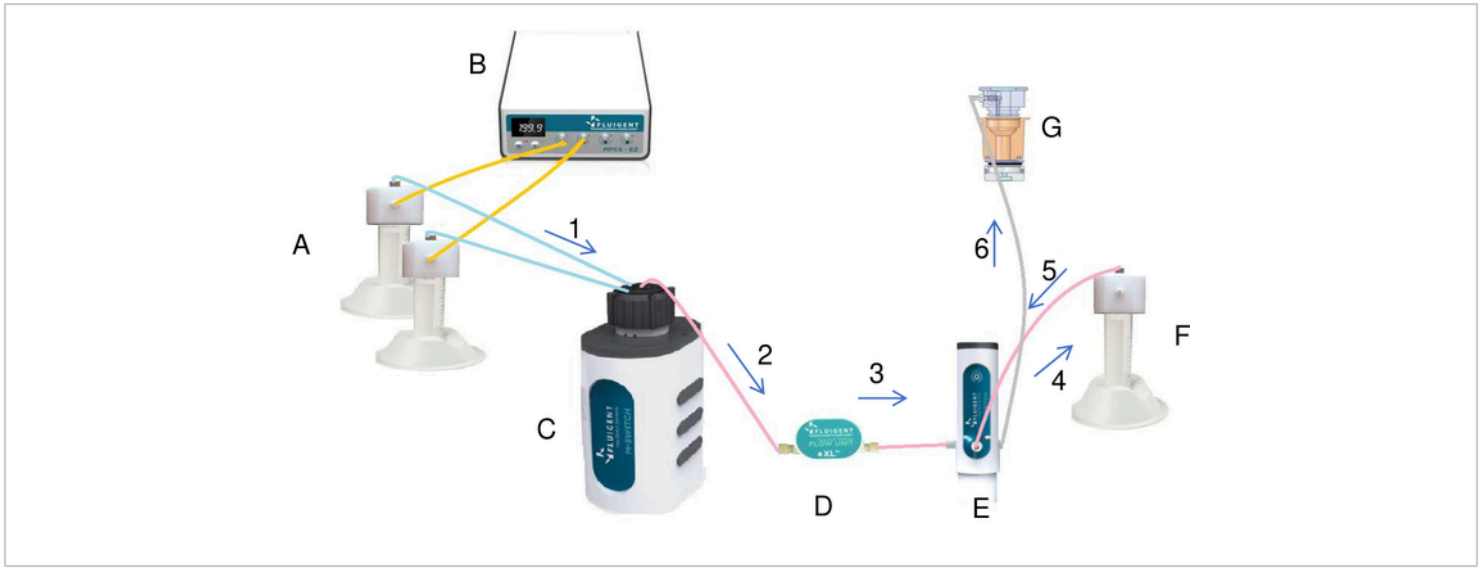


Figure 7: Preparation and injection of the crystallization solution in the reservoir by the fluidic system (A). Tubes containing salt and water are connected to the pressure/vacuum controller (B) and to the rotary valve (C). By using the pressure, pressure/vacuum controller creates a constant flow of the liquids from the tubes to the rotary valve. Each liquid passing through the flow meter (D) and the switch is injected into the mixing tube (F). Once all the liquids have been added to the mixing tube, the switch by some modifications injects the final solution from the mixing tube into the reservoir (G). The liquid flows through the system in the direction of the arrows in the diagram marked in ascending order (from 1 to 6). Please click here to view a larger version of this figure.

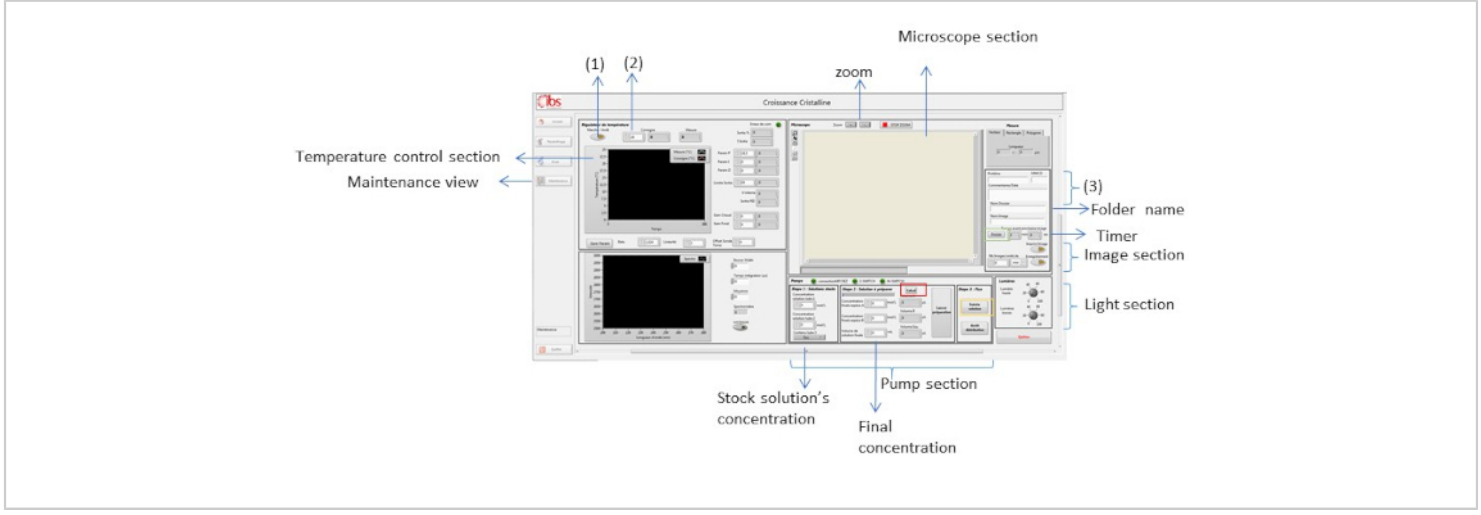


Figure 8: Maintenance view of the supervision software. This view is used to control different parameters like temperature, light, crystallization solution and zoom. Please click here to view a larger version of this figure.

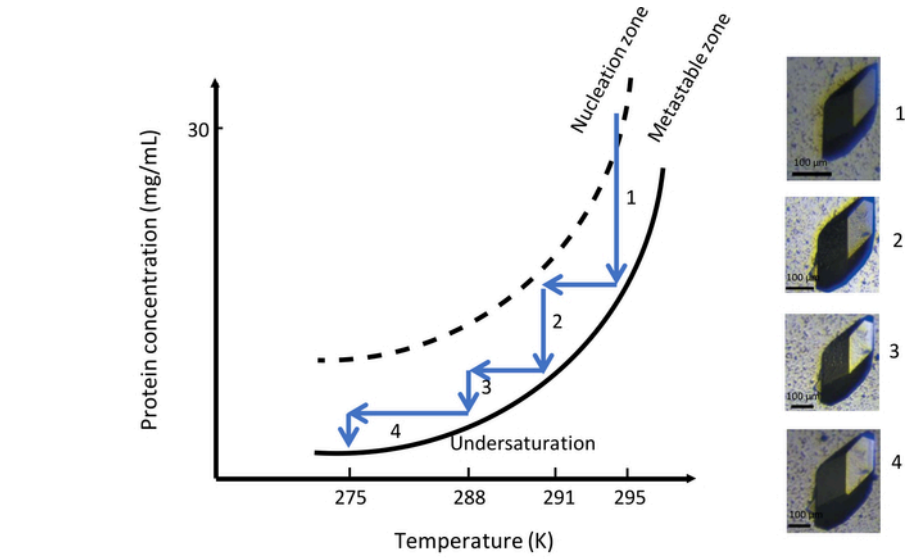


Figure 9: The phase diagram as a function of the temperature (selected images are to be tracked in ascending order). A single large lysozyme crystal is obtained by systematically changing the temperature from 295 K to 275 K. At each step, crystal growth is stopped upon reaching the solubility curve. Reducing the temperature by keeping the solution in the metastable zone restarts crystal growth. The images have different levels of magnification. This Figure is adapted from Junius *et al.*[8,18] Please click here to view a larger version of this figure.

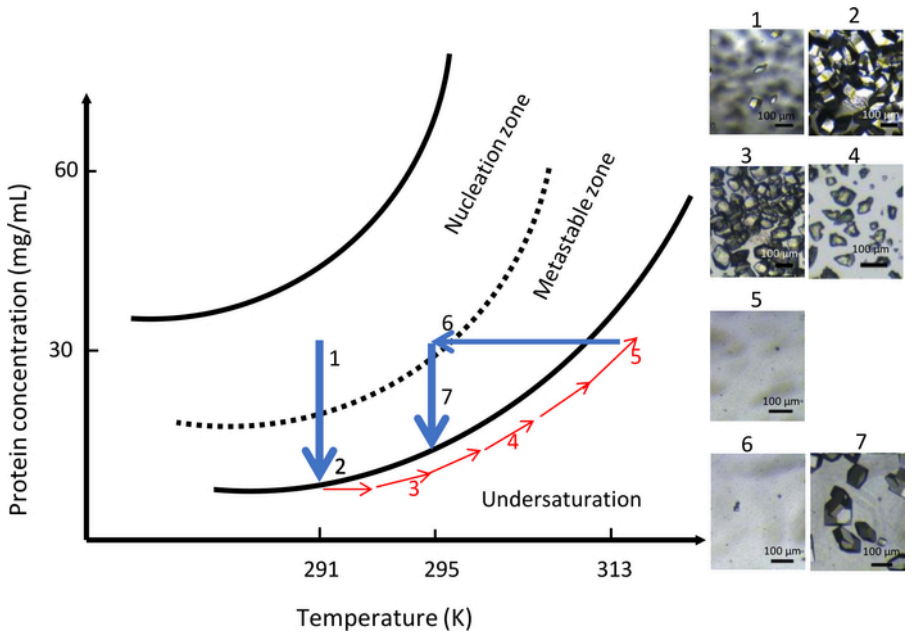


Figure 10: Optimization of crystal growth at constant chemical composition using temperature control (selected images are to be tracked in ascending order). Starting the nucleation process in the nucleation zone at 291 K, far from the metastable zone, results in the formation of numerous crystals. Increasing the temperature to 313 K then dissolves the crystals until no visible nuclei are seen in the dialysis chamber. Finally, decreasing the temperature to 295 K restarts the nucleation process for the second time leading to a limited number of larger crystals. This Figure is adapted from Junius *et al.*[8,18] Please click here to view a larger version of this figure.

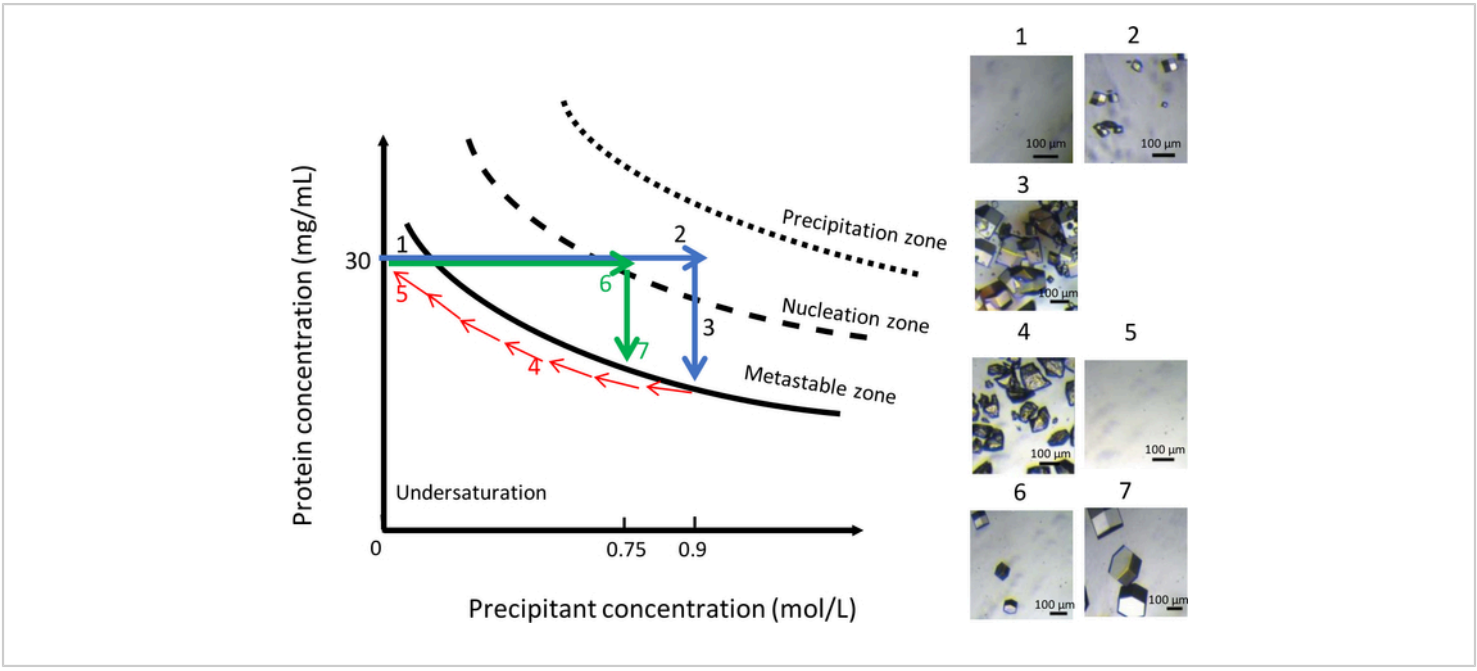


Figure 11: Optimization of crystal growth at constant temperature using variations in precipitant concentration (selected images are to be tracked in ascending order). Decreasing the precipitant concentration from 0.9 M to 0 M dissolves the crystals obtained during the first nucleation event. The crystallization process is restarted by the injection of the same precipitant but at lower ionic strength, 0.75 M, which leads to the formation of a few larger crystals. This Figure is adapted from Junius *et al.*[8,18] Please click here to view a larger version of this figure.

Discussion

Different physical, chemical and biological variables influence protein crystallization by affecting protein solubility[21]. Among these variables, temperature and chemical composition of the crystallization solution are used here in combination with dialysis technique to improve and grow large high-quality crystals of biomacromolecules for neutron diffraction studies. By using knowledge of phase diagrams, crystallization is made more predictable. Although screening of different crystallization conditions in a serial approach is also possible, the main aim of using the rational approaches presented is to separate and control the kinetics of crystal nucleation and growth.

Similar to all crystallization studies, high quality pure and homogeneous protein samples, and dust-free crystallization solutions increase the success rate of the experiment. Filtration and centrifugation of solutions are essential steps in the described protocols. Knowing the physicochemical properties of the proteins studied such as the molecular weight (to choose the appropriate dialysis membrane), the isoelectric point, and the protein solubility are crucial for the design of an optimal crystal growth experiment. Also, consideration must be made for protein stability at different temperatures or with different chemicals to prevent sample loss and increase the likelihood of success. Considering the temperature range of OptiCrys (233.0–353.0 ± 0.1 K), a broad range of proteins can be crystallized using it. But it is

worth to stress that proteins that are primarily thermo-stable, such as proteins from thermophilic sources, would benefit the most in temperature-controlled large-volume crystal growth experiments offered by this instrument.

Using a low-volume dialysis chamber (when using OptiCrys) or microdialysis buttons and screening several temperatures and crystallization conditions (e.g., grids of precipitant concentration or pH), it is possible to gain information on the location of the limit of the metastable zone (kinetic equilibrium between nucleation and metastable zones). This is invaluable when designing a successful crystal growth experiment especially for new protein candidates in crystallization. Without this information experiments can start from an area of the phase diagram with high supersaturation, too far from the limit of the metastable zone to easily control crystal nucleation. Although dissolution of the protein precipitate may be attempted, for example by increasing the temperature in the case of direct solubility, for proteins with reduced thermostability, keeping the sample at high temperature for a longer period of time may render the protein precipitation irreversible. Thus, the best strategy consists of using an initial condition with lower supersaturation located near the limit of metastability, where nucleation can be controlled and protein precipitation avoided. In line with this, crystallization prescreening decreases the chance of having a protein precipitate in the dialysis chamber and increases the success rate of the experiment.

After designing an experiment, preparing dialysis chambers (OptiCrys) or microdialysis buttons is another important step. Preventing air bubble formation in the dialysis chamber/ button increases the chance of successful crystallization especially when small volumes are used. The presence of air bubbles in dialysis chamber may also change the kinetics of the crystallization process and reduce the reproducibility of the experiment (because the protein/solution contact surface has been modified). Not only protein but also crystallization solution can affect the success of the experiment. Using new 50 mL tubes for the pumping system each time one wants to start a new experiment and washing tubing after each experiment decreases the chance of contamination and avoids the creation of salt crystals in the apparatus.

The use of microdialysis buttons is an alternative when OptiCrys is not available. The strategies for optimizing crystallization and monitoring crystal growth mentioned above, must be carried out manually. Typically this necessitates being outside a thermoregulated incubator, which can be problematic when temperature regulation is a critical step in the methodology described. This does not facilitate changing the chemical composition of the crystallization solutions, or monitoring crystal growth by imaging, so the crystal growth process cannot be controlled in real-time.

Knowledge of the phase diagram is the basis of using the crystallization bench, OptiCrys, to systematically grow large, high-quality crystals in an automated fashion. Control of physicochemical parameters like temperature, precipitant concentration, and pH during crystallization moves the protein-solution equilibrium in a well-defined kinetic trajectory across the phase diagram. This is complemented by the use of a dialysis membrane to adjust mass transport and create a controlled gradient in the crystallization chamber that affects the size and quality of the crystals. Therefore, using both thermodynamic data and kinetic trajectories is essential to control the crystallization process in order to grow high-quality crystals. Thanks to OptiCrys, systematic phase diagrams in a multidimensional space can be studied with a serial approach

using significantly less material than before. To demonstrate this methodology, we provide here a case study with a model protein, chicken egg-white lysozyme. By using and mastering the protocol presented here one can adapt it for real protein systems[5,14,17,18].

Disclosures

The authors have nothing to disclose.

Acknowledgments

MBS acknowledges the support from the LABEX VALO GRAL under the contract 2015. NJ acknowledges CEA's International Doctoral Research Program (Irtelis) for the PhD Fellowship. Authors acknowledge funding from the European Union's Horizon 2020 Research and Innovation Program under Marie Skłodowska-Curie grant agreement number 722687. Authors are also grateful to Dr Esko Oksanen (ESS, Lund) and Dr Jean-Luc Ferrer (IBS, Grenoble) for helpful conversations and insights. IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA).

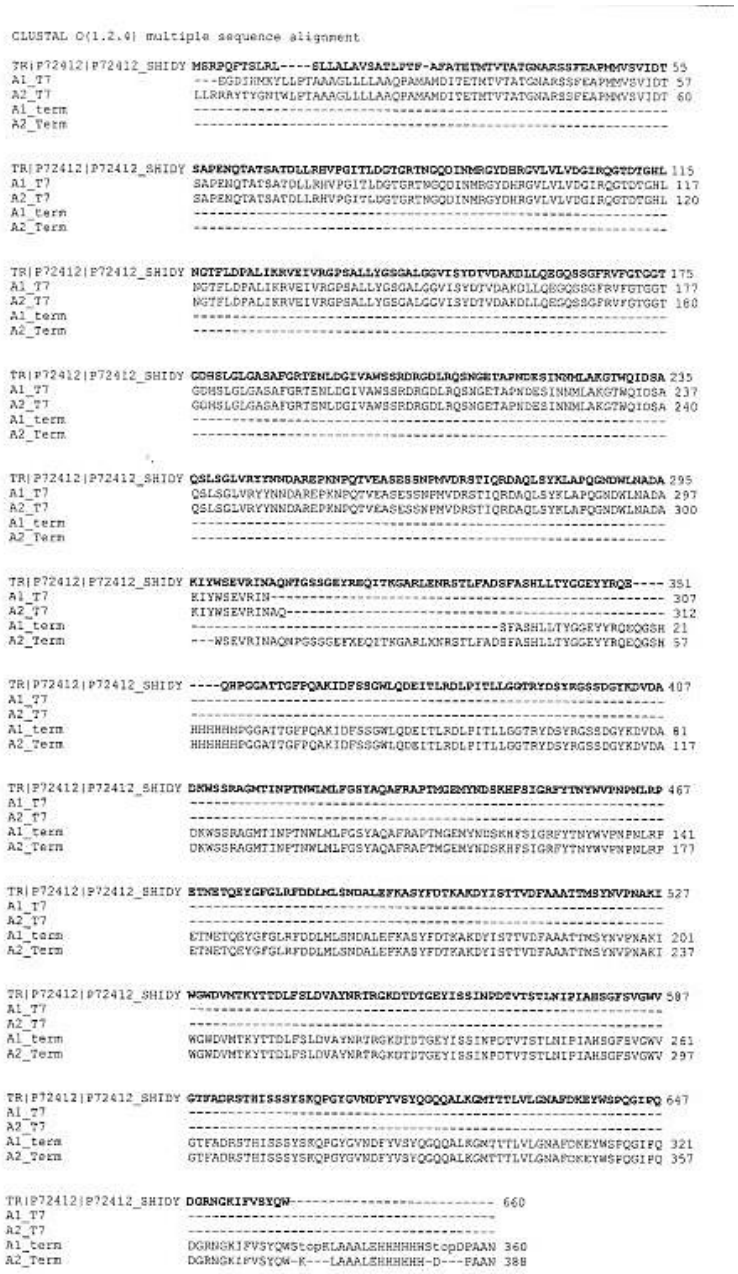
References

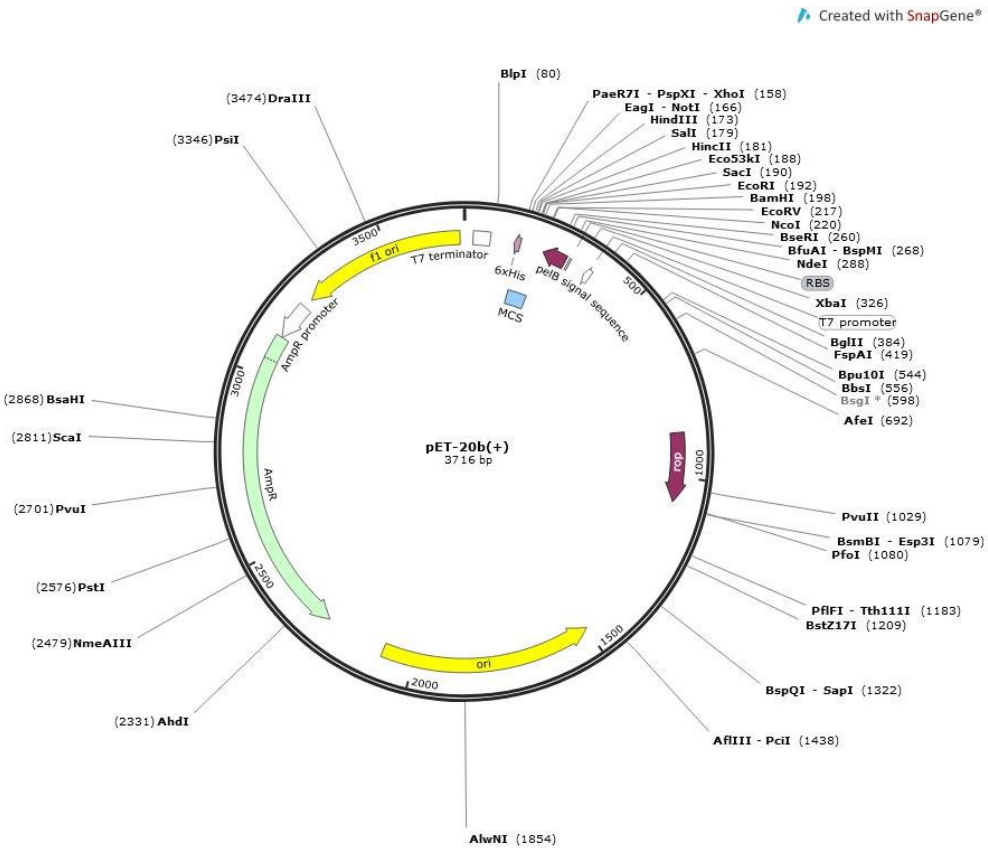
1. Blakeley, M.P., Langan, P., Niimura, N., Podjarny, A. Neutron crystallography: opportunities, challenges, and limitations. *Current Opinion in Structural Biology.* **18** (5), 593-600 (2008).
2. Snell, E.H., Van Der Woerd, M.J., Damon, M., Judge, R.A., Myles, D.A.A., Meilleur, F. Optimizing crystal volume for neutron diffraction: D-xylose isomerase. *European Biophysics Journal.* **35** (7), 621-632 (2006).
3. Ng, J.D., Baird, J.K., Coates, L., Garcia-Ruiz, J.M., Hodge, T.A., Huang, S. Large-volume protein crystal growth for neutron macromolecular crystallography. *Acta Crystallographica Section F:Structural Biology Communications.* **71**, 358-370 (2015).
4. O'Dell, W.B., Bodenheimer, A.M., Meilleur, F. Neutron protein crystallography: A complementary tool for locating hydrogens in proteins. *Archives of Biochemistry and Biophysics.* **602**, 48-60 (2016).
5. Oksanen, E., Blakeley, M.P., Bonneté, F., Dauvergne, M.T., Dauvergne, F., Budayova-Spano, M. Large crystal growth by thermal control allows combined X-ray and neutron crystallographic studies to elucidate the protonation states in Aspergillus flavus urate oxidase. *Journal of the Royal Society Interface.* **6** (2009).
6. Krauss, I.R., Merlino, A., Vergara, A., Sica, F. An overview of biological macromolecule crystallization. *International Journal of Molecular Sciences.* **14** (6), 11643-11691 (2013).
7. Chayen, N.E. Comparative Studies of Protein Crystallization by Vapour-Diffusion and Microbatch Techniques. *REVIEW Acta Cryst.* (1998).
8. Junius, N., Oksanen, E., Terrien, M., Berzin, C., Ferrer, J.L., Budayova-Spano, M. A crystallization apparatus for temperaturecontrolled flow-cell dialysis with real-time visualization. *Journal of Applied Crystallography.* **49**, 806-813 (2016).
9. Salemme, F.R. A free interface diffusion technique for the crystallization of proteins for X-ray crystallography. *Archives of Biochemistry and Biophysics.* **151** (2), 533-539 (1972).
10. Chayen, N.E., Saridakis, E. Protein crystallization: From purified protein to diffraction-quality crystal. *Nature Methods.* **5** (2), 147-153 (2008).

11. Otálora, F., Gavira, J.A., Ng, J.D., García-Ruiz, J.M. Counterdiffusion methods applied to protein crystallization. *Progress in Biophysics and Molecular Biology.* **101** (1-3), 26-37 (2009).
12. Garciá-Ruiz, J.M. Counterdiffusion Methods for Macromolecular Crystallization. *Methods in Enzymology.* **368**, 130-154 (2003).
13. Budayova-Spano, M., Koruza, K., Fisher, Z. Large crystal growth for neutron protein crystallography. *Methods in Enzymology.* **634**, 21-46 (2020).
14. Budayova-Spano, M., Dauvergne, F., Audiffren, M., Bactivelane, T., Cusack, S. A methodology and an instrument for the temperature-controlled optimization of crystal growth. *Acta Crystallographica Section D: Biological Crystallography.* **63** (3), 339-347 (2007).
15. Asherie, N. Protein crystallization and phase diagrams. *Methods.* **34** (3), 266-272 (2004).
16. Astier, J.P., Veesler, S. Using temperature to crystallize proteins: A mini-review. *Crystal Growth and Design.* **8** (12), 4215-4219 (2008).
17. Budayova-Spano, M. *et al.* A preliminary neutron diffraction study of rasburicase, a recombinant urate oxidase enzyme, complexed with 8-azaxanthin. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications.* **62** (3), 306-309 (2006).
18. Junius, N., Vahdatahar, E., Oksanen, E., Ferrer, J.L., Budayova-Spano, M. Optimization of crystallization of biological macromolecules using dialysis combined with temperature control. *Journal of Applied Crystallography.* **53** (3) (2020).
19. Grimsley, G.R., Pace, C.N. Spectrophotometric Determination of Protein Concentration. *Current Protocols in Protein Science.* **33** (1), 3.1.1-3.1.9 (2003).
20. McPherson, A., Gavira, J.A. Introduction to protein crystallization. *Acta Crystallographica Section F:Structural Biology Communications.* **70** (1), 2-20 (2014).
21. McPherson, A. Introduction to protein crystallization. *Methods.* **34** (3), 254-265 (2004).

Appendix II

Multiple sequence alignment for two plasmids contained ShuA gene sequence purified from *E.coli*.





Plasmid map of pET-20b vector

Appendix III

Lipidic cubic phase

The protocol used here is adapted from that used in Caffrey et.al.[81], [191]. Crystallization solutions were prepared and added to the 24 well crystallization plate. Monoolein(Molecular Dimensions) was prepared by adding solid lipid to a 1.5 mL Eppendorf tube which was then heated to ~40 °C for 5 min to melt the lipid. After removing the plunger of a Hamilton gas-tight syringe, 31 µL (for lysozyme) and 15 µL (for ShuA experiment) of the molten lipid was pipetted into the syringe. By inserting the plunger, the lipid was moved slowly to the end of the syringe. In the next step 21 µL (for lysozyme) and 10 µL (for ShuA experiment) of the protein solution was transferred to a second 100 µL syringe. A coupler connected both syringes and the protein solution was moved to the syringe containing the lipid. Both solutions were mixed until they became homogeneous and transparent. After pushing the entire sample into one syringe, it was attached to a syringe dispenser. The coupler was exchanged for a needle and then 0.2 µL of the cubic phase was transferred to a glass coverslip. 1 µL of the crystallization solution was added to the sample and the glass coverslip was flipped on top of the well containing the crystallization solution. Following sample preparation regular 24 well crystallization plates and coverslips were used.

Lysozyme solution was prepared by dissolving lyophilized protein in milli-Q water for a final concentration of 60 mg/mL. Monoolein was mixed with the protein in a 3:2 volume ratio. Two sets of crystallization solutions were prepared. In group 1, 0.7 to 1.2 M of sodium chloride in the buffer of sodium acetate (0.1 M pH 4) were used. Although crystals were formed in all conditions with this group, samples with 0.7 and 0.8 M (Figure III,A) sodium chloride contained cubic crystals. In group 2, solutions contained 1 M sodium chloride, 0.1 M sodium acetate pH4 and 0 to 25% PEG 400. In this group, crystals were observed when the PEG concentration was less than 10%.

As shown in (Figure III, A) crystals of lysozyme formed both inside and outside the lipidic cubic phase (LCP), in the LCP the crystals being significantly smaller. Following this observation, the next experiment with ShuA was operformed to grow small crystals for microED. This aim was successfully achieved using this rational crystallization method and uniformly sized small ShuA crystals were grown in LCP and are shown in Figure III B.

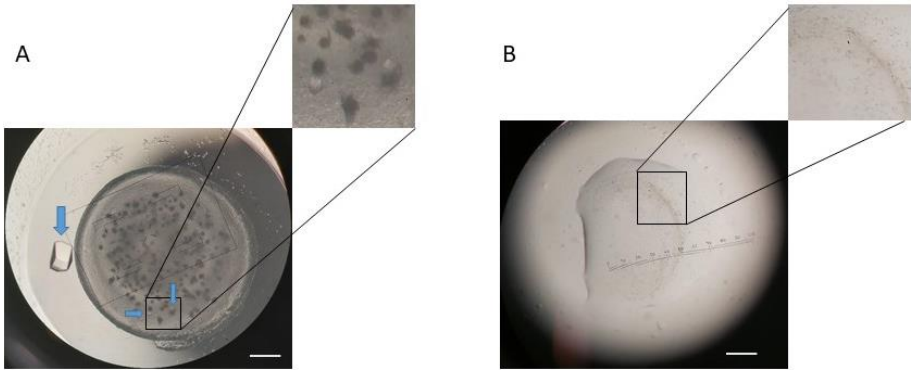


Figure III. Crystallization of protein in the lipidic cubic phase (LCP) with monoolein. A) Crystals of lysozyme, crystallization solution is 0.8 M sodium chloride and 0.1 M sodium acetate pH 4.0 and B) ShuA crystals in 0.1 M sodium chloride, 0.1 M lithium sulfate, 0.1 M sodium citrate and 40% PEG 200 (communication Cristech days 2020, Frank Brachet). Scale bar represents 200 µm.

Appendix IV

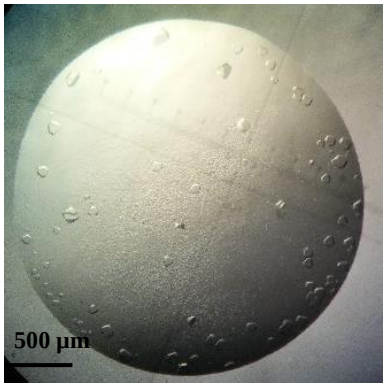
Some of the crystallization conditions used for AcrB

	1	2	3	4	5	6
A	8-10%PEG 3350 100mM $LiSO_4$	8-10%PEG 3350 100mM $MgSO_4$	8-10%PEG 3350 100mM $AmSO_4$	8-10%PEG 3350 100mM $MgCl_2$	8-10%PEG 3350 100mM KSCN	8-10%PEG 3350 100mM NaCl
B	6-16% PEG4K 100mM Naphosphate 50mM NaCl	6-16% PEG4K 100mM Naphosphat 100mM NaCl	16%-20 PEG 400 100mM KSCN 100mM HEPES pH7.5	12-16% PEG 200 50,100mM $MgCl_2$ 100mM HEPES pH7.5	16-20% PEG 200 100mM KSCN 100mM HEPES pH7.5	16-20% PEG 200 200mM KSCN 100mM HEPES pH7.5
C	6-16% PEG4K 50mM ADA 5% GLY 200mM ammonium sulfate	6-16% PEG4K 100mM ADA 5% GLY 200mM ammonium sulfate	6-16% PEG4K 50mM ADA 10% GLY 200mM ammonium sulfate	6-16% PEG4K 50mM ADA 5% GLY 100mM ammonium sulfate	16-20% PEG 400 200mM KSCN 100mM HEPES pH7.5	12-16% PEG 400 50,100mM $MgCl_2$ 100mM HEPES pH7.5

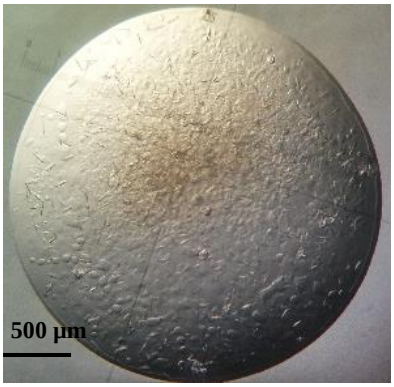
Appendix V

Crystallization condition used for HiLiDe method

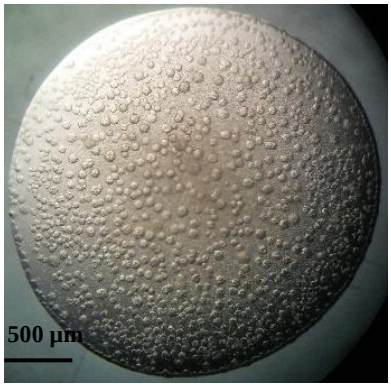
A. AcrB



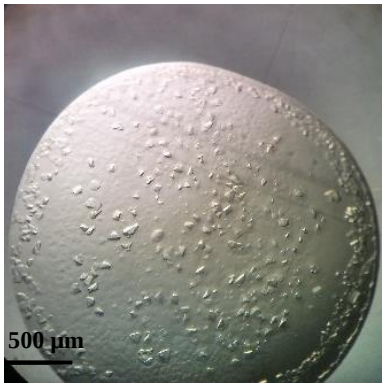
AcrB
100 mM Na citrate
100mM MES pH 6.5
12 % PEG 400
Ratio 2:1



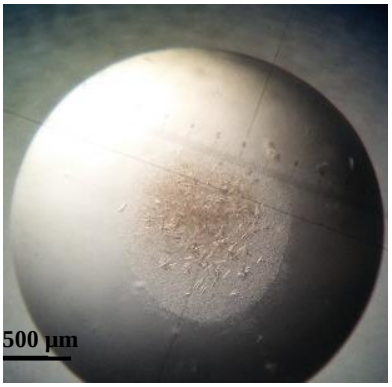
AcrB
100 mM LiCl
100mM MES pH 6.5
12 % PEG 400
Ratio 2:1



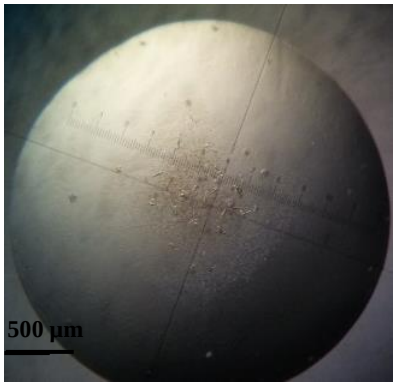
AcrB
100 mM LiCl
100mM MES pH 6.5
14% PEG 400
Ratio 2:1



AcrB
100 mM Na citrate
100mM MES pH 6.5
14% PEG 400
Ratio 2:1

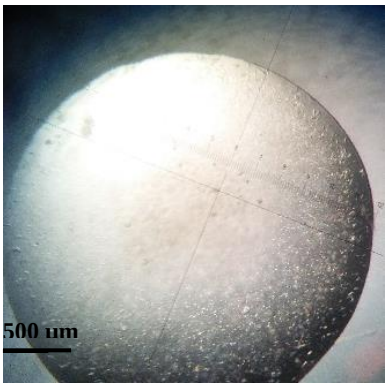


AcrB
200 mM LiCl
100mM MES pH 6.5
10 % PEG 400
Ratio 2:1

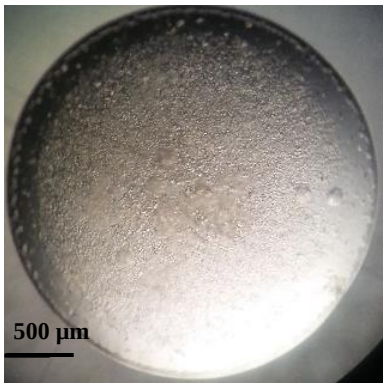


AcrB
200 mM NaCl
100mM MES pH 6.5
12 % PEG 400
Ratio 2:1

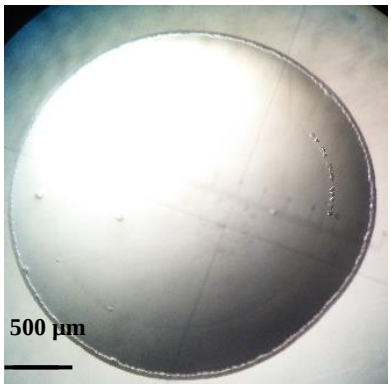
AcrB
200 mM LiCl
100mM MES pH 6.5
12 % PEG 400
Ratio 2:1



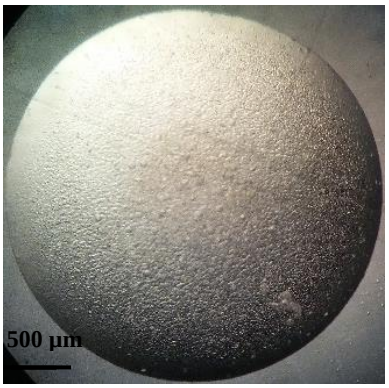
AcrB
200 mM Na acetate
100mM MES pH 6.5
12 % PEG 400
Ratio 2:1



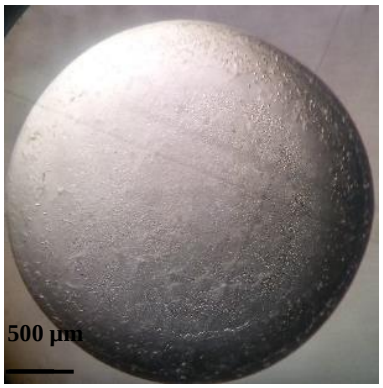
AcrB
200 mM Na acetate
100mM MES pH 6.5
14 % PEG 400
Ratio 2:1



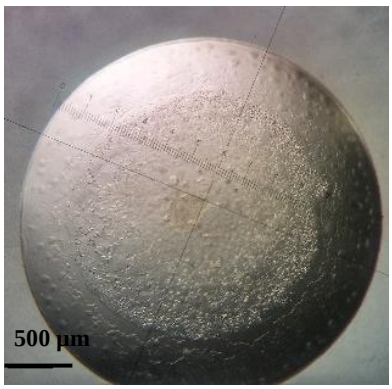
AcrB
200 mM Ammonium sulfate
100mM MES pH 6.5
14 % PEG 400
Ratio 2:1



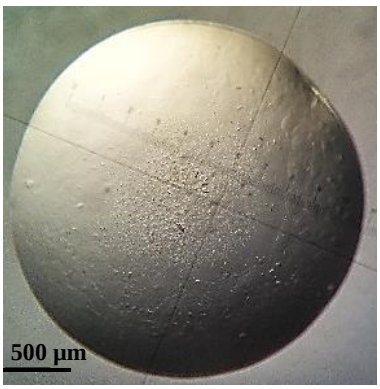
AcrB
200 mM NaCl
100mM MES pH 6.5
14 % PEG 400
Ratio 2:1



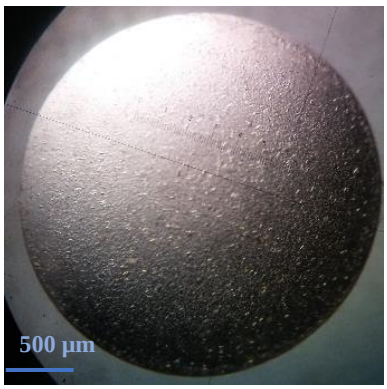
AcrB
200 mM Na citrate
100mM MES pH 6.5
14 % PEG 400
Ratio 2:1



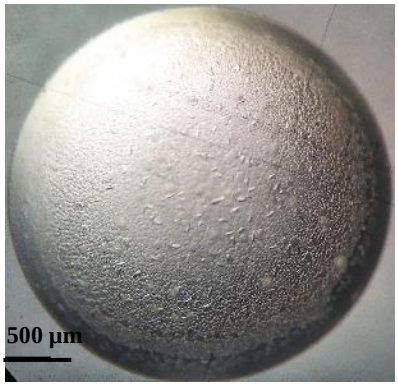
AcrB
200 mM NaCl
100mM MES pH 6.5
8% PEG 550 MME
Ratio 2:1



AcrB
200 mM LiCl
100mM MES pH 6.5
8% PEG 550 MME
Ratio 2:1

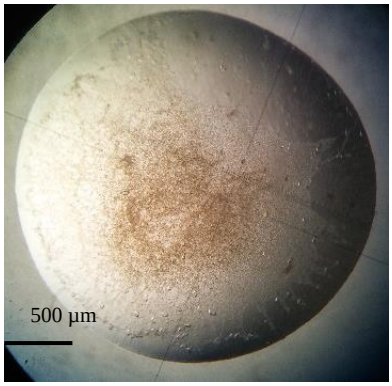


AcrB
200 mM NaCl
100mM MES pH 6.5
12 % PEG 550 MME
Ratio 2:1

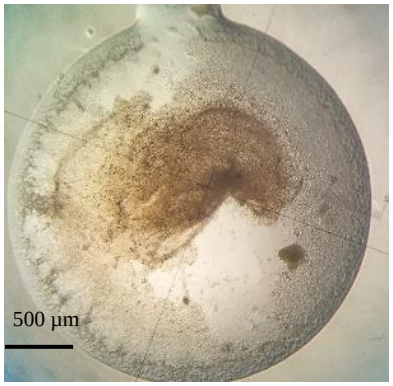


AcrB
200 mM $MgSO_4$
100mM MES pH 6.5
14% PEG 550 MME
Ratio 2:1

B. ShuA

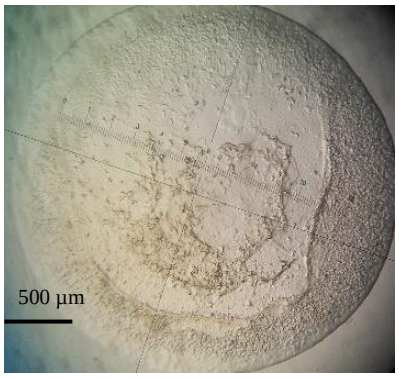


100 mM Na citrate
100mM MES pH 6.5
8% PEG 2000
Ratio 3:1

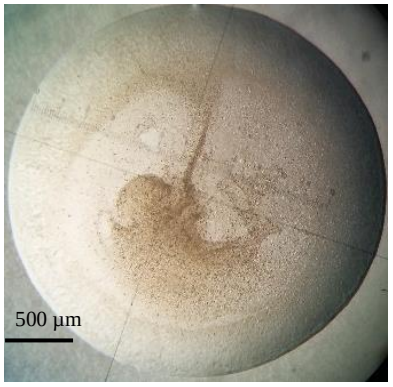


100 mM NaCl
100mM MES pH 6.5
8% PEG 2000
Ratio 2:1

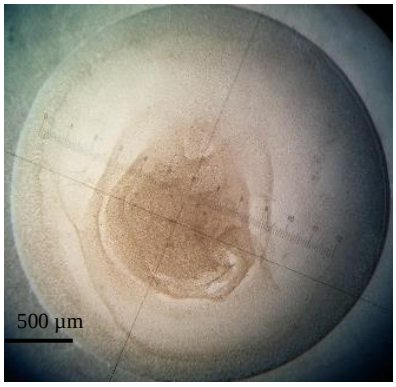
100 mM Ammonium sulfate
100mM MES pH 6.5
8% PEG 2000
Ratio 3:1



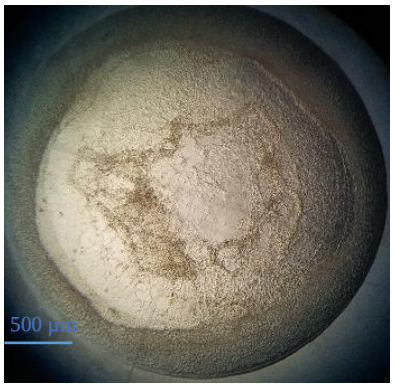
100 mM LiCl
100mM MES pH 6.5
8% PEG 2000
Ratio 3:1



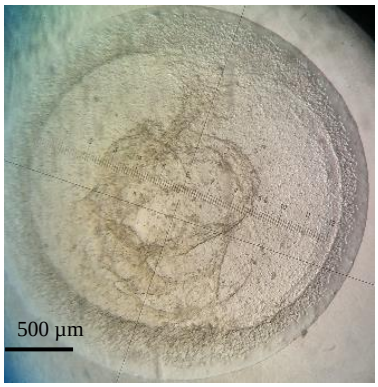
100 mM Na acetate
100mM MES pH 6.5
8% PEG 2000
Ratio 3:1



100 mM Na acetate
100mM MES pH 6.5
8% PEG 2000
Ratio 2:1

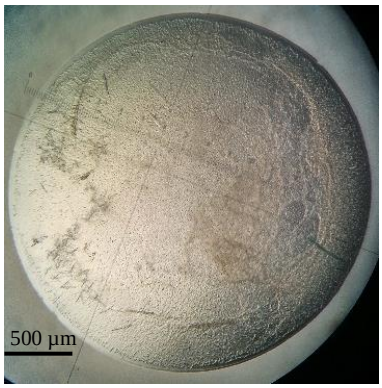


100 mM $MgSO_4$
100mM MES pH 6.5
10% PEG 2000
Ratio 3:1



100 mM Na acetate
100mM MES pH 6.5
10% PEG 2000
Ratio 3:1

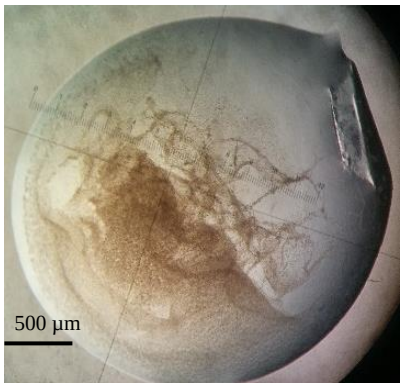
100 mM $MgSO_4$
100mM MES pH 6.5
12% PEG 2000
Ratio 2:1



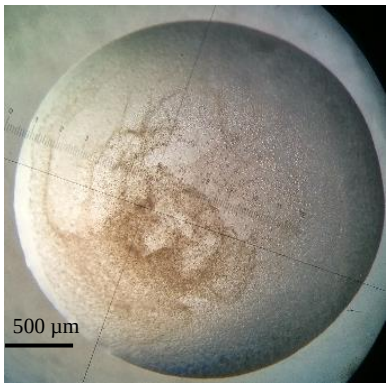
100 mM $MgSO_4$
100mM MES pH 6.5
12% PEG 2000
Ratio 3:1

100 mM Na citrate
100mM MES pH 6.5
12% PEG 2000
Ratio 3:1

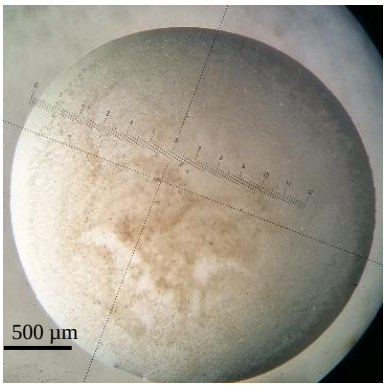
200 mM $MgSO_4$
100mM MES pH 6.5
12 % PEG 2000
Ratio 3:1



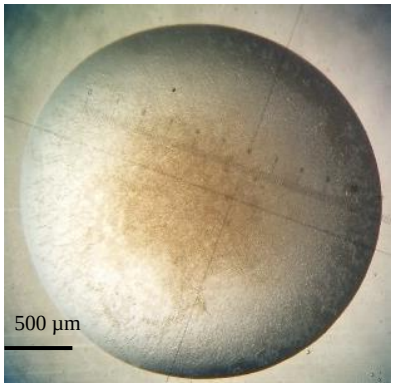
200 mM LiCl
100mM MES pH 6.5
10 % PEG 2000
Ratio 2:1



200 mM LiCl
100mM MES pH 6.5
10 % PEG 2000
Ratio 3:1



200 mM Ammonium sulfate
100mM MES pH 6.5
10 % PEG 2000
Ratio 3:1



200 mM Ammonium sulfate
100mM MES pH 6.5
8 % PEG 2000
Ratio 3:1

200 mM $MgSO_4$
100mM MES pH 6.5
8 % PEG 2000
Ratio 3:1

Appendix VI

Modeling of mass transport of detergent for optimization of membrane protein crystallization

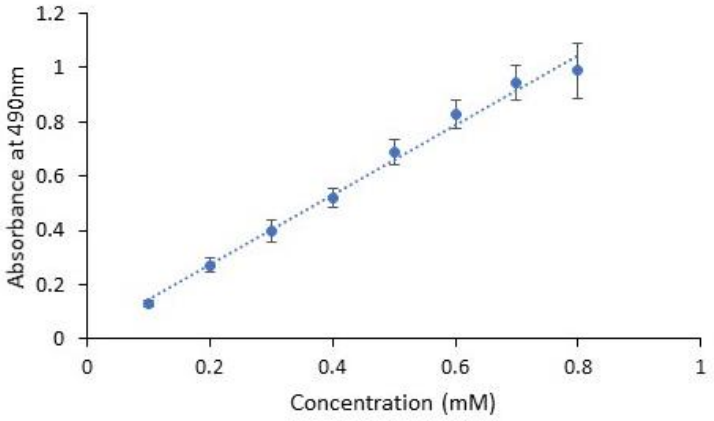
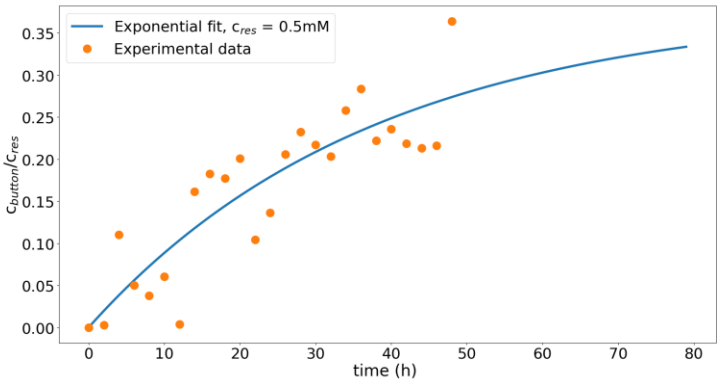


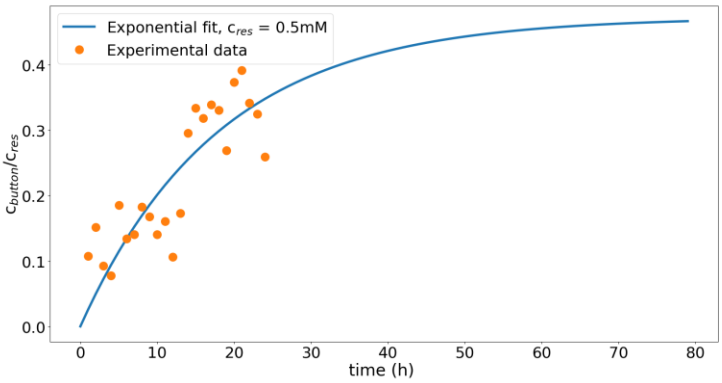
Figure VI.1. DDM calibration curve

Measured detergent concentration inside the dialysis button as a function of the time.

A



B



C

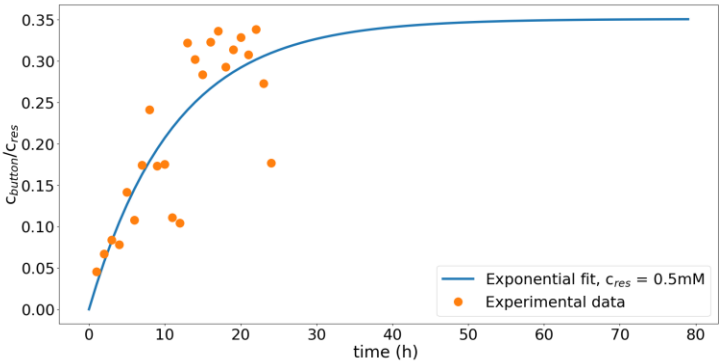


Figure VI.2. Measured DDM concentration inside the dialysis button as a function of the time. A after 48h, B and C after 24h. Dots are experimental data and the curve is the exponential fit. Each dot is the mean value of three measurements from one dialysis button.

Appendix VII

Distance between Cα of the identical residues in both refined and the search model

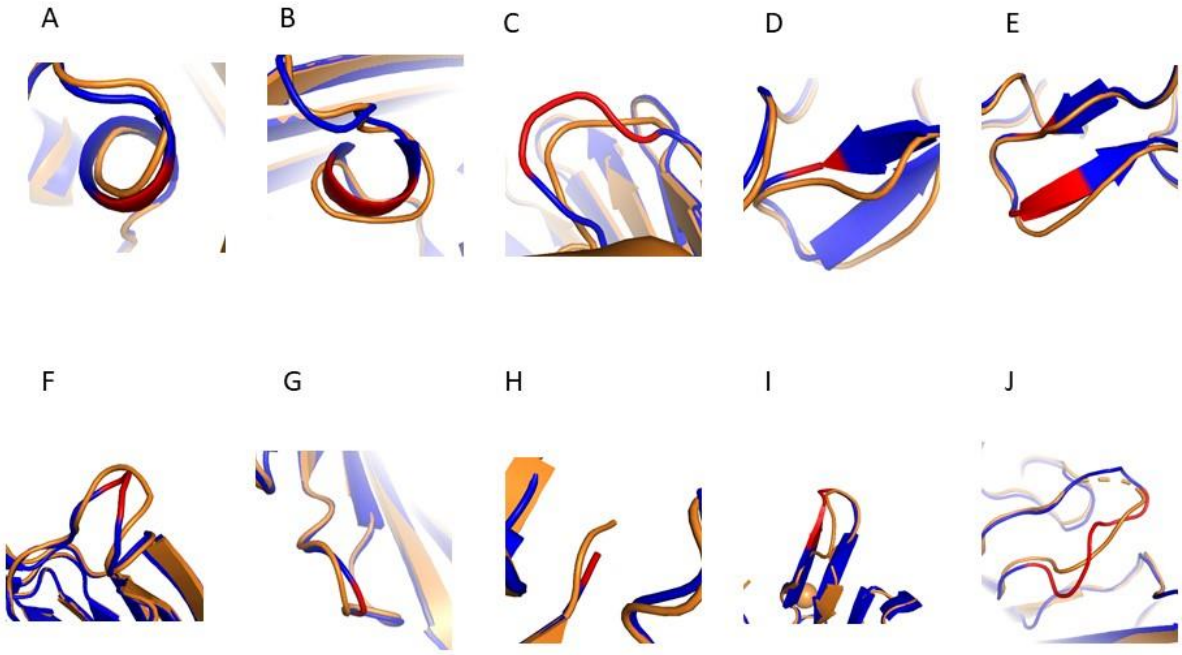


Figure VII. Comparison of the distances between the amino acid Cα in the refined model (blue) and the search model (3FHH) (orange). A to J are related to the 10 regions illustrated in figure 62.

Appendix VIII

Effect of pH on crystallization

Crystallization experiments using AcrB were performed across pH values ranging from 6.0 to 7.5 using the hanging drop vapor diffusion technique.

Firstly, two protein concentrations (5 and 10 mg/mL) and three PEG 400 concentrations (20%, 30% and 40%) with drop ratios 1:1 and 3:2 in presence of MES with pH from 6 to 7 were tested.

Crystals were observed in the presence of 20% PEG 400 and MES pH > 6.5. When 30% PEG was used only samples with 5 mg/mL, drop ratio 3:2 and 10 mg/mL, drop ratio 1:1 at pH 6.0 contained crystals or crystalline precipitate. No crystals grew (only precipitate) at highest PEG 400 concentration (40%) at the pH tested in these series of experiments.

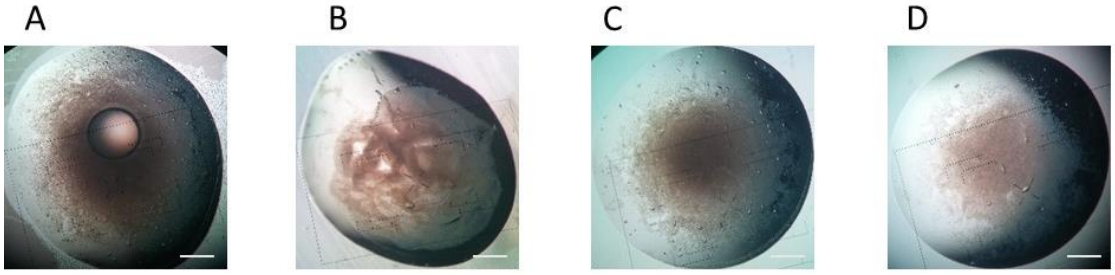


Figure VIII. Crystallization of AcrB with crystallization solution contains 20% PEG 400 and 0.1 M MES pH 6.6 with hanging drop vapor diffusion method. Protein concentration in A and B is 10 mg/mL and C and D is 5 mg/mL. Protein:crystallization solution concentration ratio for A and C is 3:2 and for B and D is 1:1. Scale bar represents 500 µm.

Further crystallization experiments were performed using various buffers, 100 mM HEPES pH 7.5, 100 mM MOPS PH 7.0 and MES pH 6.5, with 0.1 M sodium chloride or 0.1 M lithium sulfate and 12% PEG 4000. In these series of experiments, crystals only formed in samples with HEPES (pH 7.5) and MES (pH 6.5) buffers. As in the previous series of experiments the crystals formed in MES buffer (pH 7.0), rather than the effect of pH, despite its structural similarity to other sulfonic acids, it is the nature of the MOPS and therefore its chemical structure that could be considered here to affect and prevent the crystallization of the AcrB-rifampicin complex.

Appendix IX

Thaumatin crystallization using OptiCrys

Few crystallization tests were carried out with two concentrations of Thaumatin (12 mg/mL and 15 mg/mL) in 0.6 M sodium/potassium tartrate, 0.05 M ADA pH 6.5. The initial temperature was set-up to 20 °C. After 4 hours, crystals formed in the condition with 15 mg/mL protein. These grew to a final size of ~ 50 µm (Figure). In contrast, no crystals were observed within a reasonable time in samples prepared at the lower protein concentration, of 12 mg/mL. The temperature was then varied by decreasing in steps from 18°C to 16°C (or 14°C depending on the sample) but no significant changes (e.g. crystal size or morphology) were observed. It should be noted that we did not replicate or perform extensive studies in the case of thaumatin.

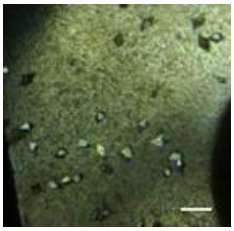


Figure IX. Crystals of thaumatin obtained using the Crystallization Bench (OptiCrys) in 0.6 M Na/K tartrate and 0.5 M ADA pH 6.5 at 20 °C. Scale bar represents 100 µm

Appendix X

Figure permissions:

Figure2: from “Membrane proteins and membrane proteomics” Request ID: 600041801 (accepted)

Figure 3: from “Crystal structure of bacterial multidrug efflux transporter AcrB” license number: 5054130880056

Figures 4: from “3D structure of AcrB: The archetypal multidrug efflux transporter of Escherichia coli likely captures substrates from periplasm” license number 5054131142655

Figure 5: from “Dissecting binding of a β-barrel membrane protein by phage display” license number: 1113658-1

Figure 6: from “Protein crystallization and phase diagrams” license number: 5062510858728

Table 1: from “Introduction to protein crystallization 2004” license number: 5054160405569

Figure7 and 8: from “Introduction to protein crystallization . 2014” license number: 5063601003195

Figure9,10 from “An overview of biological macromolecule crystallization”

No special permission is required to reuse all or part of article published by MDPI, including figures and tables.

Figure 12 from “ Protein crystallization: From purified protein to diffraction-quality crystal” license number 5139330078520

Figure13 from “Phase separation in the isolation and purification of membrane proteins”

Express permission is not required for this purpose, authors are able to re-use or adapt their article for use in their thesis/dissertation, provided a suitable acknowledgement to the original publication is included (Future Science)

Table 2: An overview of the top ten detergents used for membrane protein crystallization (table1, page10)

No special permission is required to reuse all or part of article published by MDPI, including figures and tables.

Figure 15,16: from “Protein Crystallography Methods and Protocols” license number: 5053590174606

Figure17: from “Computational Tools in Protein Crystallography” license number: 5054170877628

Figure 18: from “A methodology and an instrument for the temperature-controlled optimization of crystal growth”

To reprint any figures or tables from IUCr Journals in a scientific publication, prior permission is not required, subject to the following conditions:

- Reproduction is intended in a journal, book, thesis or similar publication.

Figure 19: from “ A crystallization apparatus for temperaturecontrolled flow-cell dialysis with real-time visualization”

From IUCr Journals

Figure 55: from “Optimization of crystallization of biological macromolecules using dialysis combined with temperature control”

From IUCr Journals

List of Publications:

1. Optimisation of Crystal Growth for Neutron Macromolecular Crystallography

Journal of Visualized Experiments 196 (2021)

E. Vahdatahar, N. Junius, and M. Budayova-Spano.

2. Optimization of Crystallization of Biological Macromolecules using Dialysis Combined with Temperature Control

Journal of Applied Crystallography 53 (2020): 1-13

N. Junius, E. Vahdatahar, E Oksanen, J-L. Ferrer and M. Budayova-Spano.

Conference:

1. Comparing crystals of AcrB obtained from different crystallization techniques

E. Vahdatahar, D. Cobessi, J.-L. Ferrer and M. Budayova-Spano.

Presented at “ANF CRISTECH 2020” taking place from 14th to 16th October, 2018 at Écully, France

2. Optimization of the Crystallization Bench for crystallizing the membrane proteins

E. Vahdatahar, N. Junius, D. Cobessi, J.-L. Ferrer and M. Budayova-Spano.

Presented at “Advanced methods in Macromolecular crystallization” course, taking place from 10th to 16th June, 2018 at Nove Hrady, Czech Republic.

JOURNAL OF APPLIED CRYSTALLOGRAPHY
ISSN 1600-5767

Optimization of crystallization of biological macromolecules using dialysis combined with temperature control

Niels Junius,‡§ Elham Vahdatahar,§ Esko Oksanen,¶ Jean-Luc Ferrer and Monika Budayova-Spano*

Univ. Grenoble Alpes, CEA, CNRS, IBS, 38000 Grenoble, France. *Correspondence e-mail: monika.spano@ibs.fr

Received 9 July 2019
Accepted 8 March 2020

Edited by J. M. García-Ruiz, Instituto Andaluz de Ciencias de la Tierra, Granada, Spain

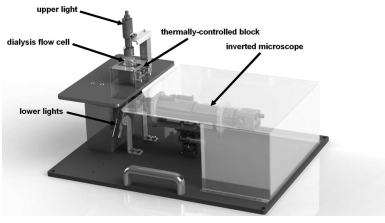
‡ Present address: ELVESYS SAS – 172 Rue de Charonne, 75011 Paris, France.
§ These authors contributed equally to this work.
¶ Present address: European Spallation Source ESS ERIC, Odarslövsvägen 113, 22484, Lund, Sweden.

Keywords: crystal growth optimization; crystal size control; temperature–concentration phase diagrams; dialysis; temperature control; biomacromolecular crystallography.

A rational way to find the appropriate conditions to grow crystal samples for bio-crystallography is to determine the crystallization phase diagram, which allows precise control of the parameters affecting the crystal growth process. First, the nucleation is induced at supersaturated conditions close to the solubility boundary between the nucleation and metastable regions. Then, crystal growth is further achieved in the metastable zone – which is the optimal location for slow and ordered crystal expansion – by modulation of specific physical parameters. Recently, a prototype of an integrated apparatus for the rational optimization of crystal growth by mapping and manipulating temperature–precipitant–concentration phase diagrams has been constructed. Here, it is demonstrated that a thorough knowledge of the phase diagram is vital in any crystallization experiment. The relevance of the selection of the starting position and the kinetic pathway undertaken in controlling most of the final properties of the synthesized crystals is shown. The rational crystallization optimization strategies developed and presented here allow tailoring of crystal size and diffraction quality, significantly reducing the time, effort and amount of expensive protein material required for structure determination.

1. Introduction

Knowledge of the phase diagram is of key importance when designing and controlling a crystallization process (Astier & Veesler, 2008; Vekilov, 2012; Zhang *et al.*, 2014). In the case of macromolecules such as proteins, the availability of accurate phase diagram data is limited owing to the diversity of their structures and the lack of suitable experimental setups to readily perform reliable measurements of protein solubility during crystallization using a small amount of protein sample (Garcia-Ruiz *et al.*, 1999; Olesberg *et al.*, 2000; Curtis *et al.*, 2001; Asherie, 2004; Dumetz *et al.*, 2007; Yin *et al.*, 2008; Y. Zhang *et al.*, 2012). The solubility of a protein depends strongly on the protein–protein interactions as well as on the protein–solvent interactions. Any slight modification of the solution composition can influence the solubility dramatically, or even alter the nature of these macromolecules. Independently of the complexity of protein behavior, the phase transformation is still governed by both the thermodynamics and the kinetics of the system. Therefore, it is still possible to describe all this information in phase diagrams. If crystallization conditions or nucleation points are identified, the information can be plotted in phase diagrams, as represented in a simplified form in Fig. 1(*a*). In this case the information provided relates to both thermodynamics and kinetics. The thermodynamic data are the solubility curves of the different



crystalline and amorphous phases in the phase diagram. They depend on multiple parameters such as temperature, pH, solvent, impurities *etc.* In addition, kinetic trajectories in the phase diagram are relevant to control most of the final properties of the synthesized crystals. The path followed in the diagram controls the nucleation and growth of the crystals, and thus their number, size and morphology.

The success of diffraction experiments in protein crystallography is directly related to the quantity and quality of the recorded data. Considering the high performance of existing X-ray sources, it is now essentially the quality of the samples that limits the quality of the crystallographic data. New and emerging uses result in specific challenges for crystallization of proteins, in which precise control of crystal size is essential. New approaches to serial X-ray and electron crystallography, and to solving structures including time-resolved studies of short-lived intermediates, require small crystals, typically in the 0.2–10 µm size range. Serial crystallographic methods are being increasingly used at synchrotron sources (serial synchrotron crystallography) owing to advances in micro- and nano-focus beamlines, as well as at rapidly developing ultra-bright free-electron laser sources (serial femtosecond crystallography; Chapman, 2015), enabling structural studies of previously intractable proteins. Electron crystallography, traditionally applied to 2D membrane protein crystals, can now solve 3D structures from thin protein crystals and provide charge information not available from X-ray crystallography. This diffraction technique can potentially deliver atomic-resolution structural information with high throughput when suitable crystals (thinner than 200 nm for a typical 200 keV transmission electron microscope) are available (Nannenga & Gonen, 2016). At the other extreme are the requirements of the next-generation flagship neutron sources, such as the European Spallation Source (ESS, Lund). Because neutrons interact very weakly with matter, much larger, and ideally bulky, crystals are needed with volumes of >0.01 mm^3 (*i.e.* 200 µm on a side) for neutron crystallography (Blakeley *et al.*, 2015). This is often the only way to visualize all of the protons in a protein structure, key information for the analysis of interactions required for drug design.

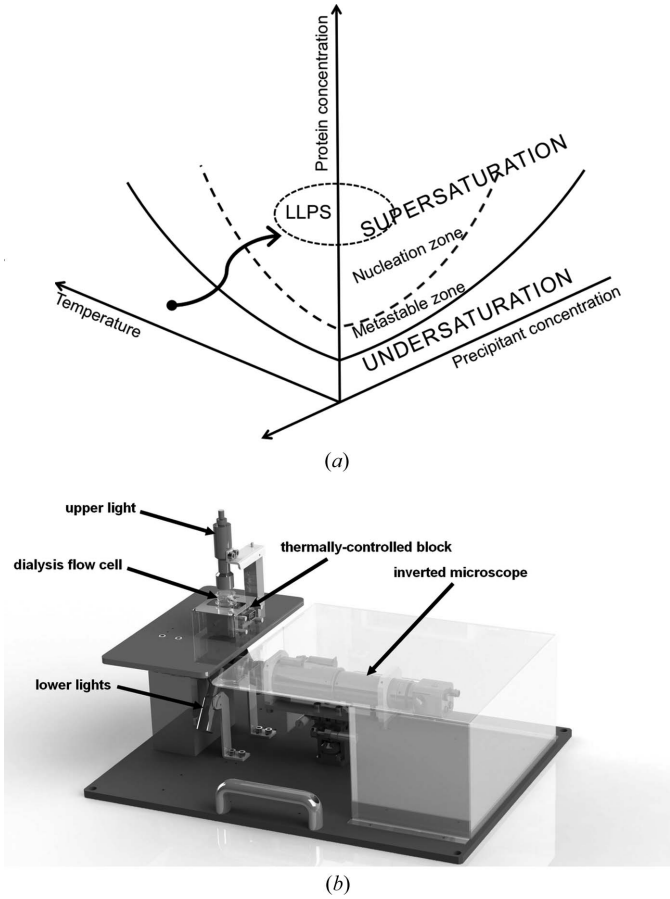


Figure 1
(*a*) Schematic view of a multidimensional phase diagram with two sections shown. The continuous curve is the solubility of the crystal phase as a function of temperature and the concentration of precipitant. The metastable zone lies between the solid and dashed curves, where the solution is supersaturated but nucleation of the crystal is either very slow or absent. Once the supersaturation is high enough, it drives nucleation and hence starts crystallization, represented by the nucleation zone, located next to the metastable zone. Finally, the process of liquid–liquid phase separation (LLPS), which occurs beyond the other zones, can both help and hinder crystallization. The arrow illustrates a specific kinetic pathway followed during crystallization. (*b*) Simplified view of the crystallization apparatus (OptiCrys) for temperature-controlled flow-cell dialysis with real-time visualization.

The need for detailed knowledge of the phase diagram is the basis of the devices (Budayova-Spano *et al.*, 2007; Budayova-Spano, 2010; Junius *et al.*, 2016) that we have developed with a focus on X-ray and neutron macromolecular crystallography. The first-generation instrument combines the use of temperature control and seeding and allows for growth of large crystals in a crystallization batch (Budayova-Spano *et al.*, 2007). A crystallization batch in the metastable zone is seeded with small protein crystals. The seeds are maintained inside this region of the phase diagram for as long as possible by adjusting the temperature each time the crystal-solution equilibrium is achieved. The temperature variations are repeated until crystals of suitable size for diffraction measurement are obtained. A dialysis button (Junius *et al.*, 2016), in addition to a crystallization batch, was later integrated into this instrument. This modification to the existing device (Budayova-Spano *et al.*, 2007) enables performing a temperature-controlled dialysis crystallization experiment.

The second-generation instrument, called the crystallization bench or OptiCrys [Fig. 1(*b*)], was built to allow the automation of the dialysis crystallization process (Junius *et al.* (2016). The instrument has recently been manufactured under license by the company NatX-ray (Saint Martin d'Hères, France). We added concentration control of the precipitant through the construction of a dialysis cell and a reservoir, allowing dialysis in continuous flow against a solution of controlled composition (Budayova-Spano, 2010; Junius *et al.*, 2016). Physico-chemical parameters such as temperature, concentration of crystallizing agents and pH can be controlled over time during the crystallization, so that the state of the substance studied moves along a well defined kinetic trajectory in the phase diagram. The dialysis membrane allows for

Table 1
Summary of some physico-chemical properties of the proteins used and their crystal growth conditions.

Protein system	MW (kDa)	pI	Crystallization condition	Protein concentration (mg ml^{-1})	Membrane MWCO (kDa)	Dialysis volume (µl)	Temperature range (K)
Urate oxidase from *A. flavus*	33.8	Basic 7.5	5%(*w*/*v*) PEG 8000, 100 m*M* NaCl, 50 m*M* Tris–HCl pD 8.5	8	6–8	100	278–298
Chicken egg-white lysozyme	14.3	Basic 11.35	0.75 *M* sodium chloride, 100 m*M* sodium acetate pH 4	30	6–8	45	278–313
Fluorescent protein EosFP from *L. hemprichii*	25.8	Acidic 6.9	2 *M* ammonium sulfate, bicine 100 m*M* pD 8	18	6–8	25	278–295
Lactate dehydrogenase from *T. thermophilus*	32.8	Acidic 5.8	5%(*w*/*v*) PEG 6000, 100 m*M* MES pH 6	2.8	6–8	15	278–293
Human carbonic anhydrase II	30	Acidic 6.8	1.2 *M* sodium citrate, 100 m*M* Tris–HCl pH 8.6	30	6–8	25	278–313
YchB kinase from *A. tumefaciens*	31.8	Acidic 6.1	20%(*w*/*v*) PEG 8000, $MgCl_2$ 5 m*M*, ATP 5 m*M*, sodium citrate pH 6	10	12–14	70	278–295

adjusting the mass transfer through the membrane during the crystallization process. As a result, the gradients created can be controlled and affect the size and quality of generated crystals. Systematic phase diagrams in multidimensional space can be investigated using far less protein material than previously. With this serial approach, we mark a break with the current paradigm of parallel experiments. We demonstrate that established rational crystallization strategies can be beneficial to provide sufficient scattering volumes for neutron studies that require large-volume well ordered single crystals as well as to generate homogeneous populations of uniformly sized protein crystals required for use by other advanced serial diffraction techniques.

To validate the method beyond model systems like chicken egg-white lysozyme we have successfully tested the setup with a number of proteins of which large, single and well diffracting crystals were previously not available or were difficult to obtain. Altogether, in our validation tests with our apparatus, we have used six different proteins in order to demonstrate how to generate controlled-size crystals. Recombinant urate oxidase from *Aspergillus flavus* (Budayova-Spano, Bonneté *et al.*, 2006; Bonneté *et al.*, 2001) catalyses the oxidation of uric acid to allantoin. Neutron structure determination (Oksanen *et al.*, 2014) was based on the large crystals (Oksanen *et al.*, 2009) grown in the temperature-controlled batch setup (Budayova-Spano *et al.*, 2007). The fluorescent protein EosFP from *Lobophyllia hemprichii* (Wiedenmann *et al.*, 2004) undergoes a photoconversion upon irradiation with near-UV light and the crystals are therefore sensitive to light. Large and well diffracting crystals have been difficult to reproduce. Human carbonic anhydrase II catalyzes the reversible hydration of carbon dioxide. Large and well diffracting crystals for neutron diffraction (Budayova-Spano, Fisher *et al.*, 2006) were obtained with the temperature-controlled batch setup (Budayova-Spano *et al.*, 2007). Lactate dehydrogenase from *Thermus thermophilus* (Coquelle *et al.*, 2007) catalyses the last step in anaerobic glycolysis, conversion of pyruvate to lactate. Well diffracting but small crystals have been grown previously. YchB kinase from *Agrobacterium tumefaciens* (Borel *et al.*, in preparation) catalyses the first step of the isoprenoid biosynthetic pathway. It has been difficult to grow large single crystals of this enzyme.

2. Experimental procedures

2.1. Description of the proteins and material used in crystallization

Chicken egg-white lysozyme was purchased from Sigma–Aldrich as a lyophilized powder, dissolved in distilled water and filtered to obtain a solution with a final concentration of about 30 mg ml^{-1}. Recombinant urate oxidase from *A. flavus* expressed in *Saccharomyces cerevisiae* was supplied by Sanofi–Aventis. A purine-type inhibitor (9-methyl uric acid) and the buffers, salts, PEGs and additives used in this study were purchased from Sigma–Aldrich. The protein complex with the inhibitor was prepared and crystallized according to the protocols described previously (Budayova-Spano, Bonneté *et al.*, 2006; Bonneté *et al.*, 2001). Crystallization solutions (Table 1) were prepared in both light and heavy water (Euriso-top, 99.92% D_2O). The pD of the buffers was adjusted with NaOD (Euriso-top, 99% D) and DCl (Euriso-top, 99.8% D) according to the formula pD = pHmeas + 0.3314n + 0.0766n^2, where n = % D_2O (Lumry *et al.*, 1951). The other proteins presented in our study were produced and purified following established protocols (Wiedenmann *et al.*, 2004; Budayova-Spano, Fisher *et al.*, 2006; Coquelle *et al.*, 2007; Borel *et al.*, in preparation). All protein concentrations were measured by UV absorbance at 280 nm (Table 1). All the solutions were filtered through 0.22 µm Millipore filters. In all the crystal growth experiments, the crystallization mixtures were obtained using dialysis techniques (Ducruix & Giégé, 1992). The cellulose membranes used in our experiments were the standard RC membrane Spectra/Por (http://spectrumlabs.com) with molecular weight cut-offs (MWCOs) of 6–8 kDa and 12–14 kDa. Before the start of the experiment, the crystallization mixtures were centrifuged and filtered to remove all solid particles (precipitates, dust or nuclei). Details of the physico-chemical properties and crystal growth conditions of the proteins studied here are summarized in Table 1.

2.2. Crystallization setups

In the semi-automated first-generation version of the instrument, the protein solution is poured into a specially designed stainless steel dialysis chamber with a transparent polycarbonate optical bottom separated from the precipitant solution by a dialysis membrane of the appropriate molecular weight cut-off (Junius *et al.*, 2016). The dialysis membrane is placed over the top of the dialysis chamber containing the sample (a variety of sizes of 25–200 µl are available) and is held in place by a groove in the dialysis button with an elastic ring. A stainless steel well is then placed over the dialysis button and plays the role of a reservoir containing the precipitant (Junius *et al.*, 2016). This dialysis setup is temperature-controlled using Peltier elements, and the variation of the chemical composition of crystallization solution in the reservoir during the experiment is performed manually (Budayova-Spano *et al.*, 2007). In the new, second-generation version of the instrument, the dialysis button is replaced by a new fluidic assembly composed of the dialysis chamber, located on the bottom, with the reservoir chamber on the top, connected to a pumping system functioning as a continuous flow cell (Junius *et al.*, 2016). This setup enables the exchange of the chemical composition in an automated way during the crystallization experiment. The user may adjust the composition of the reservoir solution and hence access different parts of the phase diagram of the molecule to be crystallized. Any combination of precipitant concentration and temperature can be explored in a systematic manner by sampling a continuum of potential crystal-producing conditions without physically perturbing the mother liquor, while the total volume of protein solution remains constant during the entire experiment. The rate of diffusion through the dialysis membrane can be controlled by using membranes with a particular molecular weight cut-off. The flow-cell dialysis setup is inserted into a brass support temperature-controlled using Peltier elements and is incorporated into the microscope table for viewing of the crystallization chamber from below by an inverted microscope [Fig. 1(*b*)]. Illumination is provided by light-emitting diodes [Fig. 1(*b*)]. The operating temperature range of 233–353 K ± 0.1 K is reached using a proportional–integral–derivative electronic temperature controller. The Peltier elements are cooled on one face with a chiller, resulting in improved temperature control, and a circuit of a dry air prevents condensation. The control software is written with *LabVIEW* (http://www.ni.com/labview/) and includes a graphical user interface for visualization and measurement of crystals, image acquisition, processing and storage as well as control of each parameter (temperature control, illumination, pumping the solutions and measuring the concentration of the different constituents of the crystallization solution).

2.3. Crystal growth optimization workflow

In this section we describe the principle of the method of crystallization with temperature changes at a constant concentration of crystallizing agent (Fig. 2) and at constant temperature with variations in concentration of crystallization agent (Fig. 3). In each optimization scheme, the case of the standard kinetic trajectory resulting in the induction of nucleation in the zone of spontaneous nucleation and respective crystal growth in the metastable zone is shown [Figs. 2(*a*) and 3(*a*)]. Since other alternative workflows are possible, the standard kinetic pathway is then supplemented by further variations in temperature (or precipitant concentration) covering the equilibrium arrival steps [Figs. 2(*b*) and 3(*b*)]. Figs. 2(*c*) and 3(*c*), respectively, illustrate the case of the induction of crystal growth of crystals seeded in the metastable zone by controlled temperature variations and concentration of precipitant. This last case is finally completed by additional variations in temperature or concentration of precipitant, covering the stages of arrival at equilibrium [Figs. 2(*d*) and 3(*d*)].

Standard trajectory A1. The process of optimizing crystal growth for a solute – such as a protein – as a function of temperature (T) at a constant precipitant concentration (P) in the case of direct solubility (solubility increases with the temperature), starting from known crystallization conditions, has the following steps (Fig. 2):

(1) The crystallization chamber is filled with protein solution at a first temperature T1 at a solute concentration C1. The reservoir is filled with a precipitant solution of concentration P1, which slowly diffuses into the crystallization chamber through the dialysis membrane [Fig. 2(*a*)].

(2) The temperature is decreased to T2 in order to drive the kinetic trajectory in the phase diagram to the spontaneous nucleation zone [Fig. 2(*a*)] or up to the vicinity of the upper limit of the metastable zone, to induce nucleation.

(3) The temperature is increased to T3, which stops nucleation, and the trajectory leads to the metastable zone of the phase diagram.

(4) The crystals are left to grow at T3 until a first equilibrium point E1 is reached [Fig. 2(*a*)], where the size of the crystals remains constant and the concentration of the solute decreases to the concentration C2.

(5) The temperature is then decreased to T4, which is within the metastable zone of the phase diagram so no new nucleation occurs.

(6) The crystals are left to grow at T4 until a second equilibrium point E2 is reached [Fig. 2(*a*)] where the crystals no longer grow and the solute concentration reaches C3.

(7) Steps 4 to 6 are repeated until crystals of the desired size are obtained.

(8) The crystals are harvested.

When a homogeneous population of small crystals (controlled by the supersaturation level chosen for nucleation) is desired, steps 3–6 are not cycled. During a cooling crystallization the crystal size is governed by the relative rates of nucleation and growth. These in turn are driven thermodynamically by the level of supersaturation. It is worth considering that the rate at which supersaturation is increasing with time (the supersaturation rate) is as important as the supersaturation value. It is also dependent on the diffusion of protein around the crystals and therefore on the mass transport in the crystallization cell (García-Ruiz *et al.*, 2016). At

high supersaturation levels, nucleation tends to dominate, giving rise to a preponderance of smaller crystals. At low supersaturation levels, growth tends to dominate, resulting in fewer but larger crystals.

Other alternative workflows are possible:

Alternative trajectory A2. All the steps previously described as the stages of temperature variation (steps 1, 2, 3 and 5) are repeated as described above, and the stages of arrival at equilibrium (steps 4 and 6) will be drawn here also with a temperature variation [shown in Fig. 2(*b*)]. In this alternative, the dialysis chamber houses a protein solution pre-equilibrated with the crystallization agent at low supersaturations corresponding to the metastable zone of the phase diagram. A small seed of crystal will be placed before the flow-cell dialysis setup is permanently closed. Then the growth of the seed crystal can also be illustrated according to the workflows described above, by inducing the temperature variations in steps 4, 5 and 6 [Fig. 2(*c*)]. It will also be possible to induce additional temperature variations covering the stages of arrival at equilibrium [shown in Fig. 2(*d*)].

Trajectory B1. A similar process can be used to optimize crystal growth at a constant temperature T as a function of a precipitant concentration P (Fig. 3). The steps here are the following:

(1) A protein solution in the crystallization chamber of the flow-cell dialysis setup, at a concentration C1, is driven to the spontaneous nucleation zone or up to the vicinity of the upper limit of the metastable zone of the phase diagram by a precipitant solution of concentration P1 in the reservoir, thus inducing nucleation [Fig. 3(*a*)].

(2) The concentration of the precipitant is lowered to P2, to stop nucleation and direct the system into the metastable zone of the phase diagram.

(3) The crystals are allowed to grow at P2 until a first equilibrium point E1 is reached [Fig. 3(*a*)], where the crystals no longer grow and the solute concentration is C2.

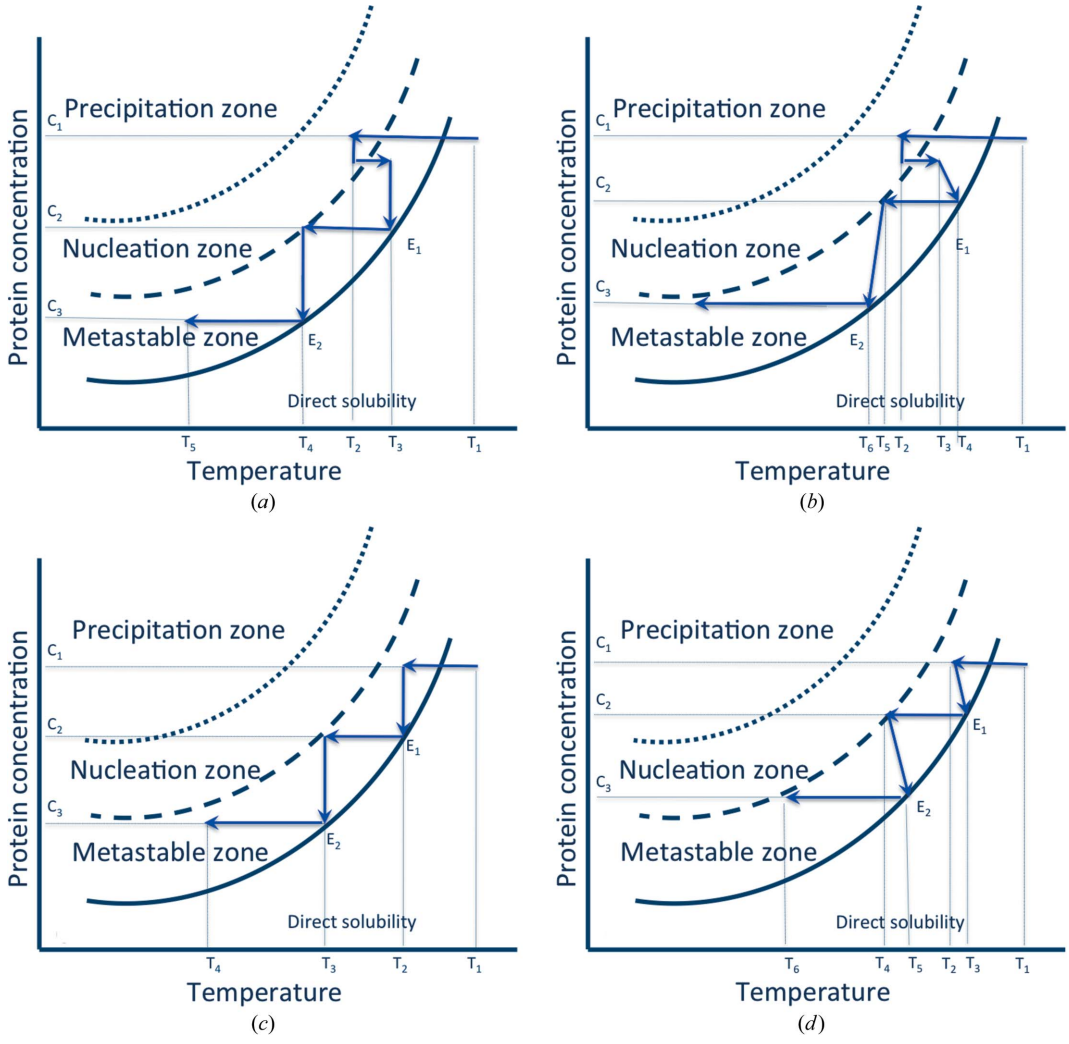


Figure 2
Principle of the method of crystallization with temperature changes at a constant concentration of crystallizing agent. Kinetic pathways schematizing the standard kinetic trajectory A1, resulting in the induction of nucleation in the zone of the spontaneous nucleation and respective crystal growth in the metastable zone due to the controlled temperature variations (*a*), and other alternative workflows A2, showing additional temperature variations covering the stages of arrival at equilibrium in workflow A1 (*b*), and induction of the crystal growth of seeded crystals in the metastable zone as a result of the controlled temperature variations (*c*), supplemented by additional temperature variations covering the stages of arrival at equilibrium (*d*).

(4) The precipitant concentration is increased to P3, still within the metastable zone so no further nucleation occurs.

(5) The crystals are allowed to grow at P3 until a second equilibrium point E2 is reached [Fig. 3(*a*)], where the crystals no longer grow and the solute concentration is C3.

(6) Steps 3 to 5 are repeated until crystals of the desired size are obtained.

(7) The crystals are harvested.

As in the previous case, when a homogeneous population of small crystals (controlled by the level of supersaturation chosen for nucleation) is desired, steps 2–5 are not cycled. During crystallization by increasing the ionic strength, the crystal size is governed by the relative rates of nucleation and growth and these are in turn controlled thermodynamically by the level of supersaturation. However, the supersaturation rate is as important as the supersaturation value and is also dependent on the diffusion of protein molecules around the crystals and therefore on the mass transport in the crystallization cell (García-Ruiz *et al.*, 2016). At high levels of supersaturation, nucleation will dominate, giving rise to a preponderance of smaller crystals. At low levels of supersaturation, growth will dominate, giving rise to larger and fewer crystals.

As in the previous case, other basic alternative workflows are possible:

Alternative trajectory B2. All the steps previously described as the stages of precipitant concentration variation (steps 1, 2 and 4) are repeated as described above and the stages of arrival at equilibrium (steps 3 and 5) will be drawn here also with a supplemental variation [represented in Fig. 3(*b*)]. In this alternative, the dialysis chamber houses a protein solution pre-equilibrated with the crystallization agent at low supersaturations corresponding to the metastable zone of the phase diagram. A small seed of crystal will be placed before the flow-cell dialysis setup is permanently closed. Then the growth of the seed crystal can also be illustrated according to the workflows described above, by inducing the concentration variations of a precipitant in steps 1, 2 and 4 [Fig. 3(*c*)]. It is

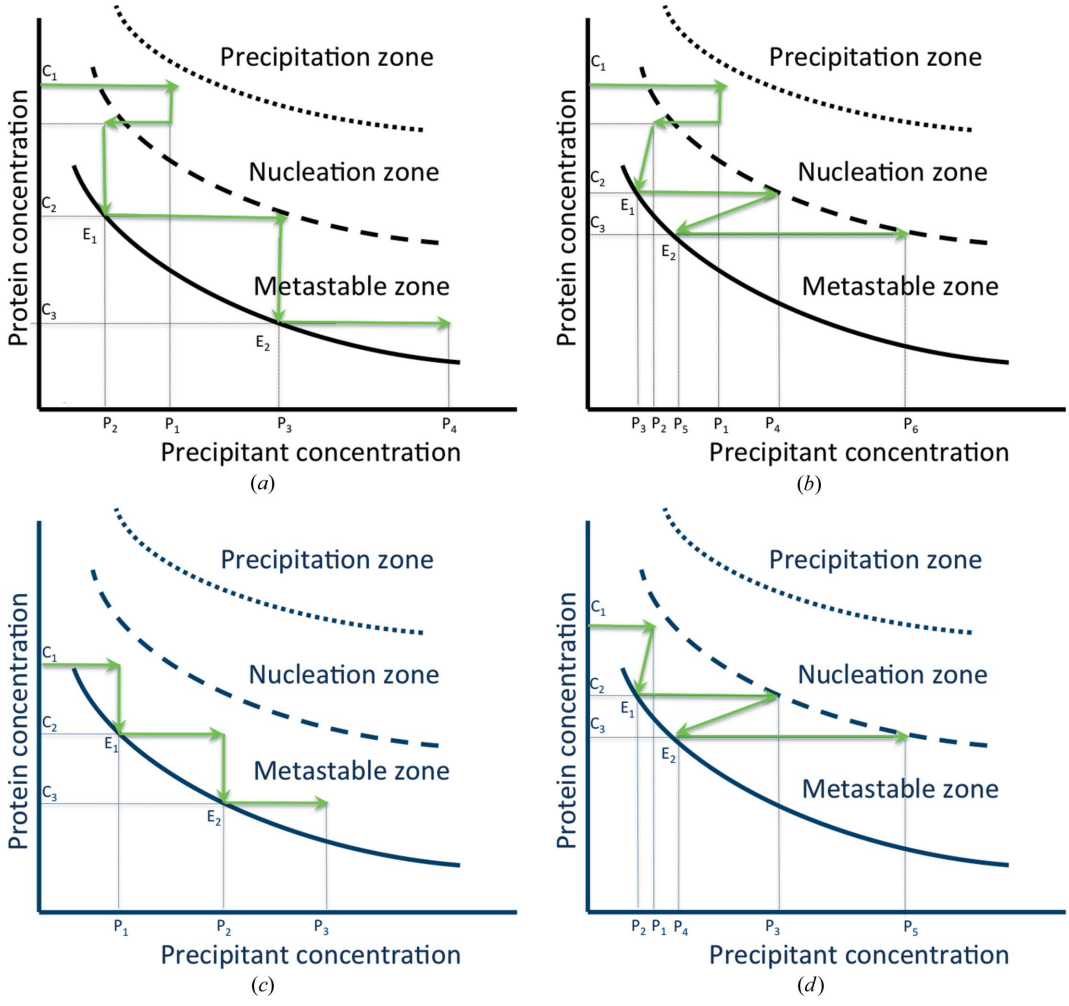


Figure 3
Principle of the method of crystallization at constant temperature with variations in concentration of crystallization agent. Kinetic pathways schematizing the standard kinetic trajectory B1, resulting in the induction of nucleation in the zone of spontaneous nucleation and respective crystal growth in the metastable zone due to the controlled variations of the precipitant concentration (*a*), and other alternative workflows B2, showing additional precipitant concentration variations covering the stages of arrival at equilibrium in workflow B1 (*b*), and induction of the crystal growth of seeded crystals in the metastable zone as a result of the controlled precipitant concentration variations (*c*), supplemented by additional precipitant concentration variations covering the stages of arrival at equilibrium (*d*).

also possible to induce additional concentration variations of a precipitant covering the stages of arrival at equilibrium [shown in Fig. 3(*d*)].

Variations of these basic workflows can be easily envisaged, for example changing both temperature and precipitant concentration either sequentially or simultaneously. In this case the kinetic trajectory in the phase diagram is best described in three dimensions. To control the growth of ordered crystals while avoiding additional nucleation we search for conditions approaching the limit of the metastable zone of the three-dimensional phase diagram. The variation in the size of the crystals is measured, for example by taking photographs of the crystals approximately every 20–30 min when the growth of the crystal begins, and then every 2–3 h towards the end. The pictures are then processed by means of an image analysis software package. The approach described above might be spread over a time period of up to two months, in particular for large crystal volumes for neutron protein crystallography. For X-ray protein crystallography, a time period from 1 to 2 weeks is typically sufficient. The selection of the temperatures during these steps depends on the nature of the molecules to be crystallized and on their thermal stability. The optimal ranges of the temperature for the different protein systems studied here are summarized in Table 1.

3. Results and discussion

Various empirical approaches, based on screening and optimization, have been proposed to generate crystals using vapor diffusion, batch crystallization, dialysis, seeding, free-interface (or counter) diffusion and temperature-induced crystallization. Some of these methods make use of high-throughput automated instrumentation and miniaturization of crystallization experiments and have huge impacts on protein crystallization in terms of saving time and conserving precious sample [*e.g.* Microcapillary Protein Crystallization System (Gerdts *et al.*, 2008); Fluidigm Corporation TOPAZ system (Segelke, 2005); *in meso* crystallization robot (Cherezov *et al.*, 2004); automated microseed matrix screening (D'Arcy *et al.*, 2007); microlytic Crystal Former (Stojanoff *et al.*, 2011)]. Several approaches have also been developed to automate crystal detection from the imaged drops (Echalier *et al.*, 2004; Forsythe *et al.*, 2006; Groves *et al.*, 2007) and others to simplify the identification of crystallization hits (Judge *et al.*, 2005; Dierks *et al.*, 2008; Kissick *et al.*, 2011).

3.1. Crystallization optimization experiments

The first crystallization experiments presented here were performed using the temperature-controlled dialysis buttons that were incorporated into the crystal growth apparatus of the first-generation instrument (Budayova-Spano *et al.*, 2007).

In agreement with the alternative workflow A2 (Section 2.3), large crystals of the recombinant urate oxidas (Uox) from *A. flavus* in complex with 9-methyl uric acid (9MUA) were grown from a few crystals (Fig. 4) seeded in the pre-equilibrated mother liquor [protein solution containing 1–2%(*w*/*v*) of PEG 8K] before final equilibration to 5%(*w*/*v*) was reached in the dialysis chamber (Table 1). After 3 days at 293 K, many small crystals grew next to the few larger crystals of the previously seeded Uox–9MUA complex [Fig. 4(*a*)]. To reduce the number of crystals in the dialysis button, the temperature was then raised to 298 K for 10 min to dissolve the excess of small crystals [Fig. 4(*b*)]. It was subsequently reduced to 293 K and to 291 K to promote the growth of 2–3 selected crystals of interest, including the selected crystal shown in the remainder of the sequence in Fig. 4. Finally, large crystals of the Uox–9MUA complex were obtained by following the kinetic pathway illustrated in the schematic equilibrium phase diagram represented in Fig. 4(*k*).

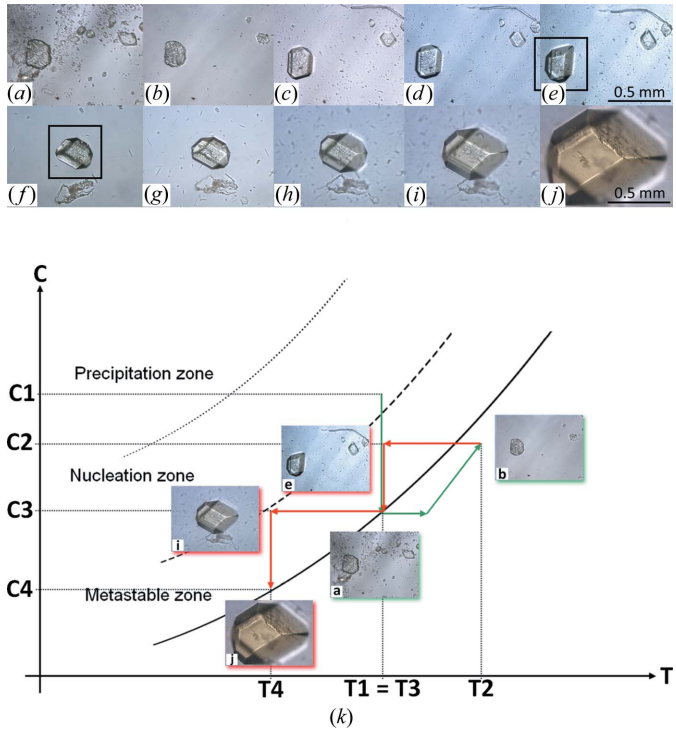


Figure 4
Dialysis experiment with variation of the temperature in the case of growth of large crystals of the urate oxidase complex with 9-methyl uric acid. (*a*) Situation at the end of the third day. Once the concentration of the crystallization agent inside the dialysis button at T1 = 293 K has reached that of the reservoir solution, the formation of many small crystals of the Uox–9MUA complex is observed next to a few seeded crystals in the dialysis button. The temperature is increased to T2 = 298 K. (*b*) Situation after 10 min from dissolution of small crystals in excess at T2 = 298 K. Growth of a few crystals observed at T3 = 293 K during the first day (*c*) and 2 days after dissolution (*d*). (*e*) Growth of a few crystals and of one selected seeded crystal observed at T3 = 293 K 5 days after dissolution. Growth of the selected seeded crystal observed (*f*) 5 days, (*g*) 7 days and (*h*) 9 days at 293 K and (*i*) 12 days and (*j*) 17 days at T4 = 291 K after dissolution. (*k*) Schematic phase diagram (protein concentration versus temperature) incorporating some selected images (to be tracked in alphabetical order) to illustrate the optimization workflow in growing large Uox–9MUA complex crystals. In accordance with alternative workflow A2 (Section 2.3), C1 represents the initial protein concentration used in the crystallization experiment (8 mg ml^{-1}). C2 is weaker than the initial protein concentration, as it results from crystal growth in the metastable zone. C3 and C4 are then protein concentrations corresponding to the relative equilibrium points reached.

With fluorescent protein EosFP we demonstrated that temperature variation can be successfully used to induce nucleation resulting from the process of (metastable) liquid–liquid phase separation (LLPS) (Fig. 5). After 2 h at 293 K, 2 *M* ammonium sulfate (Table 1) had completely diffused into the dialysis button and the demixing of two liquids (protein-rich and protein-poor liquid phases) was observed [Fig. 5(*a*)]. The temperature was then raised to 295.5 K. Photographs (*b*), (*c*), (*d*) and (*e*) in Fig. 5 show the transformation of the dense liquid droplets of the protein-rich phase during the first 2 min that followed the increase in temperature in the crystallization chamber. The dense liquid droplets dissolved and after 6 h the first crystal appeared [Fig. 5(*f*)]. The sequence of photographs [Figs. 5(*g*), 5(*h*), 5(*i*) and 5(*j*)] as well as the corresponding schematic conceptual construct of the phase diagram [Fig. 5(*k*)] show the crystal growth at 295.5 K resulting from the LLPS process.

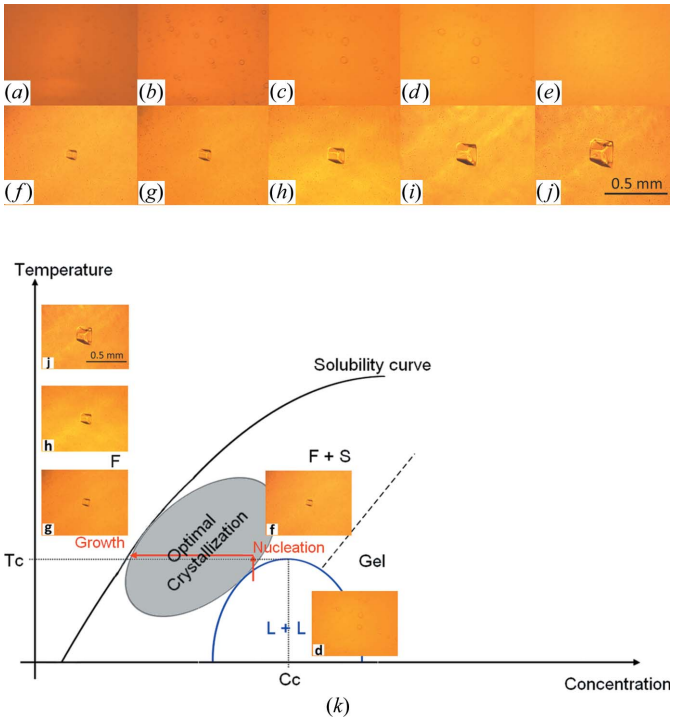


Figure 5
The dialysis experiment with variation of the temperature in the case of crystal growth of EosFP from the process of LLPS. (*a*) Approximately 2 h after injection of the protein into the dialysis button at 293 K, the formation of the dense liquid droplets of the protein-rich phase emerging from the LLPS in the dialysis is observed. The temperature is increased to 295.5 K. Dissolution of the dense liquid droplets at 295.5 K (*b*) 30 s later, (*c*) 1 min later, (*d*) 90 s later and (*e*) 2 min later. (*f*) Approximately 6 h later at 295.5 K, the appearance of the first crystal formed as a result of the LLPS process is observed. Growth of the formed crystal at 295.5 K (*g*) 12 h later, (*h*) 36 h later, (*i*) 3 days later and (*j*) and 6 days later. (*k*) Schematic, temperature versus protein concentration, conceptual construct of the phase diagram incorporating some selected images (to be tracked in alphabetical order), showing the optimization workflow in growing EosFP crystals from the LLPS process. Various zones are shown: areas of the fluid (protein solution or F), coexistence of the protein-rich and protein-poor liquid phases (LLPS or L + L), coexistence of the crystal (S) and the protein solution (F + S), possible gelation (Gel), and optimum condition for protein crystallization (Optimal Crystallization). T_C–C_C represents the critical point beyond which LLPS is not longer possible.

Two other proteins that crystallized successfully using the temperature-controlled dialysis buttons were human carbonic anhydrase II (hCA II) and YchB kinase from *A. tumefaciens* (Table 1). The crystals were obtained in about 1 week (Fig. 6). In the case of hCA II – by following the standard workflow B1 described in Section 2.3 – the concentration of ammonium sulfate was varied by dialysis from 0.8 to 1.2 *M* every 2 days with increments of 0.2 *M* at a constant temperature of 278 K [Fig. 6(*a*)]. In the case of YchB, the situation was the opposite since the standard workflow A1 described in Section 2.3 was used. In this case, the concentration of the precipitant PEG 8000 was kept constant, at 20%(*w*/*v*), and the temperature was changed from 295 to 293 K at the end of the fourth day, 2 days after the equilibrium had been reached in the dialysis button [Fig. 6(*b*)].

Other examples of crystallization trials presented here were carried out with the prototype temperature-controlled dialysis flow-cell setup (second-generation instrument) to demonstrate the control of nucleation and crystal growth. The first example (Fig. 7) is a proof-of-principle experiment with chicken egg-white lysozyme (Table 1), demonstrated in our previous work (Junius *et al.*, 2016) and completed here with one extra experiment. As described previously, protein solution was placed in the crystallization chamber of the temperature-controlled dialysis flow cell. Fig. 7(*a*) reproduces the controlled process demonstrated for large-crystal growth of a single lysozyme crystal (Junius *et al.*, 2016). The pictures incorporated in the qualitative crystallization phase diagram show the crystal habit and volume of a nucleated crystal observed before equilibration at 295 K (after 3 days), at 291 K

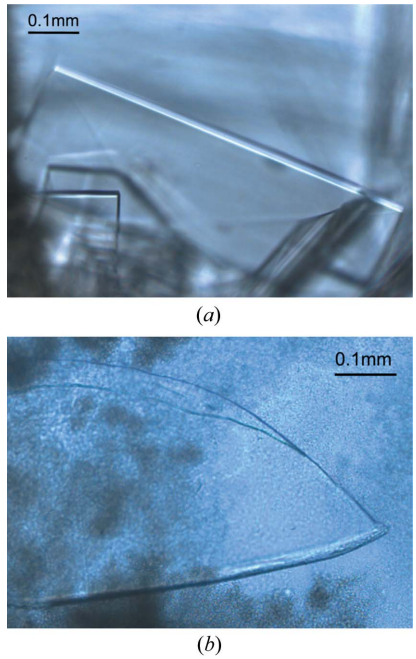


Figure 6
Crystals of human carbonic anhydrase II (*a*) and YchB kinase from *A. tumefaciens* (*b*) obtained in about 1 week in the temperature-controlled dialysis button with volumes of 25 and 70 µl, respectively.

(after 1 day), at 288 K (after 4 days) and at 285 K (after 2 days). Ten days after the start of the experiment, the overall crystal growth process was complete, leading to a large single crystal of the volume typically needed in neutron protein crystallography when a protein is perdeuterated (around 0.1 mm^3). Next, Figs. 7(*b*) and 7(*c*) demonstrate the reversibility of the dialysis experiments for nucleation, crystal growth, dissolution and re-growth, where fewer but larger lysozyme crystals resulted from temperature and precipitant concentration variations, respectively (Junius *et al.*, 2016). This is schematically illustrated from our previous experiment with the help of qualitative crystallization phase diagrams that incorporate pictures showing the crystal habit and size of nucleated crystals in first and second nucleation events. In these two cases, the lysozyme crystals obtained during the first nucleation event are more numerous and smaller [photographs surrounded by red in Figs. 7(*b*) and 7(*c*)] than those obtained during the second nucleation event [photographs surrounded by green in Figs. 7(*b*) and 7(*c*)] induced by the variation of the temperature [Fig. 7(*b*)] or the concentration of precipitant [Fig. 7(*c*)], the latter being visibly fewer and larger. Finally, the last experiment with lysozyme [Fig. 7(*d*)], carried out under the conditions detailed in Table 1 and launched at 291 K, illustrates the opposite case: nucleation, crystal growth, dissolution, and re-nucleation and growth of a large number of very small crystals of lysozyme. After 120 min (the estimated time for diffusion of NaCl to the crystallization chamber is 90 min) the first crystals appeared, and a few days later, near to the equilibrium at 291 K, crystals with a size of around 50 µm had grown in the dialysis chamber. Then the temperature was increased to 308 K in order to dissolve all crystals [photographs surrounded by red in Fig. 7(*d*)]. After more than 24 h at 308 K, when dissolution was almost complete, we again lowered the temperature to the initial value at 291 K. This time, after a few hours, a large number of crystals with very small volume, with size less than 10 µm, appeared at the bottom of the crystallization chamber [photograph surrounded by green in Fig. 7(*d*)].

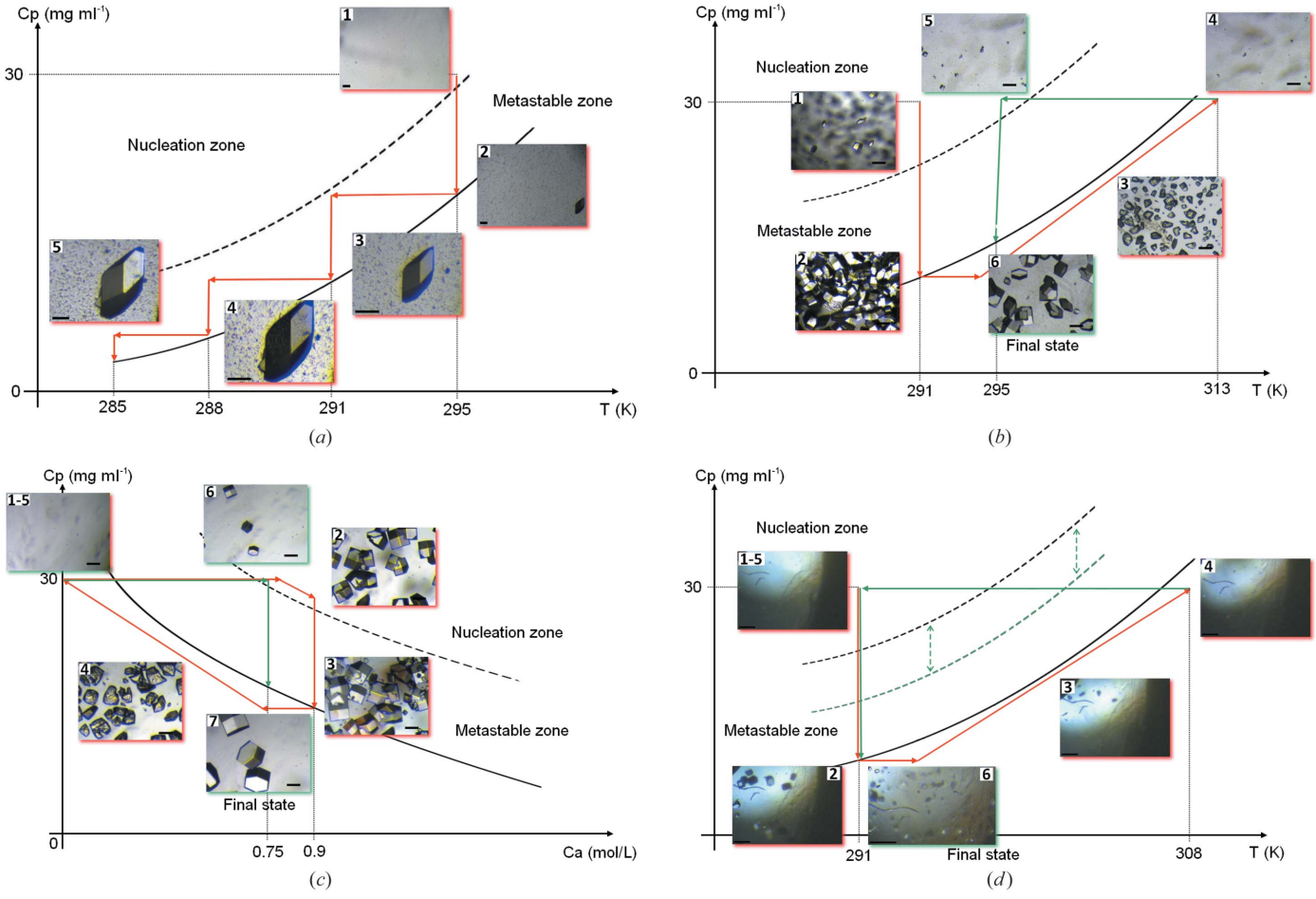
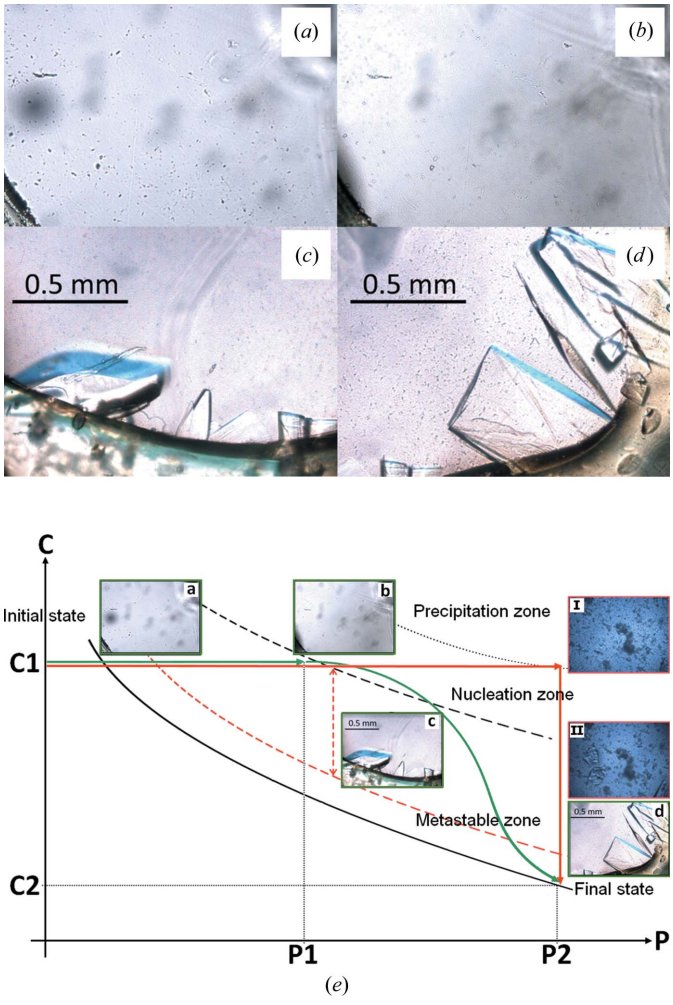


Figure 7
Schematic phase diagrams (protein concentration versus temperature or precipitant concentration) incorporating some selected images (to be tracked in ascending order) and illustrating the crystallization optimization workflow in growing chicken egg-white lysozyme crystals: (*a*) Generating the large single lysozyme crystal obtained at constant chemical composition using the control of the temperature variations. (*b*) Generating uniformly sized lysozyme crystals at constant chemical compositions using the control of the temperature variations. (*c*) Generating uniformly sized lysozyme crystals at constant temperature using the control of the concentration of crystallizing agent. (*d*) Generating a large number of small uniformly sized lysozyme crystals obtained at constant chemical composition using the control of the temperature variations. The shrinkage of the metastable zone by increasing the rate generating the supersaturation is shown schematically. (Scale on photographs represents 100 µm.)

The last example presented here is lactate dehydrogenase from the hyperthermophilic bacterium *T. thermophilus*. In this case, the dialysis experiment was carried out at constant temperature, $T = 293$ K, but with a variation in composition of the crystallization agent. Protein solution at a concentration of *ca* 2.8 mg ml^{-1} was placed in the crystallization chamber of the temperature-controlled dialysis flow-cell setup. The composition of the reservoir solution was varied from 2.5 to 5%(*w*/*v*) of PEG 6000 in 100 m*M* MES buffer at pH 6 with an increment of 2.5%(*w*/*v*) of PEG (Fig. 8). Initially [Fig. 8(*a*), time 0] the PEG concentration inside the crystallization chamber was 0%. Three days later, the concentration of PEG inside the dialysis button reached that of the reservoir at 2.5%. No crystals were observed at this point [Fig. 8(*b*)]. The PEG concentration in the reservoir was then increased to 5%(*w*/*v*) and 2 days later large crystals appeared [Fig. 8(*c*) and 8(*d*)].



(*e*)

Figure 8
The dialysis crystallization experiment with lactate dehydrogenase from *T. thermophiles* carried out at 293 K. (*a*) The beginning of the crystallization experiment (time 0). (*b*) Three days later, when the PEG 6000 concentration in the crystallization chamber is P1 = 2.5%(*w*/*v*). (*c*), (*d*) Two days after the PEG 6000 concentration was increased to P2 = 5%(*w*/*v*), crystals of the enzyme are observed and grow to large volume. (I) Situation observed in the comparative dialysis experiment carried out directly at P2 = 5% once the concentration difference between the compartments has been reached and (II) a few days later at equilibrium. (*e*) Schematic phase diagram (protein concentration versus precipitant concentration) incorporating selected images (to be tracked in ascending or alphabetical order) and illustrating the crystallization optimization workflow in growing large lactate dehydrogenase crystals. The enlargement of the metastable zone by reducing the rate generating the supersaturation is shown schematically. In accordance with a variant (not shown in Fig. 4) of the standard workflow B1 (Section 2.3), C1 represents the initial protein concentration used in the crystallization experiment (2.8 mg ml^{-1}) at P1 and C2 is the protein concentration corresponding to the relative equilibrium point reached at P2.

In order to rationalize these experiments, we performed some comparative dialysis experiments with human carbonic anhydrase II and lactate dehydrogenase from *T. thermophilus*. The same batches of proteins were used (30 mg ml^{-1} in the case of hCA II and 2.8 mg ml^{-1} in the case of lactate dehydrogenase). The main difference between these experiments and those presented previously is the fact that here we do not impose an extra precipitant concentration or a temperature gradient. The precipitant concentration in the reservoir was kept constant and corresponds to the final supersaturation where the nucleation took place in previous experiments [1.2 *M* ammonium sulfate, 100 m*M* Tris–HCl pH 8.6 in the case of hCA II and 5%(*w*/*v*) PEG 6000, 100 m*M* MES pH 6 in the case of lactate dehydrogenase], and the temperature was maintained at 293 K. After a few days, a crystalline precipitate was observed at the bottom of the dialysis chamber in the case of lactate dehydrogenase, as shown in the pictures I and II, surrounded in red, in Fig. 8(*e*), illustrating a schematic kinetic pathway imposed during the crystallization process in the corresponding phase diagram. In the case of hCA II a huge number of small crystal clusters grew from such conditions. Finally, another interesting comparison concerns crystallization experiments conducted under the same crystallization conditions with YchB kinase from *A. tumefaciens* using the temperature-controlled dialysis button with temperature variation and traditional vapor diffusion techniques at constant temperature. We were unable to obtain single crystals of YchB kinase using the traditional vapor diffusion techniques. Only large twinned crystals or crystal clusters could be grown by vapor diffusion, whereas the temperature-controlled dialysis method yielded large single crystals. The bound adenosine triphosphate molecules were also clearly visible in the electron density map, unlike the crystals grown with vapor diffusion (Borel *et al.*, in preparation).

3.2. Temperature effect

Temperature is recognized as a non-invasive control parameter for protein crystallization. Before temperature-induced crystallization can be routinely used as a method of preparation of protein crystals, qualitative data on the temperature-dependent solubility of the protein should be obtained. However, it is only available for a limited number of proteins, because of the labor-intensive techniques required for the experimental determination of solubility curves. Knowledge of the phase diagram and the specific control of the crystallization parameters such as the temperature and the concentration of crystallization agents and/or additives as described

here allow the number of crystals and their macroscopic defects to be reduced, as well as selection of the nucleation and/or growth of the desired phase. Even when the precise position of the solubility curve of the phase diagram is not experimentally known, the ability to control the crystallization parameters in a reversible manner together with real-time observation of the crystals allows the phase diagram to be explored in a qualitative way.

Temperature or precipitant gradients can be used to precisely and reversibly control the relative supersaturation levels of protein solutions. Temperature controls the balance between enthalpy and entropy effects on free energy, which are comparable in magnitude. Depending on whether crystallization is enthalpy driven or entropy driven, proteins become either more soluble at higher temperatures (direct solubility) or less soluble at higher temperatures (reverse solubility) (Budayova-Spano *et al.*, 2007; Oksanen *et al.*, 2009). The temperature influence is due to variation of the acid/base constants of the protein side chains (Chernov & Komatsu, 1995). In addition, the effective pK_a values of the ionizable groups are related to the ionic strength of the medium. As a result, the solubility increases with temperature when the ionic strength is low (*i.e.* the solution contains low dielectric constant components) and *vice versa*. The temperature–solubility function is not a property of the protein itself, but is linked to the protein-solution system. Therefore, the choice of the buffer is crucial and should not alter the ionic strength of the system as far as possible. In addition, depending on the buffer substance, its pH may vary with temperature. As an example, the Tris buffer, being probably the most frequently used buffer in biological experiments, is not always the best choice since it has a significantly high degree of temperature sensitivity. The effects are therefore very different when Tris is used at 277 K, at room temperature or at 310 K. This means that the pH value has to be set for the temperature around which it is used.

Here we demonstrated the kinetic ripening method, previously successfully employed in growth of large protein crystals by the batch crystallization technique (Budayova-Spano *et al.*, 2007) and reported also by other authors (Astier & Veesler, 2008) (Fig. 4), in an experiment in which the crystal size distribution is wide due to nucleation [Fig. 4(*a*)]. In contrast to an isothermal Ostwald ripening (a phenomenon occurring at constant temperature), here the temperature fluctuations imposed in the neighborhood of the equilibrium temperature induce dissolution of the smallest crystals and growth of the largest with the same phase. The pictures of crystals in Fig. 4 and the corresponding theoretical phase diagram illustrated in Fig. 4(*k*) present the complete kinetic ripening process for the Uox–9MUA complex. In the first stage, temperature is increased by a few degrees. Small and large crystals dissolve [Fig. 4(*b*)] in the undersaturated zone in the vicinity of the solubility curve [pictures surrounded by green in Fig. 4(*k*)], but as small crystals have less matter to be transferred they dissolve faster and the process is stopped before complete dissolution of the larger crystals by a temperature decrease (second stage). Finally, large faceted crystals grow [Figs. 4(*c*)–4(*j*)] inside the metastable zone [pictures surrounded by red in Fig. 4(*k*)].

Protein precipitates are often considered as a dead-end product that cannot evolve towards monocrystals, but in fact many macromolecules have been observed to crystallize from precipitates. It is now clearly established that these precipitates consist of aggregates or gels produced by metastable LLPS (Broide *et al.*, 1996; Grouazel *et al.*, 2002; Asherie, 2004; Vivarès & Bonneté, 2004; Dumetz *et al.*, 2008). The conceptual construct of the phase diagram (temperature versus protein concentration) illustrates the liquid–liquid coexistence curve (blue line) in the case of fluorescent protein EosFP [Fig. 5(*k*)]. The coexistence curve shows the boundary between the region where the protein solution remains homogeneous and the region where demixing occurs and dense droplets with protein-rich phase form, leading to LLPS. According to Ostwald's rule of stages, LLPS occurs prior to crystal nucleation: for kinetic and thermodynamic reasons, liquid nucleation, which proceeds by density fluctuation alone, is faster and easier than crystal nucleation, which requires both density and structure fluctuation (Vekilov, 2010). In practice, an increase in temperature or a decrease in protein concentration leads to supersaturated conditions in which droplets of the dense phase dissolve [Figs. 5(*a*)–5(*e*)]. This zone in the T–C phase diagram, below the solubility curve and above the LLPS curve, represents the location where the right crystallization conditions can be found. Fine-tuning the temperature therefore leads to better control of nucleation and growth. This example shows that even qualitative knowledge of the phase diagram and the respective positions of the phase boundaries for LLPS and crystal nucleation allow identification of optimal conditions for protein crystallization. Our experimental design allowed us to induce crystal nucleation from the liquid–liquid separation of metastable phases by changing the temperature, which is consistent with the literature (Broide *et al.*, 1996). Another way to induce crystallization from the LLPS process would be to change the composition of the crystallization solution at constant temperature and at constant protein concentration (F. Zhang *et al.*, 2012).

The described approaches can be also beneficial in the case of twinned crystals or crystal clusters as mentioned for YchB kinase from *A. tumefaciens*, where the temperature-controlled dialysis crystallization enabled us to obtain large single crystals unlike the traditional method of vapor diffusion. The presence of multiple phases such as polymorphs, solvates, microcrystalline solids of other phases and amorphous liquids may complicate growth of the desired crystal phase. Here again, the controlled fluctuations of temperature and/or of concentrations of crystallization agents and/or additives may be used to drive the transition phenomenon of the phases (disappearance of one phase to the benefit of another phase). Such solution-mediated phase transitions can be used to grow large crystals of the stable phase at the expense of the metastable phase as demonstrated previously (Budayova *et al.*, 1999; Oksanen *et al.*, 2010). The same applies for the macroscopic defects of the crystals such as satellite crystals bonded onto single crystals which can be completely dissolved to the

benefit of the growth of large single crystals (Budayova-Spano *et al.*, 2007; Astier & Veesler, 2008). The crystal habit and crystal quality are also improved because different growth conditions (*e.g.* temperature and supersaturation) induce different crystal growth rates for each face (Budayova *et al.*, 1999; Budayova-Spano *et al.*, 2007; Junius *et al.*, 2016).

3.3. Crystallization kinetics

The metastable zone width is an essential parameter for the growth of large-size crystals from solution, since it is a direct measure of stability of the solution in its supersaturated region. The larger the zone width, the higher the stability (Buckley, 1951; Zaitseva *et al.*, 1995). It appears that our apparatus and crystallization procedure enabled us to modify crystallization kinetics by reducing the equilibration rates and thus to provide better control of crystal nucleation and crystal growth compared with the current methods widely used in protein crystallization. The comparative crystallization experiments with and without variation of the chemical composition of the reservoir in the cases of lactate dehydrogenase and of hCA II demonstrate clearly that precipitant concentration gradients affect the nucleation kinetics. Successively increasing the crystallizing agent concentration in the dialysis crystallization process, in contrast to the traditional dialysis experiment in which the gradient is not imposed, will result in a decrease of the rate of generating the supersaturation and therefore the reduction of nucleation rate and the growth of fewer and larger crystals. The interpretation of this phenomenon with the phase diagram is that the kinetic extent of the metastable zone is enlarged and therefore the spontaneous nucleation rate is decreased at higher supersaturations by the precipitant concentration gradient. This is illustrated by the theoretical phase diagram representing the case of lactate dehydrogenase from *T. thermophilus* [Fig. 8(*k*)]. As mentioned in our previous work, this kind of crystal-growth process is very beneficial for neutron studies that require large single crystals to provide sufficient scattering volumes (0.1–1.0 mm^3, depending on the unit-cell volume and deuteration approach).

On the other hand, the crystal growth rate is a function of supersaturation, that is, the higher the supersaturation, the higher the growth rate. As already mentioned, nucleation or formation of tiny crystals is also driven by the supersaturation. Consequently, monitoring the level of supersaturation over time helps to control the growth and nucleation rates achieved, thereby controlling the size of the crystals produced. This has been demonstrated here with a model protein, chicken egg-white lysozyme. We were able to drive the crystallization process to generate the desired number and size of crystals in all crystallization experiments. In particular, here we emphasize the case where a large number of very small crystals have been generated [Fig. 7(*d*)]. In agreement with the theory, the temperature variations that we induced during this experiment led to the shrinkage of the metastable zone (unlike that observed in the case of hCA II or lactate dehydrogenase) of the corresponding phase diagram [Fig. 7(*d*)]. Therefore, the relatively rapid decrease in temperature to restart nucleation [picture surrounded by green in Fig. 7(*d*)] leads to an increase in the rate generating the supersaturation. Nucleation is fast, many crystals form nearly simultaneously and the majority of crystals grow to approximately identical size. This kind of crystal growth process may be beneficial for free-electron laser and synchrotron serial crystallography, which require large numbers of uniformly sized small crystals (<50 µm).

Note that all the examples shown in this work have a direct solubility as a function of temperature and make use of the salting-out regime of the protein–precipitant phase diagram. However, our methodology is well suited to also deal with inverse solubility and make use of the salting-in regime, which is somewhat underutilized in growing large crystals.

4. Conclusion

Given the time and effort involved in material preparation and structure determination, it is justified to devote more time to optimizing the crystalline material, since the quality or size of the crystals often limits what can be learned. Here we describe rational protocols based on multidimensional crystallization phase diagrams that we established for *in situ* generation of crystals of specific sizes and morphology optimized for different downstream structure determination approaches using our recently developed device (OptiCrys). This approach is illustrated by the crystallization of several soluble proteins. OptiCrys consists of crystal growth apparatus that, in addition to precise temperature monitoring, also allows the control and change of the crystallization solution components (*e.g.* precipitant concentration, buffer, additive, ligands) in an automated manner through dialysis. The crystallization process can be monitored and controlled in real time with a video microscope and a PC via supervision software. The system allows the survey of multiple crystallization conditions in a systematic way with the same biological sample. The sample is not consumed in the experiment, and if the sample is not damaged (*e.g.* denatured), the conditions can be changed reversibly. This enables one to optimize the kinetic path through the phase diagram, which controls the nucleation and growth of the crystals, and thus their number, size and morphology. Tailoring of crystal number, size, phase and diffraction quality reduces the time, protein material and efforts required for structure determination. The described strategy differs from the current paradigm in performing serial instead of parallel experiments. The dialysis membrane acts as a precision dosing device to control the composition of the crystallization solution, the level of supersaturation and its generation rate by tuning the mass transfer across the membrane. Mass transport in crystallization experiments is the process that forms supersaturation gradients and therefore affects both crystal size and diffraction quality. As demonstrated from diffraction quality of lysozyme and YchB kinase crystals grown as described here and already tested (Junius *et al.*, 2016), the transfer of solutes and solvent through semi-permeable membranes seems to provide the slowly varying

conditions required for early crystal nucleation and relatively undisturbed structure formation in crystal growth. In the future, we aim to carry out these studies on other protein targets in order to assess crystal mosaicity as well as homogeneity of the crystal structure by appropriate methods, such as X-ray diffraction topography.

Acknowledgements

The authors thank Dr Dominique Bourgeois and Dr Virgile Adam (IBS, Grenoble) for the pure EosFP fluorescent protein, Dr Dominique Madern (IBS, Grenoble) for the lactate dehydrogenase, Dr Franck Borel (IBS, Grenoble) for the YchB kinase and Dr Zoë Fisher (ESS, Lund) for the human carbonic anhydrase. We are grateful to Professor Bertrand Castro and Dr Mohamed El Hajji of Sanofi–Aventis (Montpellier, France) for supplying the pure hydrogenated recombinant urate oxidase. We acknowledge the European Molecular Biology Organization for a Long-Term Fellowship to EO. IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA).

Funding information

MBS acknowledges support from the MRCT CNRS under the contract 2010–2011 and from LABEX VALO GRAL under the contract 2015. MBS thanks the CEA's International Doctoral Research Program (Irtelis) for the PhD Fellowship to NJ, and the Horizon 2020 Research and Innovation Program of the European Union under grant No. 722687 for the fellowship MARIE SKŁODOWSKA – CURIE to EV.

References

Asherie, N. (2004). *Methods*, **34**, 266–272.
Astier, J. P. & Veesler, S. (2008). *Cryst. Growth Des.* **8**, 4215–4219.
Blakeley, M. P., Hasnain, S. S. & Antonyuk, S. V. (2015). *IUCrJ*, **2**, 464–474.
Bonneté, F., Vivarès, D., Robert, C. & Colloc'h, N. (2001). *J. Cryst. Growth*, **232**, 330–339.
Broide, M. L., Tominc, T. M. & Saxowsky, M. D. (1996). *Phys. Rev. E*, **53**, 6325–6335.
Buckley, H. E. (1951). *Crystal Growth*. New York: Wiley.
Budayova, M., Astier, J. P., Veesler, S., Czjzek, M., Belaich, A. & Boistelle, R. (1999). *J. Cryst. Growth*, **196**, 297–304.
Budayova-Spano, M. (2010). Patent FR10/57354, UJF (extension: EP117730945, US13821053, JP2013528746).
Budayova-Spano, M., Bonneté, F., Ferté, N., El Hajji, M., Meilleur, F., Blakeley, M. P. & Castro, B. (2006). *Acta Cryst.* F**62**, 306–309.
Budayova-Spano, M., Dauvergne, F., Audiffren, M., Bactivelane, T. & Cusack, S. (2007). *Acta Cryst.* D**63**, 339–347.
Budayova-Spano, M., Fisher, S. Z., Dauvergne, M.-T., Agbandje-McKenna, M., Silverman, D. N., Myles, D. A. A. & McKenna, R. (2006). *Acta Cryst.* F**62**, 6–9.
Chapman, H. N. (2015). *Synchrotron Rad. News*, **28**, 20–24.
Cherezov, V., Peddi, A., Muthusubramaniam, L., Zheng, Y. F. & Caffrey, M. (2004). *Acta Cryst.* D**60**, 1795–1807.
Chernov, A. A. & Komatsu, H. (1995). *Science and Technology of Crystal Growth*, edited by J. P. van der Eerden & O. S. L. Bruinsma, pp. 329–353. Dordrecht: Springer.
Coquelle, N., Fioravanti, E., Weik, M., Vellieux, F. & Madern, D. (2007). *J. Mol. Biol.* **374**, 547–562.
Curtis, R. A., Blanch, W. & Prausnitz, J. M. (2001). *J. Phys. Chem. B*, **105**, 2445–2452.
D'Arcy, A., Villard, F. & Marsh, M. (2007). *Acta Cryst.* D**63**, 550–554.
Dierks, K., Meyer, A., Einspahr, H. & Betzel, C. (2008). *Cryst. Growth Des.* **8**, 1628–1634.
Ducruix, A. & Giégé, R. (1992). Editors. *Crystallization of Nucleic Acids and Proteins: A Practical Approach*, The Practical Approach Series, Vol. 86. Oxford University Press.
Dumetz, A. C., Chockla, A. M., Kaler, E. W. & Lenhoff, A. M. (2008). *Biophys. J.* **94**, 570–583.
Dumetz, A. C., Snellinger-O'Brien, A. M., Kaler, E. W. & Lenhoff, A. M. (2007). *Protein Sci.* **16**, 1867–1877.
Echalier, A., Glazer, R. L., Fülöp, V. & Geday, M. A. (2004). *Acta Cryst.* D**60**, 696–702.
Forsythe, E., Achari, A. & Pusey, M. L. (2006). *Acta Cryst.* D**62**, 339–346.
Garcia-Ruiz, J. M., Novella, M. L. & Otalora, F. (1999). *J. Cryst. Growth*, **196**, 703–710.
García-Ruiz, J. M., Otálora, F. & García-Caballero, A. (2016). *Acta Cryst.* F**72**, 96–104.
Gerdts, C. J., Elliott, M., Lovell, S., Mixon, M. B., Napuli, A. J., Staker, B. L., Nollert, P. & Stewart, L. (2008). *Acta Cryst.* D**64**, 1116–1122.
Grouazel, S., Perez, J., Astier, J.-P., Bonneté, F. & Veesler, S. (2002). *Acta Cryst.* D**58**, 1560–1563.
Groves, M. R., Müller, I. B., Kreplin, X. & Müller-Dieckmann, J. (2007). *Acta Cryst.* D**63**, 526–535.
Judge, R. A., Swift, K. & González, C. (2005). *Acta Cryst.* D**61**, 60–66.
Junius, N., Oksanen, E., Terrien, M., Berzin, C., Ferrer, J.-L. & Budayova-Spano, M. (2016). *J. Appl. Cryst.* **49**, 806–813.
Kissick, D. J., Wanapun, D. & Simpson, G. J. (2011). *Annu. Rev. Anal. Chem.* **4**, 419–437.
Lumry, R., Smith, E. L. & Glantz, R. R. (1951). *J. Am. Chem. Soc.* **73**, 4330–4340.
Nannenga, B. L. & Gonen, T. (2016). *Curr. Opin. Struct. Biol.* **40**, 128–135.
Oksanen, E., Blakeley, M. P., Bonneté, F., Dauvergne, M.-T., Dauvergne, F. & Budayova-Spano, M. (2009). *J. R. Soc. Interface*, **6**, S599–S610.
Oksanen, E., Blakeley, M. P., El-Hajji, M., Ryde, U. & Budayova-Spano, M. (2014). *PLoS One*, **9**, e86651.
Oksanen, E., Dauvergne, F., Goldman, A. & Budayova-Spano, M. (2010). *J. Appl. Cryst.* **43**, 1113–1120.
Olesberg, J. T., Arnold, M. A., Hu, S. B. & Wiencek, J. M. (2000). *Anal. Chem.* **72**, 4985–4990.
Segelke, B. (2005). *Expert Rev. Proteomics*, **2**, 165–172.
Stojanoff, V., Jakoncic, J., Oren, D. A., Nagarajan, V., Navarro Poulsen, J.-C., Adams-Cioaba, M. A., Bergfors, T. & Sommer, M. O. A. (2011). *Acta Cryst.* F**67**, 971–975.
Vekilov, P. G. (2010). *Cryst. Growth Des.* **10**, 5007–5019.
Vekilov, P. (2012). *J. Phys. Condens. Matter*, **24**, 193101.
Vivarès, D. & Bonneté, F. (2004). *J. Phys. Chem. B*, **108**, 6498–6507.
Wiedenmann, J., Ivanchenko, S., Oswald, F., Schmitt, F., Röcker, C., Salih, A., Spindler, K.-D. & Nienhaus, G. U. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 15905–15910.
Yin, D. C., Inatomi, Y., Luo, H. M., Li, H. S., Lu, H. M., Ye, Y. J. & Wakayama, N. I. (2008). *Meas. Sci. Technol.* **19**, 045303.
Zaitseva, N. P., Rashkovich, L. N. & Bogatyreva, S. V. (1995). *J. Cryst. Growth*, **148**, 276–282.
Zhang, F., Roosen-Runge, F., Sauter, A., Wolf, M., Jacobs, R. M. J. & Schreiber, F. (2014). *Pure Appl. Chem.* **86**, 191–202.
Zhang, F., Roth, R., Wolf, M., Roosen-Runge, F., Skoda, M. W. A., Jacobs, R. M. J., Stzucki, M. & Schreiber, F. (2012). *Soft Matter*, **8**, 1313–1316.
Zhang, Y., Zhao, J., Di, J., Jiang, H., Wang, Q., Wang, J., Guo, Y. & Yin, D. (2012). *Opt. Express*, **20**, 18415–18421.