Amélioration et automatisation des étapes de préparation des cristaux de protéines à la diffraction aux rayons X

Mohammad Yaser Heidari Khajepour

UNIVERSITÉ DE GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Physique appliqué**

Arrêté ministériel: 7 août 2006

Presentée par

« Mohammad Yaser HEIDARI KHAJEPOUR »

Thèse dirigée par **« Jean-Luc FERRER »**

préparée au sein de l'**Institut de Biologie Structurale (CNRS/CEA/UJF), Grenoble, France**

dans l'**École Doctorale de Physique**

Amélioration et automatisation des étapes de préparation des cristaux de protéines à la diffraction aux rayons X

Thèse soutenue publiquement le **« 19 Septembre 2012 »**,
devant le jury composé de :

Pr. Arnaud DUCRUIX **Président / Rapporteur**
Professor à l'Université Descartes Paris V, France

Dr. Florence POJER **Rapporteur**
Responsable plateforme crystallography et chercheur à l'Ecole Polytechnique Fédérale de Lausanne, Suisse

Dr. François HOH **Membre**
Chercheur au Centre de Biochimie Structurale, Montpellier, France

Dr. Uwe MUELLER **Membre**
Responsable de groupe à BESSY, Berlin, Allemagne

Pr. Roger FOURME **Membre**
Professor à l'Université Paris-Sud, France

Dr. Jean-Luc FERRER **Membre / Directeur**
Responsable de groupe à l'Institut de Biologie Structurale, Grenoble, France
Université Joseph Fourier / Université Pierre Mendès France / Université Stendhal / Université de Savoie / Grenoble INP

Improving and automating preparation steps of protein crystals for X-ray diffraction

Acknowledgments

In this manuscript are presented three years of everyday challenging studies and developments that could have never been realized without the help and support of my colleagues from IBS, especially my colleagues at Synchrotron Group and all people at MetalloProteins Group.

I would like to thank my PhD advisor, Doctor Jean-Luc Ferrer, and also Xavier Vernède for supporting me during these past three years. They have been supportive and have given me the freedom to pursue various projects.

I am also very grateful to Doctor Franck Borel and Elodie Barbier from Synchrotron Group, for their kind office neighborship and advices in laboratory techniques and protein crystallography.

I would like to thank Doctor David Cobessi and Doctor Monika Spano from Synchrotron Group, for sharing their knowledge and helping me in drafting this manuscript.

I would like to thank Doctor Michel Pirocchi, Christope Berzin and Maxime Terrien for helping me in technical aspects in my developments and experiments at FIP-BM30A beamline at ESRF.

I also thank Doctor Christine Cavazza and Hugo Lebrette from MetalloProteins Group, for many insightful exchanges and also for their collaboration with the NikA-FeEDTA protein.

I thank Doctor Juan Carlos Fontecilla-Camps, head of MetalloProteins Group, for his help and advices in drafting my scientific publications.

I also thank Doctor Florence Pojer for hosting me at Protein Crystallography Core Facility at EPFL and allowing me to use their equipments for my developments and experiments.

I thank the members of the jury for accepting this task, especially Doctor Florence Pojer and Doctor Arnaud Ducruix for examining my manuscript.

Finally I would like to thank my wife, Afsaneh and my family for supporting me during the past three years.

Preface

The studies and developments made during this PhD were mainly on physics and engineering aspects in methodologies of the X-ray macromolecular crystallography. Understanding the general science of macromolecular crystallography and X-ray diffraction was mandatory for my thesis works presented in this manuscript. Nevertheless, a deep understanding of macromolecular crystallography was not crucial. Hence this science is introduced in this manuscript, to better understand the context in which these works has been achieved and also why these studies and developments have been led.

This manuscript starts with an introduction of macromolecular crystallography as the first chapter. The second chapter contains one of my scientific publications which has been submitted to *Acta Crystallographica section D*. It tackles the automation of *in situ* X-ray diffraction of protein crystals for laboratory and synchrotron macromolecular crystallography diffraction facilities. My second publication, submitted to *Acta Crystallographica section D*, is presented in chapter III. It reports the development of Robotic Equipment for Automated Crystal Harvesting. Chapter IV presents the possibilities of completely automated pipelines by implementing the two instrumentation developments of this thesis, with new studies and developments to be done.

Contents

"Science is a wonderful thing if one does not have to earn one's living at it."

Albert Einstein

Chapter I: Introduction to Protein Crystallography

Here by are presented the main aspects of protein crystallography, in order to better understand the context of my PhD and also the aims of my studies in high throughput protein crystallography. Firstly is given a brief introduction to structural biology with its experimental techniques, outlining the advantages of X-ray diffraction crystallography. Secondly, protein crystallization principles and techniques are described with emphasis on vapor diffusion technique allowing high throughput crystallization. Thirdly, X-ray diffraction crystallography is briefly presented with different diffraction techniques of protein crystals. Then, preparation steps of protein crystals for X-ray diffraction are detailed for the two most important methods, *in situ* and frozen sample X-ray diffraction. At last, a statement of automation advances allows introducing automation developments needed towards fully high throughput structural crystallography and my PhD issues.

"Anyone who stops learning is old, whether at twenty or eighty."

Henry Ford

1. Protein Crystallography

1.1. Structural Biology

This science concerns the molecular structure of biological macromolecules, specially proteins and nucleic acids. The aim is to understand the role of macromolecules structures on their functions. The determination of three-dimensional structure of macromolecules has earned its importance as macromolecules fulfill most of cell functions. For example, most of therapeutically molecules are compounds that target a protein. They interact with implicated functional domains to modify or block a physio-pathological activity of the protein or to induce an activity with therapeutic effects.

Proteins are highly diverse macromolecules, fundamental to cells functioning and their physiological interaction with tissues and the whole organism. They accomplish crucial tasks in structural (cytoskeleton) metabolism (chemical reaction catalyze, enzymes), signal exchange (hormones), cellular recognition (surface receptors), motility (myofibrils), immune defense (antibody), etc. Regarding to their composition and concatenation of their amino acids, each protein has a tridimensional structure which determines its chemical properties and biological functions related to its environment. This structure-function relation is revealed by the existence of functional domains with which the protein and its structural and/or functional partners (ligands) interact (small molecules, proteins, DNA, etc). One example is enzymes active sites which manage chemical catalyze of specific reactions by fixing adequate substrates. Another example is signal transduction initiated by small molecules which fixes on membrane or nuclear receptors. In extreme cases, a protein folded differently could have completely different functions (e.g. Prion).

Sequence analysis is likely to outfit insightful information in many problems linked to proteins properties and activities (membrane insertion, protein interaction sites highly probable, potential antigenic sites, etc). However structural information is often essential for understanding action mechanisms, functions and protein-ligand interaction modes (protein, nucleic acid, small molecule, etc). Accordingly tridimensional protein structural studies provide insights for further investigation such as the analysis of functional domains, stability criteria definition, epitopes prediction, understanding enzymatic mechanisms, etc.

The resolution of an experiment method in structural biology is related to the minimum distance between two points as to distinguish them separately. The higher the resolution of an experimental method, the lower distance between two points are observable. The distance between atoms constituting a protein is around 1 Å (0.1 nm or 10^{-10} m) and an average protein size is about 100 Å (10 nm). For observing the form of a molecule or few proteins arrangement a resolution near to 100 Å is sufficient. On the other hand, if atoms organization of inside the molecule is of interest, a higher resolution of about 1 Å is required. This is called the atomic resolution.

Technological developments of structural biology with atomic resolution reveal details of protein-ligand interaction beyond the simple shape complementarily of two molecules (e.g. in enzymes active sites). Atomic resolutions also allow determining the nature of interactions (hydrogen links, hydrophilic interactions, lipophilic, electrostatic, dipoles, etc) implicated in ligand binding, activities of considered functional domain and possible induced structural modifications that condition their functional properties. Therefore access to the structure of protein-ligand complex at atomic scale allows prior rational design of new active molecules with sought functional/therapeutic properties (e.g. ability to block the reaction of a specific active site of an enzyme).

In the following, we will discuss about the different experimental techniques of studying biological macromolecular structures.

1.2. Experimental Methods

Three types of radiations are used for obtaining atomic resolutions: electromagnetic radiations (X-rays, high frequency electromagnetic waves), electrons and neutrons. The properties of each of these radiations are detailed below:

- **X rays:** Discovered in 1895 by Roentgen[1], this electromagnetic radiation has a wavelength from 0.1 Å to 1000 Å. X-rays are generated by home laboratory sources or by synchrotrons. Home laboratory sources are of two kinds: sealed tube and rotating anode. Rotating anode sources generate more intense radiation than sealed

[1] Wilhelm Conrad Roentgen, 1845-1925, discovered X-rays in 1895 in Würzburg in Germany.

tube sources. Other technologies for laboratory sources are under development, such as liquid metal targets. On the other hand, a synchrotron X-ray source is much more powerful providing highly intense and focused beam compared to laboratory sources. These X-ray beams can be used to conduct diffraction experiment, following various techniques, the most commonly used being the single crystal monochromatic beam diffraction.

- **Electromagnetic radiation in NMR:** This technique uses electromagnetic radiation emission and measurement on solid-state and solution samples placed in a static and very high frequency electromagnetic field (60 to 1000 MHz). It gives information on inter-atomic distances in a molecule which are used to solve the three-dimensional structure. This method is limited to molecules with a molecular mass lower than 50 kDa.

- **Electrons:** Discovered in 1897 by Thomson[1], they are generated in electronic microscopes with a wavelength of about 0.01 Å. Samples should be prepared in very thin layer (about 1 nm).

- **Neutrons:** Discovered in 1920 by Chadwick[2], are generated thanks to nuclear reactors or spallation nuclear sources with wavelengths between 1 Å and 10 Å. These neutron beams can be used to conduct diffraction experiments. Due to the limited neutron flux, compared to X-ray beams, larger crystallized samples (~1 mm x 1 mm x 1 mm) and longer exposure time are needed. So duration of experiment is significantly longer than for X-ray diffraction.

These radiations are complementary for the study of a molecular structure, due to their differences in their properties, providing different experimental methods in structural biology: X-ray Diffraction, Solution/Solid-state NMR, Electron Microscopy/Crystallography, Neutron Diffraction, Solution Scattering mostly known as **S**mall **A**ngle **X**-ray **S**cattering (SAXS), Fiber Diffraction, etc.

[1] Joseph John Thomson, 1856-1940, discovered electrons in 1897 at Cambridge University.
[2] James Chadwick, 1891-1974, discovered neutrons in 1932 at Cambridge University.

Hits	%	Experimental Method
72405	87,74	**X-ray Diffraction**
9417	11,41	**Solution NMR**
436	0,53	**Electron Microscopy**
51	0,06	**Solid-State NMR**
50	0,06	**Hybrid**
38	0,05	**Neutron Diffraction**
37	0,04	**Fiber Diffraction**
33	0,04	**Electron Crystallography**
32	0,04	**Solution Scattering**
23	0,03	**Other**
82522		**Total hits**

Table 1: PDB Experimental Method statistics (http://www.rcsb.org/pdb)

PDB statistics (Table 1 and Figure 1) show clearly that X-ray diffraction and Solution NMR contribute to the larger part of the solved structures. The gap between the two methods is essentially due to limiting molecule sizes that can be studied by the NMR and also the time required for the experiments. Nevertheless, this method has the advantage of not needing crystals. X-ray diffraction method requires protein crystal samples, which often are not easy to produce. We will discuss this point in Chapter I:2.2. Principles of Protein Crystallization. In spite of that X-ray diffraction has the advantage, with all technological and instrumental development (see 1.3. Crystallography and 3.2. Methodologies), of acquiring data rapidly, up to resolutions as high as 0.45 Å (PDB reference: 3NIR).

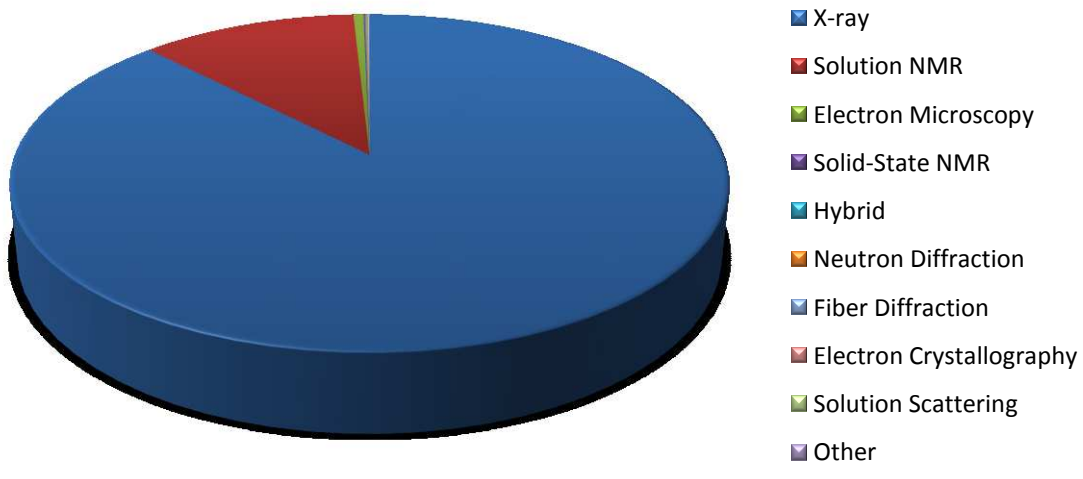

Figure 1: Experimental Method pie chart from PDB statistics (http://www.rcsb.org/pdb)

1.3. Crystallography

With molecules exposed to X-rays, the scattered radiation contains precious information on molecules structure. Nevertheless, over 99% of the X-rays pass through the molecule without being scattered. So to emphasis the scattering signals, a large number of same molecules should be arranged in a well known spatial configuration. This arrangement of molecules constitutes what is called a crystal (Figure 2). The science related to the arrangement of molecules is called crystallography. It defines a crystal as a unique form of arrangement of molecules or a three-dimensional repetition of molecules creating a lattice.

Mathematical studies on crystal morphology showed a limited number of ways of arranging molecules within a crystal lattice. In 1845, Bravais[1] concluded that the number of lattices is limited to 14 due to the symmetrical criteria, in crystalline structures. Thereby spatial configuration of crystals can be known.

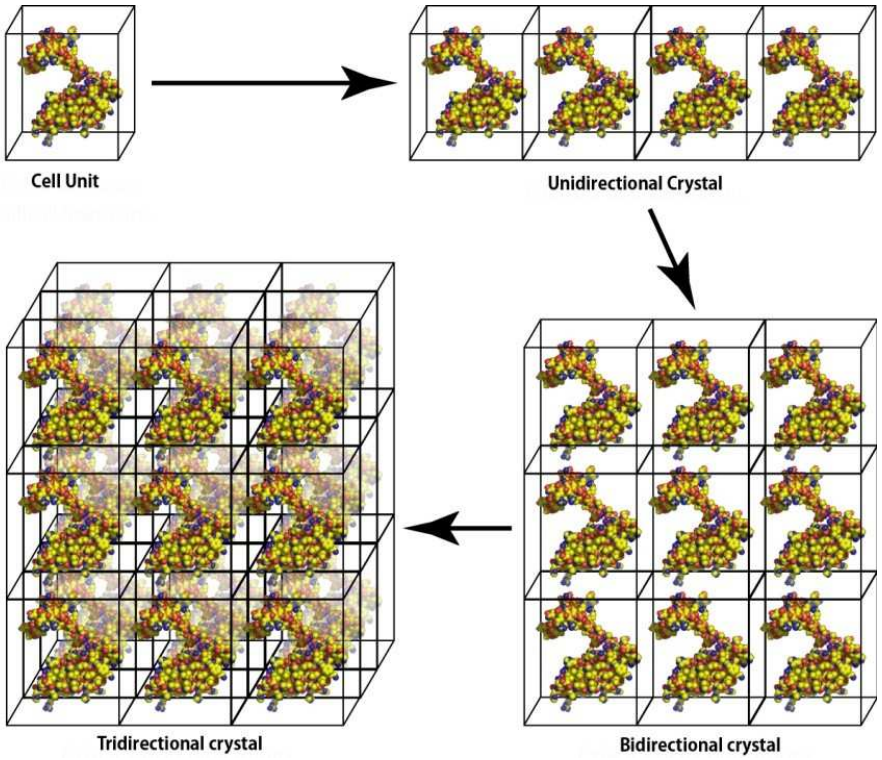

Figure 2: Tridimensional Crystal (Cherrier, 2006)

In structural biology, single crystal X-ray diffraction technique is used to solve the structure of macromolecules. Nucleic acids and proteins are first produced in monocrystalline structures (see 2. Protein Crystallization) and then exposed to X-ray for diffraction

[1] Auguste Bravais, 1811-1863, showed the 14 Bravais lattices in 1845 at Ecole Polytechnique de Paris in France.

measurements (see 3. X-ray Diffraction). Processing of diffraction measurements lead to tridimensional structure of the concerning macromolecule. Myoglobin (PDB reference: 1MBN) was the first protein to have its tri-dimensional structure solved by X-ray crystallography, in 1958 by Kendrew[1] (Kendrew *et al.*, 1958).

Several major steps marked the history of evolution of X-ray crystallography. One of the most important one was the creation of large facilities, synchrotrons with high intensity X-ray beams and protein X-ray crystallography dedicated beamlines. Today more than 30 synchrotrons with more than 70 X-ray crystallography beamlines are available all around the world. This gives easier access to X-ray intense beams and makes experiments faster. Another major step in X-ray crystallography was cryo-cooling of crystals (see Chapter I:4.2. b) Cryo-protection and flash-cooling) for X-ray diffraction. Room temperature exposure of macromolecular crystals causes serious radiation damage to crystals. Thus, crystals perish into X-ray beam and diffraction patterns quality decreases rapidly to unexploitable. Cooling crystals help reducing radiation damage so that complete dataset for structure resolution can be collected from a single sample. A third considerable advance was automation of sample preparation, data collection and data processing procedures (see 5. Why high-through put crystallography). Complete high-throughput crystallography pipeline is the logical next step, which has been discussed widely by the crystallography community since few years. In this manuscript, few developments are made towards this ambition.

[1] John Cowdery Kendrew, 1917-1997, was a biochemist and crystallographer who shared the 1962 Chemistry Nobel Prize with Max Perutz, at Cambridge University.

2. Protein Crystallization

X-ray crystallography being the key to a rapid and high resolution tri-dimensional structure of macromolecules, the protein crystallization is a necessary step for it. Nevertheless, obtaining good diffraction quality crystals remains the major bottleneck to structure resolution. There are several parameters and methods improving the crystallization process, but there is no method of predicting the best conditions of good diffraction quality crystal growth for a specific protein. In the following macromolecular crystals are presented and crystallization principles and techniques for an optimum crystal growth are detailed.

2.1. Protein Crystals

Protein crystals are macroscopic objects composed of regular arrangement of molecules. Macromolecule crystals are grown in aqueous solutions in which the concerning protein is solubilized. Due to inter-molecular space in crystalline lattice, protein crystals contain from 27 % to over 78 % solvent (Matthews, 1968). Their dimensions are quite random and could vary from 5 µm to more than 500 µm and they can grow in rather any unpredictable shape (see Figure 3). Due to their content, dimensions and often their unstable equilibrium state, protein crystals are very sensitive objects to mechanical stress, humidity, temperature, etc.

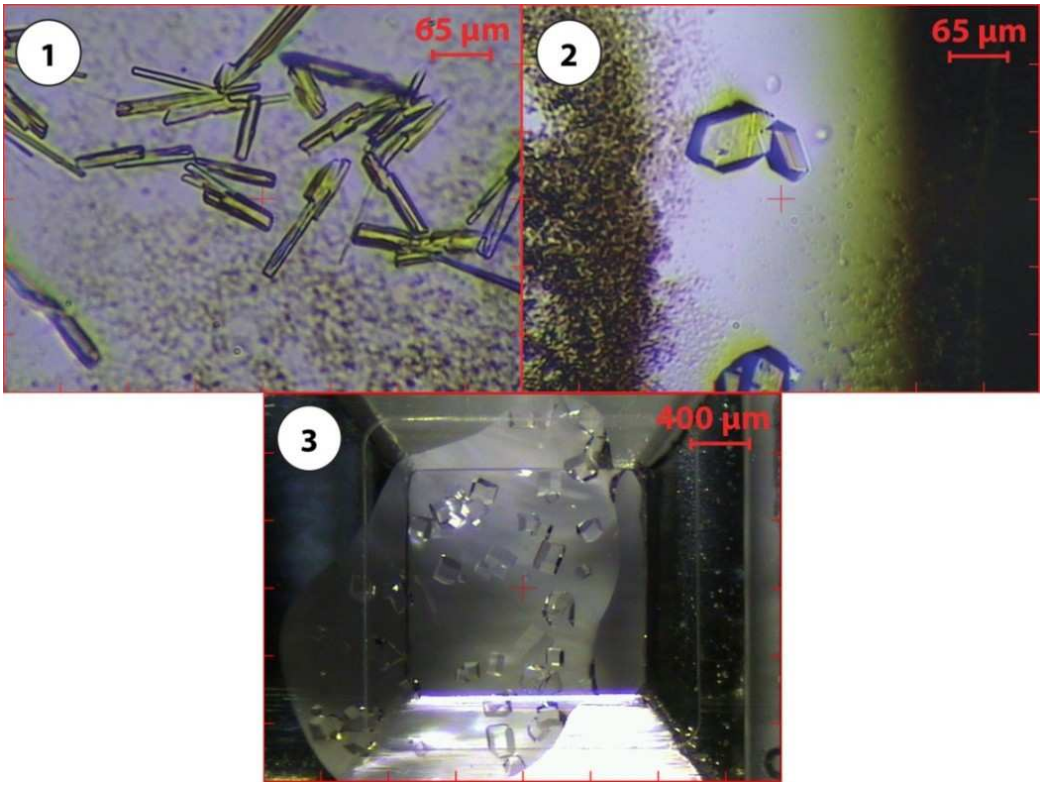

Figure 3: Protein crystals ① Aspartate Amino Transferase, ② YCHB, ③ Hen egg-white lysozyme

2.2. Principles of Protein Crystallization

The crystal growth physics is a quite complex knowledge which today is highly advanced (McPherson, 1999; Ducruix *et* Riès-Kautt, 1990; Asherie, 2004; Vekilov, 2004; Veesler *et* Boistelle, 1999). Even though instrumental developments allow better controlling the kinetics of crystal growth (Budayova-Spano *et al.*, 2007) nevertheless this knowledge is poorly controlled in mostly used and also in high throughput crystallization techniques. Thus crystallization still remains mostly experimental.

Here by a few key expressions are defined to better comprehend the crystallization process.

- **Supersaturation:** This term refers to a solution that contains more of the dissolved material than could be theoretically dissolved by the solvent under the solubility amount (i.e. Solubility Curve, see Figure 4).
- **Metastable:** It refers to a physical or a chemical stable state that could last long. *"This corresponds to the metastable zone, where the supersaturation level is too low for nucleation, so that no new crystals form in any reasonable amount of time."* (Budayova-Spano *et al.*, 2007).
- **Nucleation:** *"A line of recent theories and simulations have suggested that the nucleation of protein crystals might, ..., proceed in two steps: the formation of a droplet of a dense liquid, ..., followed by ordering within this droplet to produce a crystal."* (Vekilov, 2004).

Crystal growth is mainly a result of precise molecular organization of a supersaturated solution in a thermodynamically adequate condition and favorable kinetics. In order to overcome very low molecular attractive force of protein molecules, highly purified and homogeneous protein samples are required for protein crystallization (Giegé *et al.*, 1986). Despite, crystal growth could take over several weeks, whereas some could grow in only few hours.

Crystallization conditions could be related to numerous features, such as: protein concentration, temperature, solvents and their concentrations, pH, etc. In theory, to obtain crystals, the crystallization solution should evolve, with a controlled kinetic and temperature, towards a supersaturated state to trigger nuclei formation. The evolution of the

crystallization solution to the supersaturated state is related to both protein concentration and precipitant concentration. Furthermore, the evaluation of the supersaturated state through a concentration ratio[1] defines the driving force[2] for nucleation and growth (Veesler *et* Boistelle, 1999). Once at supersaturated state and when nucleation points start to appear, protein concentration would rather diminish to avoid numerous nuclei formation and also to enhance the crystal growth. The reduction of protein concentration to the metastable zone in the phase diagram induces slow crystal growth in order to let crystals reaching maximum degree of order in their structure (see Figure 4).

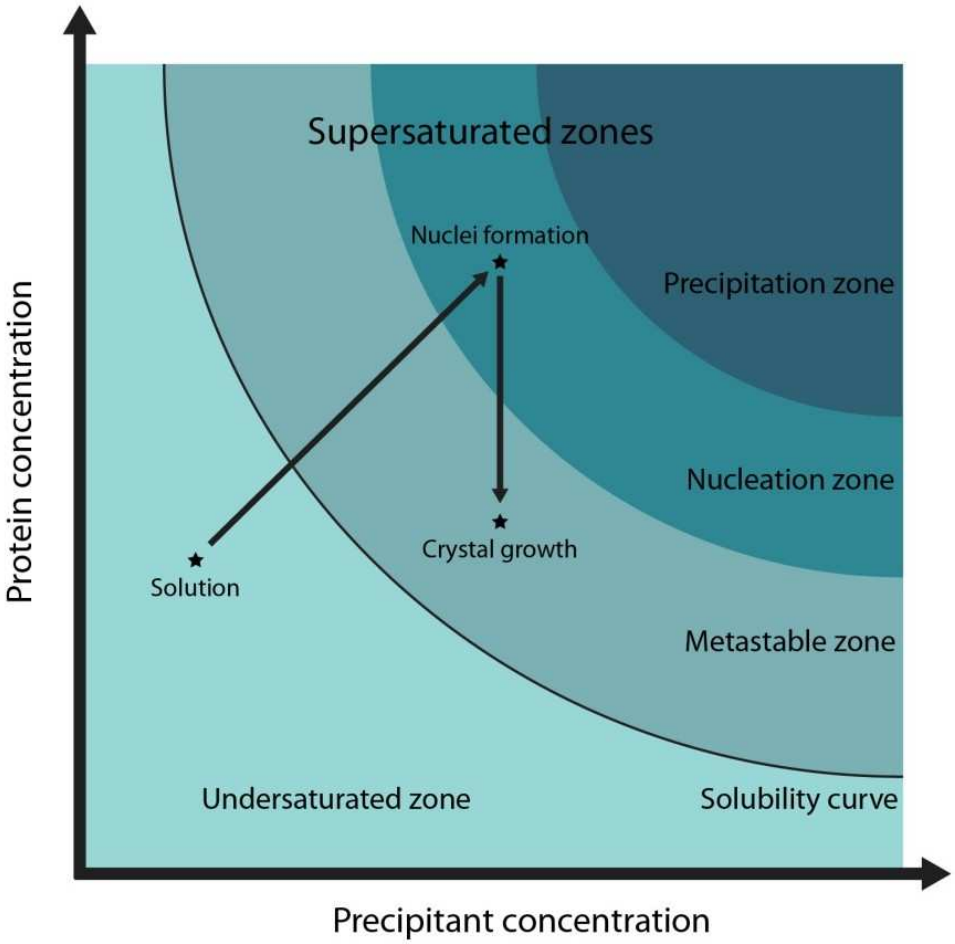

Figure 4: Crystallization phase diagram

In practice, the effective protein concentration is doped in solvents by the addition of precipitant agents such as salt (e.g. ammonium sulfate) or PEG (Polyethylene Glycerol). Thereby several techniques are used to reach the supersaturated state and favor crystalline

[1] Supersaturation ratio $\beta = C/C_s$ where C and Cs are the actual concentration and the saturation concentration respectively.

[2] Driving force is the difference between the chemical potential of the solute molecules in the supersaturated state (μ) and saturated state (μ_s) respectively: $\Delta\mu = \mu - \mu_s = k_B T \ln \beta$ where kB is the Boltzmann constant, T the absolute temperature and β the supersaturation ratio.

precipitation (see 2.3. Crystallization techniques). Nuclei formation induces protein concentration decrease. Depending on the initial protein concentration and the nuclei formation, the concentration should reach the metastable zone over the solubility curve (see Figure 4) were crystal growth could continue.

2.3. Crystallization techniques

As mentioned above, the key point in protein crystallization is to control the thermodynamics and kinetics of supersaturation evolution. Two major techniques are used: liquid-liquid diffusion and vapor diffusion, producing different kinetics of the equilibrium. In this manuscript both techniques are described with emphasize on vapor diffusion techniques which is the most commonly used methods allowing high throughput crystallization and thus the one used in this thesis studies.

a) Liquid-liquid diffusion crystallization

The diffusion is made, either through a direct liquid-liquid interface (Crystallization batch, Counter diffusion), either through a dialyze membrane.

- **Crystallization batch:** This method is the oldest crystallization method. A precipitant reagent drop, of around 1 µL, is dispensed directly into a protein solution of the same volume. This brings instantly the solution to supersaturated state. The drop is covered by an oil (e.g. paraffin) to avoid evaporation. Hopefully nuclei formation and crystal growth will follow. This method is the simplest but not the most productive method.

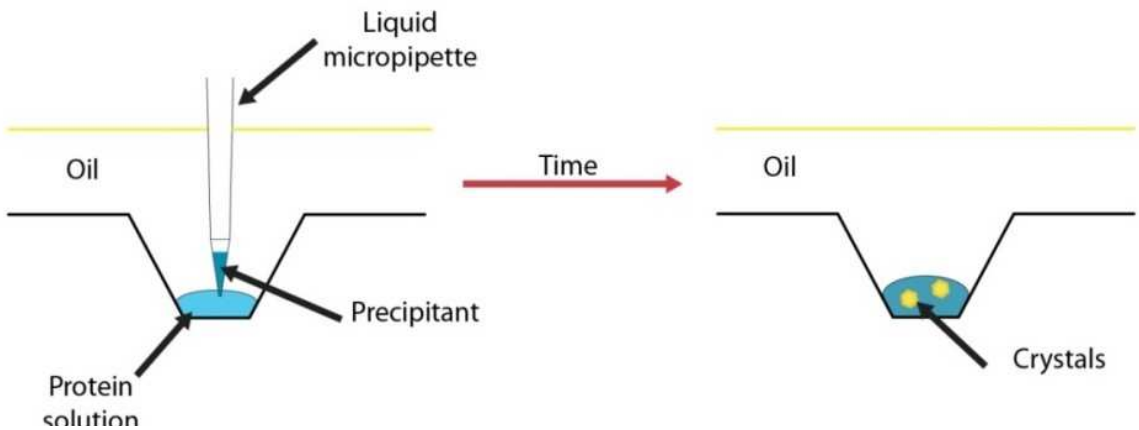

Figure 5: Crystallization batch technique

- **Counter diffusion:** The method uses small bore capillaries in which, first the precipitant solution of about 5 µL is dispensed. A volume of the protein solution is then added over and the tube is sealed with grease and kept vertically. The small diameter of the capillaries allows slow diffusing of the two solutions into one another, creating a continuous gradient of supersaturation. The supersaturation ratio decreases towards the bottom. This will allow the nucleation and crystal growth at different height of the capillary.

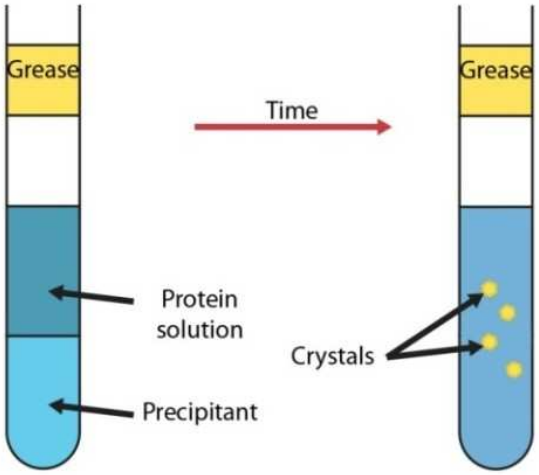

Figure 6: Counter diffusion in capillaries for protein crystallization

- **Dialysis:** Many variation of dialysis technique for crystallization exist, but the most convenient and common one is the dialysis buttons. The protein solution is dispensed in a button covered with adapted membrane. Different dialysis buttons with different volumes and different membranes with different molecular weight cut off range are also commercially available. The membrane is held thanks to an elastic rubber ring. The button is thus plunged into precipitant solution, where the membrane avoids protein extraction and allows precipitant diffusion into the button.

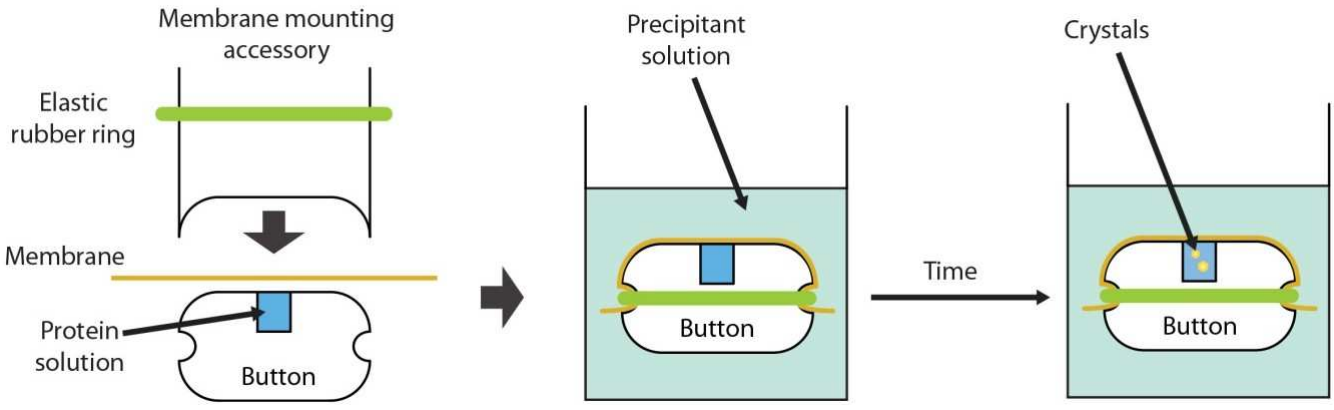

Figure 7: Dialysis crystallization button

b) Vapor diffusion

As shown in Figure 8, two different methods, but with the same vapor diffusion principle, are used: hanging drops and sitting drops. Crystallization drops of 50 nL to over 4 µL, containing a mixture of protein solution and precipitant are dispensed next to a reservoir containing larger volume of precipitant (25 µL to 1 mL). The whole is kept confined in 100 µL to few milliliter spaces. Natural vapor diffusion between the two solutions slowly evolves the concentration in protein mixture to an equilibrium state. This evolution of concentrations the crystallization drop in precipitant and in protein leads to supersaturation state and hence to the nucleation and crystal growth (Hampel *et al.*, 1968).

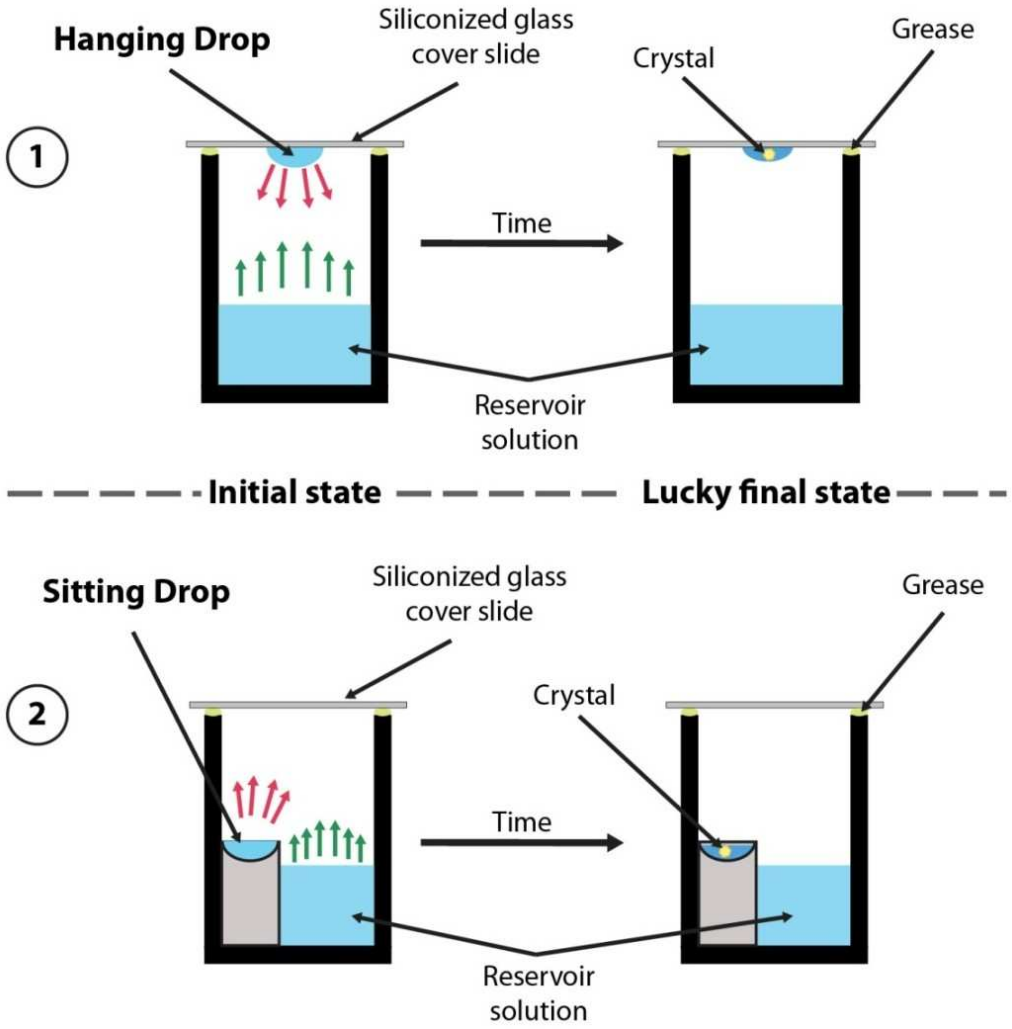

Figure 8: Vapor diffusion crystallization techniques ① with hanging drops and ② with sitting drops

Crystallization drops can be dispensed on glass cover slides or sealing trays and are disposed over reservoir trays. A grease layer between the cover and the reservoir prevents evaporation, and so drops are hanging. Crystallization plates with 16 to 96 reservoirs are commercially available for hanging drops vapor diffusion method.

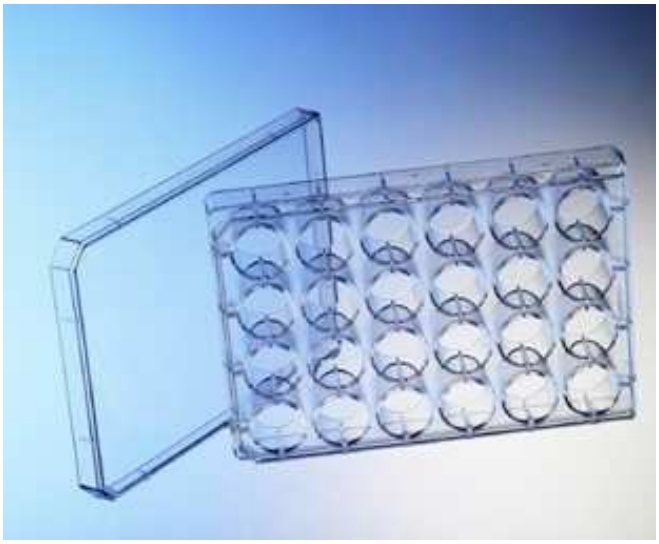

Figure 9: Greiner Bio-One 24 well crystallization plate for hanging drops

Special crystallization plates with few wells (1 to 3) per reservoir for sitting drops are used for vapor diffusion crystallization technique. Since automation of liquid dispensers, these crystallization plates are the most used in protein crystallography. In the last few years many crystallization robots and plates have been developed to reduce the volume of sample used for each crystallization drop and also accelerate the liquid dispensing. Nowadays robots can manage dispensing few nano-litter drop sitting drops on 96 well microplates (see Figure 10). A complete 96 well microplate can be prepared with these robots in matters of seconds, paving the way to high-throughput protein crystallography.

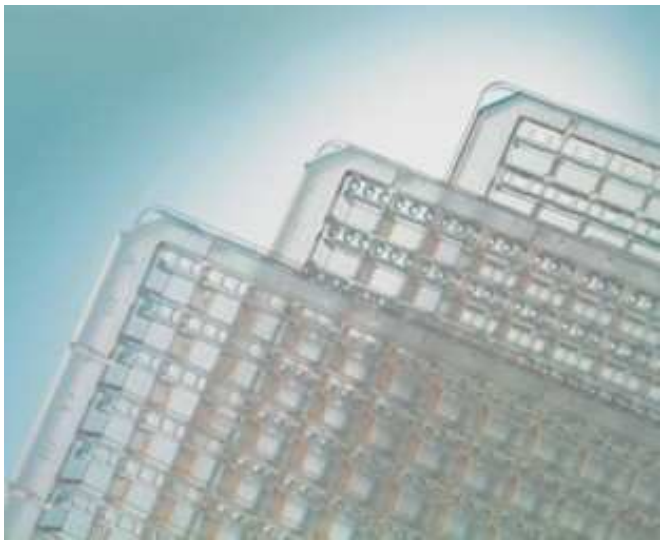

Figure 10: Greiner Bio-One 96 well crystallization microplates (with SBS standard geometry) for sitting drops

Once crystal growth has succeeded and protein crystal samples are available, they should be prepared for X-ray diffraction. Preparation steps are detailed in the following.

3. X-ray Diffraction

As mentioned above, X-ray diffraction represents the most contributive experimental method for structure resolution in structural biology. In order to present the possible improvements that can be introduced to the preparation procedures and the quality of collected data during crystallography experiments, the principles of this method and the different experimental methodologies are introduced.

3.1. Principles

X-ray wavelength is adapted to observe atomic details, as inter-atomic dimensions are about 1 to 2 Å and X-ray wavelength is in the range of 0.1 to 1000 Å. Nevertheless, using X-ray for direct observation at the atomic scale is not possible, considering that the refractive index of X-ray is so small that an optic lens for X-ray microscope is impossible to make. Therefore, analysis of the atomic structure of macromolecules requires another method. An alternative solution is to collect the X-ray diffraction measurements from a single crystal (see Figure 11). By processing these collected data, we are able to deduce the atomic structure of the crystallized macromolecule.

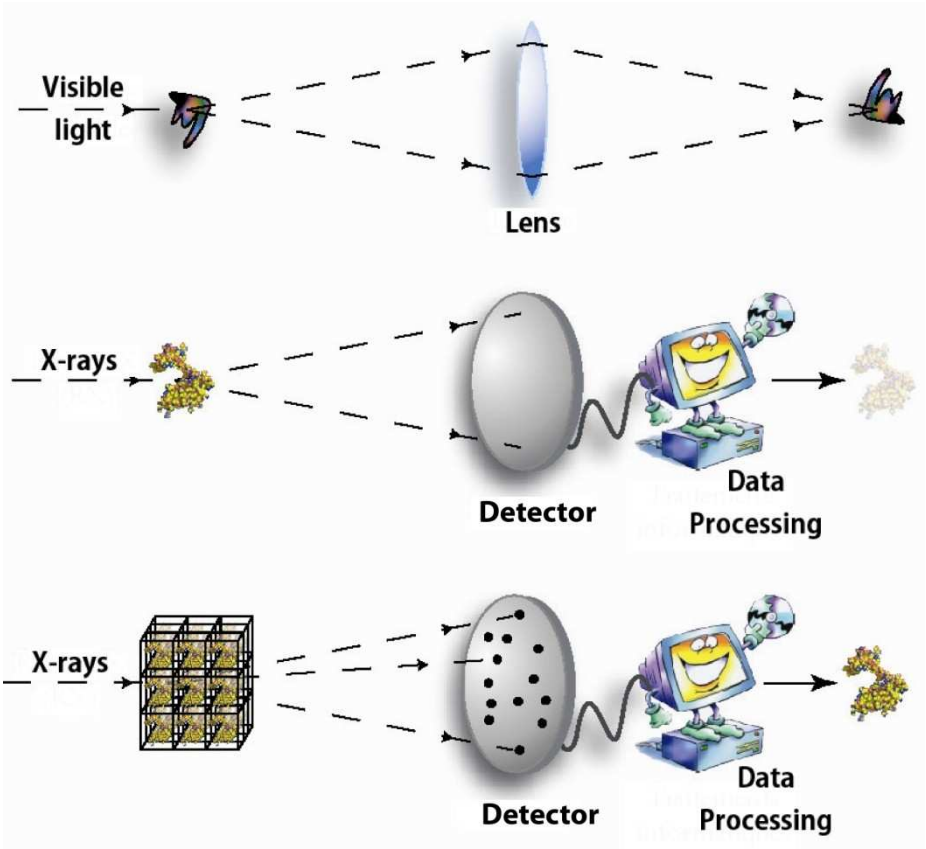

Figure 11: X-ray diffraction (Cherrier, 2006)

X-ray diffraction is considered as a scattering technique. X-ray photons from an incident beam are reflected when encountering atoms of the exposed sample, giving birth to a scattered beam. As mentioned before, over 99% of the X-rays pass through the molecule without being scattered. So to emphasize the scattering signals, a large number of same molecules should be arranged in a well known spatial configuration, which is called a crystal. In 1912, Bragg[1] discovered that precious information could be revealed by measuring the intensity and the angle of the scattering beam on a crystalline sample. Bragg law relates the incident wavelength to the scattering angle and the distance between atomic planes of a crystal lattice (see Figure 12). A discrete atomic model of a crystal in Figure 13 shows the distance *d*, and the θ angle as half of the angle between the incident and the scattered beam.

$$n\lambda = 2d \sin\theta$$

Figure 12: Bragg law

Moreover, Bragg discovered that reflected photons from the incident beam could interfere constructively (overlapping one another producing a more intense scattered wave) or unconstructively (neutralizing one another or decreasing the intensity of the wave) (see Figure 13). The results of these interferences of scattered beam are the spots observed on diffraction patterns. Crystals, as a three-dimensional periodic repetition of molecules, allow increasing the constructive interferences to give more intense spots and thus generate usable information for structure resolution.

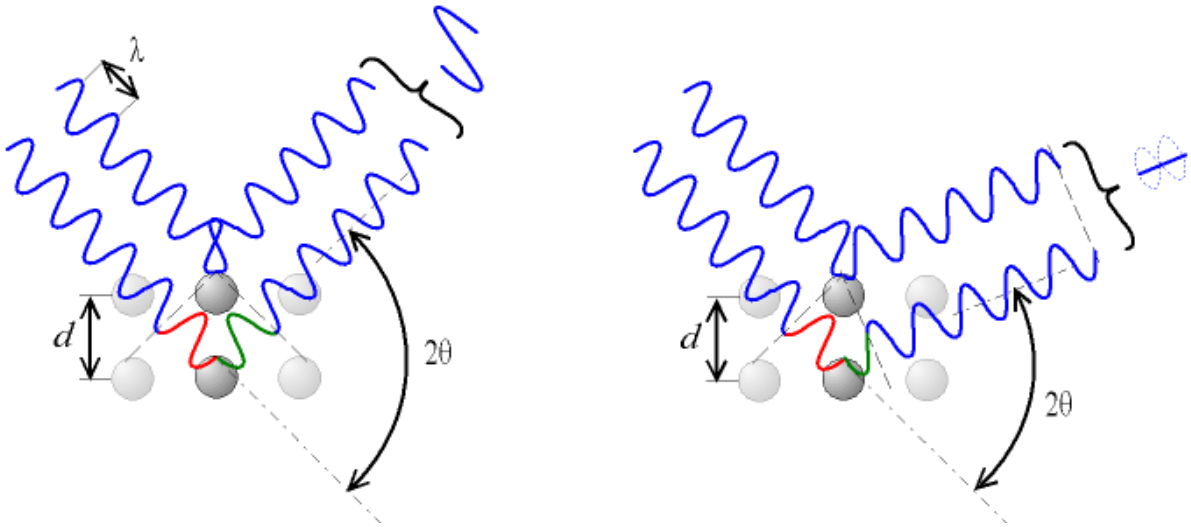

Figure 13: Constructive (on the left) and deconstructive (on the right) interferences in X-ray diffraction of a crystal sample (Image from Wikipedia web site http://en.wikipedia.org)

[1] William Lawrence Bragg, 1890-1971, discovered the Bragg law in X-ray diffraction in 1912 and was joint winner of Nobel Prize in Physics in 1915.

The scattered X-ray from the crystal allows the measurement of a large number of Bragg reflections in a single exposure. A crystal lattice is three-dimensional, only a fraction of the crystal lattice points are in diffracting position at any given orientation of the crystal. Therefore, the crystal is also rotated through an angle of 0.1 to 2° during the exposure to bring more reflections to diffracting position. A diffraction pattern represents an instant image of the crystal. Thus, in order to be able to reconstruct the three-dimensional structure of the crystal, exposures at different orientations of the crystal are required. The crystal lattice has a rotational symmetry allowing limited orientation in diffraction data collection. Thus crystals with higher symmetry require fewer diffraction images to cover the entire crystal lattice.

A whole data set for a protein crystal contains several diffraction patterns and a diffraction pattern contains many reflections. The processing of these data to structure determination is a quite complex physical and mathematical knowledge which has made considerable advances in automation in the last two decades allowing high throughput data processing (Kabsch, 1988) and refinement to structure determination (Holton *et* Alber, 2004; Perrakis *et al.*, 1999).

3.2. Methodologies

The predominant method in scattering protein crystals is the frozen sample X-ray diffraction. Also a new method is gaining in importance for screening crystals and sometimes collecting dataset at room temperature, called *in situ* X-ray diffraction.

a) Frozen sample X-ray diffraction

Until two decades ago protein crystals were exposed at room temperature to X-rays. After the crystallization step of targeted proteins, crystal sample were sucked into micro-capillaries from their crystallization drops. In order to reduce the background scattering the mother liquid sucked with the crystal was removed from around the crystal in the capillary. Thus the crystal was exposed in the capillary to X-rays at room temperature. Even though X-ray diffraction is considered as non-destructive scattering technique, macromolecular crystals suffer in the X-ray beam due to radiation damage (Garman *et* Schneider, 1997). Due to the radiation damage induced to the crystal, only few diffraction patterns with good

quality scattered spots were exploitable. Experiments at 4°C showed reduction of the radiation damage and improvement of the quality of data collected. Later on, cryo-crystallography allowed scattering crystals at cryogenic temperatures below 140 K with very limited radiation damage and thus improvement of the quality of data collected, despite an increase of the crystal mosaicity. As a result, frozen sample X-ray diffraction has become the most common technique used in collecting protein X-ray diffraction data.

In this method, frozen samples are sat-up on a goniometer to allow rotating crystals for single wavelength rotation X-ray diffraction measurements. The goniometer axis is called the spindle axis, which is usually perpendicular to the beam axis. The spindle and the beam axis intersect at the sample position. The crystal has to be centered carefully on this position to avoid the crystal exiting the beam while rotating. This configuration in rotating crystals is considered as Kappa = 0° along with Phi rotation for the goniometer (see Figure 14). Accordingly samples are exposed to X-rays while rotating the crystal. Diffraction patterns are saved for every rotation step, classically from 0.1° to 2° but essentially 1°. Total angular sector to be collected depends on the symmetry of the crystal lattice.

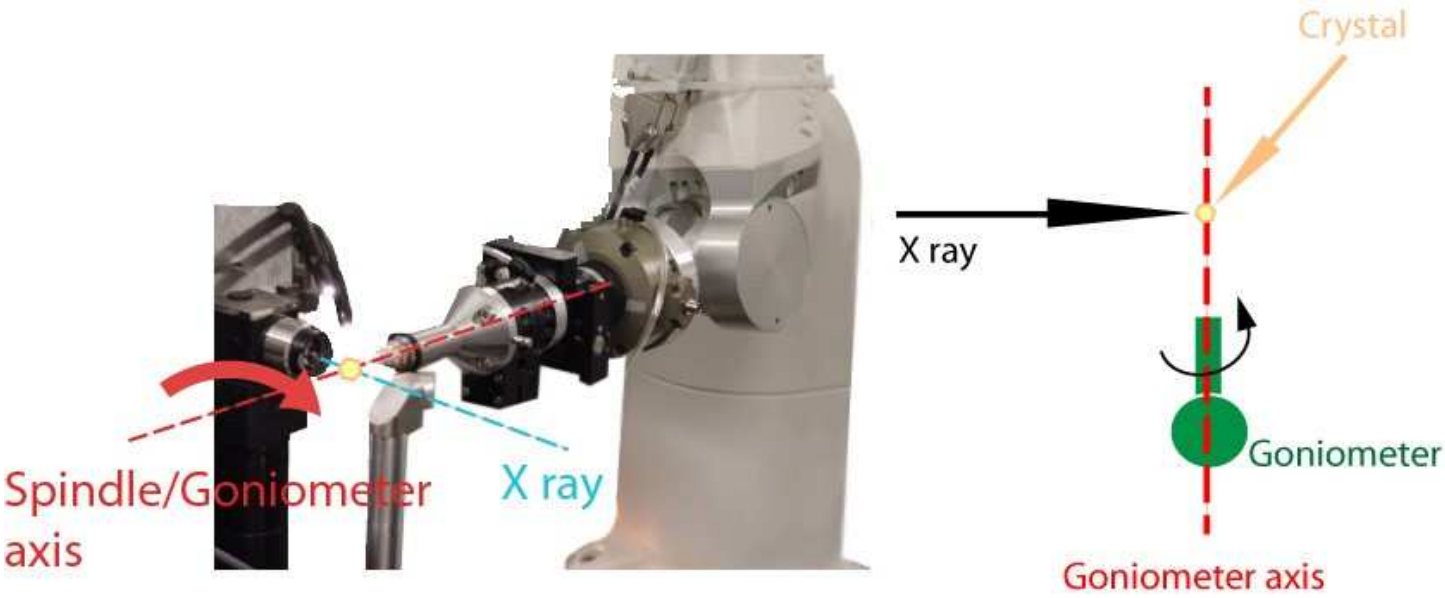

Figure 14: Frozen sample X-ray diffraction set-up with Kappa = 0 and Omega rotation

Different strategies (Dauter, 1999) are possible in rotating the crystal and exposing the crystal lattice for X-ray diffraction. To have complete data of some crystals thus the orientation of the crystal or of the spindle axis of the goniometer could be changed (see Figure 15).

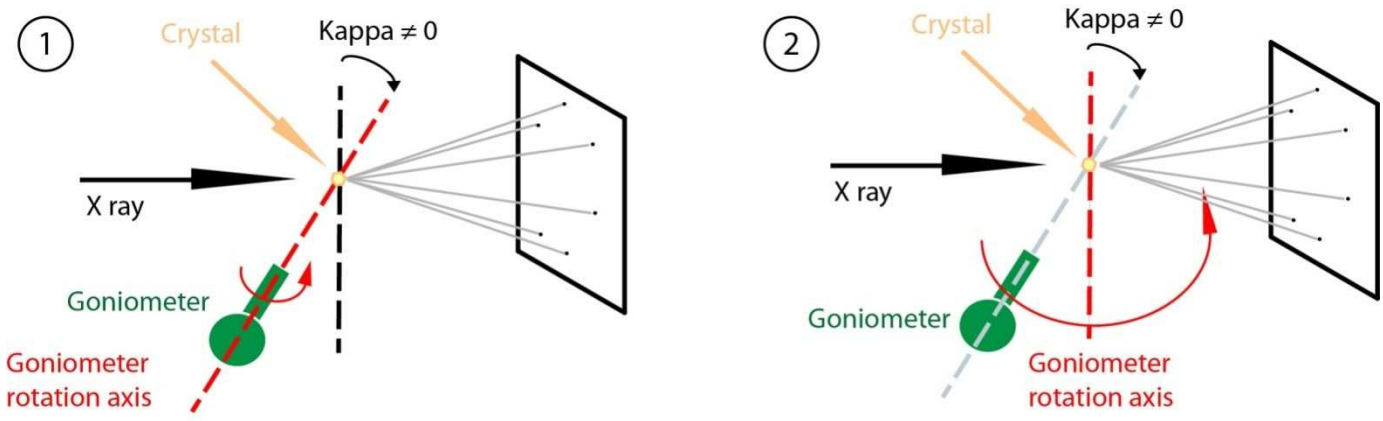

Figure 15: Goniometer Kappa ≠ 0 configurations in diffraction strategies, ① Kappa rotation, ② Omega rotation

b) Room temperature in situ X-ray diffraction

Today in macromolecular crystallography, crystals are more often prepared in drops dispensed in crystallization plates (see 2. Protein Crystallization). Each crystallization plate could contain 24 to over 380 drops. Depending on each protein, the drops volume, contents and concentrations, from 0 to hundreds of crystals could appear in the same plate. Sometimes the crystals formed in the drops are not made of the targeted molecules, but they are instead made of a molecule from the crystallization solution (e.g. very commonly NaCl or ammonium sulfate crystals). In order to analyze crystals before preparing them for frozen sample X-ray diffraction experiments, crystals could be diffracted *in situ*. With this technique, developed in 2004 on FIP-BM30A beamline at ESRF, crystals can be analyzed directly in their crystallization plates (Jacquamet *et al.*, 2004). Further than the discrimination between protein and salt crystals, as mentioned above, diffraction patterns from *in situ* exposure could reveal precious information on the crystals: protein crystal or not, diffracting quality, diffracting resolution, mono-crystalline or poly-crystalline, point group, mosaicity, etc. Nowadays, this method is more and more used to screen crystallization plates for good diffraction quality crystals and is even quite automated (Bingel-Erlenmeyer *et al.*, 2011).

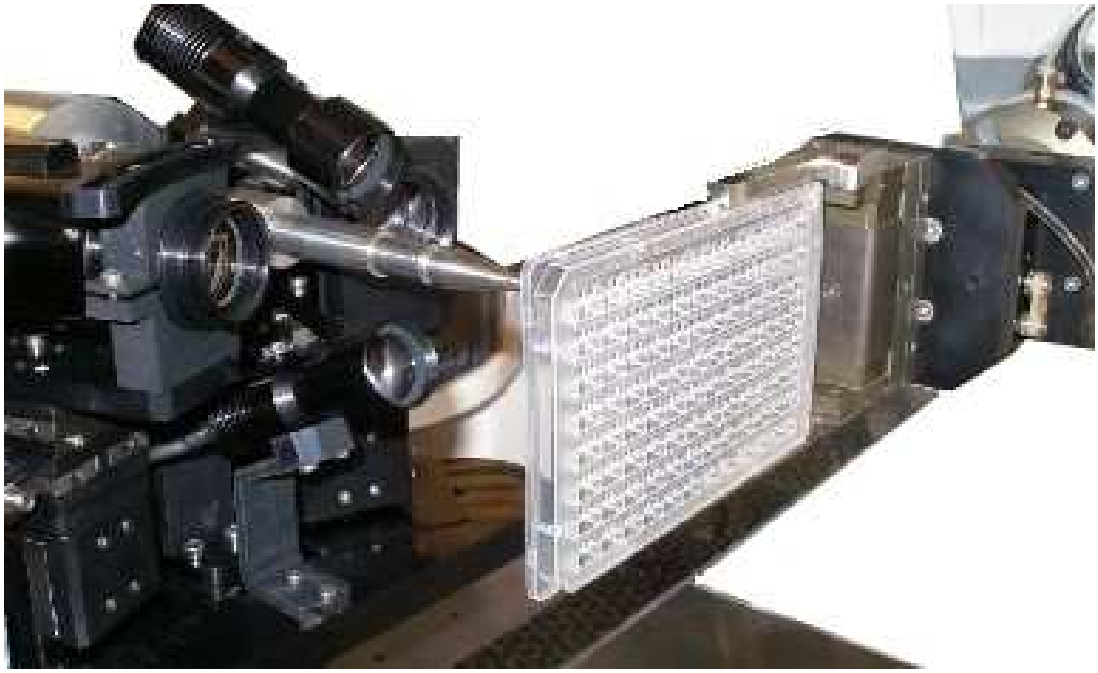

Figure 16: Robotic *in situ* X-ray diffraction with G-Rob

Nevertheless, for *in situ* diffraction method, crystals are exposed among their crystallization solution and also the crystallization plate. This induces higher background scattering in the diffraction patterns comparing to frozen sample method. Furthermore, to solve the tri-dimensional structure of studied macromolecules, a complete diffraction dataset is needed. The large angular sector required for a complete dataset may be challenging too, considering the geometrical limitations of the crystallization plates and the rapid decay of the crystal at room temperature. Several crystals may be needed to achieve a sufficient completeness of data (see Chapter II: Crystal Listing for automated *in situ* crystal centering and data collection). In spite of all, the *in situ* method has grown in importance with the possibility to solve structures. This method is highly recommended for protein crystals hardly cryo-protected or cryo-cooled (see 4.2. b) Cryo-protection and flash-cooling).

4. Preparations for X-ray diffraction

Depending on the X-ray diffraction strategy chosen to analyze crystals, preparations are different. Even though *in situ* X-ray diffraction is still not as widespread as frozen sample X-ray diffraction, yet both methodologies are detailed due to the potential of the *in situ* X-ray diffraction crystal analysis.

4.1. *In situ* X-ray diffraction

For crystals to be diffracted in their crystallization plates, the preparations should be done upstream the crystallization. Vapor diffusion sitting drop crystallization microplates[1] are the most adapted to *in situ* diffraction, as their geometry enables holding plates vertically without mixing the crystallization drop with the reservoir solution (see Figure 17).

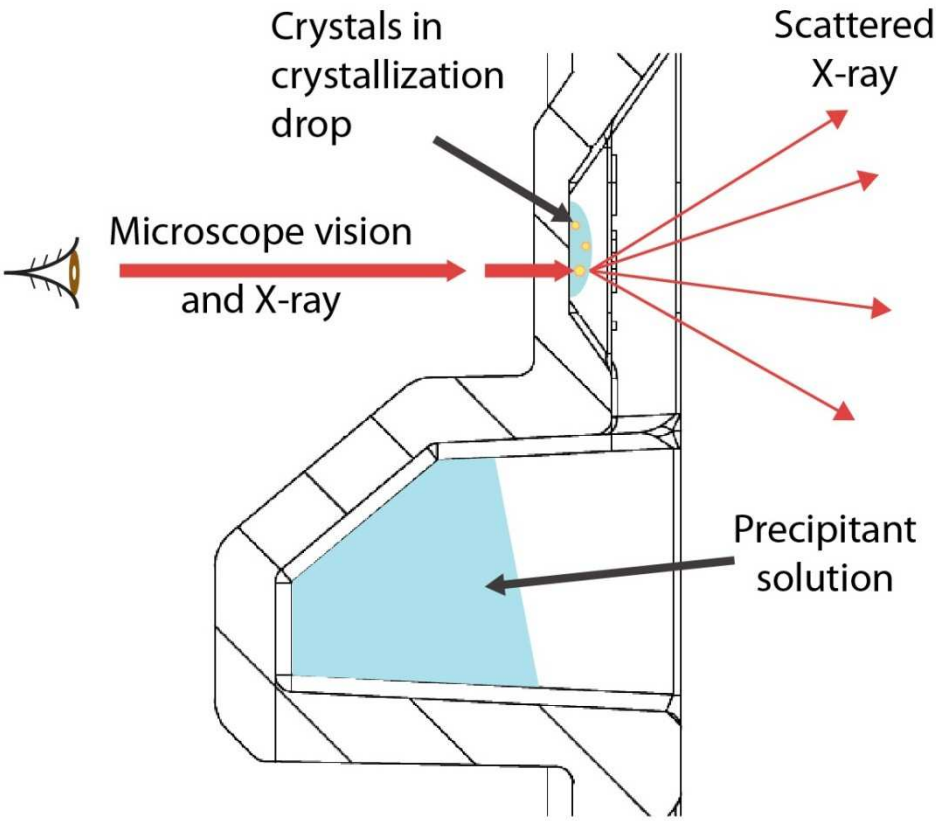

Figure 17: *In situ* X-ray diffraction in microplates

In order to reduce background scattering due to microplate's material, best crystallization plates with the lowest background scattering should be chosen. Depending on the geometry

[1] Standard dimensions are defined for microplates by American National Standards Institute (ANSI), http://www.slas.org/education/standards/ANSI_SBS_2-2004.pdf

and the material of micoplates, different scattering values are observed (see Figure 18). CrystalQuick™ X microplates show improved performance in this field with till about three times lower background scattering comparing to other classical crystallization plates.

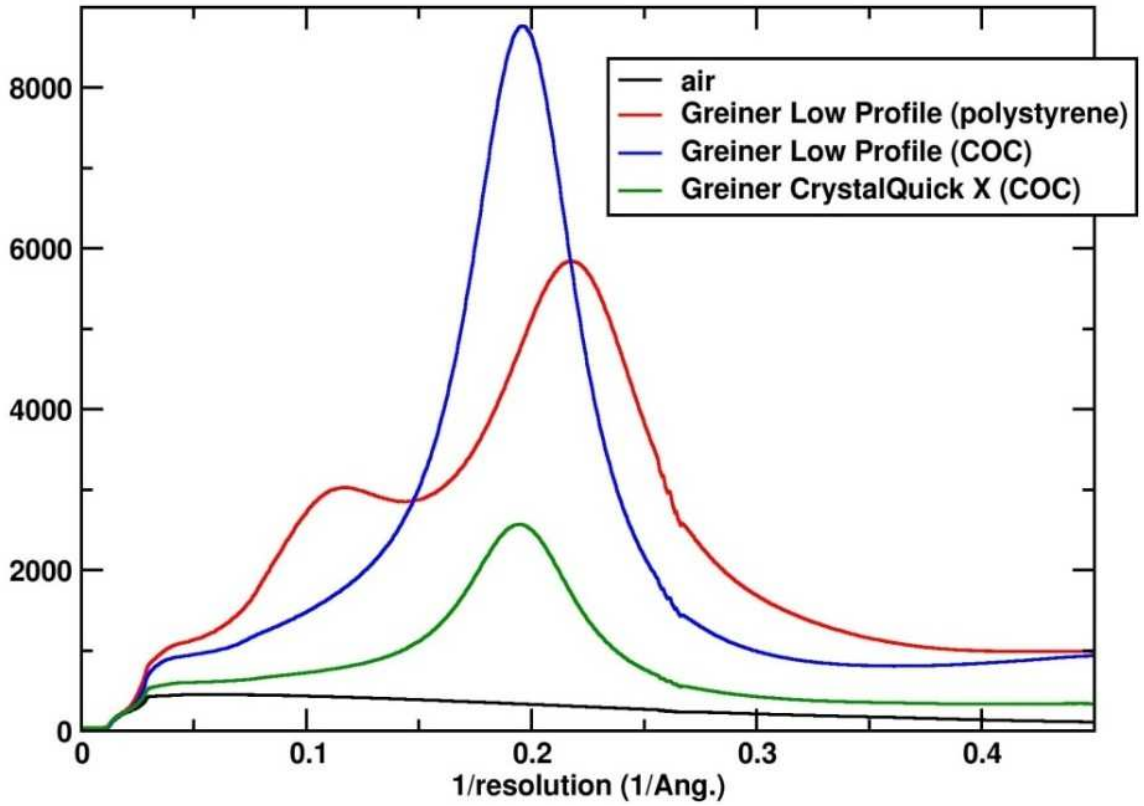

Figure 18: Background scattering curves of different crystallization microplates in arbitrary unit measured at FIP-BM30A beamline at ESRF

With robots capable to collect data *in situ*, crystallization plate geometries should be adapted for oscillation around crystals without obstructing the incoming X-rays to the crystal and also the scattered X-rays by the crystal (e.g. CrystalQuick™ X plates allow ±40° rotation around crystals, see Chapter II:2.3.).

After choosing best adapted microplate and crystallization, crystals are centered manually or through motorized human controlled instruments and X-ray diffraction data can be collected.

4.2. Frozen sample X-ray diffraction

Radiation damage in macromolecular X-ray crystallography is an age-old issue (Garman, 2010). The root cause of this damage is the energy lost by the beam in the crystal owing to either the total absorption or the inelastic scattering of a proportion of the X-rays as they

pass through the crystal. Cryo-cooling samples for X-ray diffraction show significant advantages in reducing the radiation damage by better preserving the crystal. Higher resolution data can more easily be obtained owing to the longer crystals order preservation and so collecting better diffracting quality data (Garman, 1999) and from fewer crystals for a complete dataset.

As for the frozen sample X-ray diffraction preparation of protein crystal grown in solution, freezing process is not straightforward. Crystals need to be transferred out of their mother liquid and prepared through different steps (see a) Harvesting). Additionally, collecting data at cryogenic temperatures could not only reduce the radiation damage but it can also reduce atomic movements and so contribute to higher resolution in collected data. As crystals contain from 27 to 78% solvent, the ice formation should be avoided. The ice formation of water molecules induces their volume expansion which damages protein molecules crystalline arrangement. Ice formation also induces crystalline water molecules that scatter X-rays, and so this is crucial to avoid. As a result, cryo-protecting solutions are diffused into crystals and fast cryo-cooling is managed (see b) Cryo-protection and flash-cooling) to turn water molecules into amorphous ice, with reduced volume expansion. Hence, as crystals are mounted on supports with transparent materials to X-ray, crystals can be exposed to X-ray for diffraction data collection.

The materials and methods used for each of these steps are described in the following.

a) Harvesting

This step concerns the transfer of crystals from their crystallization mother liquid into handy support for other preparative operations and X-ray diffraction. This task is more difficult than it seems as crystals are quite small, difficult to see and so fragile objects. The most common tool and method used nowadays is the use of micro-loops (Teng, 1990). When socking the loop into a liquid, a thin liquid film covers the loop by capillarity. The principle is to hang crystals into the liquid film on the loops (See Figure 19). Since few years, several commercialized loops in different material (e.g. Nylon and Kapton®) and dimensions (20 µm to 500 µm) are available.

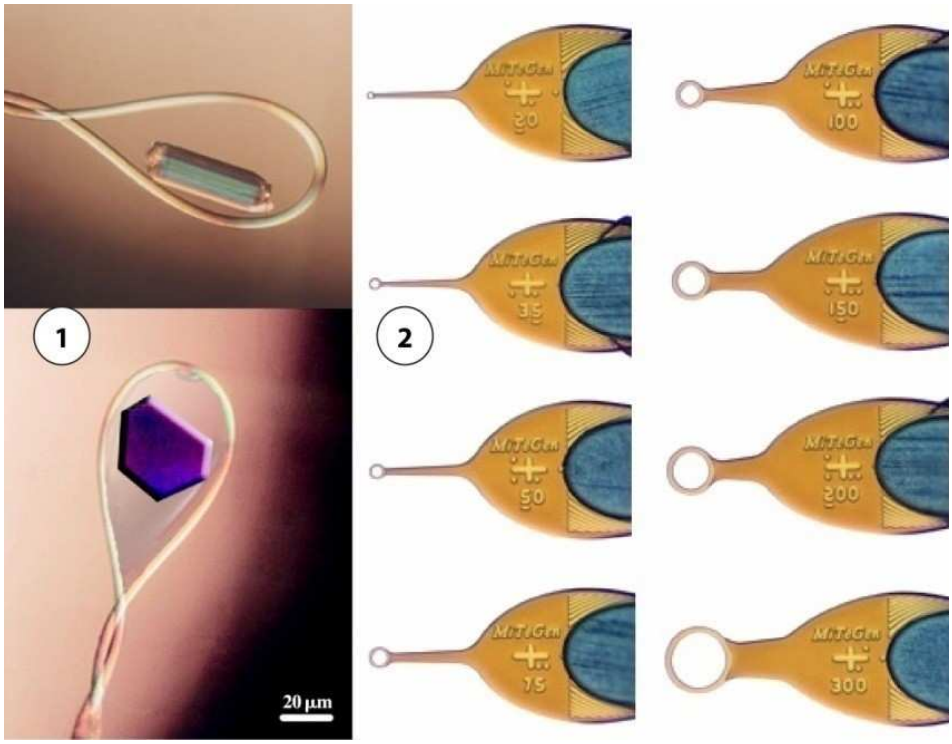

Figure 19: Harvesting loops, ① Nylon CryoLoop™ from Hampton Research, ② Kapton® MicroLoops™ from MiTeGen.

Nylon and Kapton® are respectively polyamide and polyimide materials with quite good transparency features to X-rays (see Figure 20). Thus, these loops are used as crystal manipulators and holders for all the preparation operations, from harvesting to X-ray diffraction.

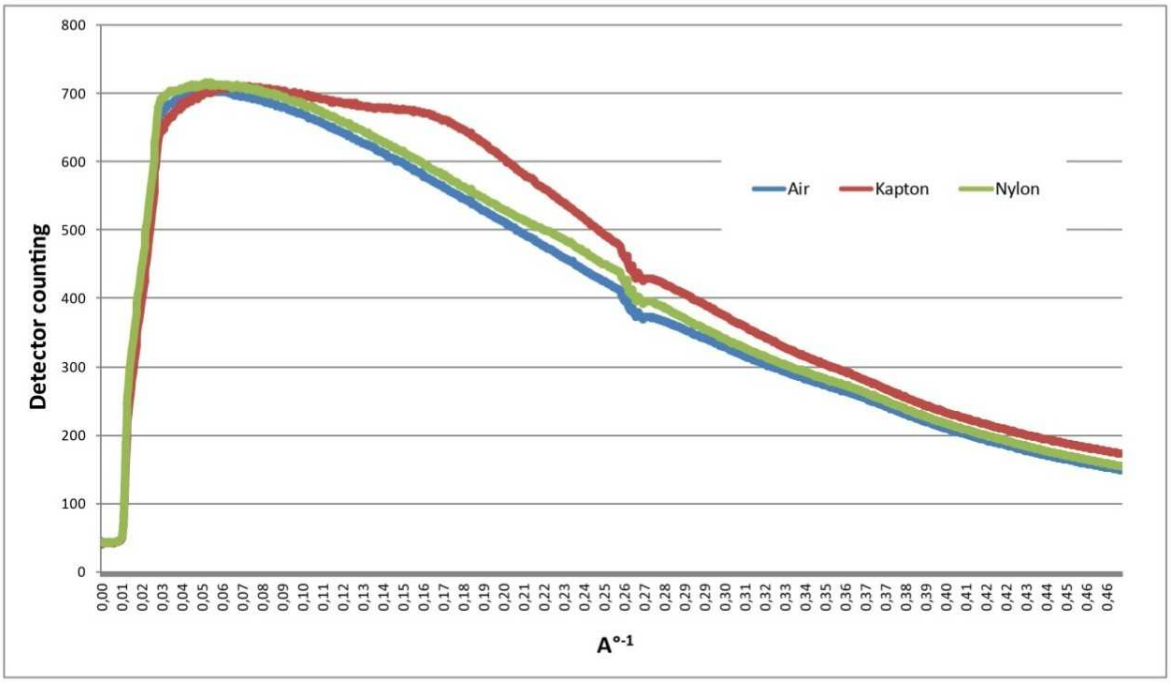

Figure 20: X-ray scattering curves of Nylon and Kapton®

In order to improve the manual handling of these loops and to adapt them to goniometer heads, loops are mounted on pins which are plugged into caps (See Figure 21). Caps manufactured with a magnetic base can be easily mounted on magnetic pens to better handle loops and also on magnetized goniometer heads.

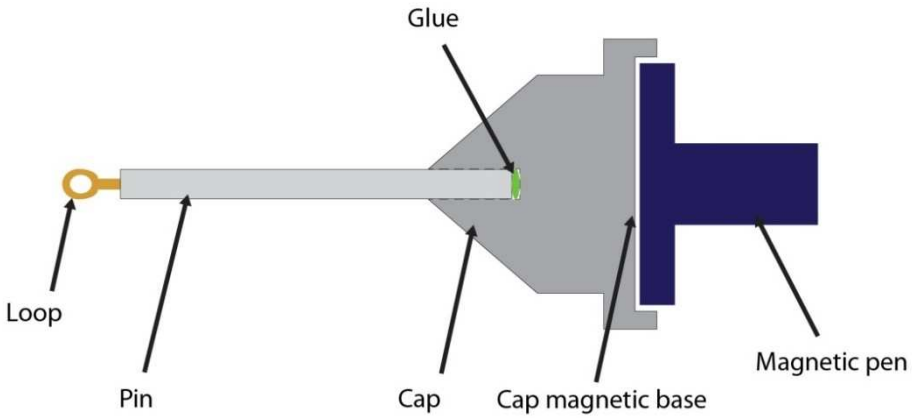

Figure 21: Loop + Pin + Cap + Magnetic Pen

In the last few decades these developments have made harvesting easier. In spite of all, this operation remains pretty difficult as crystals dimensions and fragility require accurate manual micromanipulation. Besides, crystallization drops states could worsen the difficulty of this task. Indeed, crystals are some time stuck to the bottom, or a thin layer of solidified solution covers the drop and many other complicated situations may be encountered. Consequently, harvesting crystals without damaging them is a challenging work.

b) Cryo-protection and flash-cooling

As mentioned above, the aim of cryo-protecting crystals followed by flash-cooling to cryo-temperatures is to prevent ice formation in crystals and also in loops' solution, for cry-temperature X-ray diffraction. Hereby we present how the addition of cryo-protecting agents and flash-cooling avoid ice formation in frozen aqueous solutions and so in crystals and mother liquid around crystals in the loops.

At atmospheric pressure, pure water melting temperature (T_m) is at 273 K, homogenous nucleation temperature (T_h) at 235 K and its glass transition temperature (T_g) is in between 130 K and 140 K (Rasmussen *et* MacKenzie, 1971). By lowering water temperature with slow cooling rates (few $K.s^{-1}$), ice nucleation points will appear and allow crystalline rearrangement of water molecules (See Figure 22). By flash-cooling to lower temperatures

than its glass transition temperature, water molecules state will change to vitreous or amorphous ice by transiting ice nucleation zone. As the transition is done fast enough, no nucleation or crystalline arrangement appears. For pure water, the required cooling rates are ~10^6 $K.s^{-1}$ (Brüggeller *et* Mayer, 1980). In the last few decades, numerous studies have been led to find best cooling rates possible in practice, with different cooling agents (Teng *et* Moffat, 1998; Walker *et al.*, 1998; Kriminski *et al.*, 2003). All studies agree in that 10^6 $K.s^{-1}$ cooling rates range is unachievable. This explains the necessity of using the cryo-protecting agents. Indeed, mixing water with Glycerol, Ethylene Glycol or MPD can reduce the required cooling-rates to lower than 10^2 $K.s^{-1}$ (Peyridieu *et al.*, 1996 and Warkentin *et al.*, 2006).

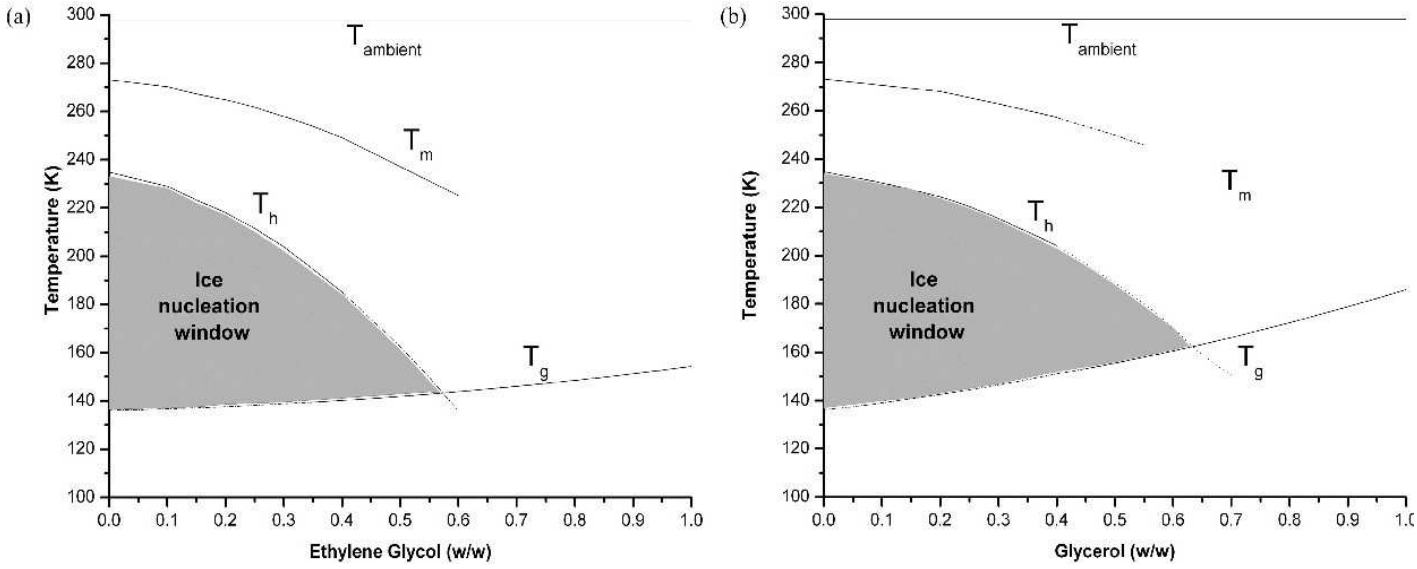

Figure 22: Phase diagrams of (a) Ethylene Glycol and (b) Glycerol at atmospheric pressure (Shah *et al.*, 2011)

For cryo-protecting, crystals are generally soaked into a cryo-protecting drop, right after the harvesting step. Crystals are very often released into the cryo-protecting drop. So it is needed to transfer the crystal out the drop once again before proceeding to the flash-cooling. Unfortunately, cryo-protecting agents can also harm crystals. They can affect proteins solubility or cause crystal cracking or dissolution. At high cryo-protecting agent concentrations crystal structures are unluckily exposed to changes (Cobessi *et al.*, 2005). Consequently, finding the optimum cryo-protecting solution is another challenge to the structure resolution at cryogenic temperature. This is of course one more reason to manage *in situ* experiments, when feasible.

Most commonly cryo temperature elements used at atmospheric pressure to improve flash-cooling crystals are liquid propane/ethane, liquid nitrogen and gaseous helium and nitrogen

stream. Even though high pressure could improve cryo-cooling, the most common methods used are at atmospheric pressure due to the complexity of high-pressure process and instruments (Kim *et al.*, 2005 and Thomanek *et al.*, 1973).

Helium gas stream instruments can reach low temperatures of about 10 to 30 K. Open flow cryo temperature helium stream has been used for cryo-crystallography (Hanson *et al.*, 1999). But helium remains expensive for random experiments. With liquid propane, quite good results have been obtained, nevertheless due to its inflammability high security precautions are needed. Propane is used rarely in very specific cases (e.g. flash-cooling in anaerobium incubators). Nitrogen gas (100 K to 120 K) and liquid (77 K) are highly popular cryo elements used in cryo-crystallography thanks to their availability, low cost and instrumental simplicity. In most cases, depending on crystals, 10 to 30% w/w Glycerol or Ethylene Glycol allows good quality flash-freezing with liquid or gas nitrogen.

For gas cryogenic elements, generally crystals on loop are exposed suddenly to the cryo temperature gas stream thanks to a shutter cutting the stream. For liquid cryogenic elements, the crystal on loop is plunged directly into the liquid.

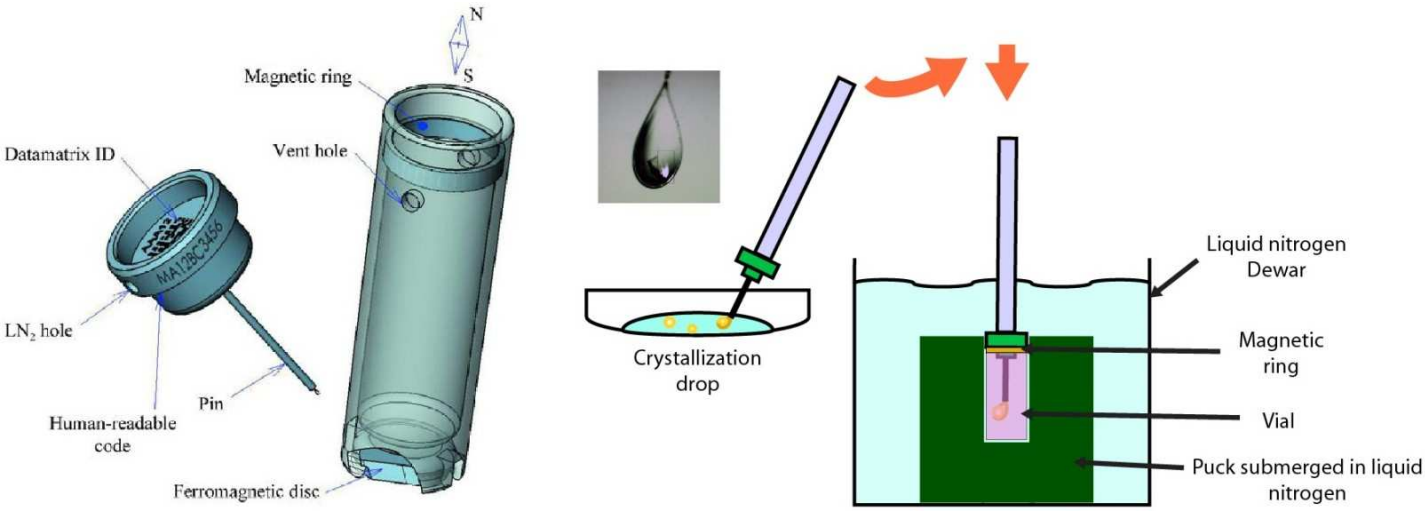

Figure 23: Harvesting, flash-cooling and storage into liquid nitrogen thanks to Pin + Vial + Cap + Puck

Once crystals frozen, they can be stored in liquid nitrogen. To keep frozen samples at cryo-temperatures while transferring them, a cylindrical reservoir called vial is used to cover the cap keeping the loop with crystal in liquid nitrogen (see Figure 23). A magnetic ring at the

top of the vial maintains the vial on the cap. Vials among caps are stored into packs which are disposed into Dewar[1] flasks.

Many different pucks for vials among caps storage have been developed, to facilitate carrying or shipping frozen samples from laboratories to synchrotrons and also for automated sample transfer to goniometer for diffraction measurements, as described in the following section.

c) Diffraction measurements

Crystals can be diffracted whether *in situ* at room temperature in crystallization microplates (Jacquamet *et al.*, 2004) or by preparing them for frozen sample diffraction at cryo temperatures. For both the aim is to collect as much as good quality data possible in order to be able to solve the structure with the highest resolution and completeness through data processing, structure model building and structure refinement.

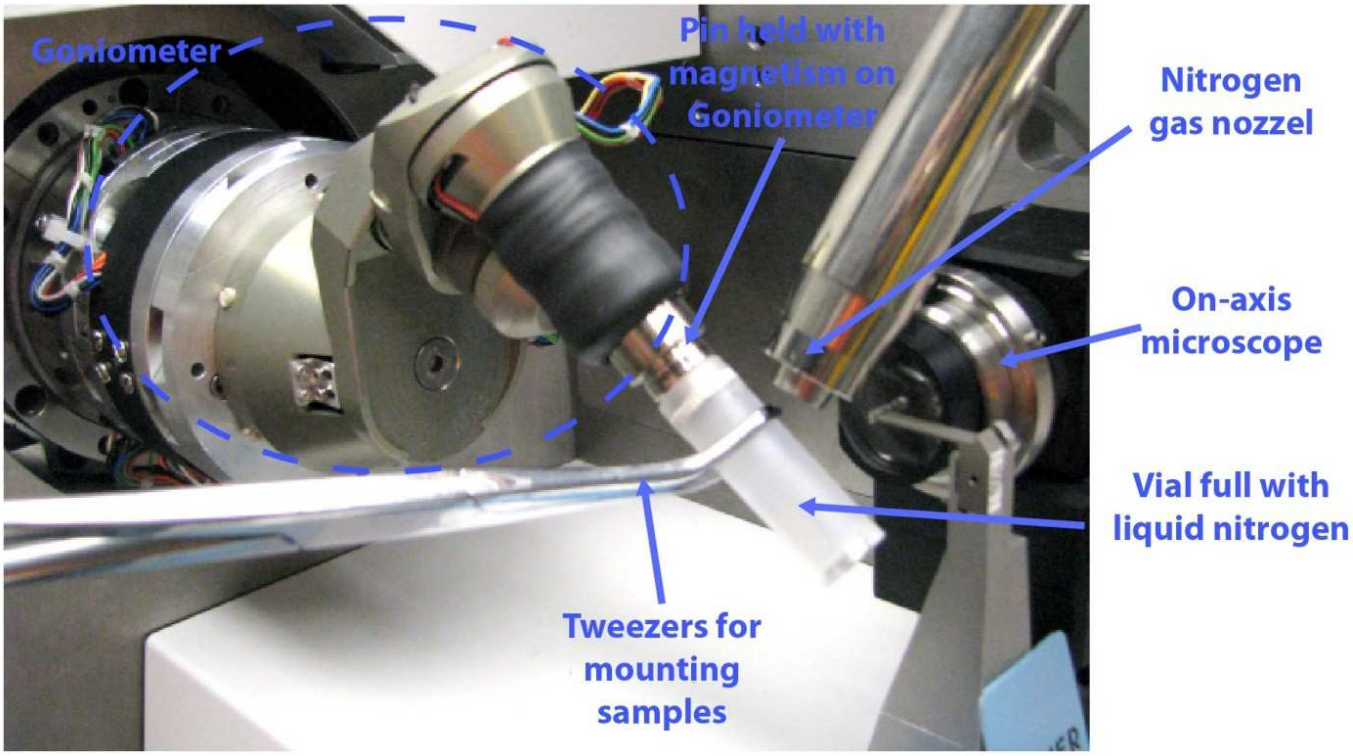

Figure 24: Manual mounting/dismounting frozen sample on MD2 goniometer in Kappa ≠ 0 configuration (Macromolecular crystallography beamline at BESSY II, Berlin)

[1] John Dewar, 1842-1923, invented Dewar flask, a reservoir with good thermal insulation, at Cambridge University.

Frozen samples can be mounted on goniometer heads manually (see Figure 24). The goniometer heads are magnetized to hold caps once in touch with the cap's base. The vial plus its cap is presented to the goniometer magnetized head. This maintains the cap in position. Then vial full with nitrogen liquid is removed. A nuzzle blows the cryogenic temperature (100 K to 120 K) nitrogen gas stream towards the sample. This keeps the sample frozen during the whole experiment. Generally, a microscopic view of the sample calibrated with the beam position and two translations on the goniometer head allows centering the sample accurately into the spindle axis and so into the beam.

X-ray sources combined with optics and detectors, play an important role on the achievable resolution and also on experimental time. The higher the beam intensity, the less exposure time is needed for intense spots at high resolution on diffraction patterns. At the other hand, electronic detectors are capable of high-throughput data collection. Today's microfocus beamlines at synchrotrons combined with highly performance electronic detectors, enable collecting a complete dataset in even less than a minute. In order to follow this rhythm and to fully benefit from these facilities, automating the sample preparation and manipulation steps are required.

5. Why high-through put crystallography

5.1. Stakes and needs

In the last two decades, interest in atomic structure of proteins continuously increased. One of the most contributing steps has been the use of anomalous signal from selenium, with selenomethionine, and the MAD (Multi-wavelength Anomalous Diffraction) method to solve the phase problem (Hendrickson *et al.*, 1990; Weis *et al.*, 1991). Moreover, in terms of means, progress in chemical and molecular biology have increased the possibility to produce more and more proteins with greater cadence. With genome sequencing developments, the number of proteins of interest has risen. Besides, the number of applications of protein structures is also increasing from the classical drug design to structure-based drug design (Williams *et al.*, 2005; Grey *et* Thompson, 2010), with pharmaceutical companies investing on macromolecular crystallography beamlines (e.g. beamline X06DA at Swiss Light Source) and plant engineering. Thus, the number of proteins to study continues to grow and the need of faster structural studies and so high throughput structural biology has become a necessity.

5.2. Responses

With arisen demands in structure resolution, more and more synchrotrons with macromolecular crystallography dedicated beamlines have been built world widely. The X-ray beam intensities along with instrumentation developments allow automating and accelerating increasing experiments. In the near future, intense synchrotron beams combined with high-performance electronic detectors could achieve a complete dataset collection in only few seconds.

5.3. State of the art in automation

a) Crystallization

As mentioned before, crystallization robots can achieve very rapid and accurate liquid dispensing. They can manage dispensing a whole 96-well plate with crystallization drop of 100nL in less than a minute. Therefore, crystallization assays become less time consuming and require now reduced amount of protein. Large screening assays are now possible, increasing the potency to obtain diffracting crystals.

b) Sample changers and electronic detectors

About two decades ago, to save one single diffraction pattern with electronic detectors, took more than 15 seconds. Today, higher resolution detectors have dead time of few milliseconds.

Many attempts have been made in developing automated frozen sample changers that transfers crystals from a liquid nitrogen storage Dewar to a goniometer. First system developed was the SAM system at Stanford Synchrotron Radiation Laboratory (SSRL). Rigaku has commercialized a robotic system developed at Abbott laboratory in the name of ACTOR™, since 2001. The Automounter has been developed also in early 2000 at Berkeley at ALS. Other systems were born in Europe as well at the same time, such as the SC3 system. This system built at EMBL at Grenoble in France and commercialized by Maatel. Two robotic systems based on 6-axis robotic arms were also built in Grenoble at ESRF, at FIP-BM30A beamline: Cryogenic Automated Transfer System (CATS, commercialized by IRELEC) and G-Rob (commercialized by NatX-ray). All these systems contributed to automating X-ray diffraction experiments and thus stimulate the speed of experiments.

c) Data processing and structure resolution

With the computing powers increasing in hardware and also software developments for data analysis (Kabsch, 1988; Leslie, 2006), structure resolution has been quite simplified. Software as Elves (Holton *et* Alber, 2004) is able to go from data processing to refinement. Using automatic procedures, Phenix (Adams *et al.*, 2010) can handle for example anomalous data to find the heavy atom positions, calculate and improve the phases, in order to rebuild and refine the structure, while ARP/WARP (Perrakis *et al.*, 1999) can build and refine the structure. With these hardware and software available on beamlines and also in laboratories, structures can come through in few hours, comparing to two decades ago when same tasks took months or years.

5.4. Missing steps in automation

In structural biology, from genome sequencing to structure resolution, almost all major steps has been automated, increasing the output of this science. Yet few essential steps remain manually operated.

Firstly, for *in situ* diffraction crystals have to be centered one after another. Knowing the high number of crystals that are needed to be centered in a row for screening, this step forms the bottleneck of a fully automated procedure. As *in situ* diffraction in microplates has shown its importance in screening crystals and even collecting complete datasets, no developments have been reported to fully automate this process. In chapter II of this manuscript, a new system completing fully automated pipelines for *in situ* analysis of crystals in screening microplates and also data collection for structure resolution is presented.

Secondly, as for the frozen samples, the preparation steps such as harvesting, cryo-protecting and flash-cooling remain manual and critical to high-through put crystallography. Many developments have been reported in the last few years in attempting to automate crystal harvesting and also cryo-protection and flash-cooling of crystals with more or less complete and adapted systems (see Chapter III:1. Introduction). In spite of all, these systems didn't succeed in filling the gap for a fully automated macromolecular crystallography pipeline. In chapter III of this manuscript, a new system (REACH: Robotic Equipment for Automated Crystal Harvesting) capable of harvesting protein crystals thanks to a micro-gripper, cryo-protection and flash-cooling is presented. The setup developed is integrated to the beamline FIP-BM30A for direct data collection or transfer on loop and storage into liquid nitrogen Dewar by local or remote users.

Chapter II: Crystal Listing for automated *in situ* crystal centering and data collection

As one of the two major developments during my PhD, the Crystal Listing allows achieving fully automated *in situ* crystal centering and data collection for samples in microplates. Based on image processing crystal centering software, this function can be easily adapted to any *in situ* X-ray diffraction apparatus. Thus another step forward has been made towards high-through put macromolecular crystallography. This work has been clearly a result of developments, studies and experiments led during my PhD. The mechanical, automation and software developments of this system and also the assessment experiments have been lead and realized as part of my PhD under supervision of Dr Jean-Luc Ferrer with some technical contributions of coauthors. As for X-ray diffraction data processing, data clustering and structure refinement and resolution, they have been managed majorly by Hugo Lebrette and also by Dr Jean-Luc Ferrer. The following scientific report has been submitted to *Acta Crystallographica section D*, on 1 August 2012.

"You gotta be pretty desperate to ... (do) it with a robot."

Homer Simpson, *The Simpsons*

Crystal Listing for automated crystal centering and *in situ* X-ray diffraction data collection

Yaser Heidari[1], Hugo Lebrette[2], Xavier Vernede[1,2], Pierrick Rogues[3], Jean-Luc Ferrer[1,4]

1 Institut de Biologie Structurale Jean-Pierre Ebel, Groupe Synchrotron; Commissariat à l'Energie Atomique et aux Energies Alternatives, Centre National de la Recherche Scientifique, Université Joseph Fourier; F-38027 Grenoble cedex 1; France.

2 Institut de Biologie Structurale Jean-Pierre Ebel, Groupe MetalloProtéines; Commissariat à l'Energie Atomique et aux Energies Alternatives, Centre National de la Recherche Scientifique, Université Joseph Fourier; F-38027 Grenoble cedex 1; France.

3 NatX-ray; 38400 Saint Martin d'Hères; France.

4 Correspondence; Email: jean-luc.ferrer@ibs.fr; Phone: +33 4 38785910; Fax: +33 4 38785122

Abstract

As High Throughput Protein Crystallography has earned its importance in crystallization platforms, the need to develop and invest in adapted and automated equipments for crystal analysis has become essential. The trend today is to use the smallest sample amounts to screen the highest possible number of conditions but it often leads to the production of very small crystals. Robots have been developed to reduce the time spent in preparing crystallization solutions and also in screening crystallization plates. Crystallization microplates have been conceived with various geometries to improve the output. Therefore the crystals to be analyzed need to be harvested, cryo-protected and flash-cooled which are quite challenging steps, as the crystals' reaction to these delicate operations is unpredictable. *In situ* X-ray diffraction analysis has become a valid option for these operations and a growing number of users apply it for crystal screening and also to solve structures. Robots and improved crystallization plates facilitate the *in situ* analysis. Nevertheless, because of radiation damage at room temperature, a large number of crystals have to be analyzed to obtain a complete dataset by merging data. In this high throughput approach, centering crystals automatically relative to the beam represents the bottle-neck of *in situ* analysis. In this article we report a new methodology that uses mostly existing

instruments to define local geometry coordinates for each crystal in the plate for an automated crystal centering into the beam, *in situ* crystal screening and data collection.

1. Introduction

To optimize the study of protein structures using X-ray crystallography, it is crucial to improve the X-ray diffraction data quality. So far, the commonest method remains cryo-crystallography, by mounting and flash-cooling the protein crystals in loops (Teng, 1990). Nevertheless, crystallographers need to screen their crystals to select the best ones regarding their diffraction quality and resolution (Bingel-Erlenmeyer *et al.*, 2011). This selection step is time-consuming because it requires to mount, cryo-protect and flash-cool each crystal. Furthermore, many macromolecular crystals and particularly membrane protein crystals get damaged by the cryo-protecting and flash-cooling procedure applied before X-ray data collection. In this context, room temperature *in situ* X-ray diffraction plate screening represents an attractive alternative approach (Jacquamet *et al.*, 2004). In the last few years, manual and automated systems for *in situ* X-ray analysis have been developed, such as PX Scanner (Agilent Technologies), jig (Hargreaves, 2012) and Cryogenic Automated Transfer System (CATS) (Ohana *et al.*, 2004). However, many experimental features in current X-ray *in situ* analysis are not optimized. Firstly, because a part of the crystallization plate and the crystallization drop are exposed to X-ray along with the crystals, a high background X-ray scattering is generated affecting the quality of the diffraction data. Secondly, a large number of crystals is often needed to complete the dataset as radiation damage limits the exposure time of each crystal.

Here, we report the use of a new crystallization plate called CrystalQuick™ X (Bingel-Erlenmeyer *et al.*, 2011), made of a specific material and with a geometry that reduces background scattering compared to other commercially available crystallization plates. Furthermore, in order to automate crystal centering for *in situ* diffraction analysis, we also report the development of a procedure we called Crystal Listing. With this method the user can create lists of crystals selected in crystallization microplates and can launch automated crystal centering and X-ray data collection. This system has been developed at the FIP-BM30A beamline at ESRF and has been tested on the G-Rob home laboratory facility of Professor Cole at Ecole Polytechnique Fédérale de Lausanne (EPFL). Our approach may have a general application because it can be easily implemented on beamlines equipped with *in situ* robotics.

2. Materials

For the experiments reported in this manuscript, a few instruments and software were developed or used which are detailed below.

2.1. Visualization Bench

This specially developed instrument is a fully automated plate screening apparatus, with an inverted microscope (Figure 25①). A microplate holder combined with three direction-motorized translations and motorized zoom, allow for automated plate visualization. Wells are inspected from below, perpendicular to the plate. Five adjustable positions are available for the two installed light sources: three for the front light and two for the back light positions. Adapted electronics, software and a Graphical User Interface (GUI) enable the automation of the Visualization Bench.

2.2. G-Rob

After the success of Cryogenic Automated Transfer System (CATS, Ohana *et al.*, 2004) developed based on a 6-axis robotic arm, its capabilities were exploited to their limits to develop a new system called G-Rob at the FIP-BM30A beamline of the ESRF. This system can transfer frozen samples from a storage Dewar to the beam and is accurate enough to be used as a goniometer for frozen loop samples and capillaries in order to collect X-ray diffraction data directly (Jacquamet *et al.*, 2009). The goniometer accuracy of G-Rob system has been shown to have a sphere of confusion with a better than 15 µm radius. G-Rob, thanks to its 6-axis arm, can also manage "in plate" *in situ* data collection (Figure 25②). G-Rob with its tool for standard crystallization microplates can transfer them from a storage hotel into the beam position. All six axis of the robot arm are then used to rotate the plate around a horizontal axis intersecting with the beam. Depending on the geometry of the crystallization plates and the instrumentation set-up, G-Rob can rotate microplates up to 80° (±40°) around crystals for *in situ* diffraction data collection.

The G-Rob home laboratory system at EPFL was used for these experiments. This system is equipped with a Rigaku 007HF microfocus rotating anode X-ray source, coupled to a Photonic Science “Image Star 165” CCD detector, allowing data collection on small crystals to 2 Å.

2.3. CrystalQuick™ X microplate

To better fulfill *in situ* diffraction requirements and improve the quality of collected data, a new crystallization plate, CrystalQuick™ X (Bingel-Erlenmeyer *et al.*, 2011) has been co-developed by Greiner Bio-One (Reference number: 609890) and the FIP-BM30A beamline at ESRF (Grenoble, France). This plate is now commercialized by NatX-ray (Saint Martin-Hères, France). CrystalQuick™ X is a 96-well microplate with two wells per reservoir for vapor diffusion experiments using sitting drops. Crystals can be made to diffract from below the wells with an oscillation range of 80° (±40°). To avoid obstructing the diffraction patterns the side walls of the wells have a 145° angle with respect to the bottom of the wells. Adapted material is used for this microplate and the thickness of the bottom of wells has been reduced to the manufacturing limits (200 µm to 300 µm thick). These plates generate a three times lower background X-ray scattering when compared to other commercialized crystallization plates. References corresponding to each well are engraved on this microplate, simplifying the well referencing while screening the plate.

2.4. Samples

Protein crystal samples were prepared for automated crystal centering assessment and also to test the automated *in situ* data collection procedure. Chicken egg-white lysozyme from Hampton Research Lysozyme Kit (Reference number: HR7-108) was crystallized by mixing 150 nL of a 20 to 100 mg/ml protein solution in 0.02 mM sodium acetate trihydrate pH 4.6 with 150 nL of the 30% (w/v) polyethylene glycol monomethyl ether 5000, 1.0 M Sodium chloride, 0.05 M sodium acetate pH 4.6 (Hampton Research, product number HR2-805) reservoir solution. The periplasmic binding protein NikA from *Escherichia coli* was also used. Its apo form was produced and purified as previously described (Cherrier *et al.*, 2008). Prior to crystallization, 10 mg/mL apo-NikA solution was incubated overnight at 4°C with a two-fold molar excess of Fe(III)-EDTA. Sitting drops were prepared by mixing 500 nL of this protein solution with 500 nL of the 0.1 M sodium acetate pH 4.7, 1.50 to 1.90 M ammonium sulfate reservoir solution (Cherrier *et al.*, 2005). Protein samples were crystallized in the CrystalQuick™ X plate. For NikA-FeEDTA, drops were dispensed manually whereas for Lysozyme, the mosquitto® liquid handling robot from TTP Labtech Ltd (UK) was used.

3. Methods

3.1. Crystal Listing software

This software was developed on the Visualization Bench using CrystalQuick™ X microplates and is integrated in the Visualization Bench GUI. It is based on an image recognition program we developed in C programming language with the OpenCV image processing library (Bradski & Kaebler, 2008), which uses crystallization plate's well number marks. Each time "Move to selected well" button of the GUI is pressed, first the local mark of the selected well is approximately centered in the image. Then the program recognizes the mark, calculates the well center and moves the plate to center the well accurately in the microscope's field (Figure 25③). This position is considered as the well local reference (0;0). With a right-click on an interested positions or crystals in the well, the coordinates of that position referring to the well center and the zoom value are saved into a data base as a square image of 500 µm a side, centered on the clicked position (Figure 25④). The stored data can be used for two purposes: automated crystal centering for automated *in situ* data collection with *in situ* X-ray diffraction apparatus such as G-Rob or automated crystal growth monitoring with an automated microscope such as the Visualization Bench system.

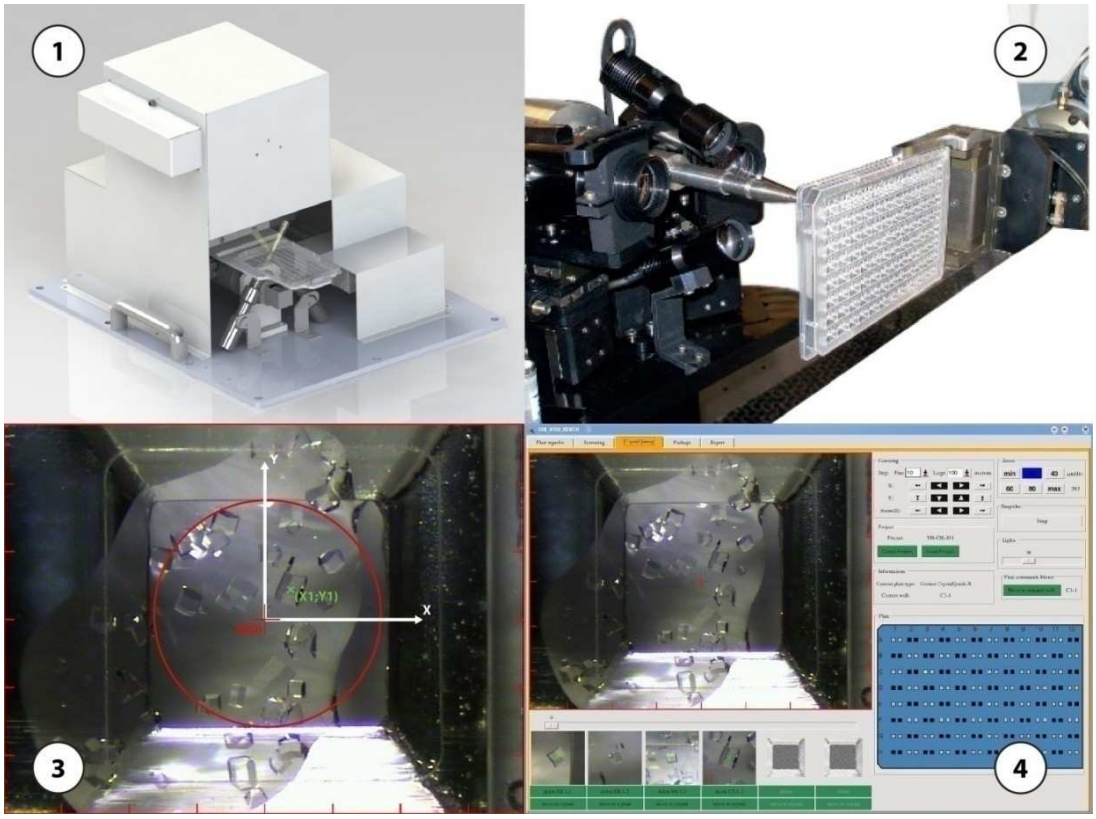

Figure 25: ① Visualization Bench, ② *In situ* X-ray diffraction with G-Rob, ③ Automatically centered well with crystal coordinates in the local reference, ④ Crystal Listing tab in Visualization Bench GUI

3.2. Automated crystal centering and *in situ* X-ray diffraction software in G-Rob

The Crystal Listing has been implemented in G-rob GUI, in order to give users the ability to save crystals on G-Rob and also to modify saved lists on Visualization Bench. Software developments are been incorporated in the GUI of G-Rob to allow loading a list of positions from the database to initiate precede automated crystal centering and X-ray diffraction measurements of each position on the list. Diffraction parameters are set before launching the procedure, for either single-image or oscillation diffraction data collection. The listed positions are then applied sequentially. For each position, the G-Rob moves the plate to view the entire well mark and performs the shape recognition analysis mentioned above, to center the well. Because the center reference (0;0) of the well is has been previously determined, the robot centers the crystal in the beam by reading the corresponding coordinates from the database. Initial auto-focusing is performed at this position, based on the Root Mean Square Deviation algorithm developed with OpenCV library (Bradski & Kaebler, 2008) in C programming language. The motorized zoom of the G-Rob microscope shifts the zoom value in which the position is saved, by reading it from the database, and further auto-focusing is performed on the crystal. In order to improve the centering, the saved image of the listed position is used for image correlation with G-Rob microscope view, to correct any possible errors in the positioning of the crystal (Figure 26). Once the correlation and crystal centering is done, data collection is initiated by the G-Rob diffraction. All these steps are run automatically on all the listed positions of the loaded list.

In all the experiments described in this manuscript, a single wavelength 1° oscillation X-ray diffraction data collection strategy has been used for data collection.

3.3. Data processing

Diffraction data reduction was carried out in a semi-automated manner using the xdsme script developed by Pierre Legrand, based on XDS (Kabsch, 2010) and Pointless (Evans, 2006; Winn *et al.*, 2011). For each crystal dataset, the optimum frame number for data processing was chosen using both R_{sym} and I/Sigma values. Using this criterion some datasets were not considered and for others 3 to 5 frames were used. Chosen data from different crystals were merged and scaled two by two, using XSCALE (Kabsch, 2010), in order to plot a clustering

tree based on the R_{merge} factor of each pair of data. The data to merged and scaled together were selected in order to generate an optimized dataset with particular attention being paid to completeness and the R_{merge} factor. The mean weighted cell parameters were obtained using Cellparm (Kabsch, 2010). The atomic coordinates of the lysozyme and NikA-FeEDTA$(H_2O)^-$ X-ray models (PDB ID codes 1LZ8 and 1ZLQ respectively) were used as starting models for molecular replacement. Phasing and crystallographic refinement were performed using Phenix (Adams *et al.*, 2010). The three-dimensional models were examined and modified using the graphics program COOT (Emsley & Cowton, 2004).

4. Results & Discussion

4.1. Saving crystal information with Crystal Listing

The name of the project is entered in a pop-up window activated by clicking on the "Create a project" button on the Crystal Listing tab of both Visualization Bench and G-Rob GUIs. This name is used to create a specific folder in the database that includes the crystal list file containing the crystal well number, the number of the crystal in the well, the crystal coordinates with respect to the local reference (0;0), and zoom and lightening values. This folder also contains crystals images.

The CrystalQuick™ X microplates prepared with Nika-FeEDTA and lysozyme crystallization conditions were screened using Crystal Listing. A few lists of crystals from each plate were created in both G-Rob and Visualization Bench. These lists contained from 30 to 80 crystals. The first goal was to assess the accuracy of the crystal centering process in the G-Rob system, when crystals were saved on Visualization Bench and on G-Rob. The centering accuracy could change depending on whether crystals were saved on either Visualization Bench or G-Rob system. The second goal was to assess the feasibility of coupling the automated crystal centering with *in situ* diffraction in G-Rob. The crystals were then centered and X-ray diffraction data were collected automatically in the in house G-Rob system at EPFL in order to solve the structure of each protein.

4.2. Accuracy assessments

A list of 44 lysozyme crystals was created using the Visualization Bench. This list was then loaded to G-Rob. An automated program on the G-Rob system centered crystals one by one following the procedure described above (3.2. Automated crystal centering and *in situ* X-ray diffraction software in G-Rob) with the only difference being that the crystal image correlation step was removed. This was due to differences in microscope angle between the Visualization Bench and the G-Rob system of EPFL used for these experiments. In this system, the angle between the viewing axis and the perpendicular axis to the bottom of the wells was 15° whereas in the Visualization Bench, the viewing axis was perpendicularly to the bottom of wells. This difference precluded the use of image correlations between the two systems. Nevertheless, even without the use of image correlation and correction step the crystal centering in G-Rob was remarkably accurate.

Centering errors were measured by comparing the center of the images (red crosses materializing their centers, see Figure 26) of taken after the G-Rob automated crystal centering with images from the database (green crosses materializing their centers, see Figure 26). The radius average error was about 40 µm with a Standard Deviation of 17 µm.

Number of samples		44 crystals
Average X error	Average Y error	Average Radius error
33 µm	21 µm	41 µm
Std Dev X error	Std Dev Y error	Std Dev Radius error
17 µm	14 µm	17 µm

Table 2: Automated centering accuracy assessment of crystals saved on the Visualization Bench and centered automatically with the G-Rob.

From the 44 automated crystal centering attempts only 2 were unsuccessful. In these two cases, it was found that the crystal moved between Crystal Listing data acquisition and the automated centering. This displacement was due to the fact that the Crystal Listing operation was done on the Visualization Bench with the microplate held horizontally, whereas the G-Rob system manipulates microplates vertically both in front of its microscope and in the X-ray beam.

In order to decimate setup differentiation and add the image correlation step correcting the centering error, a list of 79 lysozyme crystals were prepared using the Crystal Listing tab of G-Rob GUI. The same automated crystal centering process as above was used. In addition, the image correlation step was added into a correction software loop. This loop was limited to 5 iterations or less than 5 µm centering error.

Number of samples		79 crystals
Average X error	Average Y error	Average Radius error
3 µm	3 µm	5 µm
Std Dev X error	Std Dev Y error	Std Dev Radius error
2 µm	2 µm	3 µm

Table 3: Automated centering assessment of crystals saved with the implemented Crystal Listing function into G-Rob and centered automatically with the G-Rob.

Centering error measurements showed an average radius error of about 5 µm with a standard deviation of 3 µm (Figure 26). Only 2 out of 79 crystals were miss-centered. The two failed cases were due to well reference blurred images which impaired the shape recognition step. A reliable procedure could be fully recovered using an auto-focusing step on well references before shape recognition image processing.

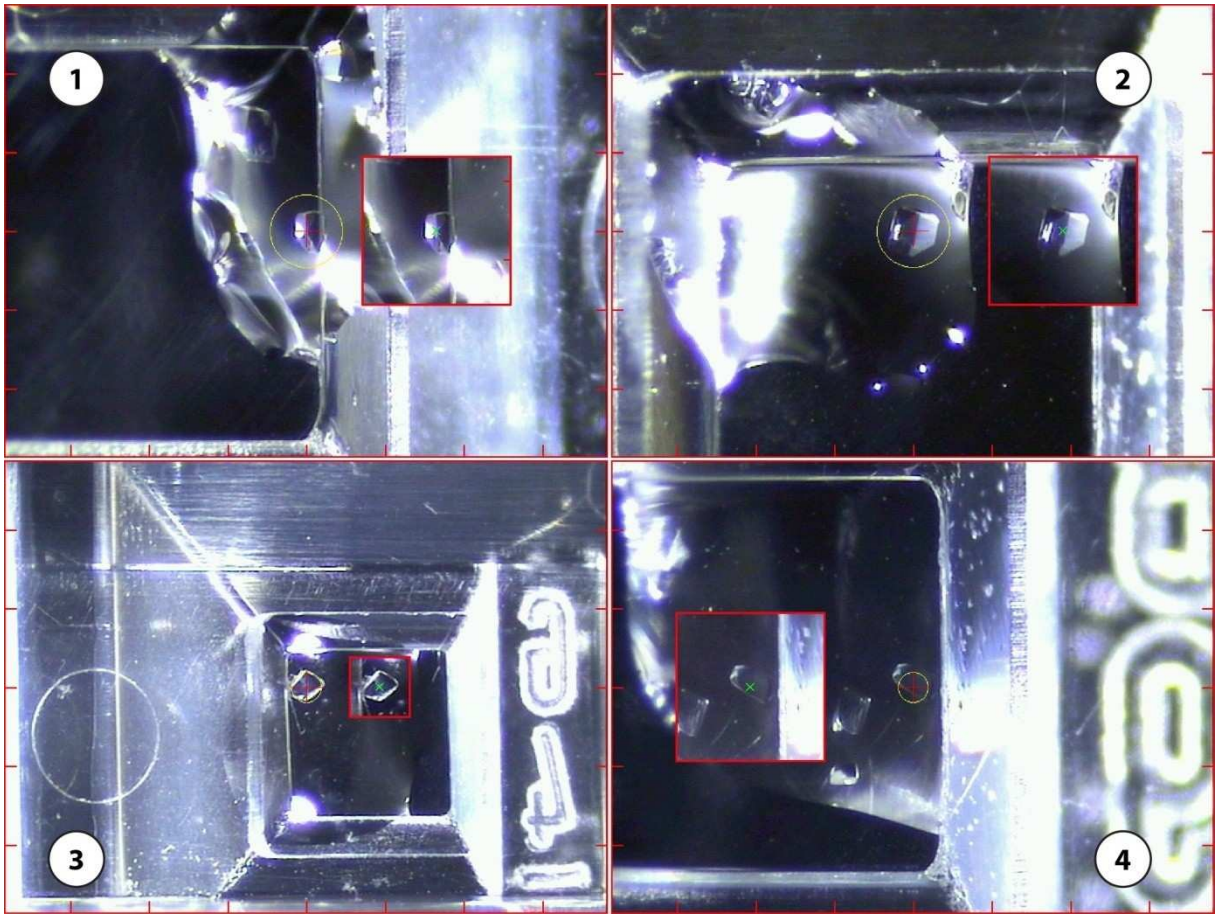

Figure 26: The four examples show small images of 500 µm x 500 µm saved during the Crystal Listing procedure with G-Rob with the green crosses materializing their centers and bigger images are taken after crystal centering with G-Rob with red crosses materializing their centers and the yellow circles materializes the beam. ① and ② correspond to NikA-FeEDTA crystals. ③ and ④ are lysozyme crystals. These images correspond to crystals diffracted and used for structure resolutions.

4.3. Automated *in situ* data collections and data analyses

Series of crystal lists, from 40 µm to 450 µm crystals from both proteins, lysozyme and NikA-FeEDTA, were saved in the "in house" G-Rob at EFPL. These lists were used in the automated crystal centering and data collection procedures in order to solve both structures by only using *in situ* data collection. For lysozyme, 24 crystals were listed and three to six frames were recorded for each one, following procedures detailed above. Three to five 1° oscillation diffraction patterns from 8 of the 24 exposed crystals were used to scale and merge data leading to a 2.1 Å resolution dataset with 71.6% completeness. Concerning NikA-FeEDTA, 59 crystals were exposed. Three 1° oscillation frames from 12 collected datasets were selected for scaling and merging, giving a final 2.45 Å resolution dataset with 68.4% completeness (Table 4). The steps of data selection and processing were performed partly manually, but the total automation of data processing is currently under development.

	Lysozyme	NikA-FeEDTA
Data collection		
Wavelength (Å)	1.542	1.542
Oscillation (°)	1	1
Data reduction		
Space group	$P4_32_12$	$P2_12_12_1$
Resolution (last shell) (Å)	39.57 – 2.10 (2.20 – 2.10)	45.15 – 2.45 (2.55 – 2.45)
Completeness (last shell) (%)	71.6 (75.0)	68.4 (71.4)
Reduction		
Total reflections (last shell)	15323 (1923)	49829 (5590)
Unique reflections (last shell)	5457 (722)	27135 (3158)
Redundancy (last shell)	2.8 (2.7)	1.8 (1.8)
R_{merge} [a] (last shell) (%)	13.9 (38.5)	13.8 (41.6)
I/σ (last shell) (I)	5.61 (2.75)	4.45 (2.20)
Unit Cell (Å)	a=79.14 b=79.14 c=38.93	a=87.93 b=95.03 c=126.40
Refinement		
Resolution range (Å)	39.57 – 2.1	45.15 – 2.45
R_{work} [b] (%)	18.82	17.39
R_{free} [c] (%)	23.11	25.07
R.m.s.d bonds (Å)	0.008	0.015
R.m.s.d angles (°)	1.099	1.259
Reflections in refinement	5457	27133

Table 4: Data and Refinement Statistics. The final data set statistics are obtained by scaling and merging a large number of diffraction data sets from data collections on different crystals: 8 crystals for lysozyme and 12 crystals for NikA-FeEDTA.

[a] $\mathbf{R_{merge}} = \frac{\sum_{hkl}\sum_i |I_i(hkl) - \overline{I(hkl)}|}{\sum_{hkl}\sum_i I_i(hkl)}$ with $\overline{\mathbf{I(hkl)}} = \left(\frac{1}{N}\right)\sum_{i=1}^{N} \mathbf{I_i\,(hkl)}$

[b] $\mathbf{R_{work}} = \frac{\sum ||Fobs| - |Fcalc||}{\sum |Fobs|}$

[c] R_{free} is the same as R_{work} but calculated for 5% data omitted from the refinement.

Datasets showed no specific difficulty for structure solution and model refinement and led to structures with reasonably good resolutions and statistics. Indeed, the Fe-EDTA ligand bound to NikA is clearly visible in the omit map (Figure 27).

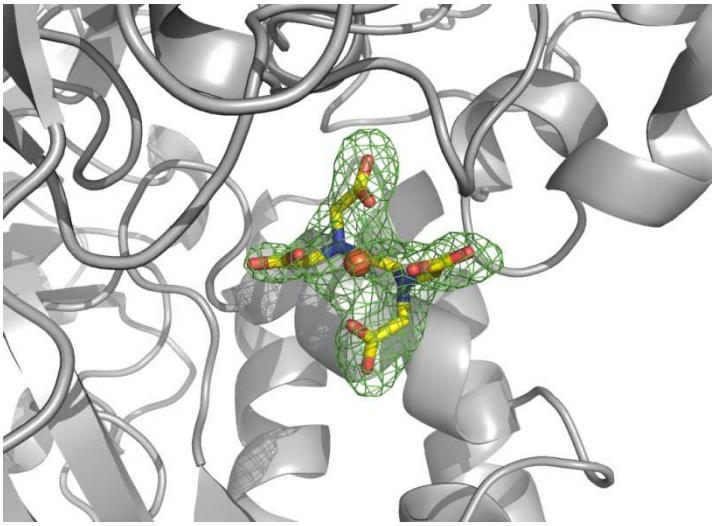

Figure 27: Fe(III)-EDTA binding site in NikA. Omit Fourier electron density map of Fe-EDTA contoured at 3 σ. This figure was prepared with The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.

5. Conclusion

The methodology we have developed can perform high accuracy fully automated crystal centering and data collection with a robotized *in situ* diffraction apparatus such as G-Rob and CATS. Crystal lists are prepared off-line using Crystal Listing on Visualization Bench optimizing beam time usage thus keeping diffraction platforms available for other users. With this approach, considerable time is saved for diffraction analysis of protein crystals and, in some cases, it may lead to structure resolution. Implementing Crystal Listing in image screening apparatus can also automate accurate crystal growth monitoring in crystallization platforms.

Acknowledgements

These developments were co-funded by the CEA, the CNRS and NatX-ray. We would like to thank Dr Florence Pojer, head of Protein Crystallography Corer Facility and scientific from Professor Stewart Cole Laboratory at EPFL (Lausanne) for her help and availability. We would also like to thank Dr Franck Borel, Dr David Cobessi, Dr Richard Kahn and Dr Christine Cavazza (IBS, Grenoble) for their help and support. We would like to thank specially Dr Juan Carlos Fontecilla-Camps and Dr Christine Cavazza, from MetalloProteins group at IBS in Grenoble, for their comments on the manuscript.

References

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B.,, Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J.(2010). *Acta Cryst.* D**66,** 213-221.

Axford, D., Owen, R. L., Aishima, J., Foadi, J., Morgan, A. W., Robinson, J. I., Nettleship, J. E., Owens, R. J., Moraes, I., Fry, E. F., Grimes, J. M., Harlos, K., Kotecha, A., Ren, J., Sutton, G., Walter, T. S., Stuart, D. I., Evans, G. (2012). *Acta Cryst.* D**68,** 592-600.

Bingel-Erlenmeyer, R., Olieric, V., Grimshaw, J. P. A., Gabadinho, J., Wang, X., Ebner, S. G. (2011). *Cryst. Growth Des.* **11,** 916–923.

Cherrier, M. V., Martin, L., Cavazza, C., Jacquamet, L., Lemaire, D., Gaillard, J., Fontecilla-Camps, J. C., (2005). *J. Am. Chem. Soc.* **127,** 10075-10082.

Cherrier, M. V., Cavazza, C., Bochot, C., Lemaire, D., Fontecilla-Camps, J. C. (2008). *Biochemistry* **47,** 9937–9943.

Emsley, P., Cowtan, K. (2004). *Acta Cryst.* D**60,** 2126–2132.

Evans, P. R. (2006). *Acta Cryst.* D**62,** 72-82.

Bradski, G., Kaebler, A., *Computing Vision with the OpenCV Library.* Edited by Mike Loukides. O'Reilly Media Inc., 2008.

Hargreaves, D. (2012). *J. Appl. Cryst.* **45,** 138–140.

Heidari Khajepour, M. Y., Vernede, X., Cobessi, D., Lebrette, H., Rogues, P., Terrien, M., Berzin, C., Ferrer, J.-L., submitted.

Jacquamet, L., Joly, J., Bertoni, A., Charrault, P., Pirocchi, M., Vernede, X., Bouis, F., Borel, F., Perin, J.-P., Denis, T., Rechatind, J.-L., Ferrer, J.-L. (2009). *J. Synchrotron Rad.* **16**, 14–21.

Jacquamet, L., Ohana, J., Joly, J., Borel, F., Pirocchi, M., Charrault, P., Bertoni, A., Israel-Gouy, P., Carpentier, Ph., Kozielski, F., Ferrer, J.-L. (2004). *Structure* **12,** 1219-1225.

Kabsch, W. (2010). *Acta Cryst.* D**66,** 125-132.

Le Maire, A., Gelin, M., Pochet, S., Hoh, F., Pirocchi, M., Guichou, J.-F., Ferrer, J.-F., Labesse, G. (2011). *Acta Cryst.* D**67,** 747-755.

Ohana, J., Jacquamet, L., Joly, J., Bertoni, A., Taunier, P., Michel, L., Charrault, P., Pirocchi, M., Carpentier, P., Borel, F., Kahn, R., Ferrera, J.-L. (2004). *J. Appl. Cryst.* **37,** 72-77.

Teng, T.Y. (1990). *J. Appl. Cryst.* **23,** 387-391.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., Wilson, K. S. (2011). *Acta Cryst.* D**67,** 235–242.

Chapter III: REACH: Robotic Equipment for Automated Crystal Harvesting

This chapter is dedicated to developments aiming to automate and robotize the harvesting, cryo-protecting and flash-cooling crystals for X-ray diffraction analysis. This new development has been one the two major projects of my PhD studies. This project has been clearly mostly conducted and realized by myself under supervision of Dr Jean-Luc Ferrer. The developments, studies and experiments led and realized for this system have been for the greater part a result of my PhD with contribution of the coauthors of the following scientific report. X-ray diffraction data processing, structure refinement and resolution has been conducted by Dr David Cobessi and Hugo Lebrette under supervision of Dr Jean-Luc Ferrer. Results from this developed system have been submitted, through a scientific report, to *Acta Crystallographica section D*, on 1 Jully 2012.

"It always seems impossible until it's done."

Nelson Mandela

REACH: Robotic Equipment for Automated Crystal Harvesting, using a 6-axis robot arm and a micro-gripper

Mohammad Yaser Heidari Khajepour[1], Xavier Vernede[1,2], David Cobessi[1], Hugo Lebrette[2], Pierrick Rogues[3], Maxime Terrien[1], Christophe Berzin[1], Jean-Luc Ferrer[1,4]

1 Institut de Biologie Structurale Jean-Pierre Ebel, Groupe Synchrotron; Commissariat à l'Energie Atomique et aux Energies Alternatives, Centre National de la Recherche Scientifique, Université Joseph Fourier; F-38027 Grenoble cedex 1; France.

2 Institut de Biologie Structurale Jean-Pierre Ebel, Groupe MetalloProtéines; Commissariat à l'Energie Atomique et aux Energies Alternatives, Centre National de la Recherche Scientifique, Université Joseph Fourier; F-38027 Grenoble cedex 1; France.

3 NatX-ray; 38400 Saint Martin d'Hères; France.

4 Correspondence e-mail: jean-luc.ferrer@ibs.fr

Abstract

In protein crystallography experiment, only two critical steps remain manual, the transfer of crystals from their original crystallization drops into the cryo-protection solution followed by flash-cooling. These steps are risky and tedious, requiring a high degree of manual dexterity. These limiting steps are a real bottleneck to high-throughput crystallography and limit the remote use of protein crystallography core facilities. To eliminate this limit, the Robotic Equipment for Automated Crystal Harvesting (REACH) was developed. This robotized system, equipped with a two-finger micro-gripping device, allows crystal harvesting, cryo-protection and freezing. With this set-up, harvesting experiments were performed on several crystals, followed by direct data collection with the same robot arm used as a goniometer. Analysis of the diffraction data demonstrates that REACH is highly reliable and efficient, and does not alter crystallography data. This new instrument fills the gap of the high-throughput crystallography pipeline.

1. Introduction

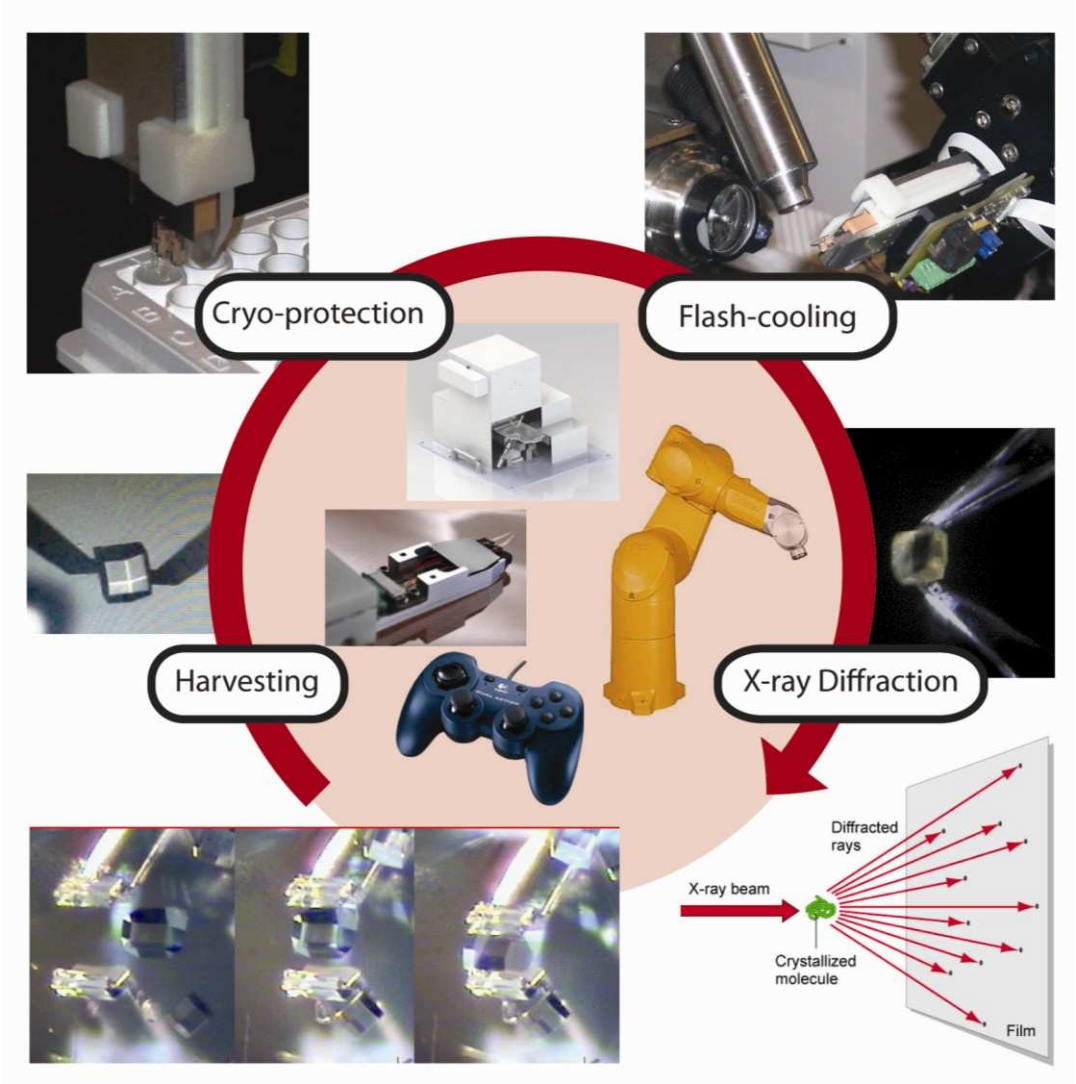

Figure 28: Graphical abstract

Protein structure determination by X-ray crystallography involves numerous steps. In recent years most of these steps such as protein purification (Kim *et al.*, 2004), crystallization (Mueller-Dieckmann, 2006) and also data collection and processing have been mostly automated (Adams *et al.*, 2011; Ferrer, 2001; Manjasetty *et al.*, 2008). The critical steps that remain are harvesting crystals from their crystallization drop, for crystals grown using the vapor diffusion method (McPherson, 1989), followed by cryo-protection and flash-cooling. These steps are still managed manually. Due to their solvent content, ranging from 20% to more than 80%, protein crystals are very fragile and may easily be damaged with variation of temperature and ambient humidity or mechanical stress. Considering also the small dimensions of protein crystals (from ~10 µm to ~500 µm), it is particularly difficult to

manipulate crystals manually through the preparation steps without causing any damage. Furthermore with high throughput "nanodrops" crystallization robots mostly used nowadays, crystals grow even smaller, rather in the ~5 µm to ~50 µm range. *In situ* diffraction in the crystallization drop at room temperature is an alternative to crystal harvesting (Jacquamet *et al.*, 2004). Nevertheless because of limitations due to crystal symmetry and crystal degradation during beam exposure at room temperature, harvesting and flash-cooling samples are very often necessary.

Over the last few decades the most common method to harvest protein crystals has been using micro loops (Teng, 1990). Crystals are visualized through a binocular microscope and manipulated manually in their crystallization drops. First of all, harvesting crystals in this configuration is very annoying as the microscope blocks an easy access to the drop. If the crystals are obtained by using the hanging drop technique, the access to the drop is a bit easier. However nowadays on high throughput protein crystallization setups, crystals are produced in sitting micro to nano-litter drops dispensed with pipeting robots on 96-well microplates. Manipulating into microplate drops requires more dexterity to access the drops, due to the geometry of the microplates. Furthermore, since the volume of crystallization drops is reduced, fast manipulation is required to avoid evaporation. Secondly, manipulating crystals requires a high degree of delicacy and sharpness, especially when crystals are smaller and smaller. Protein crystals with all their fragility have to be hanged in the loop liquid while taking out the loop from their crystallization drop. Crystals may sometimes be trapped in a skin at the surface of the drop or may be stuck to the bottom of the well. In the latter case crystals are tapped to be removed from the bottom. In these difficult situations, harvesting done manually stresses the crystal and could harm or even destroy the crystal. Thirdly, in most cases, once the crystal is harvested on a loop it has to be transferred into a cryo-protecting solution before flash-cooling (Parkina & Hope, 1998). Consequently, in most cases, the crystal will be released into the cryo-protecting drop and it has to be harvested once again. All these manual operations increase the difficulty of the task and also increase the risk of damaging the crystal. Finally, crystals should be flash-cooled to avoid ice formation (Kriminski *et al.*, 2002) and will need to be kept at a temperature below 140K (Garman & Schneider, 1997). The most traditional methods are to plunge the loop into liquid nitrogen (77K) or to expose the loop to a 100K nitrogen gas

stream. The reproducibility of these operations is quite random as they are managed manually (Warkentin *et al.*, 2006).

At least four different robotic harvesting systems for protein crystals have been developed in the last decade: one with a two-finger manipulator system (Ohara *et al.*, 2004), another with a traditional harvesting loop on a 6-axis robot arm (Viola *et al.*, 2011), the "Crystal Harvester", with two motorized loops (Bruker AXS), and a series of micro-manipulators for seeding and harvesting protein crystals (Georgiev *et al.*, 2004; Vorobiev *et al.*, 2006). These systems have better accuracy and minimal vibration compared to human manipulation. In spite of their numerous advantages compared to traditional methods, they haven't found success because of lack of reliability and compatibility issues to standard materials and procedures. These harvesting systems have nevertheless paved the way to new methods automating crystal harvesting and preparation for X-ray diffraction.

In this manuscript we present the new Robotic Equipment for Automated Crystal Harvesting (REACH) recently developed on beamline FIP-BM30A at the European Synchrotron Radiation Facility (ESRF), in order to achieve more robust and reliable macromolecular crystal harvesting and preparation operations compatible with standard macromolecular crystallography materials.

2. Materials & Methods

2.1. Samples

The 14.4 kDa lysozyme protein from hen egg-white (Roche, Reference number: 10837059001), was crystallized by mixing 500 nL of a 50 mg/mL protein solution in 0.24% (w/w) acid acetic with 500 nL of 5% NaCl (w/v) reservoir solution. The 56.3 kDa NikA protein from *E. coli* was also used. Its cytoplasmic apo form was expressed and purified as previously described in Cherrier *et al.*, 2008. A 10 mg/mL apo-NikA solution was pre-incubated overnight at 4°C with 2 molar equivalent of FeEDTA and this protein-ligand complex was crystallized by mixing 0.5 µL of this solution with 0.5 µL sodium acetate 0.1 M pH 4.7, ammonium sulfate 1.5 to 1.95 M reservoir solution (Cherrier *et al.*, 2005). Protein samples are crystallized on CrystalQuick™ X plates (Figure S2), a vapor diffusion sitting drop microplate (Bingel-Erlenmeyer *et al.*, 2011). CrystalQuick™ X has been developed by Greiner Bio-One and the FIP-BM30A group and is commercialized by NatX-ray. This microplate is an SBS-standard 96-Well plate, with two flat wells for sitting drops per reservoir (Figure 29). The geometry of this plate provides better access to drops for crystal manipulation (Figure 30B). Wells are 1.3 mm deep in CrystalQuick™ X plate whereas the wells in other plates are from 3 mm to 4 mm deep. Plates were filled manually. They were then screened for pairs of crystals grown in the same drop. For each pair, one of the two crystals was harvested, cryo-protected and flash-cooled manually using LithoLoops™ (Molecular Dimensions) and the other one with the REACH system.

Figure 29: Crystal cryo-protection. The micro-gripper soaks the crystal in a cryo-protectant solution prior to flash-cooling.

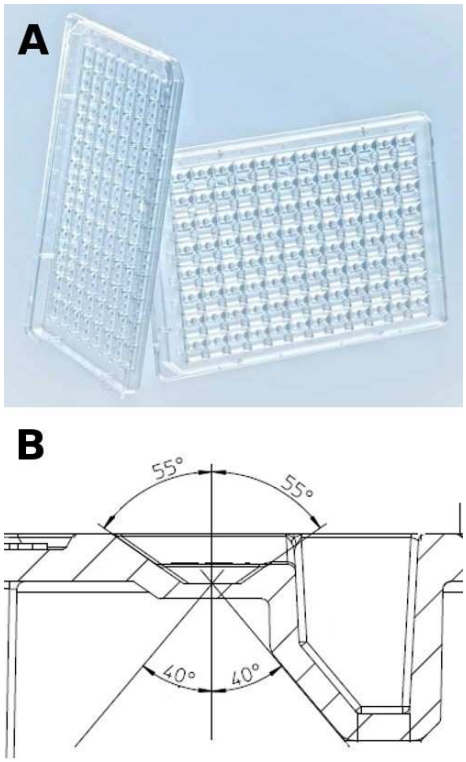

Figure 30: The CrystalQuick™ X plate. (A) Picture of the CrystalQuick™ X plate. (B) Improved geometry of the CrystalQuick™ X plate wells, for a larger oscillation range during *in situ* data collection.

2.2. Beamline

Experiments were carried out on beamline FIP-BM30A (Roth *et al.*, 2002) at the ESRF. This beamline uses a bending magnet as a source and delivers a monochromatic beam with an intensity of $5e^{11}$ photons/(0.3x0.3mm^2)/s and $2x10^{-4}$ energy resolution at 12.5 keV. In these experiments the beam size was defined at 0.2 mm x 0.2 mm. An ADSC Q315r CCD detector was used for recording the diffraction frames. Two goniometers were available on the beamline: a MD2 with on-axis microscope (Maatel) and the G-Rob system. For these experiments the G-Rob system was used as the goniometer and the MD2 on-axis microscope was used to define the spindle position and to visualize the sample for centering into the X-ray beam. The two centering translations on the robot arm were used to center the ending elements of the micro-gripper on the G-Rob spindle axis. For each sample, X-ray diffraction data were collected with 1° oscillation at 0.98 Å wavelength.

2.3. Manual method and diffraction with G-Rob goniometer

In the manual method, crystals were visualized by a classical laboratory binocular microscope and were harvested with SPINE standard loops (Hampton Research, reference

number: HR8-124). Crystals were then soaked into the cryo-protecting solution (25% w/w Glycerol and reservoir solution) for about 20 to 30 seconds and flash-cooled into a 100K temperature nitrogen gas stream generated by a Cryostream 700 system (Oxford Cryosystem).

2.4. Using REACH with direct data collection

In the robotic method, crystallization plates are screened using the Visualization Bench. Its graphical user Interface (GUI) displays the microscope image. A drop of the appropriate cryo-protecting solution is dropped over the crystallization drop. A button on the GUI sends the micro-gripper over the visualized well. Thus the control of the robot and micro-gripper is enabled through the GUI and a game pad. Thanks to the 6-axis arm of G-Rob (Stäubli), the micro-gripper is capable of three translations and two rotational movements. Furthermore opening and closing control of the micro-gripper is integrated into the GUI and into the game pad buttons. First, the motorized translations and zoom of the Visualization Bench are used to center crystals in the microscope and to adjust the focus. Then the user drives the movements of the G-Rob arm to approach the micro-gripper to crystals. The lights are also controlled from the GUI to optimize vision quality. Once the crystal is captured between the two SU-8 ending elements of the micro-gripper, a button on the GUI transfers the crystal with a safe but fast trajectory to the spindle position, into the nitrogen gas stream. The trajectory of the robot in approach of the spindle position is programmed perpendicular to the 100K stream with the robots fastest speed to optimize the flash-cooling. The trajectory ends at a position where the crystal is already centered into the spindle position. Since the G-Rob does the goniometer task and the ending elements of the micro-gripper are transparent to X-ray, it was possible to proceed with data collection, without releasing the crystal or any human manipulation.

2.5. Using REACH for crystal transfer on loop

Alternatively, for sample storage, a robotic transfer on loop was also tested. A manual goniometer head was placed in front of the Visualization Bench. An empty loop was plunged into suitable cryo-protecting solution and placed on the goniometer head. The HC1 system (Sanchez-Weatherby *et al.*, 2009) was used to blow a humidified room temperature nitrogen stream on the loop while harvesting and transferring the crystal to avoid dehydration. Once

the crystal is grabbed between the two ending elements of the micro-gripper, an automated robotic trajectory takes out the crystal from its crystallization drop and transfers it to about 1 mm above the loop. A microscope and adapted lightening have been place towards the loop in order to have a magnified view. The robot was then controlled with the game pad to release the crystal in the loop. Crystals were then frozen manually into liquid nitrogen.

2.6. Diffraction data collection

Diffraction data were processed using XDS (Kabsch, 2010) and scaled with SCALA (Evans, 2006) from CCP4 (CCP4, 1994) or XSCALE from XDS. Phasing was performed by molecular replacement with PHASER (McCoy *et al.*, 2007) from CCP4 using 1LZ8 and 1ZLQ from the Protein Data Bank (PDB) as starting models for lysozyme and NikA-FeEDTA, respectively. Refinement was performed using PHENIX (Adams *et al.*, 2010). Root mean square deviation (RMSD) values were calculated on main chains using COOT (Emsley & Cowtan, 2004).

3. Results

3.1. G-Rob

The REACH system takes advantage of the G-Rob robot arm accuracy and its goniometer capability. G-Rob is a multi-task robotic system based on a 6-axis robot arm, developed on beamline FIP-BM30A at the ESRF (Grenoble, France). G-Rob is accurate enough to operate as a goniometer (Jacquamet *et al.*, 2009). This system is commercially available since 2009 (NatX-ray). It is able to collect X-ray diffraction data with a sphere of confusion better than 15 µm radius for frozen samples and capillaries. This setup is completed by a fully motorized Visualization Bench equipped with an inverted microscope and a three-direction motorized microplate holder.

On G-Rob, two motorized translations are installed at the end of the robot arm to center each sample on the 6th axis of the robot which is used as the spindle axis. In the following experiments, this centering operation is done only once, when G-Rob holds its micro-gripper tool before the harvesting operation. In so doing, once the crystal is transferred to the spindle position, it is already centered into the beam with a positioning error less than 10 µm. Thus X-ray diffraction data can be collected right away.

3.2. The micro-gripper

REACH equips the G-Rob robot arm with a specially designed micro-gripper for crystal handling in order to harvest samples from microplates, perform cryo-protection and flash-cooling, and expose them to the X-ray beam (Figure 31).

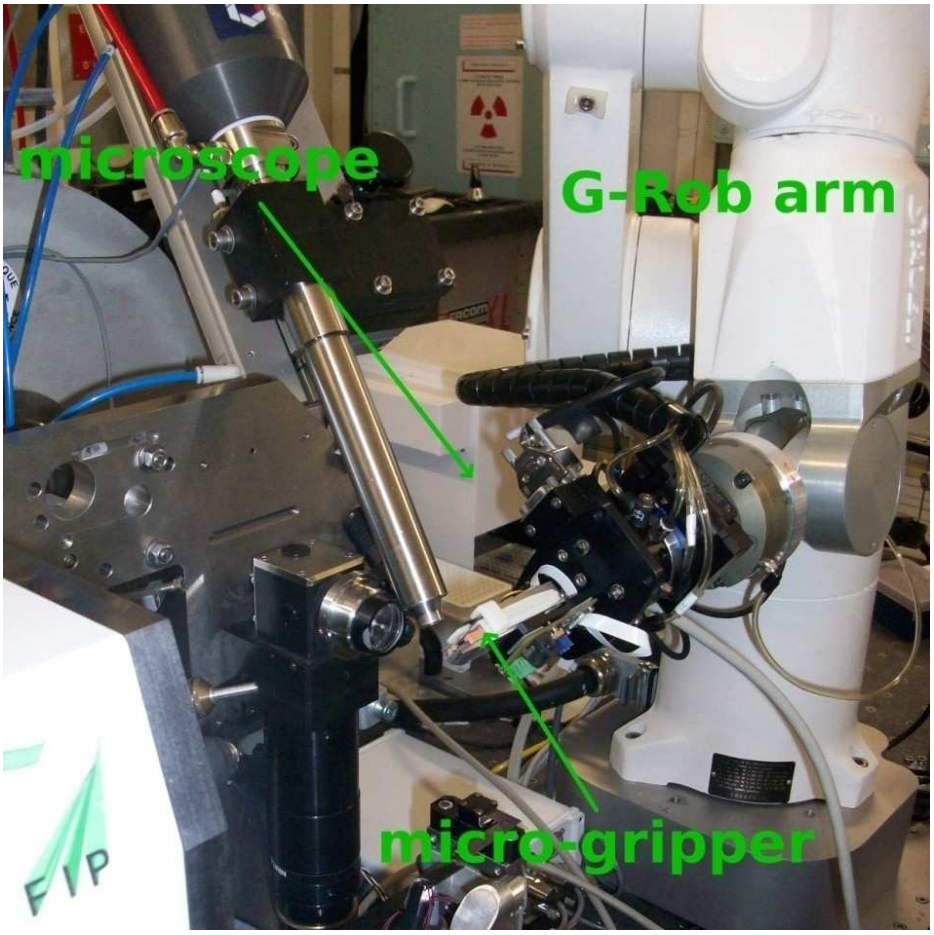

Figure 31: Experimental Setup on FIP-BM30A beamline. The G-Rob robot arm (top right) presents the sample in the beam for data collection. The on-axis camera (center) is used for the sample centering.

This new tool for the G-Rob system has been developed in order to achieve harvesting crystals in their crystallization drop by grabbing them with a two-finger piezo-electric micro-gripping device (Agnus *et al.*, 2009). Each finger has two degrees of freedom controlled with a resolution of 1.0 µm and a reproducibility of 0.1 µm. By combining symmetrical translations of both piezo-electrical fingers, an opening gap range from 0 µm to 500 µm is possible. This micro-gripping device has been developed by Femto-ST (Besançon, France) and is now commercialized (Percipio-Robotics). Ending elements in touch with crystals are fabricated separately from the two-finger micro-gripper (Figure 32). These ending elements are composed of an epoxy based polymer called SU-8. This material has a remarkable stiffness (Ling *et al.*, 2009). The ending element geometry was designed to have the best possible grip on crystals, first to avoid crashing the crystals by spreading the griping pressure on several points or in best cases on facets of the crystals, and secondly to improve the chance of grabbing the crystals in the first attempt or in difficult cases were crystals are stuck to the bottom of the crystallization well. The ending element geometry has been also

designed for the lowest volume of SU-8 exposed to the X-ray, in order to minimize scattering (~20 µm of thickness for each ending element, see Figure 32B). Thus crystals are exposed to X-ray in the micro-gripper by the G-Rob robot arm for crystallography data collection. SU-8 produces a very low scattering background in X-ray (data not shown). Compared to other common materials used for the fabrication of crystal harvesting loops, the SU-8 shows a background scattering in X-ray exposure between Kapton® and nylon. The geometry and the thickness of the arms of the ending elements combined with the elasticity of the SU-8 bring enough flexibility to limit the stress on crystals and thus yield before crushing the crystals.

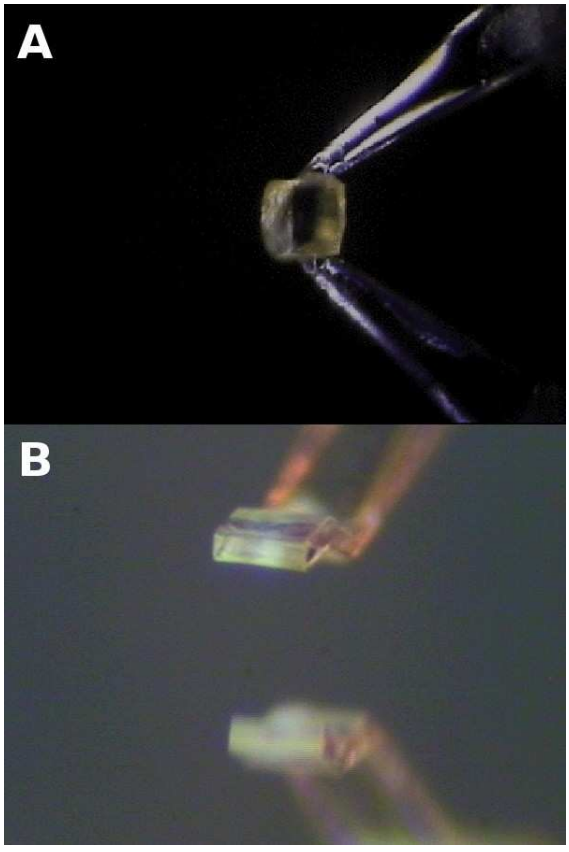

Figure 32: Image of the Micro-Gripper. (A) A lysozyme crystal handled by the micro-gripper. (B) Last generation of the ending elements, made of SU-8.

3.3. Comparison experiments

In order to assess the impact of the stress inflicted on crystals with the micro-gripper, series of tests of harvesting, cryo-protection and flash-cooling were led manually and with the REACH system. In robotic method, crystals are directly exposed ("direct data collection") after being grabbed by the micro-gripper, in order to evaluate the gripping influence on crystals structure. Crystals from lysozyme and NikA-FeEDTA proteins were used for this

experiment (see Materials and Methods). Two pairs of crystals from same wells of each protein were chosen and prepared for diffraction data collection with G-Rob, using both manual and robotic method, as mentioned in Experimental Procedures bellow.

Analysis of data reduction showed no significant differences in mosaicity, resolution limits and unit cell dimensions (Table 5). Unit cell volume comparisons of both manual and robotic harvested samples (Table 6) also show no significant difference. Nevertheless their comparison with PDB structures 1ZLQ and 1LZ8, respectively for NikA-FeEDTA and for lysozyme, show variations from 1.4% to 3.6%. Diffraction data for lysozyme (PDB entry: 1LZ8) were collected at 120K and not at 100 K. Thermal expansion cannot account for this difference. Indeed, calculations based on Tanaka, 2001, considering the crystal and solvent as water, show only 0.15% volume variation of each unit cell. Therefore we assume that the unit cell volume differences are due to the experimental setup discrepancy.

Data and refinement statistics are similar whatever the crystal harvesting method, robotic or manual. The RMSD values (Table 6) between the structures, based on main chain comparison, are low and do not exceed 0.46 Å for both proteins. Thus, we can confirm that the stress on the crystals is controlled and that there is no structural rearrangement due to the use of the micro-gripper.

Although there was not visible improvement in the data statistics, certainly due to the reduced number of tested crystals, we observed a reduced amount of solvent around the crystal when harvested with the robot. This results in a reduced background scattering. The average background measured by XDS (INIT step) and normalized to 1 sec exposure time and 1 mA current in the ESRF ring, is 0.154 and 0.174, respectively for lysozyme and NikA-FeEDTA, when harvested manually, to be compared to 0.126 and 0.071 respectively, when crystals are harvested with the robot.

	lysozyme				NikA-FeEDTA			
Dataset	**Manual 1**	**Manual 2**	**Robotic 1**	**Robotic 2**	**Manual 1**	**Manual 2**	**Robotic 1**	**Robotic 2**
Data collection								
Wavelength (Å)	0.97955	0.97955	0.9795	0.9797	0.97969	0.97968	0.97967	0.97967
Oscillation (°)	1	1	1	1	1	1	1	1
Range	60	90	69	110	75	110	90	90
Data reduction								
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Resolution (last shell) (Å)	38.65-1.50 (1.58-1.50)	36.78-1.80 (1.90-1.80)	38.62-1.75 (1.84-1.75)	38.99-1.60 (1.69-1.60)	47.01- 2.65 (2.75- 2.65)	40.70-1.85 (1.95-1.85)	44.25-2.30 (2.40-2.30)	44.22-1.95 (2.05-1.95)
Completeness (last shell) (%)	84.7 (88.7)	100 (100)	99.9 (100)	99.7 (100)	97.4 (98.3)	97.9 (98.4)	98.6 (98.6)	97.3 (98.2)
Reduction								
Total reflections	84948	73949	59671	125316	90500	380011	163609	267767
(last shell)	(11509)	(10330)	(8306)	(16597)	(9330)	(54747)	(19205)	(37037)
Unique reflections	15560	10887	11761	15436	29023	83913	44557	71555
(last shell)	(2324)	(1548)	(1671)	(2201)	(3015)	(12171)	(5294)	(9924)
Redundancy (last shell)	5.5 (5.0)	6.8 (6.7)	5.1 (5.0)	8.1 (7.5)	3.1 (3.1)	4.5 (4.5)	3.7 (3.6)	3.7 (3.7)
R_{sym} [a] (last shell) (%)	4.9 (37.9)	5.5 (46.4)	8.8 (42.0)	5.8 (42.7)	12.4 (39.2)	4.7 (35.9)	5.6 (33.5)	5.3 (32.9)
R_{pim} [b] (last shell) (%)	2.2 (18.2)	2.3 (19.2)	4.3 (20.7)	2.2 (16.5)	8.7 (26.1)	2.6 (19.2)	3.7 (20.7)	3.5 (20.1)
I/σ (last shell) (I)	17.2 (3.9)	21.5 (4.1)	10.8 (4.5)	17.7 (3.8)	7.34 (2.92)	19.23 (4.40)	16.68 (4.35)	16.45 (4.51)
Mosaicity	0.247	0.401	0.331	0.376	0.190	0.317	0.318	0.234
Unit Cell (Å)	a=77.31 b=77.31 c=36.97	a=77.51 b=77.51 c=36.78	a=77.30 b=77.30 c=36.89	a=77.98 b=77.98 c=36.71	a=86.28 b=94.02 c=123.3	a=86.24 b=93.64 c=123.2	a=86.24 b=93.74 c=123.4	a=86.33 b=93.88 c=123.1
Refinement								
Resolution range (last shell) (Å)	38.65-1.50 (1.59-1.50)	34.66-1.80 (1.89-1.80)	34.57-1.75 (1.84-1.75)	33.21-1.60 (1.65-1.60)	47.01-2.65 (2.74-2.65)	40.70-1.85 (1.87-1.85)	40.71-2.30 (2.35-2.30)	43.17-1.95 (1.98-1.95)
R_{work} [c] (last shell) (%)	18.16 (22.45)	16.90 (21.34)	16.25 (20.0)	17.25 (21.72)	17.40 (22.95)	17.53 (27.20)	18.51 (25.34)	17.17 (25.63)

R_{free}[d] (last shell) (%)	20.21 (25.77)	21.61 (26.09)	19.74 (27.11)	19.37 (22.08)	26.91 (33.81)	21.55 (32.57)	25.47 (35.85)	21.65 (31.81)
R.m.s.d bonds (Å)	0.006	0.007	0.008	0.008	0.008	0.007	0.008	0.008
R.m.s.d angles (°)	1.063	1.062	1.187	1.125	1.150	1.124	1.087	1.117
Reflections in refinement	15534	10856	11725	15394	29015	83910	44550	71549
B factor average ($Å^2$)	19.1	26.6	21.9	25.1	32.94	30.23	41.51	30.04

Table 5: Data and Refinement Statistics. Comparison of dataset statistics for lysozyme and NikA-FeEDTA crystals harvested either manually (named "Manual X") or with the REACH system (named "Robotic X").

[a] $R_{sym} = \Sigma|I_i - \langle I \rangle|/\Sigma I_i$, where I_i is the intensity of a reflection and $\langle I \rangle$ is the average intensity of that reflection.

[b] $R_{pym} = (\Sigma(1/(n-1))\Sigma|I_i - \langle I \rangle|)/\Sigma\langle I \rangle$, where n is the number of observation of the reflection.

[c] $R_{work} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$.

[d] R_{free} is the same as R_{work} but calculated for 5% data omitted from the refinement.

	Comparative RMSD on main chain (Å)				Unit Cell volume changes (%)		
Lysozyme	**1LZ8**	**Manual 1**	**Manual 2**	**Robotic 1**	**Manual 1**	**Manual 2**	**Robotic 1**
Manual 1	0.202	-	-	-	-	-	-
Manual 2	0.259	0.162	-	-	0.00	-	-
Robotic 1	0.223	0.083	0.123	-	0.24	0.24	-
Robotic 2	0.246	0.181	0.090	0.156	1.03	1.02	1.27
NikA-FeEDTA	**1ZLQ**	**Manual 1**	**Manual 2**	**Robotic 1**	**Manual 1**	**Manual 2**	**Robotic 1**
Manual 1	0.321	-	-	-	-	-	-
Manual 2	0.364	0.219	-	-	0.57	-	-
Robotic 1	0.470	0.289	0.210	-	0.24	0.33	-
Robotic 2	0.332	0.207	0.124	0.243	0.25	0.32	0.01

Table 6: Comparative RMSD on main chains and Unit Cell Volume Changes between Manually and Robotically Harvested Crystals.

3.4. Transfer-to-loop experiments

To store harvested crystals on standard loops using the REACH system, series of tests were led (see 2. Materials & Methods). Three crystals of both NikA and lysozyme protein crystals, prepared as mentioned in Experimental Procedures, were harvested by REACH and transferred in a humidified stream and released into a loop containing the cryo-protecting solution (25% w/w Glycerol and reservoir solution). Crystals on loops were then frozen manually in liquid nitrogen. From the same drops, similar crystals were harvested, cryo-protected and flash-cooled manually. Diffraction data were then collected on each crystal at beamline BM30A at ESRF with G-Rob (see 2. Materials & Methods). However, even if the transfer of crystals from the micro-gripper to the loop was successful, it was not possible to push this comparison further. Indeed, it appeared that transferred crystals were significantly damaged, likely due to the long exposure to a non-optimal humidified stream at uncontrolled temperature.

4. Discussion

4.1. Advantages of the robotic harvesting

The use of REACH shows high accuracy and stability in manipulating crystals in their crystallization drops. In particular, this instrument significantly helps with the harvesting of crystals stuck to the bottom of crystallization plates. Crystals from 40 µm to 400 µm were manipulated and harvested successfully, even when grown in 96 well microplates in nano-drops. These tests revealed that harvesting very small crystals is significantly facilitated, compared to manual harvesting. Except for these difficult cases, harvesting time for both manual and robotic methods was comparable. Time benefits for the REACH system comes rather from the following steps. Indeed, when using the robot, once the harvested, the crystal is already mounted on the “goniometer” G-Rob and centered into the beam, ready for data collection. Using the manual method, the sample holder has to be transferred to the goniometer head, and the crystal centering operation is mandatory as loop dimensions and crystal position in the loop are random. This operation took from one to two minutes per crystal. Therefore the robotic method provides a higher reliability and repeatability, facilitates harvesting of difficult crystals, and saves time when coupled to direct data collection.

Also, the crystals harvested using REACH were transferred with a reduced amount of mother liquor and cryo-protecting solution compared to crystals harvested with a loop. We didn't experience any ice formation in the flash-cooling step. Lower background, due to reduced diffusion rings, was noticed with the crystals using REACH in comparison with crystals on loop.

Remote access to crystallography setup is becoming an important challenge, considering the emergence of beamlines coupled to crystallization platforms, or core facilities shared by several laboratories. Our robotized harvesting method paves the way for fully remote controlled protein crystallography experiments, from crystallization assays to structure solution. The REACH system, completed with the tape punching, sample cryo-protection, and flash-cooling steps described below, will provide such a complete high throughput automated pipeline.

4.2. Film punching

To completely automate the procedure, the sealing film on the crystallization plate should be removed before crystal harvesting, and sealed back once the micro-gripper is out. An automated cutting system is under development to punch a portion of the sealing film using a heated wire. Experiments of cutting the sealing film at room temperature on a 500 nL drop disposed on a CrystalQuick™ X plate with a circle wire (about 4 mm in diameter) heated to 220°C, show a temperature increase for less than 5 seconds with a pick variation of less than 4°C. A smooth air sucking around the heating wire could avoid the heat reaching the drop and therefore could reduce the temperature increase of the drop. At the end of the harvesting operation an automated sealing device will stick a patch of tape over the hole to prevent evaporation and to save the drop.

4.3. Cryo-protection and flash-cooling with the micro-gripper

For the experiments presented above, cryo-protectant was added to the drop prior to harvesting. Alternatively, an automated procedure is being developed to soak the crystal using the micro-gripper into a cryo-protecting drop, without releasing the crystal (Figure 29). The soaking time can be specified on the GUI, so the robot will transfer the crystal to the spindle position automatically at the end of soaking period. This procedure was tested, and the geometry of the ending elements has been improved to avoid releasing the crystal into the cryo-protecting drop (Figure 32B). The quality of frozen crystals was not assessed by diffraction measurement.

4.4. Improved transfer to a loop

As described above, the transfer of crystals with the REACH system on a loop in a humidified gas stream is not the proper way to preserve crystals. In the future, we will instead consider the two following scenarios.

On a classical goniometer / sample changer system, commonly available on many beamlines, a loop soaked into a cryo-protectant solution will be first transferred on the goniometer head and will be frozen and kept frozen with the cryogenic stream at the spindle position. Upon the crystal harvesting with the REACH system, the cryogenic stream is suspended with a shutter for the time to transfer robotically the crystal on the loop previously installed at

the spindle position. Then the cryogenic stream is restored and the crystal on loop is flash-cooled. The sample can then be analyzed with X-ray diffraction or transferred back to the storage Dewar by the sample changer.

Alternatively, when REACH is used alone, a G-Rob dedicated goniometer head tool can be used for handling a loop. G-Rob places this tool on a specific tool rack with a cryogenic stream blowing towards the loop position. Then it transfers a loop with adapted cryo-protectant solution to this goniometer head tool, as a classical sample changer. G-Rob picks up the micro-gripper tool from the tool rack for crystal harvesting. Then it releases the crystal on the loop in a trajectory synchronized with the interruption of the cryogenic stream. Next step, the robot can release its micro-gripper tool in its tool rack before taking its goniometer head tool, either for direct data collection or for transfer of the sample to the storage Dewar.

Acknowledgements

We thank the FIP-BM30A staff for their help for the data collection. We would also like to thank Dr Richard Kahn (IBS, Grenoble), Dr Christine Cavazza (IBS), Dr Juan Sanchez-Weartherby (Diamond Light Source Ltd, Oxfordshire) and Dr Florence Pojer (EPFL, Lausanne) for help and support. The present project was funded by the Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), the Centre National de la Recherche Scientifique (CNRS) and NatX-ray. Y.H. benefited from a PhD grant funded by CNRS and NatX-ray. J.-L.F. and X.V. are co-founders of NatX-ray.

References

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Echols, N., Headd, J .J., Hung, L.-W., Jain, S., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, N. D., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2011). *Methods* **55**, 94–106.

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve R. W., McCoy, A. J., Moriarty N. W., Oeffner, R. & Read, R. J. (2010). *Acta Cryst.* D**66**, 213-221.

Agnus, J., Hériban, D., Gauthier, M. & Pétrini, V. (2009). *Precis Eng.* **33**, 542-548.

Bingel-Erlenmeyer, R., Olieric, V., Grimshaw, J. P. A., Gabadinho, J., Wang, X. & Ebner, S.G. (2011). *Cryst. Growth Des.* **11**, 916–923.

CCP4, C. C. P. N. (1994). The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst.* D**50**, 760-763.

Cherrier, M. V., Martin, L., Cavazza, C., Jacquamet, L., Lemaire, D., Gaillard, J. & Fontecilla-Camps, J. C. (2005). *J. Am. Chem. Soc.* **127**, 10075-10082.

Cherrier, M. V., Cavazza, C., Bochot, C., Lemaire, D. & Fontecilla-Camps, J. C. (2008). *Biochemistry* **47**, 9937–9943.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.

Ferrer, J.-L. (2001). *Acta Cryst.* D**57**, 1752–1753.

Garman, E. F. & Schneider, T. R. (1997). *J. Appl. Cryst.* **30**, 211-237.

Georgiev, A., Allen, P. K. & Edstrom, W. (2004). *IEEERSJ International Conference on Intelligent Robots and Systems IROS*.

Jacquamet, L., Joly, J., Bertoni, A., Charrault, P., Pirocchi, M., Vernede, X., Bouis, F., Borel, F., Perin, J.-P., Denis, T., Rechatin, J.-L. & Ferrer, J.-L. (2009*). J. Synchrotron Rad.* **16**, 14–21.

Jacquamet, L., Ohana, J., Joly, J., Borel, F., Pirocchi, M., Charrault, P., Bertoni, A., Israel-Gouy, P., Carpentier, P., Kozielski, F. & Ferrer, J.-L. (2004). *Structure* **12**, 1219-1225.

Kabsch, W. (2010). *Acta Cryst.* D**66**, 125-132.

Kim, Y., Dementieva, I., Zhou, M., Wu, R., Lezondra, L., Quartey, P., Joachimiak, G., Korolev, O., Li, H. & Joachimiak, A. (2004). *J. Struct. Funct. Genomics* **5**, 111–118.

Kriminski, S., Caylor, C. L., Nonato, M. C., Finkelstein, K. D. & Thorne, R. E. (2002). *Acta Cryst.* D**58**, 459-471.

Ling, Z., Liu, C. & Lian, K. (2009). *Microsyst. Technol.* **15**, 429–435.

Manjasetty, B. A., Turnbull, A. P., Panjikar, S., Büssow, K. & Chance, M. R. (2008). *Proteomics* **8**, 612–625.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M., Storonia, L. C. & Read, R. J. (2007). *J. of Applied Crystallogr.* **40**, 658-674.

McPherson, A. (1989). Preparation and analysis of protein crystals (Malabar, USA: Krieger Publishing Company).

Mueller-Dieckmann, J. (2006). *Acta Cryst.* D**62**, 1446–1452.

Ohana, J., Jacquamet, L., Joly, J., Bertoni, A., Taunier, P., Michel, L., Charrault, P., Pirocchi, M., Carpentier, P., Borel, F., Kahnd, R. & Ferrer, J.-L. (2004*). J. Appl. Cryst.* **37**, 72-77 .

Ohara, K., Ohba, K., Tanikawa, T, Hiraki, M., Wakatsuki, S., Mizukawa, M. & Tanie, K. (2004). *Proceedings of the 2004 International Symposium on Micro-Nanomechatronics and Human Science*, 301-306.

Parkina, S. & Hope, H. (1998). *J. Appl. Cryst.* **31**, 945-953.

Roth, M., Carpentier, P., Kaïkati, O., Joly, J., Charrault, P., Pirocchi, M., Kahn, R., Fanchon, E., Jacquamet, L., Borel, F., Bertoni, A., Israel-Gouy, P. & Ferrer, J.-L. (2002). *Acta Cryst.* D**58**, 805–814.

Sanchez-Weatherby, J., Bowler, M. W., Huet, J., Gobbo, A., Felisaz, F., Lavault, B., Moya, R., Kadlec, J., Ravelli, R. B. G. & Cipriani, F. (2009). *Acta Cryst.* D**65**, 1237–1246.

Tanaka, H. (2001*). J. Mol. Liquids* **90**, 323-332.

Teng, T. Y. (1990). *J. Appl. Cryst.* **23**, 387-391.

Viola, R., Walsh, J., Melka, A., Womack, W., Murphy, S., Riboldi-Tunnicliffe, A. & Rupp, B. (2011). *J. Struct. Funct. Genomics* **12**, 77–82.

Vorobiev, S., Edstrom, W., Song, T., Laine, A., Hunt, J. & Allen, P. K. (2006). *Acta Cryst.* D**62**, 1039-1045.

Warkentin, M., Berejnov, V., Husseini, N. S. & Thorne, R. E. (2006). *J. Appl. Cryst.* **39**, 805–811.

Chapter IV: Concluding remarks for complete automated pipelines

Robotic instrumentation solutions described in the last two chapters can be integrated to macromolecular crystallography facilities, beamlines and laboratories systems, to complete fully automated pipelines from crystallization to structure resolution. This chapter describes several scenario and strategy possibilities in implementing these solutions for obtaining complete good quality datasets automatically, starting from crystals in their crystallization plates.

"Imagination is more important than knowledge."

Albert Einstein

1. Filling the gap in automated crystallography

1.1. Generalities

In Structural Biology, as mentioned in the first chapter (see Chapter I:5.) more and more proteins are to be studied. Thus more and more structures are to be solved, which can be better achieved with automated macromolecular crystallography pipelines. As crystallization robots have achieved fairly improvements in productivity and performances, today automation requirements for macromolecular crystallography consist mostly in the individual crystal handling, from harvesting to flash freezing.

In this manuscript we presented two major automation developments for *in situ* crystal screening and data collection (see chapter II) and also preparing and crystallography data collection on frozen samples (see chapter III).

1.2. Test platforms for REACH and Crystal Listing

REACH was implemented on beamline FIP-BM30A at ESRF, which is a crystallography French facility. This facility offers a fully automated beamline with two robotic systems: CATS and G-Rob. CATS, based on a 6-axis robotic arm, can manage sample transfer from a liquid nitrogen Dewar were frozen crystal samples are stored, to the MD2 goniometer (Maatel). The MD2 goniometer, equipped with an on-axis microscope, procures sample vision with motorized sample centering capability. Different configurations are possible for this goniometer allowing Kappa = 0 and Kappa ≠ 0 with Phi rotation (see Chapter I:3.2. b)) for different X-ray exposure strategies. The G-Rob system, also based on a 6-axis robotic arm, can manage sample transfer. The accuracy of its arm allows, by changing its ending tool, to manage the goniometer task for frozen samples. The robot can manage various goniometer configurations (see Figure 14 and Figure 15). In Kappa = 0 and Kappa ≠ 0 with Phi rotation, only the sixth axis of the robot rotates giving a sphere of confusion better than 15 µm radius. Using all six axis of the robot offers the possibility to rotate crystals in Kappa ≠ 0 with Omega rotation (see Figure 15②).

Crystal Listing function has been implemented and tested, as reported in the article in chapter II, on a laboratory version of G-Rob. Thanks to this function data collection is possible on series of crystals *in situ* and so crystals can be screened and even protein

structure can be resolved. With the in house G-Rob version at EPFL, frozen samples can be mounted manually on G-Rob with its goniometer tool for diffraction or automatically from a liquid nitrogen Dewar where samples are stored. Crystallization microplates can be transferred to the spindle position from a storage shelf in order to analyze crystals *in situ*. The implementation of the Crystal Listing function occurred quite easy. With small software modifications and calibrations, it was possible to perform fully automated plate screening and data collection with CrystalQuick™ X crystallization microplates. Following this experience, we can confirm that Crystal Listing function can be easily adapted for other crystallization plates and implemented in various visualization apparatus for off-line Crystal Listing for crystal growth investigation or automated crystal centering and data collection on diffraction apparatus. It can also be simply implemented on motorized *in situ* apparatus such as CATS, PX Scanner, motorized jig, etc.

1.3. To be done

Some software developments in processing diffraction data, combined with Crystal Listing function, could lead to automated good diffracting crystal investigation. Diffraction pattern analysis can be carried out to exclude precipitant crystals. A scoring program could classify crystals with respect to their diffraction resolution quality, based on resolution, mosaicity, etc. Many other possibilities in ranking crystals on their screening diffraction patterns can be considered. When several crystals are required to reach a good completeness, clustering of the datasets collected on these different crystals can be used to select the best datasets to be merged.

Regarding REACH, an ultimate step toward full automation will be the automated crystal recognition in the drop. Such a tool will be used to drive the harvesting tool in order to make possible the harvesting without the assistance of an operator. However, based on the limited success of the previous attempts, such a goal remains so far unreachable.

In situ diffraction and robotic harvesting for in house facilities and also synchrotron beamlines contribute to close the gap in the automation of protein crystallography between crystallization and structure resolution. Combination of these instruments can offer various strategies to crystallographers for analyzing their crystal samples. Different combinations of these techniques are presented below.

2. Automated pipelines

As facilities in different steps of macromolecular crystallography structural resolution are quite complex and expensive, collaboration between different core facilities are mandatory to offer scientists complete automated pipelines. For example, FIP beamline could benefit ESRF neighboring the EMBL, to collaborate with High Throughput Crystallization Laboratory (HTX Lab) facility to expand automated services offered to crystallographers. Thus proteins are sent to EMBL-HTX Lab for crystallization, and then crystallization plates are transferred to FIP beamline (less than 400 meters separates HTX Lab from FIP-BM30A), were diffraction analysis can be managed remotely and with different strategies.

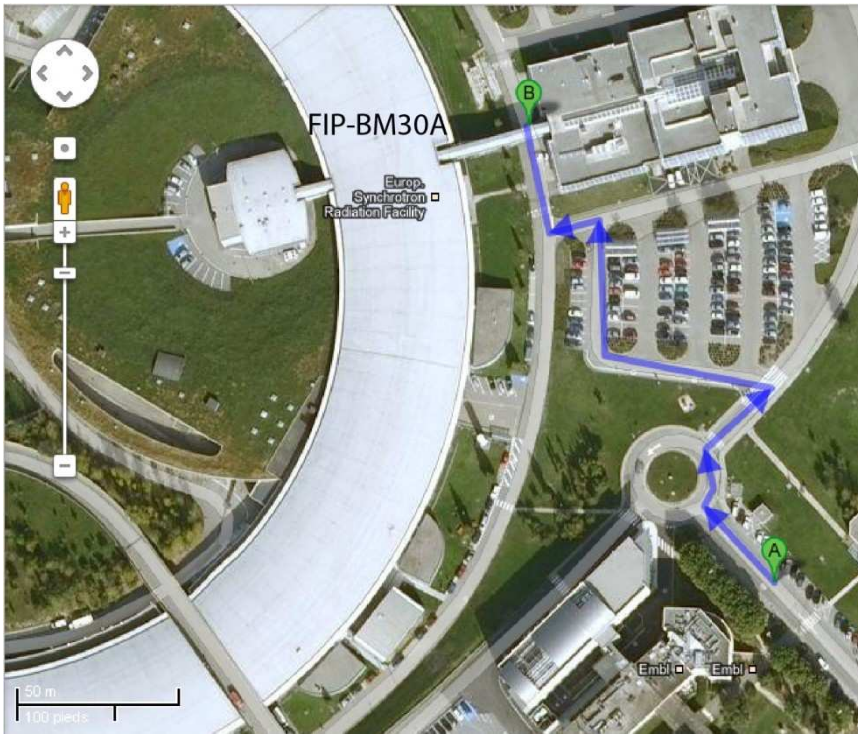

Figure 33: EMBL HTX Lab to ESRF FIP-BM30A beamline

Instruments available at FIP-BM30A completed, as mentioned above, with Crystal Listing and REACH functions, offers the following possibilities to crystallographers (see Figure 34), depending on their crystals features and diffraction results.

Firstly, crystallization plates can be screened automatically with Crystal Listing and G-Rob or CATS system. Once good diffraction quality crystals are located, *in situ* or frozen sample diffraction strategies are possible. If cryo-protection could be performed without harming crystals, one will have the choice to harvest, cryo-protect and flash-cool its crystals into nitrogen gas stream, with REACH, and proceed to a direct X-ray diffraction experiment with

the crystal in the micro-gripper. On the other hand, with REACH, crystals can be also transferred on loops for flash-cooling into nitrogen 100 K temperature stream or into liquid nitrogen. Crystals can then be transferred remotely into a liquid nitrogen storage Dewar, awaiting transfer to the goniometer thanks to the sample changer. In both cases data collection of crystals on loop can be carried out. If cryo-protection step harms crystals, then good diffracting crystals can be listed thanks to Crystal Listing function and an automated crystal centering and X-ray data collection on each of these crystals can be launched. Hence, by merging datasets of all exposed crystals, a complete dataset can be gathered to solve the structure.

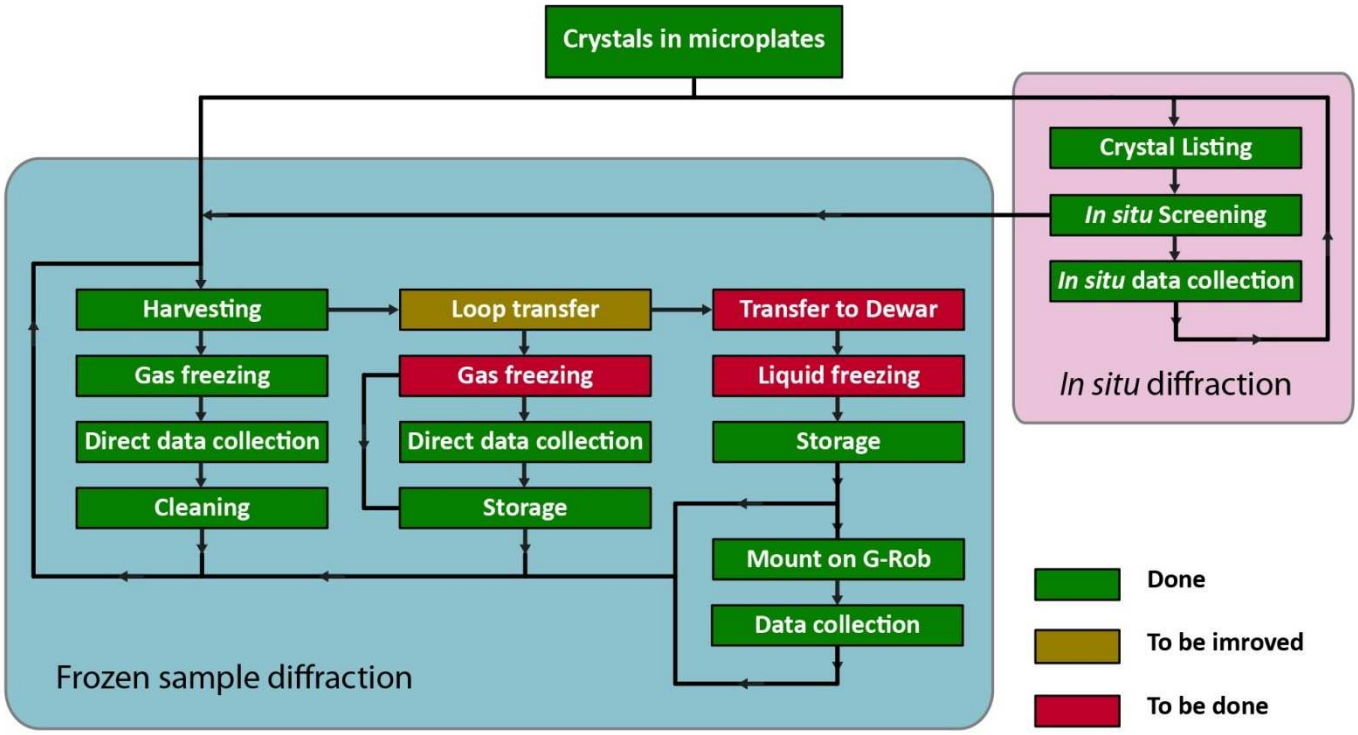

Figure 34: Automated X-ray diffraction strategy possibilities

To go further in developing automated pipelines, collaboration of different facilities could bring other new solutions. Crystallization platforms often allow an automated image screening of crystallization plates. Images are often available through a secure website, as it is the case for the HTX Lab. By integrating Crystal Listing to the imaging instruments of these facilities, wells could be accurately centered, as done in Crystal Listing function, before saving an image. Hence, by adding the possibility to click on crystals and so save crystals images and coordinates, lists of crystals to be diffracted in an automated screening process can be created online. Thus, crystallization plates with their listed crystals data can be

transferred to FIP-BM30A to conduct automated crystal centering and *in situ* X-ray diffraction for screening or collection data on saved crystals.

With such highly automated and remotely controlled instruments, possibilities are endless. We can hope that it will bring in the future an improvement in the quality of the data and a reduction of time required for structure resolution.

Abbreviations

ALS	Advanced Light Source
ANSI	American National Standards Institute
BESSY	Berlin Electron Storage ring society for Synchrotron radiation
CATS	Cryogenic Automated Transfer System
CEA	Commissariat D'Energie Atomique
EMBL	European Molecular Biology Laboratory
EPFL	Ecole Polytechnique Fédérale de Lausanne
ESRF	European Synchrotron Radiation Facility
FIP	French beamline for Investigation of Proteins
GUI	Graphical User Interface
HTX Lab	High Throughput Crystallization Laboratory
IBS	Institut de Biologie Structurale
LN2	Liquid Nitrogen
MAD	Multi-wavelength Anomalous Diffraction
MX	Macromolecular Crystallography
PDB	Protein Data Base
PEG	Polyethylene Glycol
REACH	Robotic Equipment for Automated Crystal Harvesting
SAM	Stanford Automated Mounting
SBS	Society for Biomolecular Sciences
SC3	Sample Changer 3
SSRL	Stanford Synchrotron Radiation Laboratory

Figures

Tables

References

Adams P. D, Afonine P. V, Bunkóczi G, Chen V. B, Davis I. W, Echols N, Headd J. J, Hung L. W, Kapral G. J, Grosse-Kunstleve R. W, McCoy A. J, Moriarty N. W, Oeffner R, Read R. J. "PHENIX: a comprehensive Python-based system for macromolecular structure solution." *Acta Crystallographica D* 66 (2010): 213-221.

Asherie N. "Protein crystallization and phase diagrams." *Methods* 34 (2004): 266–272.

Bingel-Erlenmeyer R., Olieric V., Grimshaw J. P. A., Gabadinho J., Wang X., Ebner S. G. «SLS Crystallization Platform at Beamline X06DA—A Fully Automated Pipeline Enabling in Situ X-ray Diffraction Screening.» *Cryst. Growth Des.* 11 (2011): 916–923.

Brüggeller P., Mayer E. "Complete vitrification in pure liquid water and dilute aqueous solutions." *Nature* 288 (1980): 569-571.

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 Cryst.* D63 (2007): 339-347.

Cherrier M. *Thèse: Etude cristallographique des protéines NikA et NikR impliquées dans le transport du nickel chez Escherichia coli.* 2006.

Cobessi D., Celia H., Pattus F. "Crystal Structure at High Resolution of Ferric-pyochelin and its Membrane Receptor FptA from Pseudomonas aeruginosa." *Journal of Molecular Biology* 352 (2005): 893–904.

Dauter Z. "Data-collection strategies." *Acta Cryst* 5 (1999): 1703-1717.

Ducruix A. F., Riès-Kautt M. M. "Solubility diagram analysis and the relative effectiveness of different ions on protein crystal growth." *Methods* 1 (1990): 25–30.

Garman E. "Cool data: quantity AND quality." *Acta Cryst* 55 (1999): 1641–1653.

Garman E. F. "Radiation damage in macromolecular crystallography: what is it and why should we care?" *Acta Cryst.* D66 (2010): 339–351.

Garman E. F., Schneider T. R. "Macromolecular Cryocrystallography." *J. Appl. Cryst.* 30 (1997): 211-237.

Giegé R., Dock A.C., Kern D., Lorber B., Thierry J.C., Moras D. "The role of purification in the crystallization of proteins and nucleic acids." *Journal of Crystal Growth* 76 (1986): 554–561.

Grey J., Thompson D. "Challenges and Opportunities for New Protein Crystallization Strategies in Structure-Based Drug Design." *Expert Opin Drug Discov.* 5 (2010): 1039–1045.

Hampel A., Labanauskas M., Connors P. G., Kirkegard L., RajBhandary U. L., Sigler P. B., Bock R. M. "Single crystals of transfer RNA from formylmethionine and phenylalanine transfer RNA's." *Science* 162 (1968): 1384-1387.

Hanson B. L., Martin A., Harp J. M., Parrish D. A., Bunick C. G., Kirschbaum K., Pinkertonb A. A., Bunick G. J. "Use of an open-flow helium cryostat for macromolecular cryocrystallography." *Journal of Applied Crystallography* 32 (1999): 814-820.

Hendrickson W. A., Horton J. R., LeMaster D. M. "Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure." *EMBO* 9 (1990): 1665-1672.

Holton J., Alber T. "Automated protein crystal structure determination using ELVES." *PNAS* 101 (2004): 1537–1542.

Jacquamet L., Ohana J., Joly J., Borel F., Pirocchi M., Charrault P., Bertoni A., Israel-Gouy P., Carpentier Ph., Kozielski F., Ferrer J-L. "Automated Analysis of Vapor Diffusion Crystallization Drops with an X-Ray Beam." *Structure* 12 (2004): 1219-1225.

Kabsch W. "Automatic Indexing of Rotation Diffraction Patterns." *J. Appl. Cryst.* 21 (1988): 67-71.

Kendrew J. C., Bodo G., Dintzis H. M., Parrish R. G., Wyckoff H., Phillips D. C. "A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis." *Nature* 181 (1958): 662-666.

Kim C. U., Kapfer R., Gruner S. M. "High-pressure cooling of protein crystals without cryoprotectants." *Acta Crystallographica* D61 (2005): 881–890.

Kriminski S., Kazmierczak M., Throne R. E. "Heat transfer from protein crystals: implications for flash-cooling and X-ray beam heating." *Acta Crystallography* D59 (2003): 697-708.

Leslie A. G. W. "The integration of macromolecular diffraction data." *Acta Cryst* D62 (2006): 48-57.

Matthews B.W. "Solvent content of protein crystals." *J Mol Biol.* 33 (1968): 491-497.

McPherson A. *Crystallization of Biological Macromolecules.* CSHL Press, 1999.

Perrakis A., Morris R., Lamzin V. S. "Automated protein model building combined with iterative structure refinement." *Nature Structural Biology* 6 (1999): 458-463.

Peyridieu J. F., Baudot A., Boutron P., Mazuer J., Odin J., Ray A., Chapelier E., Payen E., Descotes J. L. "Critical Cooling and Warming Rates to Avoid Ice Crystallization in Small Pieces of Mammalian Organs Permeated with Cryoprotective Agents." *Cryobiology* 33 (1996): 436–446.

Rasmussen D. H., MacKenzie A. P. "Glass transition in amorphous water. Application of the measurements to problems arising in cryobiology." *Journal of Physical Chemistry* 75 (1971): 967–973.

Shah B. N., Chinte U., Tomanicek S. J., Hanson B. L., Schall C. A. "Flash Cooling Protein Crystals: Estimate of Cryoprotectant Concentration Using Thermal Properties." *Crystal Growth and Design*, 2011.

Teng TY. "Mounting of Crystals for Macromolecular Crystallography in a Free-Standing Thin Film." *J. Appl. Cryst.* 23 (1990): 387-391.

Teng TY., Moffat K. "Cooling Rates During Flash Cooling." *Journal of Applied Crystallography* 31 (1998): 252-257.

Thomanek U. F., Parak F., Mossbauer R. L. "Freezing of Myoglobin Crystals at High Pressure." *Acta Crystallophica* A29 (1973): 263-265.

Veesler S., Boistelle R. "Diagnostic of pre-nucleation and nucleation by spectroscopic methods and background on the physics of crystal growth." In *Crystallization of Nucleic Acids and Proteins*, by Gigegé R., Ducruix A., 311-340. Oxford University Press, 1999.

Vekilov P. G. "Dense Liquid Precursor for the Nucleation of Ordered Solid Phases from Solution." *Crystal Growth & Design* 4 (2004): 671-685.

Walker L. J., Moreno P. O., Hope H. "Cryocrystallography: effect of cooling medium on sample cooling rate." *Journal of Applied Crystallography* 31 (1998): 954-956.

Warkentin M., Berejnov V., Husseini N. S., Thorne R. E. "Hyperquenching for protein cryocrystallography." *J. Appl. Cryst.* 39 (2006): 805–811.

Weis W. I., Kahn R., Fourme R., Drickamer K., Hendrickson W. A. "Structure of the calcium-dependent lectin domain from a rat mannose-binding protein determined by MAD phasing." *Science* 254 (1991): 1608-1615.

Williams Sh. P., Kuyper L. F., Pearce K. H. "Recent applications of protein crystallography and structure-guided drug design." *Current Opinion in Chemical Biology* 9 (2005): 371–380.

Abstract

Crystallography is from far the most contributing technique for the structure analysis of macromolecules at atomic resolution. In this thesis, instrumentation development issues to improve and accelerate experimental procedures for X-ray diffraction experiments are tackled. Indeed the preparation steps of protein crystals for X-ray diffraction data collection are the main causes of forming a bottleneck towards automated pipelines from protein crystallization to structure resolution.

Firstly, an emerging method in today macromolecular crystallography is the room temperature *in situ* X-ray diffraction of protein crystal samples in their crystallization drops, with proven benefits in crystal screening and also structure resolution. However, it requires a great number of crystals to be centered and diffracted in a row. Thus a fully automated system providing a solution to this requirement is presented and assessed in this manuscript as one of the results of this PhD studies. Secondly, in this manuscript, studies and developments on automating harvesting, cryo-protecting and flash-cooling steps of protein crystals preparation for X-ray diffraction are reported, as well as assessment experiments and results. With a new robotic approach, crystals are manipulated with a micro-gripper on a 6-axis robotic arm to prepare and to analyze crystals with 360° rotation possibility for cryo-temperature single wavelength X-ray diffraction. Lysozyme and NikA Fe-EDTA protein crystals has been prepared and diffracted with this new method. Structural comparisons show no differences between the new methodology and the manual one, while robustness, repeatability and experimental time are significantly improved. At last, different integration scenarios of the presented methodologies, highlights their interest in fully automated macromolecular crystallography pipelines.

Résumé

La cristallographie est la technique qui contribue le plus à l'analyse des structures des macromolécules biologiques à la résolution atomique. Dans ce manuscrit de thèse nous abordons des développements instrumentaux pour l'amélioration et l'accélération des étapes expérimentales dans la procédure de mesure de la diffraction aux rayons X. En effet, les étapes de préparation des cristaux de protéine à la diffraction aux rayons X constituent la cause principale du goulot d'étranglement dans les plateformes à haut débit de la cristallisation des protéines jusqu'à la résolution des structures.

Premièrement, la diffraction *in situ* aux rayons X des cristaux à la température ambiante, dans les plaques de cristallisation, est une méthodologie émergeante dans la cristallographie des protéines avec des capacités bénéfiques dans le criblage des cristaux mais aussi dans la résolution de structures. Cependant, un grand nombre de cristaux devront être centrés puis analysés par la diffraction aux rayons X automatiquement l'un à la suite de l'autre. Ainsi, un système automatisé répondant à cette exigence est présenté et évalué dans ce manuscrit comme étant l'un des résultats des études menées au cours de cette thèse. Deuxièmement, des études et des développements d'automatisation des étapes d'extraction et de micromanipulation, de cryo-protection et de congélation rapide pour la préparation des cristaux à la diffraction aux rayons X sont décrits dans ce manuscrit, ainsi que les résultats des expériences et des évaluations. Avec une approche nouvelle, les cristaux sont manipulés grâce à une micro-pince montée sur un bras robotique 6-axes pour les préparer et les analyser avec la possibilité de rotation de 360° pour la diffraction aux rayons X à longueur d'onde constate et à température cryogénique. Des cristaux des protéines lysozyme et NikA Fe-EDTA ont été préparés et diffractés avec cette nouvelle méthode. La comparaison structurale ne montre pas de différence entre la nouvelle méthode et celle manuelle, cependant la robustesse, la répétabilité et le gain de temps d'expériences sont significativement améliorés. Finalement, différents scénarios d'intégration des méthodologies présentées, met en évidence leurs intérêts dans les plateformes tout automatisés de cristallographie des macromolécules biologiques.