HHS Public Access

Author manuscript

Methods Enzymol. Author manuscript; available in PMC 2024 August 16.

Published in final edited form as:

Methods Enzymol. 2023 ; 688: 307–348. doi:10.1016/bs.mie.2023.07.009.

Chapter 10. Room temperature crystallography and X-ray spectroscopy of metalloenzymes

Hiroki Makita[1], **Miao Zhang**[1], **Junko Yano**[1], **Jan Kern**[1]

[1]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.

Abstract

The ultrashort (10's of femtoseconds) X-ray pulses generated by X-ray free electron lasers allow scientists to measure X-ray diffraction and spectroscopic data from radiation sensitive metalloenzymes at room temperature mostly avoiding the effects of radiation damage usually encountered when performing such experiments at synchrotron sources. Here we discuss an approach to measure both X-ray emission and X-ray crystallographic data at the same time from the same sample volume. The droplet-on-tape setup described allows for efficient sample use and the integration of different reaction triggering options in order to conduct time-resolved studies with limited sample amounts. The approach is illustrated by two examples, photosystem II that catalyzes the light driven oxidation of water to oxygen, and isopenicillin N synthase, an enzyme that catalyzes the double ring cyclization of a tripeptide precursor into the β-lactam isopenicillin and can be activated by oxygen exposure. We describe the necessary steps to obtain microcrystals of both proteins as well as the operation procedure for the drop-on-tape setup and details of the data acquisition and processing involved in this experiment. At the end we present how the combination of time resolved X-ray emission spectra and diffraction data can be used to improve the knowledge about the enzyme reaction mechanism.

Keywords

X-ray free electron laser; X-ray emission spectroscopy; protein crystallography; metalloenzymes; photosystem II; isopenicillin N synthase

1. Introduction

Metalloenzymes are central players in many biosynthetic and energy converting pathways in Nature. They are used to perform challenging chemical reactions with efficiency, speed, and complexity under ambient conditions and in a highly controlled manner in both time and space, by utilizing the flexibility of the protein, cofactors, and water network. To understand the mechanism of such enzymes at the atomic level, solving their macromolecular structures and combining these with spectroscopic studies of the catalytic site has been an important approach. X-ray crystallography (X-ray diffraction, XRD) at synchrotron facilities (SR) has played a central role for this purpose, along with nuclear magnetic resonance (NMR), mass spectroscopy, and especially more recently cryo electron microscopy (EM). Spectroscopic approaches for transition metal containing metalloenzymes are also an ideal target for X-ray spectroscopic methods, especially X-ray absorption and emission measurements

to observe transitions around the transition metal K-edge region and have been often combined with theoretical approaches to gain a deeper understanding of the details of the electronic structure of the active sites. Structural biology has seen important advances in two new methods, cryo-EM and X-ray free electron lasers (XFELs). The former enables the visualization of large complexes at the atomic scale, at resolutions that approach or are equal to those accessible by X-ray crystallography from ensemble measurements of single molecules without the need of crystals (Kuhlbrandt, 2014; Saibil, 2022). The latter, while requiring crystals, can give a view of macromolecules and their time-dependent structural changes under physiological conditions (Barends et al., 2022; Chapman et al., 2011; Neutze, 2014). In addition, the femtosecond pulses from XFELs allow scientists to obtain X-ray spectroscopic information either using separate specialized experiments or in a combined approach, where both the spectroscopic and diffraction signal is collected at the same time from the same sample volume. In this chapter, we describe the X-ray crystallography and X-ray spectroscopy methods at XFELs relevant to the structural biology for mechanistic studies of metalloenzymes.

Over the decades, synchrotron X-ray crystallography has been a standard tool for structural biology. The introduction of cryopreservation methods in the 1980's was a major advance in the field. By reducing the rate of secondary radiation damage caused by the diffusion of radicals, it improved the useful X-ray exposure lifetime of crystals, and enabled the collection of high-resolution data. Cryo-trapping is likewise useful for other synchrotron-based structural techniques like X-ray Absorption Spectroscopy (XAS) for the same reasons (Yano et al., 2005). However, cryo-trapping methods can only reveal intermediates within the time resolution of the freezing process (usually several milliseconds, but highly dependent on the method used) and there has been an increasing desire to study systems under functional conditions (Keedy et al., 2015; Thompson et al., 2019) in order to understand processes that control the directionality of reactions. XFELs have opened up the possibility of such approaches, by enabling room temperature studies, through the ability to collect data before the onset of radiation damage. Also, the method development for XFEL experiments has recently led to a technology transfer to synchrotron sources, enabling some of these experiments (for longer time scales and less radiation sensitive samples) at synchrotron sources and hence allowing for wider access and utilization of such approaches (see for example Schulz et al., 2022).

The X-ray pulses generated by an XFEL are extremely intense with short pulse widths, containing as many photons ($10^{12\text{-}13}$) in one pulse (typically < 50 fs, 10^4-fold shorter than pulses created by X-ray storage ring facilities) as synchrotrons generate in one second, focused into a μm-size spot. These beam properties, together with pulse-by-pulse “diffract and replace” data collection, make it possible to overcome radiation damage caused by the diffusion of hydroxyl radicals (~ ps) and even by the diffusion of solvated electrons (sub ps); the XFEL pulses are generally shorter (fs) than the time scale of these secondary damage processes, and therefore enable X-ray diffraction data collection at room temperature without resorting to cryo-techniques (Chapman et al., 2011; Kern et al., 2013; Spence, 2018). This is also true for X-ray spectroscopy data collection, which probes electronic state information and structure with high resolution in the local vicinity of a given element (Bergmann et al., 2021). By combining diffraction methods and X-ray spectroscopy,

scientists can examine the dynamics of the local chemical environment of the catalytic metal center of metalloenzymes in the context of the atomic structure of the entire enzyme (Fig. 1).

2. Sample Preparation

Protein crystallization can be considered as a ‘trial and error’ procedure because it depends on the interplay of multiple parameters such as protein and precipitant concentrations, pH, temperature, ionic strength, etc. Particularly, the three-dimensional structure with appropriate degrees of flexibility and dynamics of proteins are required to maintain their functionalities. But some of these characteristics – especially those that increase heterogeneity – impede the crystal formation procedure as the protein molecules in a crystal lattice are held by weak and non-covalent intermolecular interactions that include ionic, dipole, van der Waals interactions and hydrogen bonds. Crystallizability of a protein is determined in part, by its surface characteristics meaning the distribution of charged and polar amino acids on the solvent-exposed protein surface. That is why certain proteins cannot be crystallized due to their surface characteristics that are incompatible with the crystal lattice (Bijelic & Rompel, 2018; Rupp, 2013). As a general workflow, determining optimal crystallization conditions through several trials would be essential steps to obtain protein crystals. However, before starting to find suitable crystallization conditions, purity and homogeneity of a protein solution play a prominent role for crystal formation and growth.

XFEL studies are usually conducted using micron size crystals (between 1 and ~50 μm in the longest dimension). As this size range is not commonly used for synchrotron-based experiments, some modifications to traditional crystallization protocols are necessary. Here, we describe sample preparation steps for microcrystal generation based on two protein examples. Their crystallization conditions are well-established and the proteins selected are photosystem II (PSII) and isopenicillin N synthase (IPNS).

2.1. Crystallization kinetics and phase diagram

To obtain single crystals, a highly purified protein solution needs to be brought to a supersaturated stage, which is required to overcome the nucleation free energy barrier (Fig. 2). A spontaneous appearance of a new phase, the so-called phase separation, can only happen when a system is in a non-equilibrium condition. During the phase separation, protein molecules are transitioned from the soluble state into the solid-phase to form critical nuclei. The protein concentration in solution decreases as the nuclei grow beyond the critical size, moving to the so-called growth phase, where the ordered aggregation of the molecules takes place. In this process, more and more protein molecules attach to the nuclei surface to form a single crystal. In the last step, a crystal stops growing when the protein molecules in solution decrease significantly (Bijelic & Rompel, 2018; Erdemir et al., 2019).

From thermodynamic perspectives, the two events of nucleation and crystal growth occur spontaneously when the Gibbs free energy (ΔG) value is negative. Crystal formation and growth lead to entropy reduction of the protein molecules as their movements are restricted upon incorporation into a crystal lattice, thus becoming a more ordered system. However, at the same time structured water molecules from the protein’s surface can either be released

in the solvent or additional water molecules are trapped during crystal formation. According to the calculations demonstrated in the previous literature (Velikov, 2004, 2003; Velikov & Chernov, 2002), around 5 to 30 water or solvent molecules are released in the solvent from each protein molecule bound to the crystal lattice. In the disordered bulk, the released water molecules have more translational and rotational degrees of freedom, leading to an increase in entropy of the water or solvent molecules that predominates the comparably small entropy loss of protein molecules attached to a crystal. In most cases, the release of water molecules around the protein molecules (upon attachment of the protein molecules to the crystal) is the main thermodynamic driving force for crystallization (Derewenda & Vekilov, 2006; Vekilov, 2019).

Protein crystal formation requires promoting a protein solution from the undersaturated state to the supersaturated states. This is illustrated by the phase diagram in Fig. 3. The solubility line divides the diagram into two zones: an undersaturated and supersaturated zone. This line represents the equilibrium between the existence of the solid phase and the free-molecule phase. Protein crystallization is not possible in an undersaturated zone as protein molecules stay in solution within this zone. The supersaturated region is subdivided, depending on the states of supersaturations, into three zones: metastable zone, nucleation zone and precipitation zone. The metastable zone is a region for optimal crystal growth, but no nucleation takes place within this region. If supersaturation proceeds, spontaneous crystal formation occurs within the nucleation zone. Here, both nucleation and crystal growth are possible. Increasing the conditions towards an even higher supersaturation, the precipitation zone is reached where proteins precipitate amorphously (Bijelic & Rompel, 2018).

2.2. Crystallization techniques

Different crystallization techniques (Shoeman et al., 2023, Schulz et al., 2022, Beale et al., 2019, Giegé, 2013, Rhodes, 2006, García-Ruiz, 2003) have been proposed to obtain single crystals. Here, we describe three most commonly used crystallization techniques: **Vapor diffusion**, **batch crystallization and crystal seeding.**

The **vapor diffusion** technique can be performed in two setups, the hanging or sitting drop format. Considering the hanging drop setup as an example, a small droplet, typically a few microliters between 0.1 and 5 μL, of sample mixed with crystallization buffer is placed on a siliconized glass cover slide. Crystallization buffer containing precipitant reagent is located in the reservoir. The initial reagent concentration is usually higher in the reservoir well than in the drop. The concentration gradient is needed to trigger the equilibrium process between droplet and reservoir. During this process wherein the drop and well are in a sealed environment, water from the droplet will evaporate towards the reservoir. As a result, the volume of the drop is decreased, the protein and precipitant concentration in the drop increase slowly until the solubility limit of the protein is exceeded to reach supersaturation. This is the requirement to initiate the crystallization process. Crystal formation and growth will occur under proper crystallization conditions.

Batch crystallization is based on the direct mixing of the protein and precipitant. There is no equilibrium with a reservoir needed. Through this method the protein is brought into the

nucleation zone. When starting from a known crystallization condition, crystals can appear within a few minutes or hours.

Crystal seeding is a powerful tool to grow a large quantity of uniform crystals. A major advantage of using seeding techniques is to control the crystallization process based on separating the nucleation and growth events. When using the microseeding method, submicroscopic seeds (crushed crystals) are brought to a crystallization agent fulfilling the optimal conditions for the metastable zone where crystals from seeds can grow but no spontaneous nucleation can occur.

2.3. PSII purification and crystallization

PSII thylakoid membranes are isolated from cells of *Thermosynechococcus vestitus* (*T. vestitus*, previously named *T. elongatus*) and prepared according to (Kern et al., 2005). Particularly at the initial steps of isolation of PSII thylakoids, washing with high salt can remove a large amount of phycobilisomes. However, isolated thylakoids can still contain further impurities such as remaining phycobilisomes and other soluble proteins. Since high degrees of purity and homogeneity of protein samples are prerequisite to obtain high-resolution crystal structures as mentioned above, further purification steps of PSII thylakoids are essential. Isolated PSII thylakoid membranes are solubilized by using octaethylene glycol monododecyl ether ($C_{12}E_8$) detergent. The solubilization step is important in order to extract crude PSII particles from the membrane when using appropriate concentrations of detergent. The crude PSII extract after solubilization is further purified through two steps of weak and one step of strong anion exchange chromatography in order to obtain pure PSII dimers (dPSII, MW = ~ 750 kDa (Zouni et al., 2005)). Detailed information about the purification steps of PSII dimers using $C_{12}E_8$ detergent is given in (Hellmich et al., 2014). The purification effect and the purity of PSII dimers are demonstrated by the Blue native PAGE analysis shown in Fig. 4. Crude PSII particles are shown in lane 1. After column chromatography, PSII dimer in lane 2 shows a weak band in position equivalent to that of PSII monomer (mPSII, MW = ~ 350 kDa (Zouni et al., 2005)). When detecting small amounts of PSII monomers present in the PSII dimers after column chromatography, recrystallization can be used as an additional purification step in order to obtain highly purified and monodisperse PSII dimers. The result after re-crystallization is shown in lane 3.

2.3.1. Crystallization in microbatch plates—Crystals were obtained using the microbatch method in 96-well IMPACT Plates (Greiner Bio-One GmbH). Highly purified PSII protein at a final concentration of 2 mM Chl (chlorophyll) is mixed 1:1 with the crystallization buffer that contains 0.1 M Tris (pH 7.5), 0.1 M $(NH_4)_2SO_4$, 14-19 % polyethylene glycol monomethyl ether 5000 (PEG 5000 MME). Crystals grown at 20 °C having a size between 100 to 200 μm could be observed after 1 or 2 days. (Adapted from (Hellmich et al., 2014))

2.3.2 Microseeding procedure—For making seed stock solution, a few microcrystals were selected from a microbatch plate. On average between 5 and 10 crystals of a size between 100 and 200 μm (longest axis) shown in Fig. 5 were transferred to 50 μL of 15% PEG 5000 buffer (0.1 M Tris pH 7.5, 0.1 M $(NH_4)_2SO_4$, 15% PEG 5000 MME (w/v),

0.013% $C_{12}E_8$). A seed bead tool kit (Hampton Research) was used to crush these crystals by vortexing for 5 minutes. The mixture was centrifuged at 13,000 r.p.m. for 1 minute and vortexed again for another 5 minutes. After vortexing, the crushed crystal suspension was mixed with 450 μL of buffer 15% PEG 5000 buffer with detergent to prepare the first seed stock.

Microseeding crystals of PSII can be made by mixing the PSII protein sample at 2 mM Chl concentration and the seed stock solution in a ratio of 1:3 (for example: 10 μL of PSII protein and 30 μL of seed stock solution). The final PEG concentration is now 11.25%. The mixed solution is sequentially placed in an incubator at a constant temperature of 20 °C. Under this condition, a clear microcrystal pellet usually appears after 1 hour of growth time. The size of microseeding crystals is on average between 20 and 50 μm (Fig. 6). (Microseeding protocol modified from (Ibrahim et al., 2015))

2.3.4. Post-crystallization treatment—According to the post-crystallization protocol in (Hellmich et al., 2014), microseeding crystals are treated by stepwise transfer to a buffer containing the final PEG concentration of 35 % (w/v), 0.1 M Tris (pH 7.5), 0.1 M $(NH_4)_2SO_4$ in steps of 5% PEG for 5–10 min. This crystal treatment is conducted at room temperature. Stepwise buffer exchange eventually results in the reduction of detergent and solvent content. Crystals are thus transformed from type II to type I showing a clear resolution improvement (Hellmich et al., 2014).

After 30 minutes incubation in the Tris buffer at pH 7.5, the crystal suspension is transferred to a buffer that contains PEG concentration of 35 % (w/v), 0.1 M MES (pH 6.5), 0.1 M $(NH_4)Cl$. The buffer exchange to MES at pH 6.5 is essential as keeping PSII samples too long in buffers at higher pH leads to reduction of the protein activity, meaning a reduction of the rate of light-dependent generation of O_2. Previous studies showed that nearly full activity can be restored in PSII samples when exchanging the sample back into the buffer at pH 6.5 (Young et al., 2016) within about 5 h after initial exposure of the sample to high pH. The final Chl concentration of PSII microcrystal suspension typically used for the XFEL experiments is between 0.5 and 1.2 mM. The number of crystals was determined by an automated cell counter (Invitrogen Countess II). 10 μL of microcrystal suspension was added to a cell-counting chamber slide and about 2×10^6 crystals ml^{-1} were estimated when the Chl concentration is approximately 0.5 mM and the average crystal size is about 10 μm.

2.4. IPNS

Highly purified (>95% by SDS–PAGE) recombinant IPNS was produced and purified as described (Roach et al., 1996) and stored in 25 mM Tris, pH 8.0. All protein handling and crystallization were conducted under anaerobic conditions (< 2 ppm O_2) and a N_2 atmosphere inside an anaerobic chamber (Belle Technologies, UK) and at room temperature. IPNS crystals were obtained through the hanging drop method by using 24-well VDX plates (Hampton Research, US). To obtain the IPNS crystallization solution, 4 μL of freshly prepared 100 mM $FeSO_4$ were mixed with 80 μL of 50–52 mg mL^{-1} IPNS, followed by 4 times addition of 5 μL of ACV (l-δ-(α-aminoadipoyl)-l-cysteinyl-d-valine) (2.0 mg in 20 μL 25 mM Tris, pH 8.5). To screen the optimum conditions for macrocrystal generation,

different conditions for the reservoir solution were generated by using 0.1 M Tris at different pH values between 8.1 and 8.7 and changing the salt concentration of Li_2SO_4 between 1.5 and 2.0 M. The crystallization droplet contained 3 μL of the reservoir solution mixed with 3 μL of protein solution and were equilibrated against 500 μL of the reservoir solution. Needle, plate- and hexagonal-shaped crystals appeared within 2 days at room temperature and a few 100 −150 μm long needle-shaped crystals (3-5 crystals harvested into 50 μL reservoir solution) obtained from these first round of crystallization setups were used to prepare seeds using the Seed Bead Kit as described by the manufacturer (Hampton Research, USA).

The batch method was applied to obtain microcrystals. In this procedure, a 96-chimney well plate (Corning, USA) was used. For each well, 6.5 μL of the IPNS–ACV–$FeSO_4$ mixture prepared as described above was mixed with 90 μL of 1.7 M Li_2SO_4 and 0.1 M Tris pH 8.5. To the mixture, 1 μL of the seed stock was added. The plate was sealed (Polyolefin StarSeal, Starlab, UK) and the mixed solution in the plate was kept under agitation using a microplate shaker (SciQuip, UK) at 700 r.p.m., until a crystal size between 40 and 60 μm (longest dimension) had been reached (Fig. 7). Microcrystal growth usually took between 16 and 30 h, depending on the temperature and the humidity of the anaerobic chamber. A final concentration of ~2×10^7 crystals ml^{-1} was typically achieved. (Adapted from (Rabe et al., 2020))

3. Sample Delivery System

The pulse-by-pulse “diffract and replace” data collection method as described in the Introduction section using fs XFEL beams is known as serial femtosecond crystallography (SFX) and requires a fresh batch of microcrystal sample to be delivered for (ideally) each pulse of XFEL beam to ensure efficient data collection. To conduct an SFX experiment that maximizes the beam usage and minimizes the sample waste, a sample delivery system that accounts for controlled infusion and ejection of the microcrystal sample is required. A variety of sample delivery methods have been developed to satisfy this need including on-demand ejection of crystal suspension droplets (Chen, 2020). This droplet-on-demand approach is versatile in accommodating different experimental conditions, and has been integrated with other sample delivery methods such as micromesh and polyamide conveyor belt systems (Soares et al., 2011; Roessler et al., 2013; Fuller et al., 2017). With proper design and installation of the sample delivery system, it has been demonstrated that both the XRD and X-ray emission spectroscopic (XES) data can be collected simultaneously. Combined with a flexible reaction triggering module, time-resolved crystallographic and spectroscopic studies on protein dynamics that probes on the structural and electronic state of protein intermediates is made possible. In this section, the operation of the Drop-on-Tape (DOT) sample delivery system (Fuller et al., 2017), one of such methods that enable simultaneous time-resolved XRD and XES measurements on metalloenzymes at room temperature, is detailed.

Full design and composition of the DOT system are described in (Fuller et al., 2017), and the core components are detailed in Fig. 8A. An image of the actual deployment of the DOT system, including the peripherals that are described in later sections, is shown in Fig.

9A. Briefly, the core of the system is composed of a conveyor belt and an acoustic droplet ejector (ADE) (Fig. 8A, a-c). ADE produces a droplet of discrete volume using an ultrasonic transducer (Fig. 8A, a) (Hadimioglu et al., 2016). The ejected droplets are deposited onto a continuously moving polyimide conveyor belt, which transports a queue of droplets through the reaction triggering module and to the XFEL interaction point (Fig. 8A, i). The droplet on tape is a hemisphere with ~200-300 μm in diameter and 0.8-6 nL in volume. The ejection rate is variable, with an upward rate of 120 Hz. The mechanism of droplet ejection by an acoustic transducer does not require a confined sample reservoir. As such, relatively large microcrystals with a longest dimension in the range of ~100 μm can be delivered with ease using this system. The speed of the conveyor belt is a key variable that controls the timing of reaction triggering and is detailed in the sections below. In standard operation, the conveyor belt is run at ~200-300 mm/s.

3.1. Sample Handling and Loading

To reduce the interaction of the XFEL beam with the atmospheric gases, and to shield O_2-sensitive samples from exposure to atmospheric oxygen in the chemical activation experiment, the main components of the DOT system, which includes the mounted core components and motorized stages and cameras for visualization (Fig. 8B), are enclosed in a large gas-tight chamber that is flushed with He. The sample is continuously infused into this chamber and to the reservoir pocket of the ADE (Fig. 8A,C) via a capillary feed line of 200 or 250 μm inner diameter connected to a 1-mL gas-tight syringe. In the case of PSII, which is light sensitive, the crystal suspension is loaded into a syringe (Hamilton Model 1001 PTFE Luer Lock, Hamilton Company, Reno, NV, USA) at ~0.5-1.2 mM Chl concentration and dark-adapted for 1 hr before transporting to the experimental hutch. An aluminum foil is applied to the syringe to shield the sample from light exposure. The DOT chamber is covered by a regular or matte black aluminum foil, and the experimental hutch is kept in dark when possible. For IPNS and other O_2 sensitive samples, the suspension is loaded into a gastight syringe with a sample lock (Hamilton Model 1001 Sample Lock termination, Hamilton Company, Reno, NV, USA), and is transported to the experimental hutch in gas-tight container with an oxygen absorber (AnaeroPack-Anaero System, Mitsubishi Gas Chemical America, Inc., New York, NY, USA).

3.2. Sample Infusion and Ejection

The sample infusion rate is controlled by a syringe pump with a remote pump head (Legato 130, KD Scientific, Hamilton, MA, USA). The remote head of the syringe pump is mounted on a syringe shaker that rotates the pump head and syringe between 0° and 180° on an axis perpendicular to the long axis of the syringe to keep the microcrystals suspended in the syringe for the duration of the experiment. The rotation rate and duration of stops at each end of the rotation cycle need to be optimized for each crystal sample, depending on crystal size/shape and buffer density. The infusion rate is dependent on the droplet ejection rate and size, which in turn is determined by the rate of detection and frequency of XFEL pulses. At 30 Hz ejection rate, an infusion rate of 7-9 μL/min is sufficient to yield stable ejection of droplets with 4-5 nL volume. When using the above-described conditions for PSII and IPNS between 8 - 20 PSII or around 60 - 80 IPNS microcrystals are contained in a single droplet. The gas tight syringe pump sitting on the syringe shaker can be enclosed in

different ways depending on the sample needs. For temperature sensitive samples the shaker can be mounted inside a refrigerator with a feedthrough for the sample capillary (care has to be taken that the syringes used for sample delivery are compatible with the temperature setting of the refrigerator. Gas-tight Hamilton syringes for example are not sealing well at temperatures below 12°C). For light sensitive samples the shaker can be placed inside a light tight box or enclosure. The capillary line feeding out of the enclosure and to the DOT chamber is shielded from light by an opaque capillary tubing sleeve (PEEK Capillary Tubing Sleeve, IDEX Health & Science, LLC, Oak Harbor, WA, USA). For O_2 sensitive samples, a top loading enclosure filled with Ar gas is used.

Sample droplets are generated using the ADE as mentioned above. The main components of the ADE system are: a focused ultrasonic transducer, a housing for the transducer and the immersion water, and a sample reservoir pocket with a connection to the sample feed line (Fig. 8A). A sample droplet is produced by focusing sound energy on the interface between the liquid containing the microcrystal suspension and air to pinch off a small volume of sample into the air. The sample is introduced to the focal point of the ultrasonic transducer by infusing a set volume of sample at the specified infusion rate (from the syringe and via the capillary line) into the reservoir pocket (Fig. 8A, c). The sample reservoir pocket is cylindrical in shape and has a dimension of 2.5 mm in diameter and 2.0 mm in depth and has a capacity of 10 μL. In addition to the sample feed line, one to two additional feed lines are guided to the reservoir to infuse reservoir cleaning water/solution. Besides this reservoir with the standard dimensions, a smaller (1.55 mm diameter x 1.25 mm depth, 2.4 μL capacity) pocket has been employed to produce stable sample infusion and droplet ejection for challenging samples.

The droplet ejection is managed by sending a 15 MHz tone burst to the ultrasonic transducer. Here, a 15-MHz focused immersion transducer is employed (Olympus V319-SU, 1.00-in spherical focus). A waveform of a tone burst is generated using a function generator (Agilent 33600A). Timing and repetition rate of the waveform are controlled by a low-jitter digital delay generator (DG645, Stanford Research Systems). The repetition rate is adjusted to match the frequency of the XFEL pulse. This waveform is amplified by a 150 W 250 MHz or 500W 250MHz RF amplifier (150A250 or 500A250C, AR RF/Microwave Instrumentation) to a driving signal of ~40 V. The amplified tone burst is sent to a 20-dB bidirectional coupler (ZFBDC20-61HP+, Mini-Circuits, Brooklyn, NY, USA) that transmits the tone burst to the ultrasonic transducer. Along with the tone burst for droplet ejection, a probing signal is sent to the transducer. The probing signal is reflected upon interacting with the liquid-air surface, and is redirected as a return signal by the coupler. The reflected probe signal functions as an “echo” that indicates the fluid level in the sample reservoir. The filling of the reservoir changes the distance between the transducer focal point and the liquid-air interface, and the maximum “echo” signal indicates that the sample is filled to the focal point. Using this “echo” signal as a probe, the ejection rate (outflow rate) is adjusted by varying the burst cycle and amplitude sent out by the function generator. Once the “echo” is stabilized at the maximum signal amplitude, the infusion and ejection rates are balanced.

3.3 Tape Drive Control

The droplets ejected from the ADE setup are deposited onto a seamless Kapton polyimide tape (BPB, Inc.) that is 31 inches in length, 1.5 mm in width, and 25 μm in thickness. Two linear stages are used to control the positioning of the ADE setup with respect to the tape (Fig. 8B). A linear stage with 25-mm travel range (Newport, MFA-CC) controls the movement in the y-axis to adjust the vertical distance between the ADE setup and the surface of the tape. A second linear stage with 25-mm travel range (Newport, UMR8.25), motorized by a high-load actuator (Newport, LTA-HL), is used to adjust the position of the ejected droplet on the width of the tape in the z-axis. Unless deemed necessary, the ejected droplets are positioned at the center of the tape width. The seamless tape is passed through a series of low friction bearings with radially crowned aluminum caps, and tape transport components. The tape transport components include a tension arm (Fig. 8A, f), a capstan roller (Fig. 8A, e), and two pinch rollers to press the tape against the capstan (Fig. 8A, d). The shaft of the capstan is coupled to a stepper motor (Lexium M-Drive LMDCE) which serves as the only motorized drive for the tape control. The speed of the conveyor belt is adjusted via the control of this stepper motor. The set of rollers, the tape transport components, the motor drive, and the ADE setup are mounted onto two aluminum optical breadboards that are assembled using an aluminum angle bracket. This construction that carries the tape drive and ADE setup is then mounted onto a motorized linear stage with 100-mm travel range (Newport, UTS100CC) to enable adjustment of horizontal positioning of the DOT setup with respect to the XFEL beam in the x-axis. Details of the mechanical setup including construction drawings are described in (Yano et al., 2017).

3.4 Delivering Droplets to the XFEL Beam

A brief procedure to bring the ejected droplets on the tape drive to interact with the XFEL beam is as follows. To facilitate each of these steps, a GigE camera with a zoom lens attachment is installed to provide a view of the X-ray interaction point nearly collinear with the direction of the XFEL beam with a sufficient magnification to visualize a ~200 μm droplet (Manta, Allied Vision, Stadtroda, Germany). The GigE camera is triggered to run at the same frequency as the XFEL beam, and its timing and exposure time is adjusted by the digital delay generator (DG645, Stanford Research Systems) and the camera's internal controls. First, to detect the position of the XFEL beam in the view of the collinear camera, a piece of reference scintillator (such as GaAs or YAG(Ce)) is brought into the predetermined XFEL/sample interaction point. This scintillator is mounted on a reference mount setup that is installed on the same 100-mm linear stage for x-axis movement as the tape drive setup at the same z-position as the droplet but at the offset x-position. The reference mount setup is brought into the XFEL interaction point using this linear stage. Using a long exposure time (to cover the period of the XFEL frequency) on the collinear camera, a beam position can be identified and marked on the camera view. In a sequential fashion, reduce the exposure time of the camera and adjust the timing of the camera acquisition such that the fluorescence from the XFEL beam on the scintillator is not blinking with 100 μs of exposure time or less. Once the position and timing of the XFEL beam is determined on the collinear camera, translate the DOT system in x-axis using the 100-mm linear stage to bring the tape to approximately 500 μm of the marked XFEL position. Generate a series of stable water (i.e., non-sample) droplets at the frequency that matches the XFEL. Further translate the DOT

but in a small (~10 μm) increment in x-axis while scanning through the timing of droplet ejection until a droplet explosion is observed on the collinear camera (Fig. 9C), or a water ring is detected on a diffraction detector. Through the course of an experiment, a buildup from the exploded sample droplet or infusion of inhomogeneous sample or air bubbles into the ADE setup may result in offset in timing and positioning. These slight offsets are corrected by adjusting the timing of droplet ejection at a 50 μs increment, and repositioning the x-axis linear stage that carries the DOT setup at a 5 μm increment.

3.5. Peripherals

The DOT system that includes the ADE setup, the tape drive setup, and the linear stage that carries these components are enclosed in a gas-tight housing. The housing is made of 0.5 inch acrylic with a cutout for the X-ray emission spectrometer on the top and for the X-ray diffraction detector on the downstream side. These cutouts are sealed by inflatable plastic bags that press against the analyzer crystals for the XES and a detector window for XRD (a protector surface in front of the actual detector elements) when inflated with He flow/circulation. The upstream panel includes a port for introducing the XFEL beam into the enclosure, and a cutout sealed by an anti-reflection coated Nd:YAG laser line window for introducing a free space excitation laser (described below). One of the side panels is fitted by a feedthrough panel to pass capillary feed lines as well as communication and signaling cables into the enclosure. To create a low-O_2/high-He environment inside the enclosure, a diaphragm pump (KNF Neuberger Inc., N1400) is used to purge the enclosure while flowing He at 50 – 90 SCFH to replenish the atmosphere. Depending on the flow rate of He, the enclosure can be purged to <2% O_2 in 10 to 90 min. Once the <2% level is reached, the helium flow is reduced to ~20 SCFH and the diaphragm pump is set to recirculate the helium within the enclosure. Alternatively, to further reduce the overall He consumption, the enclosure can be purged by flowing N_2. In this configuration, the flight tube collinear to the beam (depicted in Fig. 8B) and the X-ray emission shielding (Fig. 9Ag) are sealed and filled with He to secure the path of the XFEL beam and the XES photons is in a low scattering He environment. In the case of non-O_2 sensitive samples such as PSII, a configuration without purged enclosure is also possible. In these cases, a beamstop or a collimator is mounted downstream of the IP to reduce the air scattering of the beam from reaching the XRD detector.

Upon interacting with the XFEL beam, the sample droplet is exploded, and the majority of the volume is removed from the tape. These sample residues, as well as any “missed” droplets, are a potential source for coating the tape surface. This has two negative effects: it reduces the hydrophobicity of the tape, leading to flattening of newly ejected droplets to a smear or "pancake" shape instead of a hemisphere, and if moisture collects on the tape, then this could lead to derailing the tape off the rollers. Additionally, coating of the tape by a sample residue is highly detrimental to the collection of XES as the accumulated residue could generate an XES signal that overlaps with a signal from the sample droplet. Therefore, proper cleaning and drying of the tape is essential to both the XRD and XES experiment. To remove the residue and “missed” droplets, the tape feeds into two cleaning stages past the XFEL/sample interaction point. In the first stage, a generic submerged water pump (TotalPond, 1200 gph waterfall pump) is used to spray water onto the tape via two 1-mm ID

tubes with combined flow of 150 mL/min. In the second stage, four high pressure water jets with a combined flow rate of 50 mL/min are directed to the tape using a low pulsation HPLC (high-performance liquid chromatography) pump (Teledyne, LD-class HPLC). The outflow of the sprayed cleaning water is collected and removed from the enclosure by a peristaltic pump (MasterFlex, L/S driver fitted with Easy-Load II pump head). Cleaning water, which uses distilled water, is divided into two separate containers each with 14-gallon capacity. Initially, 12 gallons of distilled water is supplied to a tank for the second stage HPLC pump, and 2 gallons of distilled water is filled to the tank for the first stage submerged water pump. The outflow guided by the peristaltic pump is fed into the first stage tank. The cleaning water, therefore, is recirculated for the first stage but unidirectional for the second stage. Beyond the two cleaning stages, the tape is passed through a drying stage which injects compressed helium with two blowoff nozzles from the top and bottom of the tape.

To improve stability of droplet deposition on the tape the static charge of the tape needs to be minimized. For this purpose, an antistatic gun (Ptec IN3425 AC Ionizing Nozzle, Transforming Technologies, LLC., Toledo, OH, USA) is directed at the tape directly upstream of the sample deposition region and provided with a continuous flow of He gas.

3.6. Reaction Triggering

Between the ADE setup and the XFEL/sample interaction point of the tape drive, a variable reaction triggering module can be installed to activate a reaction in the protein with controlled delay times (Fig. 8B, j). Here, two types of reaction triggering systems, the photoactivation module and the O_2-incubation module, are described.

3.6.1. Photoactivation—For samples that require multiple, precisely timed actinic flashes to advance the reaction cycle (such as PSII), a laser excitation system is installed in the module space (Fig. 8D). In this system, fiber-coupled lasers are guided with a specific spacing on the tape to enable up to three events of photoactivation. In order to align the droplet both in space and in timing to multiple excitations, two fiber-guided IR light sources (>800 nm) are positioned equidistant between the first/second excitation sources and second/third excitation sources. The droplets passing through the IR lights scatter the light, and the intensity and timing of the scattering are probed by photodiodes. These scattered lights serve as gates that indicate the positioning and timing of the droplets on the tape. In total, the laser excitation system holds three fiber-coupled lasers for the actinic source and two fiber-guided IR lights that function as the IR gates. The output of all five sources of light emitted through the tape are monitored remotely via an oscilloscope.

The timing of the excitation laser on the photoactivation module needs to be synchronized in relation to the timing of droplet ejection as well as the tape speed. The timing of the droplet ejection (deposition delay) itself is synchronized to the timing of the XFEL beam in a manner described above. To synchronize the excitation lasers on the droplet for a given deposition delay, the tape speed is first adjusted to create a temporal overlap on the IR gate signals on the oscilloscope. The overlap indicates that a droplet is present on each of the IR gates at the same time (the separation between two droplets is equal to the distance between two gates). Because the two gates are separated by the same distance as the space between

two adjacent excitation laser sources, and a separation between a gate and excitation laser is half of the separation between two gates (or two laser sources), a combination of the tape speed that results in a droplet being present at all three laser source points can be calculated. By adjusting the delay of the laser flashes such that the timing overlaps with the timing of the IR gates and using the tape speed calculated by the spacing between the gates and laser sources, spatial and temporal overlap between droplets, excitation laser flashes, and XFEL beam can be achieved. A series of combinations and graphical representation of the tape speed calculation are detailed in (Fuller et al., 2017). Briefly, for a standard operation of the PSII experiments, a droplet ejection rate is set at 20 Hz and the tape speed of 300 mm/s is applied.

For time-resolved photoactivated measurements, a free space laser source is introduced near the XFEL/sample interaction point, separate from the laser excitation system described above. Depending on the laser source, the delay between photoactivation and XFEL interaction can range between fs-μs timescale or ns-ms timescale. The free space laser is introduced into the enclosure and DOT via a laser line window installed on the upstream panel of the enclosure, and is guided onto the tape by the kinematic mirror mount with piezoelectric controls. The spot size on a tape is typically set to 400 μm in diameter, which is sufficient to cover a ~250 μm droplet. To properly guide the free space laser onto the tape and ensure a sufficient overlap with a droplet, another GigE camera is installed to provide a view of the surface of the tape around the interaction point. The camera, like the collinear one, has delay and repetition rate controlled by a digital delay generator to match the frequency and timing of the XFEL/sample interaction. Defining the timing of the XFEL/sample interaction as t = 0, to measure the XRD and XES data of a photoactivated sample of, for example, 500 μs after photoactivation, the following steps are followed. First a good temporal and spatial overlap between free space laser, and XFEL beam, and a sample droplet is confirmed at t = 0. Second, the delay on the side-viewing GigE camera and the free space laser is decreased by 500 μs. The camera shows the position of the droplet and free space laser 500 μs prior to the arrival of the XFEL beam. Using the piezo-controlled kinematic mirror mount, the laser spot is adjusted to overlay the spot onto a droplet. Lastly, the delay on the camera is reset to 0 to confirm that a temporal and special overlap between a droplet and XFEL beam (but not the free space laser which delay stays at −500 μs) is unaffected by this procedure.

3.6.2. O_2-activation—Many of the catalytic reactions of metalloenzymes are activated by the presence and uptake of O_2 in the surrounding atmosphere. To investigate the protein dynamics of such O_2-activated reactions, a reaction module can be replaced by an O_2-incubation platform (Fig. 8E). The platform is composed of an orifice with a series of chambers for the tape to pass. The central chamber extends for ~60 mm in length and is sandwiched by a pair of two 5 mm chambers, which are flanked by two chambers with 25/39 mm in length. An isolated region filled with O_2 gas is created in the central chamber by introducing a total flow of 1-2 SCFH into the chamber. The O_2- content in the chamber is monitored by an oxygen meter directly attached to this chamber. To limit O_2 from infiltrating the rest of the enclosure, the other chambers are used as a differential pump: the 5 mm chambers on either side of the O_2 chamber are connected to a vacuum line

which creates negative pressure, while the outermost chambers have helium flow to create positive pressure. Using the O_2-incubation platform, the main chamber can be saturated with O_2 (>90%) while retaining a low level of O_2 in the rest of the DOT enclosure (<2%). The exposure time is mainly controlled by adjusting the tape speed. The length of the incubation chamber (60 mm) and the distance from the end of the chamber to the X-ray interaction point (IP) (90 mm) can be used to calculate the O_2 incubation time as well as the total delay time between start of O_2 incubation and probing the sample with X-rays as a function of the belt speed. At the standard belt speed of 200 mm/s, this results in an incubation time of 300 ms and a total delay between start of incubation and measurement of 750 ms. By varying the belt speed, these times can be changed over a long range from about 200 and 500 ms (belt speed of 300 mm/s) up to 6 and 15 s (belt speed of 10 mm/s), respectively. For very long incubation times, care has to be taken that the sample is not drying out during the passage to the IP. For this purpose, the relative humidity in the He enclosure of the system and the O_2 incubation gas can be regulated to ~20 - 30% by bubbling either or both gasses through a wash bottle before feeding them in the experimental enclosure. Depending on the buffer composition, it is still common to observe formation of dried out PEG or salt in a form of powder diffraction, at these conditions and this is often the limiting factor in determining the longest possible incubation time. Also, it is important to note that the spacing between individual droplets becomes very small for slow belt speeds (e.g., at 250 μm drop diameter, 10 mm/s belt speed and 30 Hz deposition frequency the gap between individual drops is only 80 μm) and the deposition frequency might need to be reduced to accommodate very long incubation times.

4. XES Data Collection

Tandem XES on the time-resolved SFX experiments on metalloenzymes using XFEL is highly advantageous due to XES's high element specificity and XFEL's peak brightness that enable shot-by-shot collection of the spectra. Given these advantages, XES has been applied concurrently in a series of experiments to provide *in situ* confirmation of the reaction kinetics, batch-to-batch sample integrity, and oxidation state of the active site. Using the DOT system, XES of two different transitions (Kα and $K\beta_{1,3}$) of three different elements (Ni, Fe, and Mn) has been collected using microcrystals of metalloenzymes in time-resolved SFX experiments.

The energy of the incoming XFEL beam must be set above the absorption edge of the element to be detected. The energy-dispersive spectrometer is based on the von Hamos geometry, and the assembly and performance of such a spectrometer is detailed in (Alonso-Mori et al., 2012). Briefly, the spectrometer consists of a series of cylindrically bent analyzer crystals that are mounted on motorized stages to control pitch, yaw, and linear translation. In the case of Kα XES, 2 to 4 analyzer crystals are utilized. For the $K\beta_{1,3}$ signal which is approximately an order of magnitude smaller than the $K\alpha_1$ signal, 16 crystals assembled as a 4 x 4 array are used. The spectrometer is mounted on two motorized linear stages and one rotational stage to enable adjustment in position horizontally and vertically (both directions are normal to the axis of the incoming XFEL beam). The diffracted energies were collected by a position sensitive detector (SLAC National Accelerator Laboratory, ePix100). Placement of the spectrometer and position sensitive detector is dependent on the direction

of polarization of the XFEL beam to avoid the largest elastic scattering signal. For the current polarization of the Linac Coherent Light Source (LCLS) facility, the spectrometer is mounted on a vertical plane intersecting the interaction point, while the detector is placed on the horizontal plane (Fig. 8C). Using multiple sets of analyzer crystals and position sensitive detectors, XES signals of multiple different metals can be collected simultaneously, along with the XRD data.

To align the spectrometer and detector, one of the two methods can be applied: 1) Use the DOT system to generate stable droplet of a solution sample of molecular standards (such as FeCl or MnCl) in ~1-5 mM concentration, or 2) Use a metal foil of <100 μm thickness. The foil is to be placed at the XFEL/sample interaction point but tilted to ~45° around the horizontal axis to ensure that the emitted radiation is not blocked by the foil itself. For this option the foil can be mounted on the reference mount setup described in the previous section and moved into the beam interaction region using the x-axis motorized translation stage of the DOT system. Using method 1, an alignment can be made with the accurate/actual droplet position. This approach is also highly reproducible as the relative droplet position is controlled by motorized stages. A solution of concentrated molecular standards, however, is a potential source of contamination of the setup as the concentration of the metals in microcrystal protein samples are an order of magnitude smaller. A metal foil in method 2, on the other hand, does not come into physical contact with the tape drive setup and will not introduce contamination to the system. However, the relative position of the foil is not coupled to the droplet position, and a slight discrepancy may exist between both the metal foil and the actual droplet sample with respect to where the signal appears on the position sensitive detector. As this can also lead to slight distortion in the energy axis, care must be taken that the foil is located as close as possible in the same position along the X-ray beam as the actual sample. To ensure that data taken during different times of an experiment are comparable it is important to regularly check the alignment of the spectrometer (Fig. 10).

For the collection of XES, the number of shots to be averaged depends on the metal concentration, the type of spectra to be collected, and number of analyzer crystals that are employed. For a typical PSII sample described in this chapter, using a spectrometer with cylindrically bent 16 Si(440) crystals (R = 500 mm), approximately 25,000 shots or, when operated at 20 Hz, 20 min of data collection yields Mn $K\beta_{1,3}$ spectra with sufficient signal-to-noise ratio. For the IPNS crystals described above, data collection of the Fe Kα spectra using four cylindrically bent (R = 250 mm) Ge(440) crystals yielded high quality spectra in about 7 minutes collection time at 30 Hz repetition rate. In most cases, however, as the XRD data collection requires a significantly higher number of shots or data collection time, data collection time for XES is usually not an issue.

Processing of the XES signals and examples of analysis of resulting spectra are outlined in Fig. 11. To process the XES signal collected on a position sensitive detector, pixels containing emission signal are defined as a region of interest (ROI) (Fig. 11A). Images from each shot are screened for “missed” shots by filtering by the relative signal intensity in the ROI. Areas surrounding the ROI are then used to extrapolate a background signal. The final XES spectrum is obtained by integrating over the spatial width of the signal in

the ROI. To calculate a difference spectrum of different timepoints in reaction kinetics, the spectra at each timepoint are first normalized by the area under the curve. Several examples of these area-normalized spectra along with associated difference spectra are shown in Fig. 11C and 11E. In Fig. 11C, a clear shift can be observed in Mn $K\beta_{1,3}$ signal of PSII at two different reaction states (dark states and after excitation by two saturating flashes) in these area-normalized spectra. This shift is visualized in a difference spectrum, also included in Fig. 11C. Similarly, in Fig. 11E and 11F, a shift in Fe Kα signal of IPNS at different reaction states are observable. Further details about the XES data processing can also be found in (Fransson et al., 2018).

To quantitatively analyze the changes in XES at different reaction states beyond consideration of difference spectra, spectral parameters such as first moment and full width at half maximum can be calculated. By plotting these spectral parameters, change the oxidation state and/or ligand environment of the active site metal(s) can be tracked across different delay times following the reaction triggering. Examples of these kinetic trends are shown in Fig. 11D and 11G. As described in the next section, these trends determined via XES are to be analyzed concurrently with the information obtained from XRD to assist the interpretation of the structural changes in the protein dynamics.

5. XRD Data Collection

X-ray diffraction data are collected with a large position sensitive detector downstream of the interaction region. Detectors that have been used with the DOT setup include the Rayonix MX 340 (at LCSL/MFX), the Rayonix MX 300 (at the SPring-8 Angstrom Compact Free Electron Laser (SACLA) BL2/EH3), the Jungfrau 16 M (at SwissFEL/Bernina) or the ePix10K (at LCLS/MFX) detector. The choice of detector is dependent on the availability at the facility but also the needs of the experiment. For large unit cell samples, the spacing between individual Bragg spots for the long cell axis can become the limiting factor, making a detector with a large number of pixels necessary in order to cover diffraction out to the required resolution. Usually, a minimum separation of 4-5 pixels between two neighboring Bragg spots is required to properly integrate the intensities for these spots. The minimum detector distance achievable with the DOT setup depends on the detector type but for most cases 125 mm minimum distance can be achieved. This translates for example to a resolution of 1.45 Å for the inscribed circle when using an X-ray energy of 9.5 keV (1.3 Å) with the Rayonix MX 340 detector or to a resolution of 1.4 Å when using an X-ray energy of 10.5 keV (1.18 Å) and the Rayonix MX300.

As the precise distance between the detector and the X-ray interaction point with the sample is not known a priori this needs to be calibrated. For a first order estimation a diffraction pattern from a silver behenate powder is recorded. For this purpose, a thin pellet of silver(I) behenate (AgBeh, Alfa Aesar) is pressed using a pellet press. A small wedge of this pellet is then glued to the reference arm of the DOT setup. By changing the horizontal position perpendicular to the beam direction this reference arm is brought into the beam IP. Using a transmission of ~0.1-1% and rastering the AgBeh constantly through the beam by 10 μm translation steps every 50-100 shots during the measurement, a powder diffraction pattern (Huang, Toraya, Blanton and Wu, 1993) can be recorded in about 1000 X-ray shots (16

s at 30 Hz). The known spacing of the diffraction rings for AgBeh (e.g. $d_{0\ 0\ 8}$=4.194 Å, $d_{0\ 0\ 10}$=3.35 Å, $d_{0\ 0\ 12}$=2.796 Å) is then used to calibrate the sample detector distance. For this, a tool in the cctbx.xfel data processing GUI can be used (Brewster, Young, et al., 2019).

While the data collection can be directly monitored visually via the online and side view cameras of the DOT system, allowing to fine tune droplet positioning and timing manually to optimize droplet hit rate, these do not give a complete picture of the success of the experiment. Using the *cctbx.xfel* GUI it is possible to track the progress of the experiment by plotting several statistics, including droplet hit rate, crystal indexing rate and diffraction quality (Fig. 12). Following the output on the GUI enables scientists to detect problems in the sample delivery. For example, Fig. 12A shows unstable droplet ejection as evidenced by the highly variable droplet hit rate (green trace in the middle row). In addition, many of the observed diffraction patterns cannot be properly indexed (gray dots in the top row), leading to an indexing rate of only a few % (blue trace in the middle row). On the positive side it is visible that nearly all indexed patterns show diffraction to better than 2.5 Å (yellow line in the bottom row) and often high-quality diffraction to 1.6 Å or better is observed (yellow dots in the bottom row). In contrast, the example in Fig. 12B shows a very stable droplet hit rate of nearly 100% and an indexing rate of 25-30% of all shots. But the diffraction quality of the indexed patterns is weaker with only very few crystals diffracting to better than 2.5 Å. In the first case, one should consider changing the droplet ejection parameters, maybe the buffer conditions, and the crystal concentration to achieve more stable operation. In the second case, these conditions are all well-tuned but the diffraction quality of the crystals is suboptimal and consequently, it would be worth changing to a different protein batch in such a situation.

The direct processing of the diffraction and spectroscopic data also provides fast feedback on the unit cell and the catalytic activity of the measured sample. Fig. 13A shows the observed unit cell distribution for two different batches of photosystem II crystals, generated with the *cctbx.xfel* GUI (Brewster, Young, et al., 2019). The small difference in the distribution for the c-axis length indicates problems with the post crystallization treatment for one batch compared to the other. Consequently, this feedback can be used to optimize the treatment procedure on the fly to achieve a very uniform distribution over many different protein batches, which is necessary to obtain highly isomorphous data. Fig. 13B,C shows the results of the XES processing during the experiment for one PSII protein batch. Even after only recording for a few minutes for each condition, it is possible to see the difference in the signal between the dark state (“0F”) and the double illuminated state (“2F”). The observed shift between the spectra for the two illumination conditions is indicative for the sample turning over properly and serves as confirmation that the illumination conditions are set correctly and that the sample is intact (see also (Fransson et al., 2018) for details).

While collecting data it is important to keep track of the number of indexed images for one condition (e.g., one illumination state) as a first indicator for the data set quality. To obtain a more complete picture of the dataset quality and estimate how much more data collection is necessary, repeated merging of the indexed data including subsequent experimental runs is very helpful. The *xfel.merge* program of the *cctbx* package (Brewster et al., 2018; Sauter, 2015) provides an informative output log with information regarding

the percentage of indexed images that diffract to a certain resolution bin as well as merging statistics, including the multiplicity and the half data set correlation coefficient ($CC_{1/2}$) for each resolution bin. Example outputs are shown in Fig. 14. From our experience, the number of accepted lattices needs to be at least 2000-3000 in the highest resolution bin to result in acceptable statistics for the resulting data set up to that resolution. Monitoring the acceptance rate at the desired resolution bin together with the indexing rate thus allows one to predict how much more time one needs to spend on collecting at a certain condition (see Fig 14 B,C).

Two other criteria used by us to decide about the usable resolution cut off for the collected data are the $CC_{1/2}$ value and the multiplicity of observations (Fig. 15). The $CC_{1/2}$ value should show a monotonous decrease from low to high resolution and should not exhibit jumps or sudden changes from resolution bin to resolution bin. While a low $CC_{1/2}$ could be indicative of weak data, sudden changes in the $CC_{1/2}$ values can also be caused by outlier images (e.g. a strong background peak not properly masked out) and excluding such outlier images can help improve the $CC_{1/2}$ statistics. When examining for example the statistics for the IPNS dataset shown in Fig. 15A, the $CC_{1/2}$ decreases continuously up to the highest resolution bin, indicating that the data is self-consistent up to 1.5 Å resolution. For the PSII data in Fig. 15B, the $CC_{1/2}$ number shows a monotonous decrease up to bin 16 and then starts to fluctuate, hence bin 16 should be used as the cut off for the merging based on $CC_{1/2}$. When considering the multiplicity (how many times each HKL value was observed in all indexed images included in the merging) as computed by the *xfel.merge* program a value of ~10 was used as an (somewhat arbitrary) cut off in our past work. Following this the IPNS data shown in Fig. 15A should be usable up to 1.53 Å resolution while the PSII data (Fig. 15B) can be used up to 2.14 Å resolution. It should be noted that multiplicity can be computed in different ways and the number obtained for the same data can be very different, depending on what reflections are included or excluded for the computation of the multiplicity (see also discussion in the SI material for (Ibrahim et al., 2020)). In addition, one should monitor the $I/\sigma(I)$ value to ensure that there is still a measurable signal in the last bin of the final merged data set. Again, it is difficult to compare these numbers directly to numbers from synchrotron rotation crystallography as the error estimates are not as well established for SFX data and hence the $\sigma(I)$ computation is highly dependent on the chosen error model. For a detailed discussion about the effect of the error model on SFX data processing see (Brewster, Bhowmick, et al., 2019; Hattne et al., 2014). Due to this uncertainty, we generally accept lower $I/\sigma(I)$ values (e.g., down to 0.3 for the highest resolution bin) compared to the standard $I/\sigma(I)$ cut off of ~1 conventionally used for rotation crystallography work. It should be noted that also for synchrotron work, recent developments indicate that including all data, even when having very weak $I/\sigma(I)$ values is beneficial for the quality of the electron density maps and one should utilize an iterative refinement process to decide up to what resolution bin should be included in the final data (Hattne et al., 2014; Karplus & Diederichs, 2015).

Depending on the unit cell dimensions, diffraction quality of the crystals, and the fraction that yield lattices to fill the desired high-resolution shell, data collection can take anywhere from 10 minutes to 6 hours for a complete diffraction data set. For example, using 30 Hz repetition rate (dictated by 4 x 4 detector binning, 177 μm pixels) we were able to collect

a 1.55 Å data set from IPNS microcrystals in about 15 minutes. The sample consumption rate was 6 uL/min, so we needed a total amount of 90 uL crystal suspension for this dataset. For soluble methane monooxygenase a collection time of 45 minutes at 20 Hz repetition rate (dictated by 3 x 3 detector binning, 133 μm pixels) resulted in a dataset of 1.9 Å resolution using a total of 360 μL of crystal suspension (8 μL/min consumption rate). For the weaker diffracting crystals of PSII a collection time of 150 minutes at 20 Hz repetition rate gave data out to 2.15 Å resolution, using a flow rate of 5 μL/min and a total sample volume of 750 μL.

Further processing of the merged diffraction data can be done using standard crystallography software like phenix (Liebschner et al., 2019) or the ccp4 suite (Agirre et al., 2023) and visualization and manual model building can be performed in PyMOL (Schrödinger LLC, 2015) and Coot (Emsley et al., 2010).

When interpreting the XRD data it is very helpful to include the information obtained from the XES part of the experiment. In the case of IPNS the observed changes in the Fe Kα emission signal measured from the crystals after different O_2 incubation times was in good agreement with observations of changes around the Fe atom in the active site of IPNS at these time points (Fig. 16A) (Rabe et al., 2021). For PSII we were able to correlate the kinetics of the change in the Mn oxidation state in the S_2-S_3 transition observed via the Mn $K\beta_{1,3}$ XES with the appearance of electron density for an additional oxygen atom in the cluster (Ibrahim et al., 2020) (Fig. 16B), allowing us to conclude that the oxidation and oxygen insertion process happen concomitantly.

6. Summary and Conclusions

Conducting combined diffraction and emission measurements on metalloenzymes utilizing the fs pulses from XFEL sources allows to obtain information at room temperature that is not affected by the usual problems of radiation damage one encounters with synchrotron measurements. Hence this approach opens the door to conduct time-resolved studies to probe details of the reaction mechanisms of many radiation-sensitive metalloenzymes. We described protocols to generate suitable crystalline samples for these kinds of experiments using seeding and batch methods. Conducting these experiments with the drop-on-tape setup described here is very efficient in terms of sample consumption compared to commonly used liquid jet approaches. It enables scientists to collect not only diffraction but also X-ray emission data concomitantly from the same sample at the same time. This allows the experimenter to confirm if the sample is in the expected starting state with respect to oxidation state and gives feedback about changes in the electronic structure e.g. due to reaction triggering under the experimental conditions. The described approach for XES data collection can be adapted to different transition metals and also to measuring signals from two different transition metals at the same time ((Fuller et al., 2017; Gul et al., 2015), e.g. for metalloenzymes that have a heteronuclear active site. The different reaction triggering options discussed (laser illumination and in-situ gas incubation) target various enzyme systems and can be modified to match the needs of the system under study (for an expansion using chemical activation via in-situ mixing see for example the drop-on-drop mixing approach described in (Butryn et al., 2021). Fast feedback during the experiment for both the

emission and diffraction data can be used to optimize the use of limited experimental time and the quality of the data. Our approach has been applied to a number of metalloenzymes, including not only IPNS and PSII but also soluble methane monooxygenase (Srinivas et al., 2020), ribonucleotide reductase R2b (John et al., 2022) and methyl coenzyme M reductase (Ohmer et al., 2022). We hope that other users can apply this approach to their systems and help to broaden the nascent field of room temperature time resolved structural studies on metalloenzymes.

Acknowledgements

We thank Isabel Bogacz, Philipp Simon and Patrick Rabe for discussion of different aspects of this manuscript and providing critical feedback. This work was supported by the Director, Office of Science, Office of Basic Energy Sciences (OBES), Division of Chemical Sciences, Geosciences, and Biosciences (CSGB) of the Department of Energy (DOE) (J.Y., J.K.) for X-ray spectroscopy and crystallography data collection and analysis, and methods development for photosynthetic systems, by the National Institutes of Health (NIH) Grants GM110501 (J.Y.) and GM126289 (J.K.) for instrumentation development for XFEL experiments. This research used resources of NERSC, a User Facility supported by the Office of Science, DOE, under Contract No. DE-AC02-05CH11231. XFEL data was collected at LCLS/SLAC, Stanford, funded by DOE OBES under Contract No. DE-AC02-76SF00515, and structural biology work at the LCLS is supported by NIH grant P41GM139687 and the Rayonix detector at the MFX instrument at LCLS was funded by a NIH grant S10 OD023453.

References

Agirre J, Atanasova M, Bagdonas H, Ballard CB, Baslé A, Beilsten-Edmands J, Borges RJ, Brown DG, Burgos-Mármol JJ, Berrisford JM, Bond PS, Caballero I, Catapano L, Chojnowski G, Cook AG, Cowtan KD, Croll TI, Debreczeni JÉ, Devenish NE, … Yamashita K (2023). The CCP4 suite: Integrative software for macromolecular crystallography. Acta Crystallographica Section D: Structural Biology, 79(6), Article 6. 10.1107/S2059798323003595

Alonso-Mori R, Kern J, Sokaras D, Weng TC, Nordlund D, Tran R, Montanez P, Delor J, Yachandra VK, Yano J, & Bergmann U (2012). A multi-crystal wavelength dispersive x-ray spectrometer. Rev Sci Instrum, 83(7), 073114. 10.1063/1.4737630 [PubMed: 22852678]

Barends TRM, Stauch B, Cherezov V, & Schlichting I (2022). Serial femtosecond crystallography. Nature Reviews Methods Primers, 2(1), Article 1. 10.1038/s43586-022-00141-7

Beale JH et al. Successful sample preparation for serial crystallography experiments. J Appl Crystallogr 52, 1385–1396 (2019). 10.1107/S1600576719013517 [PubMed: 31798361]

Bergmann U, Kern J, Schoenlein RW, Wernet P, Yachandra VK, & Yano J (2021). Using X-ray free-electron lasers for spectroscopy of molecular catalysts and metalloenzymes. Nature Reviews Physics. 10.1038/s42254-021-00289-3

Bijelic A, & Rompel A (2018). Polyoxometalates: More than a phasing tool in protein crystallography. ChemTexts, 4(3), 10. 10.1007/s40828-018-0064-1 [PubMed: 30596006]

Brewster AS, Bhowmick A, Bolotovsky R, Mendez D, Zwart PH, & Sauter NK (2019). SAD phasing of XFEL data depends critically on the error model. Acta Crystallogr D Struct Biol, 75(Pt 11), 959–968. 10.1107/S2059798319012877 [PubMed: 31692470]

Brewster AS, Waterman DG, Parkhurst JM, Gildea RJ, Young ID, O'Riordan LJ, Yano J, Winter G, Evans G, & Sauter NK (2018). Improving signal strength in serial crystallography with DIALS geometry refinement. Acta Cryst. D, 74(Pt 9), 877–894. 10.1107/S2059798318009191

Brewster AS, Young ID, Lyubimov A, Bhowmick A, & Sauter NK (2019). Processing serial crystallpgrahpic data from XFELs or synchrotrons using the cctbx.xfel GUI. Comp. Cryst. Newsletter, 10, 22–39.

Butryn A, Simon PS, Aller P, Hinchliffe P, Massad RN, Leen G, Tooke CL, Bogacz I, Kim IS, Bhowmick A, Brewster AS, Devenish NE, Brem J, Kamps J, Lang PA, Rabe P, Axford D, Beale JH, Davy B, … Orville AM (2021). An on-demand, drop-on-drop method for studying enzyme catalysis by serial crystallography. Nat Commun, 12(1), 4461. 10.1038/s41467-021-24757-7 [PubMed: 34294694]

Chapman HN, Fromme P, Barty A, White TA, Kirian RA, Aquila A, Hunter MS, Schulz J, DePonte DP, Weierstall U, Doak RB, Maia FRNC, Martin AV, Schlichting I, Lomb L, Coppola N, Shoeman RL, Epp SW, Hartmann R, … Spence JCH (2011). Femtosecond X-ray protein nanocrystallography. Nature, 470(7332), 73–77. 10.1038/Nature09750 [PubMed: 21293373]

Cheng RK (2020). Towards an Optimal Sample Delivery Method for Serial Crystallography at XFEL. Crystals, 10(3)215. 10.3390/cryst10030215

Derewenda ZS, & Vekilov PG (2006). Entropy and surface engineering in protein crystallization. Acta Crystallographica Section D: Biological Crystallography, 62(1), Article 1. 10.1107/S0907444905035237 [PubMed: 16369088]

Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010). Features and development of Coot. Acta Cryst. D, 66(Pt 4), 486–501. 10.1107/S0907444910007493 [PubMed: 20383002]

Erdemir D, Lee AY, & Myerson AS (2019). Crystal Nucleation. In Lee AY, Myerson AS, & Erdemir D (Eds.), Handbook of Industrial Crystallization (3rd ed., pp. 76–114). Cambridge University Press. 10.1017/9781139026949.003

Fransson T, Chatterjee R, Fuller FD, Gul S, Weninger C, Sokaras D, Kroll T, Alonso-Mori R, Bergmann U, Kern J, Yachandra VK, & Yano J (2018). X-ray emission spectroscopy as an in situ diagnostic tool for X-ray crystallography of metalloproteins using an X-ray free-electron laser. Biochemistry, 57(31), 4629–4637. 10.1021/acs.biochem.8b00325 [PubMed: 29906115]

Fuller FD, Gul S, Chatterjee R, Burgie ES, Young ID, Lebrette H, Srinivas V, Brewster AS, Michels-Clark T, Clinger JA, Andi B, Ibrahim M, Pastor E, de Lichtenberg C, Hussein R, Pollock CJ, Zhang M, Stan CA, Kroll T, … Yano J (2017). Drop-on-demand sample delivery for studying biocatalysts in action at X-ray free-electron lasers. Nature Methods, 14(4), 443–449. 10.1038/nmeth.4195 [PubMed: 28250468]

García-Ruiz JM (2003). Counterdiffusion Methods for Macromolecular Crystallization. Methods in Enzymology, 130–154. doi:10.1016/s0076-6879(03)68008-0 [PubMed: 14674272]

Giegé R. (2013) A historical perspective on protein crystallization from 1840 to the present day. The FEBS journal. 10.1111/febs.12580

Gul S, Ng JW, Alonso-Mori R, Kern J, Sokaras D, Anzenberg E, Lassalle-Kaiser B, Gorlin Y, Weng TC, Zwart PH, Zhang JZ, Bergmann U, Yachandra VK, Jaramillo TF, & Yano J (2015). Simultaneous detection of electronic structure changes from two elements of a bifunctional catalyst using wavelength-dispersive X-ray emission spectroscopy and in situ electrochemistry. Phys. Chem. Chem. Phys, 17(14), 8901–8912. 10.1039/c5cp01023c [PubMed: 25747045]

Hadimioglu B, Stearns R, Ellson R (2016) Moving Liquids with Sound: The Physics of Acoustic Droplet Ejection for Robust Laboratory Automation in Life Sciences. J. Lab. Autom, 21(1), 4–18. 10.1177/2211068215615096 [PubMed: 26538573]

Hattne J, Echols N, Tran R, Kern J, Gildea RJ, Brewster AS, Alonso-Mori R, Glockner C, Hellmich J, Laksmono H, Sierra RG, Lassalle-Kaiser B, Lampe A, Han G, Gul S, DiFiore D, Milathianaki D, Fry AR, Miahnahri A, … Sauter NK (2014). Accurate macromolecular structures using minimal measurements from X-ray free-electron lasers. Nat Methods, 11(5), 545–548. 10.1038/nmeth.2887 [PubMed: 24633409]

Hellmich J, Bommer M, Burkhardt A, Ibrahim M, Kern J, Meents A, Muh F, Dobbek H, & Zouni A (2014). Native-like photosystem II superstructure at 2.44 Å resolution through detergent extraction from the protein crystal. Structure, 22(11), 1607–1615. 10.1016/j.str.2014.09.007 [PubMed: 25438669]

Huang TC, Toraya H, Blanton TN, & Wu Y (1993). X-ray Powder Diffraction Analysis of Silver Behenate, a Possible Low-Angle Diffraction Standard. J. Appl. Cryst 26, 180–184. 10.1107/S0021889892009762

Ibrahim M, Chatterjee R, Hellmich J, Tran R, Bommer M, Yachandra VK, Yano J, Kern J, & Zouni A (2015). Improvements in serial femtosecond crystallography of photosystem II by optimizing crystal uniformity using microseeding procedures. Struct. Dyn, 2(4). 10.1063/1.4919741

Ibrahim M, Fransson T, Chatterjee R, Cheah MH, Hussein R, Lassalle L, Sutherlin KD, Young ID, Fuller FD, Gul S, Kim IS, Simon PS, de Lichtenberg C, Chernev P, Bogacz I, Pham CC, Orville AM, Saichek N, Northen T, … Yano J (2020). Untangling the sequence of events during the $S_2 \rightarrow S_3$ transition in photosystem II and implications for the water oxidation mechanism. Proc. Natl. Acad. Sci. U. S. A, 117(23), 12624–12635. 10.1073/pnas.2000529117 [PubMed: 32434915]

John J, Aurelius O, Srinivas V, Saura P, Kim I-S, Bhowmick A, Simon PS, Dasgupta M, Pham C, Gul S, Sutherlin KD, Aller P, Butryn A, Orville AM, Cheah MH, Owada S, Tono K, Fuller FD, Batyuk A, … Högbom M (2022). Redox-controlled reorganization and flavin strain within the ribonucleotide reductase R2b–NrdI complex monitored by serial femtosecond crystallography. ELife, 11, e79226. 10.7554/eLife.79226 [PubMed: 36083619]

Kamps J. (2020). Biophysical studies on enzymes related to antibiotics [Http://purl.org/dc/dcmitype/Text, University of Oxford]. https://ora.ox.ac.uk/objects/uuid:c6b3089c-ff7c-408d-a8a3-c5e7a673746e

Karplus PA, & Diederichs K (2015). Assessing and maximizing data quality in macromolecular crystallography. Curr Opin Struct Biol, 34, 60–68. 10.1016/j.sbi.2015.07.003 [PubMed: 26209821]

Keedy DA, Kenner LR, Warkentin M, Woldeyes RA, Hopkins JB, Thompson MC, Brewster AS, Van Benschoten AH, Baxter EL, Uervirojnangkoorn M, McPhillips SE, Song J, Alonso-Mori R, Holton JM, Weis WI, Brunger AT, Soltis SM, Lemke H, Gonzalez A, … Fraser JS (2015). Mapping the conformational landscape of a dynamic enzyme by multitemperature and XFEL crystallography. Elife, 4. 10.7554/eLife.07574

Kern J, Alonso-Mori R, Tran R, Hattne J, Gildea RJ, Echols N, Glockner C, Hellmich J, Laksmono H, Sierra RG, Lassalle-Kaiser B, Koroidov S, Lampe A, Han G, Gul S, Difiore D, Milathianaki D, Fry AR, Miahnahri A, … Yano J (2013). Simultaneous femtosecond X-ray spectroscopy and diffraction of photosystem II at room temperature. Science, 340(6131), 491–495. 10.1126/science.1234273 [PubMed: 23413188]

Kern J, Loll B, Luneberg C, DiFiore D, Biesiadka J, Irrgang KD, & Zouni A (2005). Purification, characterisation and crystallisation of photosystem II from Thermosynechococcus elongatus cultivated in a new type of photobioreactor. Biochim Biophys Acta, 1706(1–2), 147–157. 10.1016/j.bbabio.2004.10.007 [PubMed: 15620375]

Kuhlbrandt W. (2014). Biochemistry. The resolution revolution. Science, 343(6178), 1443–1444. 10.1126/science.1251652 [PubMed: 24675944]

Liebschner D, Afonine PV, Baker ML, Bunkoczi G, Chen VB, Croll TI, Hintze B, Hung LW, Jain S, McCoy AJ, Moriarty NW, Oeffner RD, Poon BK, Prisant MG, Read RJ, Richardson JS, Richardson DC, Sammito MD, Sobolev OV, … Adams PD (2019). Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. Acta Crystallogr. D, 75(Pt 10), 861–877. 10.1107/S2059798319011471

Neutze R. (2014). Opportunities and challenges for time-resolved studies of protein structural dynamics at X-ray free-electron lasers. Philos Trans R Soc Lond B Biol Sci, 369(1647), 20130318. 10.1098/rstb.2013.0318 [PubMed: 24914150]

Ohmer CJ, Dasgupta M, Patwardhan A, Bogacz I, Kaminsky C, Doyle MD, Chen PY, Keable SM, Makita H, Simon PS, Massad R, Fransson T, Chatterjee R, Bhowmick A, Paley DW, Moriarty NW, Brewster AS, Gee LB, Alonso-Mori R, … Ragsdale SW (2022). XFEL serial crystallography reveals the room temperature structure of methyl-coenzyme M reductase. J Inorg Biochem, 230, 111768. 10.1016/j.jinorgbio.2022.111768 [PubMed: 35202981]

Rabe P, Beale JH, Butryn A, Aller P, Dirr A, Lang PA, Axford DN, Carr SB, Leissing TM, McDonough MA, Davy B, Ebrahim A, Orlans J, Storm SLS, Orville AM, Schofield CJ, & Owen RL (2020). Anaerobic fixed-target serial crystallography. IUCrJ, 7(5), 901–912. 10.1107/s2052252520010374

Rabe P, Kamps J, Sutherlin KD, Linyard JDS, Aller P, Pham CC, Makita H, Clifton I, McDonough MA, Leissing TM, Shutin D, Lang PA, Butryn A, Brem J, Gul S, Fuller FD, Kim IS, Cheah MH, Fransson T, … Schofield CJ (2021). X-ray free-electron laser studies reveal correlated motion during isopenicillin N synthase catalysis. Sci Adv, 7(34), eabh0250. 10.1126/sciadv.abh0250 [PubMed: 34417180]

Rhodes G (2006) Crystallography made crystal clear: a guide for users of macromolecular models. Elsevier, Burlington

Roach PL, Clifton IJ, Hensgens CMH, Shibata N, Long AJ, Strange RW, Hasnain SS, Schofield CJ, Baldwin JE, & Hajdu J (1996). Anaerobic Crystallisation of an Isopenicillin N Synthase · Fe(II) · Substrate Complex Demonstrated by X-Ray Studies. European Journal of Biochemistry, 242(3), 736–740. 10.1111/j.1432-1033.1996.0736r.x [PubMed: 9022704]

Roessler CG, Kuczewski A, Stearns R, Ellson R, Olechno J, Orville AM, Allaire M, Soares AS, Heroux A (2013). Acoustic methods for high-throughput protein crystal mounting at next-generation macromolecular crystallographic beamlines. J. Synchrotron Rad 20, 805–808. 10.1107/S0909049513020372

Rupp B. (2013). Macromolecular Crystallography: Overview. In Roberts GCK (Ed.), Encyclopedia of Biophysics (pp. 1346–1353). Springer. 10.1007/978-3-642-16712-6_655

Saibil HR (2022). Cryo-EM in molecular and cellular biology. Molecular Cell, 82(2), 274–284. 10.1016/j.molcel.2021.12.016 [PubMed: 35063096]

Sauter NK (2015). XFEL diffraction: Developing processing methods to optimize data quality. J Synchrotron Radiat, 22(2), 239–248. 10.1107/S1600577514028203 [PubMed: 25723925]

Schrödinger LLC. (2015). The PyMOL molecular graphics system, version 2.0.

Schulz EC, Yorke BA, Pearson AR & Mehrabi P Best practices for time-resolved serial synchrotron crystallography. Acta Crystallographica Section D 78, 14–29 (2022). 10.1107/S2059798321011621

Shoeman RL, Hartmann E & Schlichting I Growing and making nano- and microcrystals. Nature Protocols 18, 854–882 (2023). 10.1038/s41596-022-00777-5 [PubMed: 36451055]

Soares AS, Engel MA, Stearns R, Datwani S, Olechno J, Ellson R, Skinner JM, Allaire M, Orville AM (2011). Acoustically Mounted Microcrystals Yield High-Resolution X-ray Structures. Biochemistry, 50(21), 4399–4401. 10.1021/bi200549x [PubMed: 21542590]

Spence JCH (2018). X-ray lasers for structure and dynamics in biology. IUCrJ, 5(Pt 3), 236–237. 10.1107/S2052252518005365

Srinivas V, Banerjee R, Lebrette H, Jones JC, Aurelius O, Kim IS, Pham CC, Gul S, Sutherlin KD, Bhowmick A, John J, Bozkurt E, Fransson T, Aller P, Butryn A, Bogacz I, Simon P, Keable S, Britz A, … Hogbom M (2020). High-Resolution XFEL Structure of the Soluble Methane Monooxygenase Hydroxylase Complex with its Regulatory Component at Ambient Temperature in Two Oxidation States. J Am Chem Soc, 142(33), 14249–14266. 10.1021/jacs.0c05613 [PubMed: 32683863]

Thompson MC, Barad BA, Wolff AM, Sun Cho H, Schotte F, Schwarz DMC, Anfinrud P, & Fraser JS (2019). Temperature-jump solution X-ray scattering reveals distinct motions in a dynamic enzyme. Nature Chemistry, 11(11), Article 11. 10.1038/s41557-019-0329-3

Vekilov PG (2019). Crystallization of Proteins. In Lee AY, Myerson AS, & Erdemir D (Eds.), Handbook of Industrial Crystallization (3rd ed., pp. 414–459). Cambridge University Press. 10.1017/9781139026949.014

Vekilov PG (2004). Nanoscale Structure and Assembly at Solid–Fluid Interfaces, Vol. II, Assembly in Hybrid and Biological Systems, edited by Lui XY, pp. 145–200. Dordrecht: Kluewer Academic Publishers.

Vekilov PG (2003). Solvent Entropy Effects in the Formation of Protein Solid Phases. Methods in Enzymology, 84–105. doi:10.1016/s0076-6879(03)68006-7

Vekilov PG & Chernov AA, Solid State Physics, edited by Spaepen F (2002), pp. 1–147. New York: Academic Press.

Yano J, Fuller F, Gul S, Chatterjee R, Kern J, & Yachandra V (2017, February 28). Droplet On Tape: Protocol. 10.1038/protex.2017.017

Yano J, Kern J, Irrgang K-D, Latimer MJ, Bergmann U, Glatzel P, Pushkar Y, Biesiadka J, Loll B, Sauer K, Messinger J, Zouni A, & Yachandra VK (2005). X-ray Damage to the Mn4Ca Complex in Photosystem II Crystals: A Case Study for Metallo-Protein X-ray Crystallography. Proc. Natl. Acad. Sci. USA, 102, 12047–12052. [PubMed: 16103362]

Young ID, Ibrahim M, Chatterjee R, Gul S, Fuller F, Koroidov S, Brewster AS, Tran R, Alonso-Mori R, Kroll T, Michels-Clark T, Laksmono H, Sierra RG, Stan CA, Hussein R, Zhang M, Douthit L, Kubin M, de Lichtenberg C, Long Vo P, Nilsson H, Cheah MH, Shevela D, Saracini C, Bean MA, Seuffert I, Sokaras D, Weng TC, Pastor E, Weninger C, Fransson T, Lassalle L, Bräuer P, Aller P, Docker PT, Andi B, Orville AM, Glownia JM, Nelson S, Sikorski M, Zhu D, Hunter MS, Lane TJ, Aquila A, Koglin JE, Robinson J, Liang M, Boutet S, Lyubimov AY, Uervirojnangkoorn M, Moriarty NW, Liebschner D, Afonine PV, Waterman DG, Evans G, Wernet P, Dobbek H, Weis WI, Brunger AT, Zwart PH, Adams PD, Zouni A, Messinger J, Bergmann

U, Sauter NK, Kern J, Yachandra VK, Yano J. Structure of photosystem II and substrate binding at room temperature. Nature. 2016 Dec 15;540(7633):453–457. doi: 10.1038/nature20161. Epub 2016 Nov 21. [PubMed: 27871088]

Zouni A, Kern J, Frank J, Hellweg T, Behlke J, Saenger W, & Irrgang K-D (2005). Size Determination of Cyanobacterial and Higher Plant Photosystem II by Gel Permeation Chromatography, Light Scattering, and Ultracentrifugation. Biochemistry 2005, 44, 4572–4581. [PubMed: 15766288]

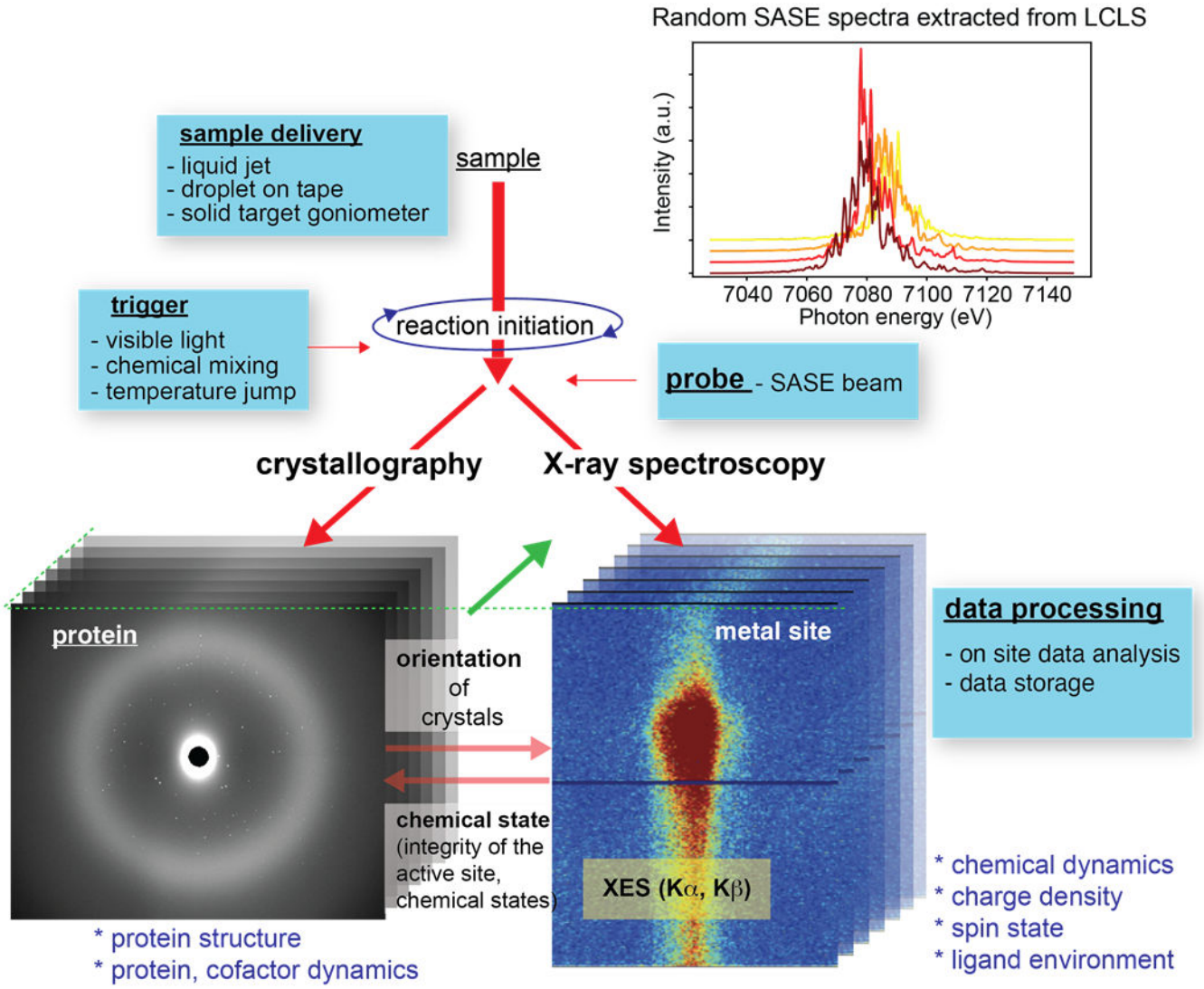

Figure 1:

Overview of the shot-by-shot data collection of X-ray crystallography and X-ray spectroscopy of biological systems at XFELs. Incoming self-amplified spontaneous emission (SASE) pulses at XFELs have a bandwidth of ~0.2% of the central energy.

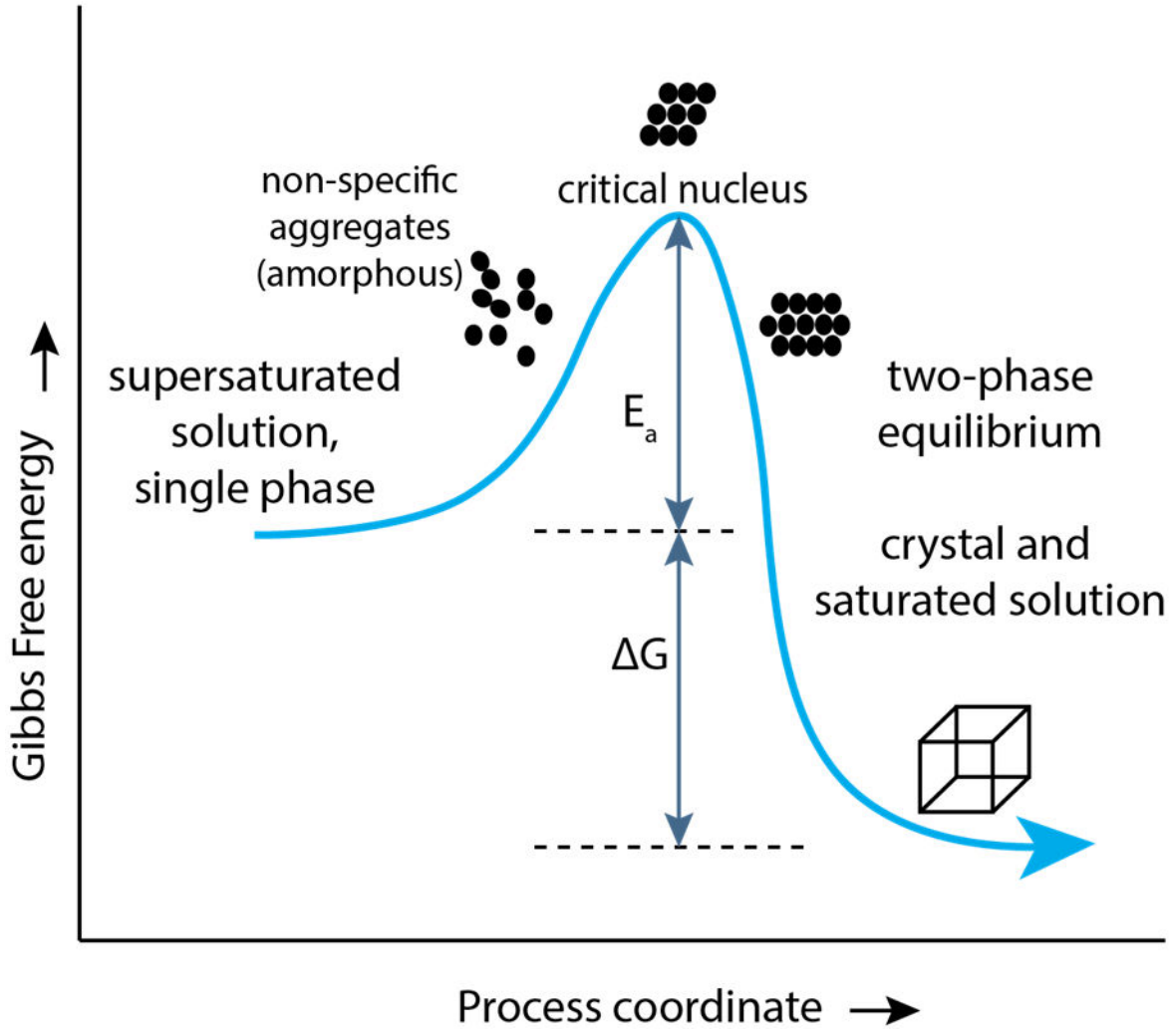

Fig 2.

Schematic representation of the activation energy barrier E_a and the driving force for crystallization **Δ**G during the crystallization process. The peak position represents a kinetic energy barrier that needs to be overcome to form a critical nucleus. Exceeding the critical nucleus size allows formation of a single crystal, inducing the desired phase separation. Small black circles represent protein molecules, and the square indicates a single crystal.

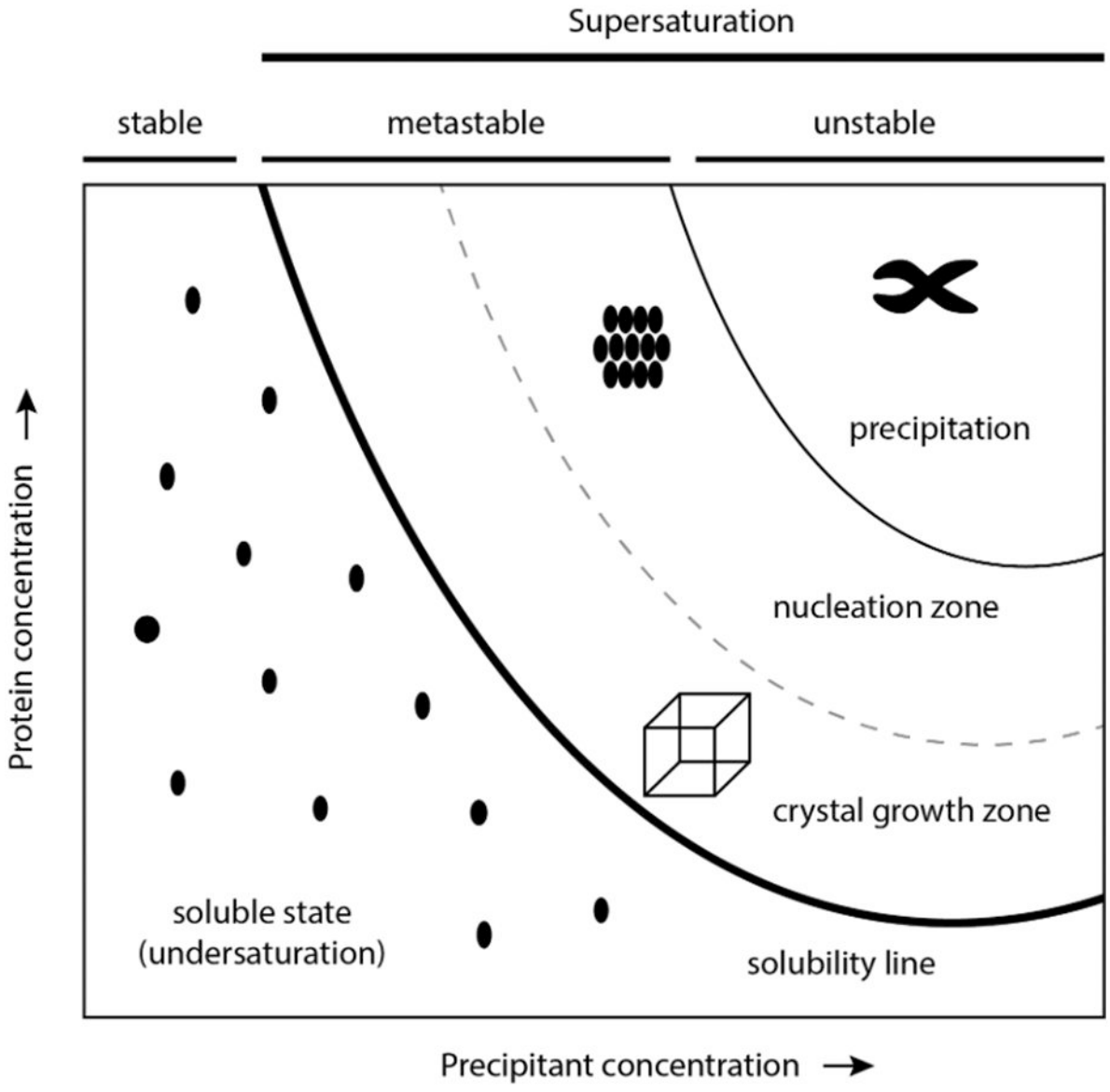

Fig. 3.

Crystallization phase diagram, which is divided by a solubility line into two regions: an undersaturated and supersaturated zone. In the soluble state, neither crystal formation nor crystal growth occurs. Protein molecules represented as black circles stay stable. When reaching supersaturation, three events can happen: precipitation, nucleation and crystal growth. Under high supersaturated conditions, only amorphous solids precipitate (precipitation zone). With decreasing saturation, spontaneous crystal formation takes place (nucleation zone). Further moving towards even lower supersaturation conditions, crystal growth or metastable zone is reached. Black circles are demonstrated as protein molecules. Single crystal is represented as a square.

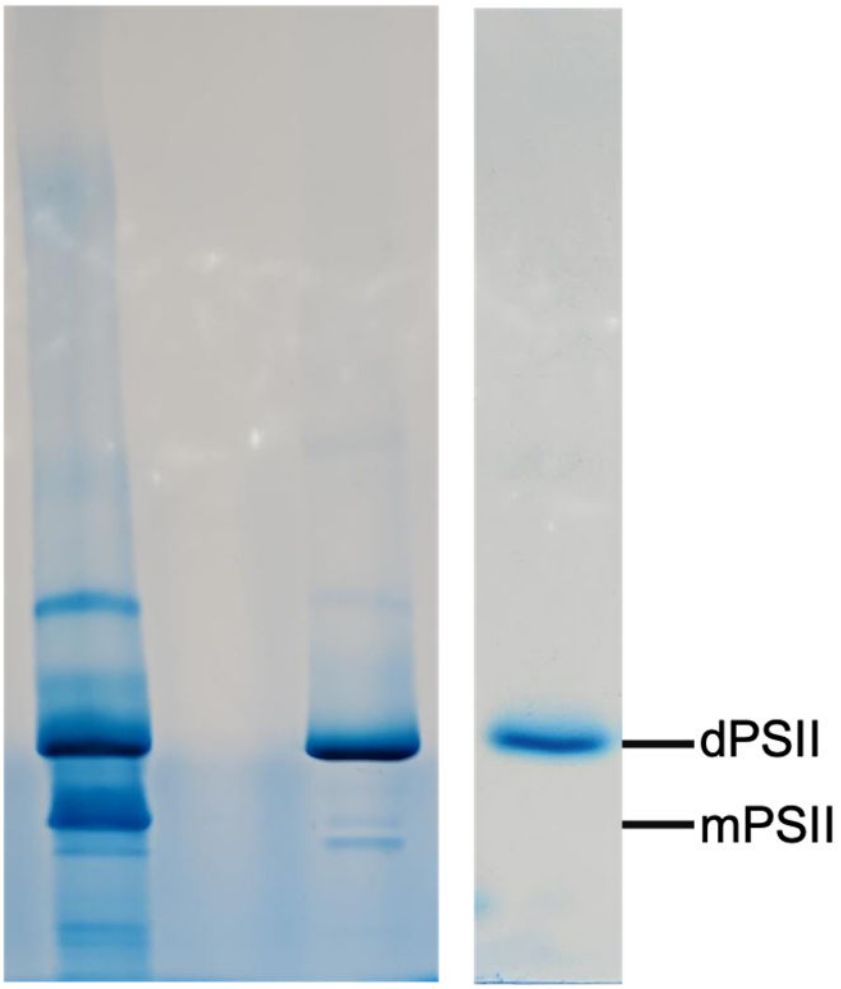

Fig. 4.

Blue native PAGE analysis of crude PSII particles (lane 1), purified PSII dimers (lane 2) and PSII dimers after recrystallization (lane 3).

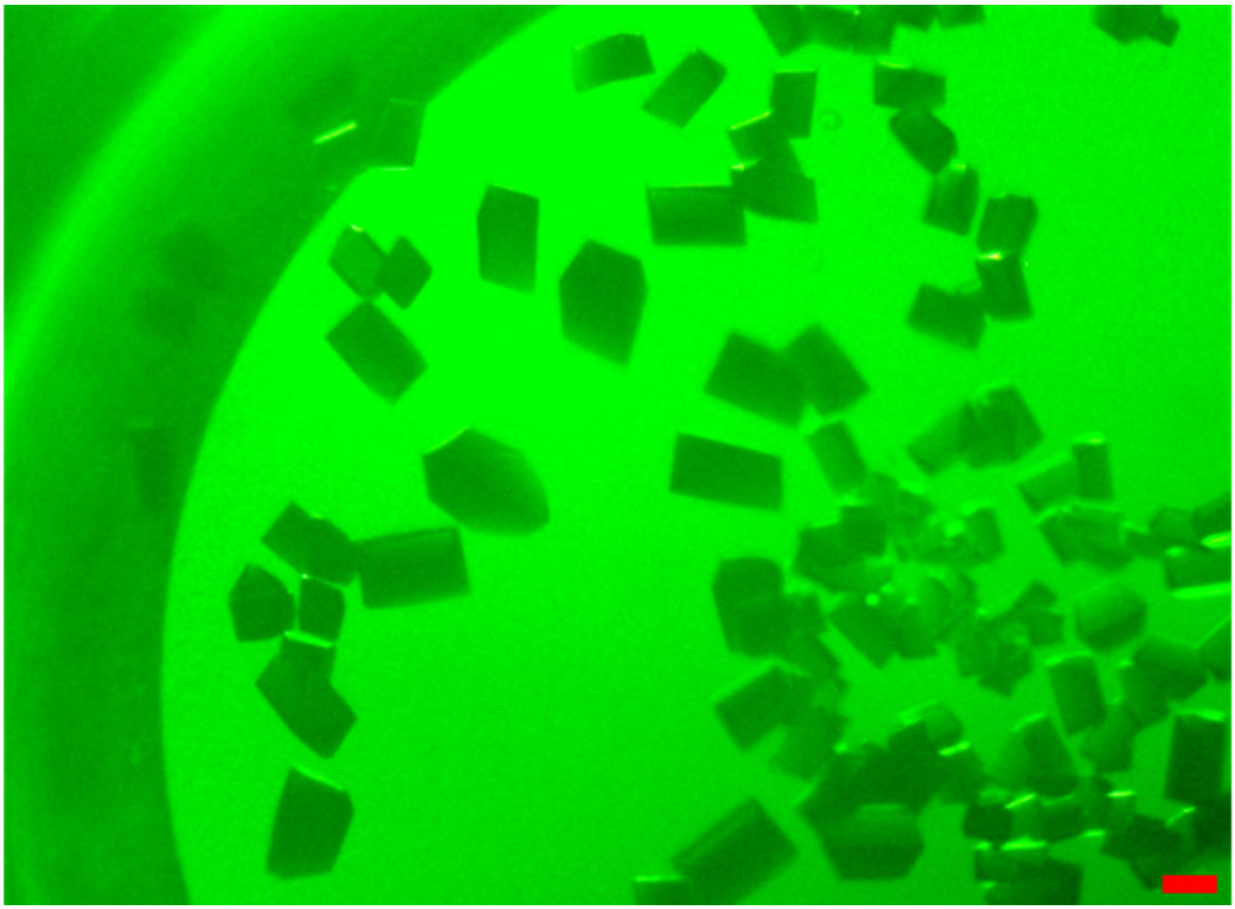

Fig 5.
Microbatch crystals were obtained from the purified PSII dimers. The scale bar represents 100 μm.

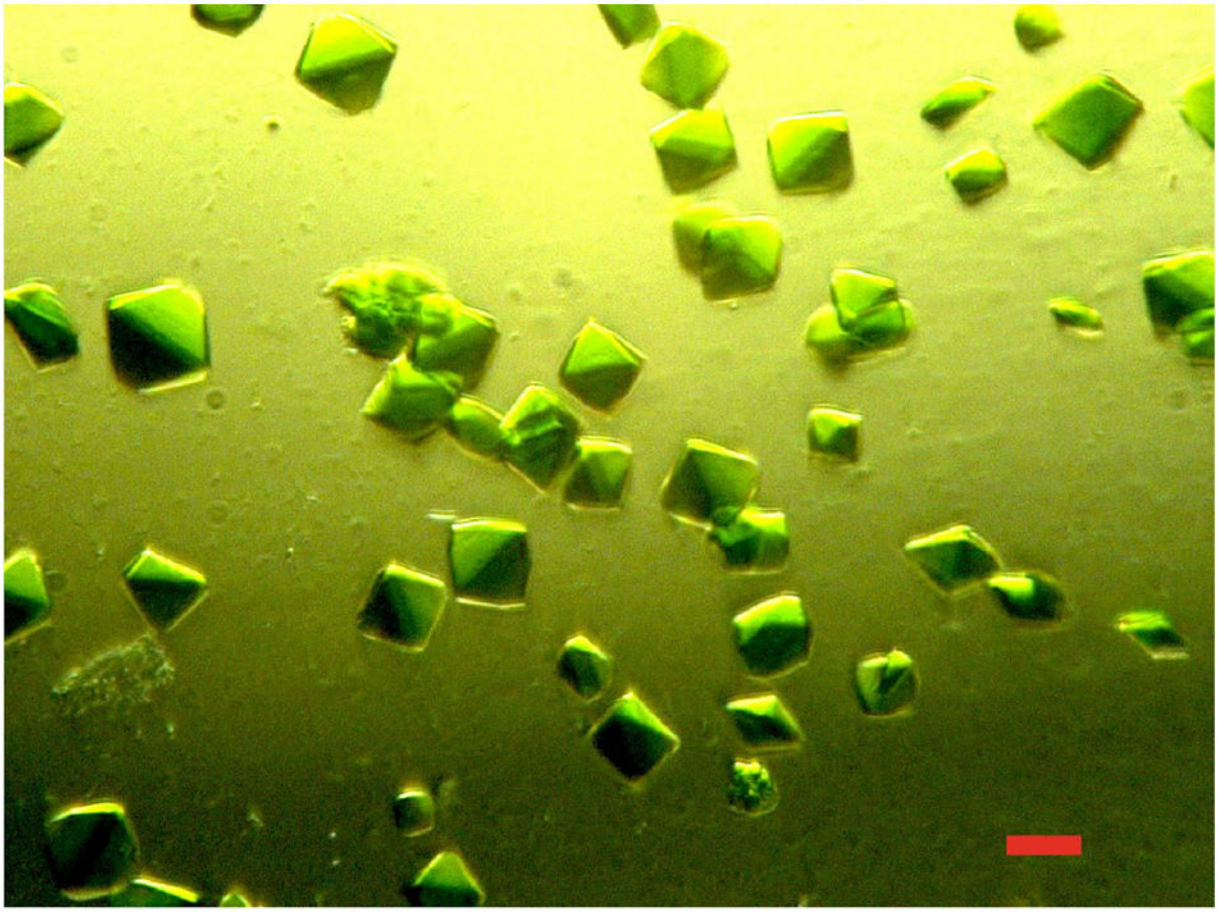

Fig 6.
Microseeding crystals of PSII dimers. The scale bar represents 20 μm.

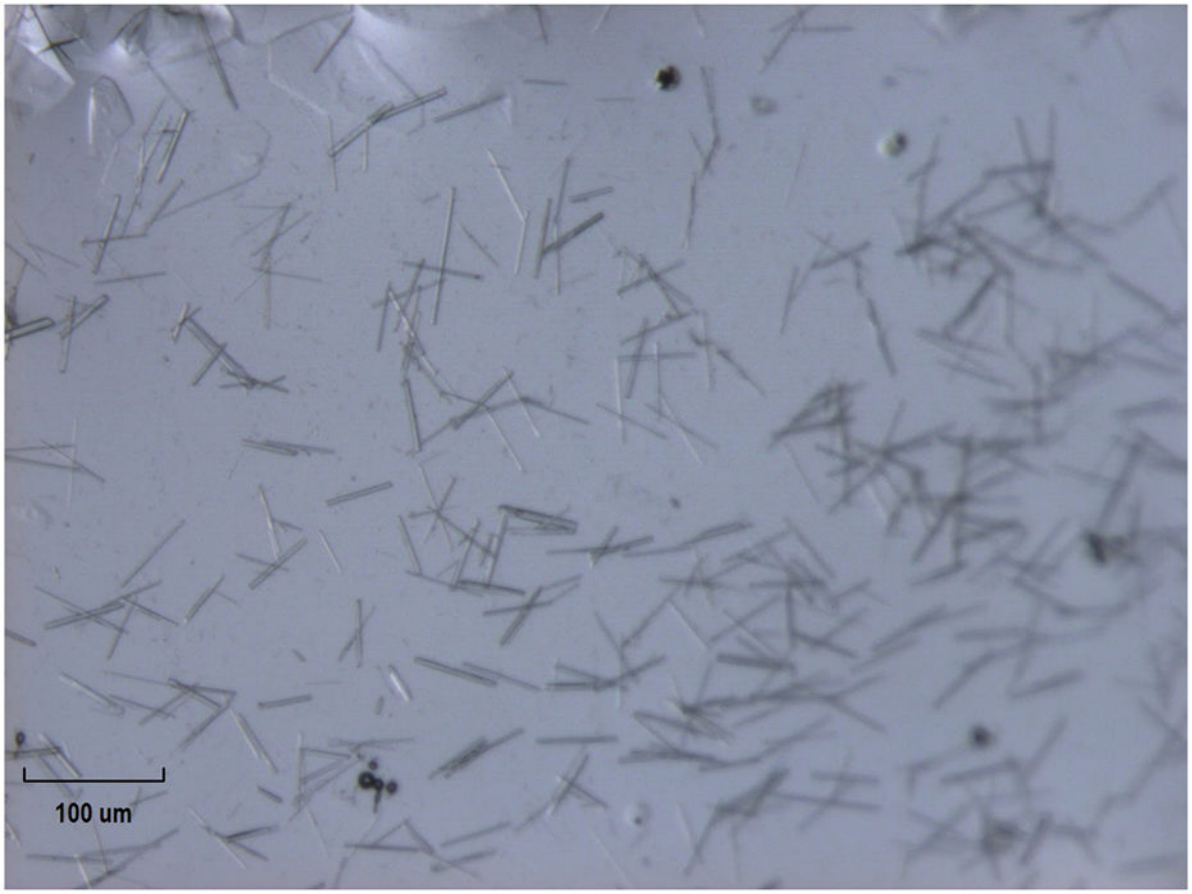

Fig. 7.
Optical microscope image of IPNS • Fe • ACV microcrystals.

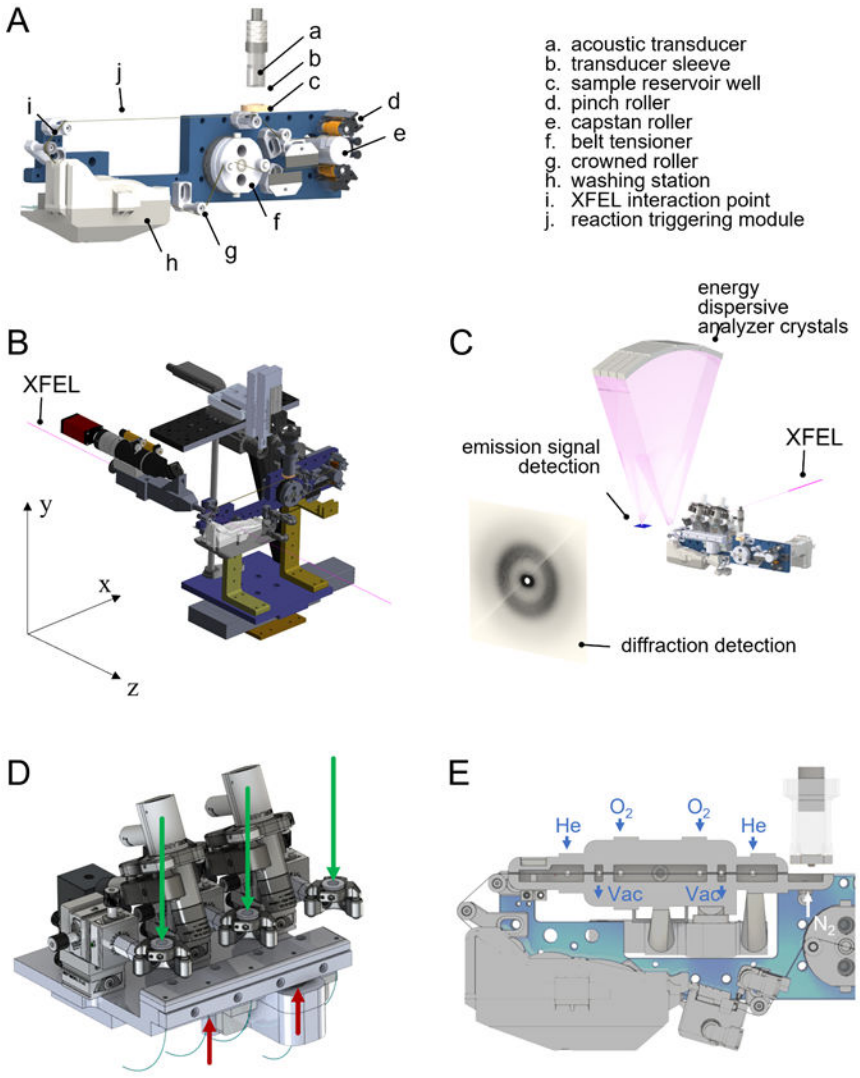

Fig. 8:

Overview of the DOT sample delivery system and deployment configurations. **A**. Core components of the DOT system. Different reaction triggering modules are inserted at position *j*. **B**. A typical mounting scheme of the DOT system with a directional reference nomenclature used in this chapter. The XFEL beam is parallel to the z-axis. The core components shown in **A** are mounted onto a translational stage for movement in the x-direction using vertical brackets. The ADE setup is mounted on multi-axis stages to enable positional adjustments in y and z. A beam enclosure to provide a camera view to the interaction point in-line with the XFEL beam is also shown. **C**. Relative positioning of the DOT setup with respect to the X-ray emission spectrometer and X-ray diffraction detector. Distance between the interaction point on the DOT system and the XRD detector is not drawn to scale. **D**. Photoactivation reaction module. A droplet travels from right to left in this diagram. A total of three photoactivation is possible using a fiber-guided laser aligned to the 1st, 3rd, and 5th apertures. A droplet position on the reaction module is tracked using a near infrared light probe aligned at the 2nd and 4th apertures. **E**. A schematic of

the O_2-activation module. A droplet travels from right to left in this diagram. The central chamber is saturated with O_2, while other chambers are used to create a differential pump.

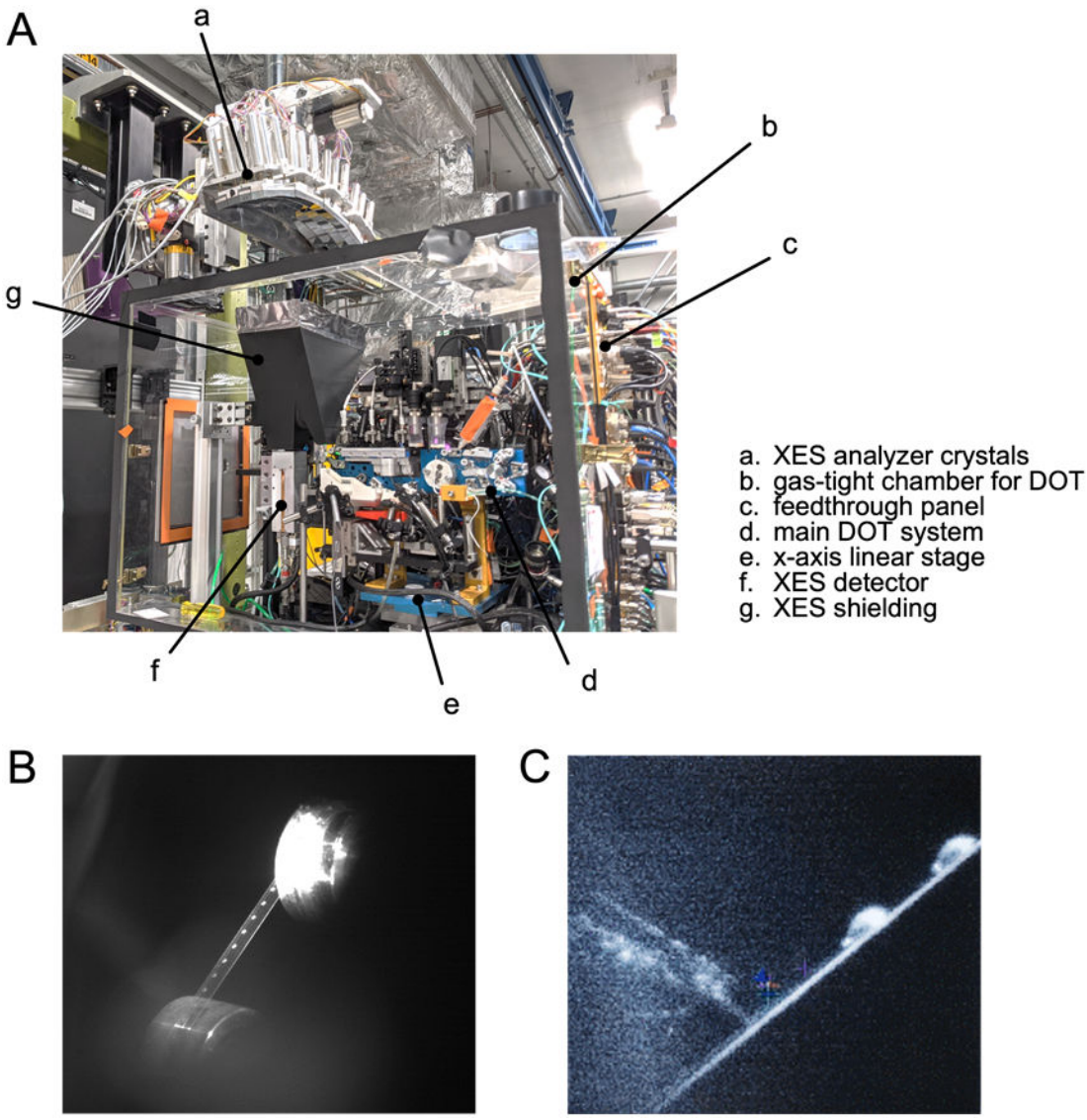

Fig. 9.

A. The DOT system deployed at the XFEL experimental hutch. In this configuration, the DOT system is enclosed in a gas-tight chamber which has a cutout for the XES spectrometer (a) and XRD detector (not shown). The cutouts are sealed by a thin inflatable plastic sheet for the XES spectrometer side, and by a window fitted by an X-ray transparent material for the XRD detector side. B. A train of droplets ejected at a set frequency on a Kapton conveyor belt. The image is taken by a GigE camera triggered at the same frequency as the droplet ejection and XFEL beam, and shows a view onto the X-ray interaction point on the DOT. In this view, the tape is transporting droplets from top to bottom. C. A droplet shot by a pulse of XFEL. The image is taken by a GigE camera with zoom lens attachment aligned to provide a collinear view as the incoming XFEL beam. In this view, the droplet is delivered from the top right corner of the image to the X-ray interaction point near the center of the image.

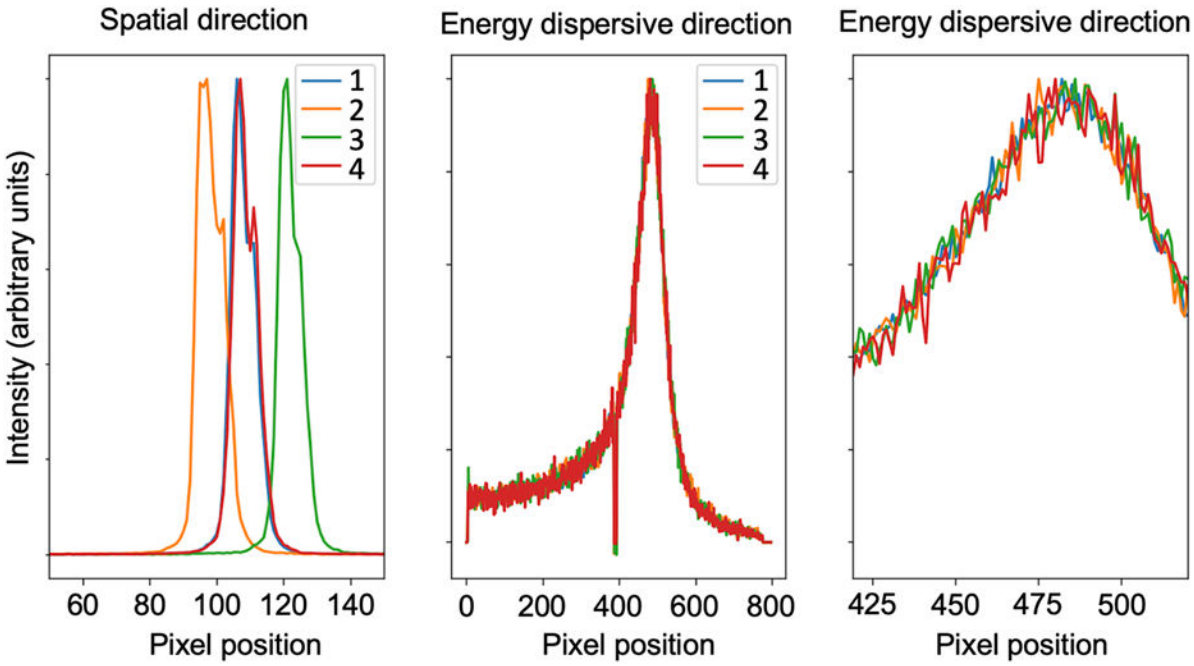

Fig 10:

Alignment of XES spectrometer, reproducibility over shifts. The Mn $K\beta_{1,3}$ spectra collected from a Mn foil for four consecutive days (1-4) of an experiment are plotted. While small changes in the position of the signal on the detector are visible (left panel, “Spatial direction”) they perfectly align in the energy dispersive direction (middle and right panel). Thereby it was confirmed that the sample spectra from different days of the experiment can be directly compared with each other.

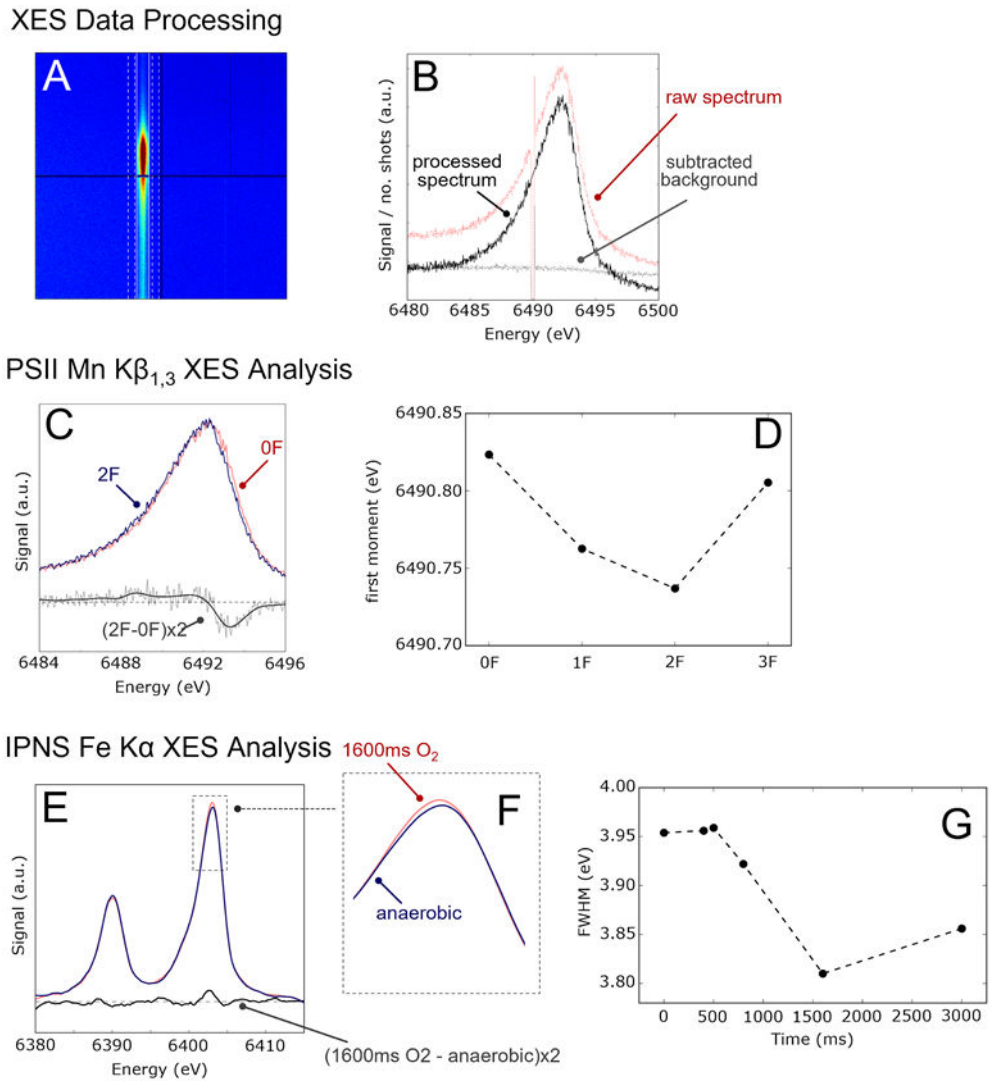

Fig. 11.

Outline of the XES data processing, from the signals on the position sensitive detector to quantitative analysis of the spectra. A. Signals collected on the position sensitive detector. The pixels containing the signals are cut out as ROI (defined by two solid lines). To correct for any drift in the background signal, two additional ROIs are defined outside the signal ROI (defined by four dashed lines). The intensities within the two background ROIs are extrapolated across the signal ROI to estimate the drift. B. A raw spectrum calculated from the signal ROI (dotted line), extrapolated background signal (dashed line), and a processed spectrum where the background signal is subtracted from the raw spectrum. C. Mn $K\beta_{1,3}$ signal of PSII, in the dark state ("0F") and after applying two actinic flashes ("2F"). A scaled DS shows a shift in the spectrum. D. First moment of the Mn Kβ1,3 spectrum calculated for dark and different flash states of PSII. A shift in the first moment as the reaction cycle is advanced indicates a change in the Mn oxidation state through the reaction. E. Fe Kα spectrum of IPNS, in the anaerobic state and after incubation with O_2 for 1600 ms. A scaled

difference spectrum is also included. F. A closer view of the shift in Fe Kα_1 peak. G. Full-width half-maximum of Fe Kα_1 peak of IPNS across different O_2 incubation time.

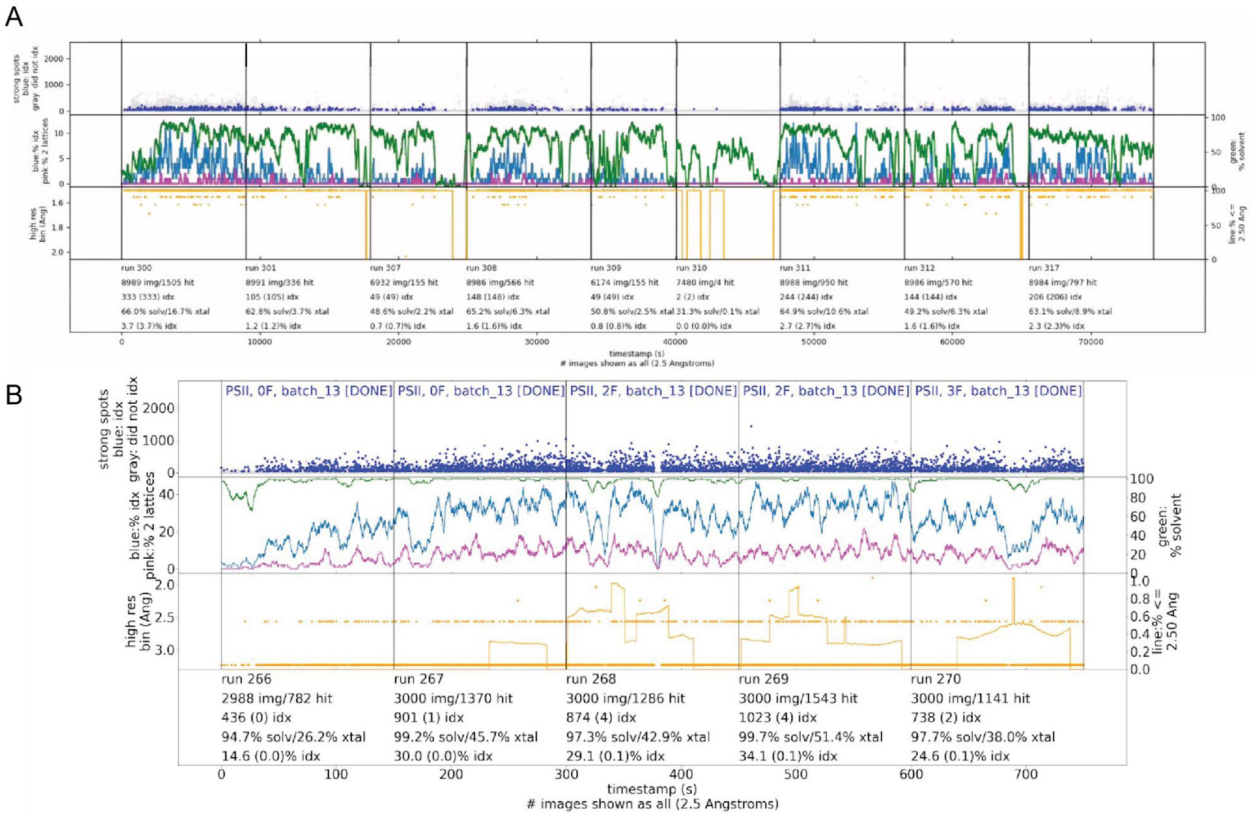

Figure 12:

Examples for direct feedback for diffraction experiments using the cctbx.xfel GUI. (A) Runs of a sample that exhibits several problems in the injection behavior and (B) sample that exhibits stable injection condition. The middle row shows the droplet hit rate (green line) and the indexing rate (blue line) as well as the percentage of double hits (purple). The top row shows the number of diffraction spots identified on each indexed image (blue dots) and on each non-indexed image (gray dots). The bottom row shows the maximum resolution quality detected (yellow dots) and the percentage of diffraction patterns that are better than 2.5 Å (yellow line). Below the graphs details about indexed images and hit rates are given for each experimental run.

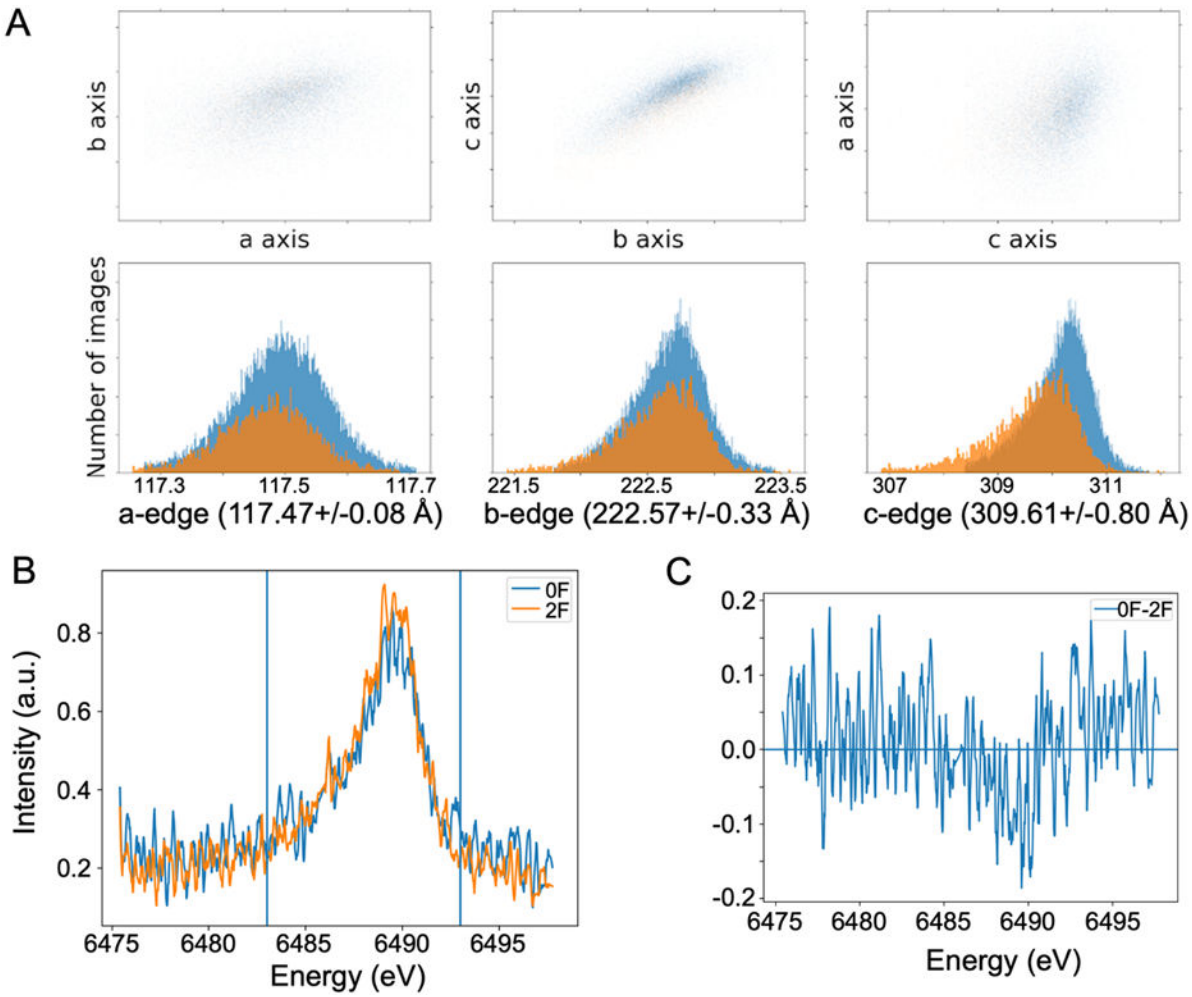

Fig. 13:

Additional feedback during the XES/SFX experiment. (A) Unit cell distribution plot for two different batches of PSII crystals (orange and blue) generated by the *cctbx.xfel* GUI. A difference especially for the c axis length is visible between the two batches. (B) Mn Kβ XES signal recorded from PSII single crystals in the dark (0F, blue trace) and doubly illuminated (2F, orange trace) state and (C) difference spectrum between the 0F and 2F XES data both showing the expected difference due to light induced oxidation of the Mn cluster in PSII.

A

Image Statistics

Bin	Resolution Range	Lattices	Accepted (%)
1	-1.0000 - 4.0716	16102	99.99
2	4.0716 - 3.2317	16076	99.83
3	3.2317 - 2.8231	16050	99.66
4	2.8231 - 2.5650	16013	99.43
5	2.5650 - 2.3811	15870	98.55
6	2.3811 - 2.2407	15552	96.57
7	2.2407 - 2.1285	15439	95.87
8	2.1285 - 2.0358	15021	93.27
9	2.0358 - 1.9574	14285	88.70
10	1.9574 - 1.8899	13776	85.54
11	1.8899 - 1.8308	12069	74.94
12	1.8308 - 1.7784	11530	71.60
13	1.7784 - 1.7316	10187	63.26
14	1.7316 - 1.6894	8925	55.42
15	1.6894 - 1.6510	8097	50.28
16	1.6510 - 1.6158	6612	41.06
17	1.6158 - 1.5835	5940	36.89
18	1.5835 - 1.5536	4917	30.53
19	1.5536 - 1.5259	3939	24.46
20	1.5259 - 1.5000	3274	20.33
All		16104	

B

Image Statistics

Bin	Resolution Range	Lattices	Accepted (%)
1	-1.0000 - 5.1574	1541	100.00
2	5.1574 - 4.0934	1540	99.94
3	4.0934 - 3.5759	1539	99.87
4	3.5759 - 3.2490	1506	97.73
5	3.2490 - 3.0161	1421	92.21
6	3.0161 - 2.8382	1200	77.87
7	2.8382 - 2.6961	912	59.18
8	2.6961 - 2.5787	785	50.94
9	2.5787 - 2.4794	675	43.80
10	2.4794 - 2.3938	512	33.23
11	2.3938 - 2.3190	448	29.07
12	2.3190 - 2.2527	387	25.11
13	2.2527 - 2.1934	282	18.30
14	2.1934 - 2.1399	231	14.99
15	2.1399 - 2.0912	193	12.52
16	2.0912 - 2.0467	124	8.05
17	2.0467 - 2.0058	94	6.10
18	2.0058 - 1.9679	73	4.74
19	1.9679 - 1.9328	39	2.53
20	1.9328 - 1.9000	21	1.36
All		1541	

C

Image Statistics

Bin	Resolution Range	Lattices	Accepted (%)
1	-1.0000 - 5.1574	19306	100.00
2	5.1574 - 4.0934	19303	99.98
3	4.0934 - 3.5759	19288	99.91
4	3.5759 - 3.2490	19082	98.84
5	3.2490 - 3.0161	18386	95.23
6	3.0161 - 2.8382	16506	85.50
7	2.8382 - 2.6961	13484	69.84
8	2.6961 - 2.5787	11770	60.97
9	2.5787 - 2.4794	10334	53.53
10	2.4794 - 2.3938	8097	41.94
11	2.3938 - 2.3190	7083	36.69
12	2.3190 - 2.2527	6230	32.27
13	2.2527 - 2.1934	4800	24.86
14	2.1934 - 2.1399	4149	21.49
15	2.1399 - 2.0912	3440	17.82
16	2.0912 - 2.0467	2356	12.20
17	2.0467 - 2.0058	1823	9.44
18	2.0058 - 1.9679	1395	7.23
19	1.9679 - 1.9328	768	3.98
20	1.9328 - 1.9000	524	2.71
All		19306	

Fig. 14:

Data statistics generated by *cctbx.xfel*. Acceptance statistics per resolution bin for data collected from IPNS (**A**) and PSII (**B**,**C**). Based on a rule of thumb of ~3000 indexed images needed in the highest resolution bin one would estimate a useable resolution of 1.53 Å for the IPNS data set, no useable dataset after 10 minutes of PSII data collection and about 2.1 Å after 75 minutes of data collection for PSII. It can be seen that the acceptance statistics obtained even after a short collection time (10 minutes, **B**) can be used as a good estimate for the data quality and the time needed to obtain a complete data set to the desired resolution (~2.1 Å, 75 minutes, **C**).

A

Bin	Resolution Range	Completeness	%	<asu multi>	<obs multi>	n_meas	asu_m_meas	Merged <I>	Merged <I/sig(I)>	CC int	N int
1	-1.0000 - 4.0716	[2796/2817]	99.25	293.07	295.27	825568	2796	177884	14.867	99.0%	2796
2	4.0716 - 3.2317	[2691/2691]	100.00	188.38	188.38	506931	2691	139082	11.818	98.6%	2691
3	3.2317 - 2.8231	[2644/2644]	100.00	156.82	156.82	414633	2644	72584	8.936	98.3%	2644
4	2.8231 - 2.5650	[2633/2633]	100.00	136.93	136.93	360537	2633	37549	7.503	98.2%	2633
5	2.5650 - 2.3811	[2622/2622]	100.00	122.19	122.19	320388	2622	25976	6.700	97.3%	2622
6	2.3811 - 2.2407	[2622/2622]	100.00	110.44	110.44	289577	2622	22324	5.933	97.3%	2622
7	2.2407 - 2.1285	[2615/2615]	100.00	102.64	102.64	268410	2615	17660	5.255	97.0%	2615
8	2.1285 - 2.0358	[2581/2581]	100.00	91.11	91.11	235158	2581	12893	4.441	96.9%	2581
9	2.0358 - 1.9574	[2589/2589]	100.00	82.29	82.29	213044	2589	9209	3.651	96.1%	2589
10	1.9574 - 1.8899	[2562/2562]	100.00	73.56	73.56	188458	2562	6789	3.030	95.6%	2562
11	1.8899 - 1.8308	[2612/2612]	100.00	61.24	61.24	159953	2612	5033	2.435	93.9%	2612
12	1.8308 - 1.7784	[2582/2582]	100.00	54.55	54.55	140837	2582	3889	2.035	91.8%	2582
13	1.7784 - 1.7316	[2575/2575]	100.00	44.71	44.71	115140	2575	3216	1.685	91.0%	2575
14	1.7316 - 1.6894	[2567/2567]	100.00	38.12	38.12	97856	2567	2559	1.396	83.1%	2565
15	1.6894 - 1.6510	[2591/2591]	100.00	31.91	31.91	82673	2591	2408	1.267	80.6%	2590
16	1.6510 - 1.6158	[2572/2572]	100.00	25.38	25.38	65288	2572	2017	1.034	72.3%	2563
17	1.6158 - 1.5835	[2570/2570]	100.00	21.17	21.17	54401	2570	1809	0.874	66.6%	2557
18	1.5835 - 1.5536	[2590/2591]	99.96	16.24	16.24	42071	2590	1655	0.755	50.8%	2560
19	1.5536 - 1.5259	[2553/2555]	99.92	12.67	12.68	32374	2553	1513	0.620	45.2%	2442
20	1.5259 - 1.5000	[2551/2573]	99.14	8.63	8.70	22198	2551	1442	0.472	28.4%	2121
All		[52118/52164]	99.91	85.03	85.10	4435495	52118	28314	4.313	99.2%	51522

B

Bin	Resolution Range	Completeness	%	<asu multi>	<obs multi>	n_meas	asu_m_meas	Merged <I>	Merged <I/sig(I)>	CC int	N int
1	-1.0000 - 5.1574	[32708/32857]	99.55	271.92	273.16	8934459	32708	3253670	17.166	99.7%	32708
2	5.1574 - 4.0934	[32085/32085]	100.00	171.29	171.29	5495728	32085	2319555	12.000	99.1%	32085
3	4.0934 - 3.5759	[31869/31869]	100.00	141.58	141.58	4512033	31869	1163158	7.649	98.1%	31869
4	3.5759 - 3.2490	[31788/31788]	100.00	119.15	119.15	3787558	31788	571541	4.739	95.4%	31788
5	3.2490 - 3.0161	[31697/31697]	100.00	95.02	95.02	3011712	31697	294067	2.927	89.3%	31697
6	3.0161 - 2.8382	[31708/31708]	100.00	71.24	71.24	2258846	31708	190364	2.071	81.4%	31708
7	2.8382 - 2.6961	[31623/31623]	100.00	52.76	52.76	1668387	31623	137412	1.551	69.4%	31623
8	2.6961 - 2.5787	[31607/31607]	100.00	41.88	41.88	1323582	31607	104786	1.191	53.0%	31604
9	2.5787 - 2.4794	[31563/31563]	100.00	32.87	32.87	1037539	31563	88101	0.962	42.9%	31554
10	2.4794 - 2.3938	[31507/31507]	100.00	25.42	25.42	800988	31507	77642	0.794	29.3%	31433
11	2.3938 - 2.3190	[31536/31537]	100.00	20.77	20.78	655163	31536	71181	0.683	19.2%	31352
12	2.3190 - 2.2527	[31500/31509]	99.97	16.44	16.44	517914	31500	66387	0.585	13.3%	30991
13	2.2527 - 2.1934	[31503/31538]	99.89	12.70	12.72	400629	31503	61106	0.487	9.8%	30063
14	2.1934 - 2.1399	[31376/31477]	99.68	9.99	10.02	314504	31376	55538	0.403	8.8%	27887
15	2.1399 - 2.0912	[30951/31506]	98.24	7.19	7.32	226566	30951	52786	0.326	6.2%	22683
16	2.0912 - 2.0467	[29148/31401]	92.83	4.76	5.12	149350	29148	49681	0.260	3.1%	14444
17	2.0467 - 2.0058	[25469/31441]	81.01	3.19	3.94	100313	25469	44705	0.207	3.4%	7913
18	2.0058 - 1.9679	[17901/31425]	56.96	1.75	3.08	55054	17901	43853	0.183	5.4%	2630
19	1.9679 - 1.9328	[9914/31442]	31.53	0.80	2.54	25224	9914	43433	0.167	3.1%	527
20	1.9328 - 1.9000	[4838/31410]	15.40	0.35	2.29	11085	4838	40558	0.142	-3.1%	86
All		[562291/632990]	88.83	55.75	62.76	35286634	562291	494331	3.091	99.0%	486645

Fig. 15:

Merging statistics for IPNS (A) and PSII data (B), showing the multiplicity (“asu multi”), $CC_{1/2}$ (“CC int”) and I/σ(I) (“Merged <I/sig((I)>”) values for each resolution bin. Indicators used to estimate the resolution cut-off for each data set are highlighted in green.

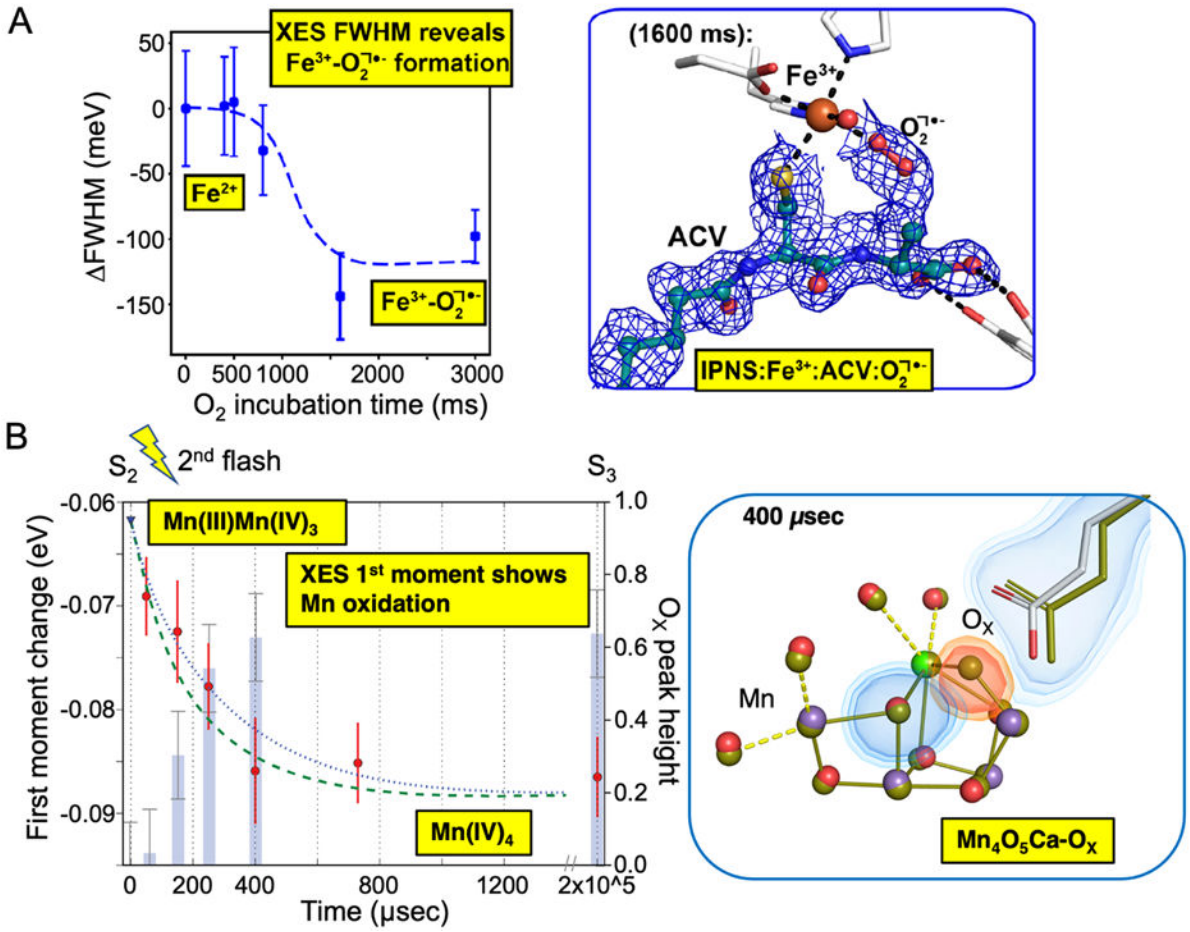

Fig. 16:

Changes in XES and XRD for different time points along the reaction process in IPNS (A) and PSII (B). (A) The full width half max (FWHM) of the Fe Kα signal from IPNS microcrystals is plotted as a function of O_2 incubation time (left) showing a clear change in Fe oxidation state between 800 and 1600 msec (left). Accordingly, a density feature is observed in the 1600 ms O_2 incubation XRD data (right), that is modeled as an O_2^- species bound to the Fe^{3+} at the active site of IPNS (adapted from (Rabe et al., 2021)). (B) The change in the first moment for the Mn Kβ XES is plotted as a function of time after the 2nd light flash given to PSII microcrystals (left). The change in the first moment is indicative of an oxidation of the Mn cluster from the $Mn(III)Mn(IV)_3$ state to the $Mn(IV)_4$ state. Concomitantly, an additional density is observed in the region of the Mn cluster starting at around 200 µs and the buildup of that density correlates well with the extent of Mn oxidation (right). This density was modeled as an additional bridging oxygen (O_X) in the Mn cluster (adapted from (Ibrahim et al., 2020)).