ISBN: 978-3-95450-207-3 doi:10.18429/JACoW-MEDSI2018-WEPH36

FROM PLATE SCREENING TO ARTIFICIAL INTELLIGENCE: INNOVATIVE DEVELOPMENTS ON PROXIMA 2A AT SYNCHROTRON SOLEIL

Damien Jeangerard[1], Martin Savko[1], Lidia Ciccone[1], Kewin Desjardins[1], Antoine Le Jollec[1], Ahmed Haouz[2], William Shepard[1†]
[1] Synchrotron SOLEIL, L'Orme des Merisiers, Saint Aubin BP 48, 91192 Gif-surYvette, France
[2] Pasteur Institute, 209-211 rue de Vaugirard, 75015 Paris, France

Abstract

PROXIMA 2A is a high performance 3rd generation synchrotron beamline dedicated to X-ray micro-crystallography on biological macromolecules. Since opening in March 2013, the experimental station has hosted a large number of users who have collected vast amounts of X-ray diffraction images from thousands of crystals. In order to streamline the throughput, enhance performance and add functionality, a number of developments have been launched on PROXIMA 2A. These cover all aspects of the beamline, from the practical to the visionary: such as the design, fabrication and implementation of a dedicated high-precision motorized stage to screen crystallization plates for *in situ* X-ray data collections, and the employment of artificial intelligence and computer vision technologies for the detection of samples under liquid nitrogen. Other notable beamline projects include the addition of a vertical translation table for the EIGER X 9M detector to permit the acquisition of ultra-high (0.6 Å) resolution X-ray data, the incorporation of a miniaturized Yttrium Aluminium Garnet (YAG) coupled photodiode within a beamstop and the determination of the sphere of confusion (SOC) of a recently added kappa arm to the goniometer.

INTRODUCTION

Modern macromolecular crystallography (MX) beamlines at synchrotron radiation centres have become highly automated systems that permit the high-throughput screening and X-ray data collection of a large number of protein crystals from academic and industrial users. However, as MX experiments have become increasingly sophisticated, there is a desire and need to record more information during the experiment (e.g. variations of the incident flux, sample shape, positions collected, X-ray dose, etc.). One of the direct consequences of maintaining higher throughput is that more complex situations need to be handled automatically without any human intervention. Given that every piece of instrumentation has its own technical limitations, there is a ceiling to the throughput efficiency obtainable. Similarly, every automated procedure is designed for a specific set of tasks, which tends to constrict the overall functionality of the experimental station. As such the automation of experiments sets its own boundaries in quality assurance, efficiency and functionality, and the mission of assuring reliability and broadening the frontiers is a continual process. Here we present some of the on-going projects on the PROXIMA 2A beamline at Synchrotron SOLEIL that improve upon the current automation available, extend the capabilities of the beamline, and add new functionalities for the user community.

† email address: william.shepard@synchrotron-soleil.fr

RESULTS

In Situ Crystallisation Plate Screener

Crystals of biological macromolecules (proteins, DNA, RNA, and their complexes) are typically grown in 96-well SBS format crystallisation plates via sitting-drop vapour diffusion experiments. When crystals appear in a drop (few 100 nL), they are physically very fragile, sensitive to changes in humidity and temperature, and consequently difficult to harvest. Furthermore, as such crystals contain 30-80% water, flash freezing them in liquid nitrogen can destroy their crystal quality. Although collecting X-ray diffraction data from crystals in crystallisation plates has been reported previously from other beamlines [1–4], the micro-focussed X-rays on PROXIMA 2A with a beam cross-section of 5 µm full width half maximum (FWHM) necessitates a high mechanical resolution over the entire dimensions of a crystallisation plate (128 mm × 86 mm). Unfortunately, such a resolution can not be delivered by the robotic arm of the sample changer system (SOC ≈50 µm), and the travel range of our version of the Micro-diffractometer (MD2, Arinax, France) is too small to cover an entire crystallisation plate. Thus, we have launched a project to custom design and build a crystallisation plate screener that is suitable for a micro-focussed X-ray beam and tailored to the environment of PROXIMA 2A.

The crystallisation plate screener (Figure 1) is composed of five motorised translation tables: Three heavy duty, long range, high-resolution (100 nm) stages (Axmo Precision, France) to move the plate horizontally (Tx), vertically (Tz) and along the X-ray beam path (Ts), which also brings the drop into the focal plane of the coaxial microscope of the MD2. Two other stages (SmarAct, Germany) will align the crystallisation drop to the centre of the air-bearing ω rotation axis, which will ensure a limited movement in angular range (±45º). The supporting frame can hold SBS format crystallisation plates. A graphical user interface, written in Python Qt4, permits the user to move to any given crystallisation well and control the various data collection parameters (zone to

ISBN: 978-3-95450-207-3 doi:10.18429/JACoW-MEDSI2018-WEPH36

screen, exposure time, detector frame rate, etc.). The images are automatically processed using the DIALS spot finder routine (dials.diamond.ac.uk), which searches for X-ray diffraction spots in the collected images. The number of diffraction spots per position is then rendered as a heat map.

At the moment the plate screener collects X-ray diffraction images as stills, but in the near future we envisage adding an ω rotation axis and incorporating other functionalities.

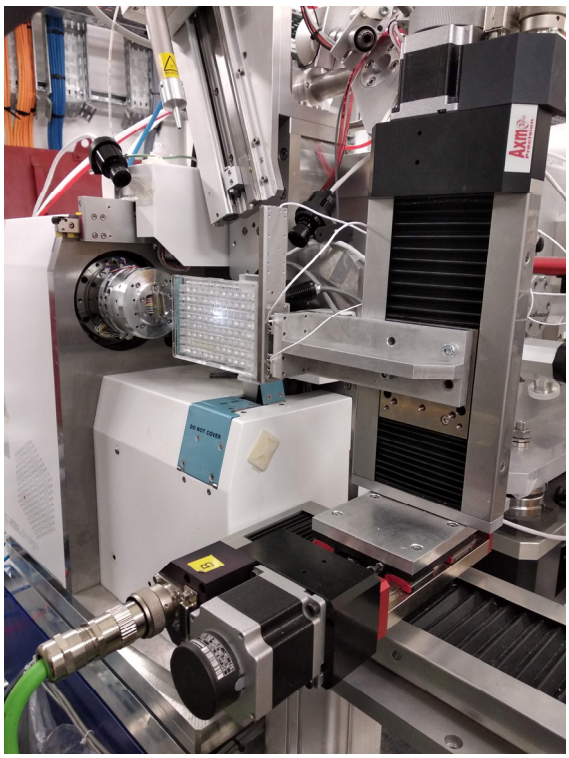

Figure 1: Photo of the Plate screener on PROXIMA 2A.

Fibre Optic Coupled YAG Beamstop

In order to record the actual flux at the sample during the experiment, we are developing a miniaturised beam monitoring system mounted inside a small beamstop of the experimental set up. For micro-crystallography experiments, it is preferable to reduce any background scatter by placing the beamstop as close as possible behind the sample. However, for MX data collections, the lowest resolution shells (>50 Å) are important in the calculation of the electron density maps. On PROXIMA 2A, the nominal beamstop is only 300 µm in height and placed 10 mm behind the sample. Making an incident beam monitor of these dimensions is a real challenge. Nevertheless, we have developed a prototype, which is a YAG crystal coupled to an optical fibre that transmits the signal to a photodiode. The prototype shows a linear response with respect to the incident flux, but it is sensitive to any movements of the crystal, which became unglued during the exposure to the intense X-rays.

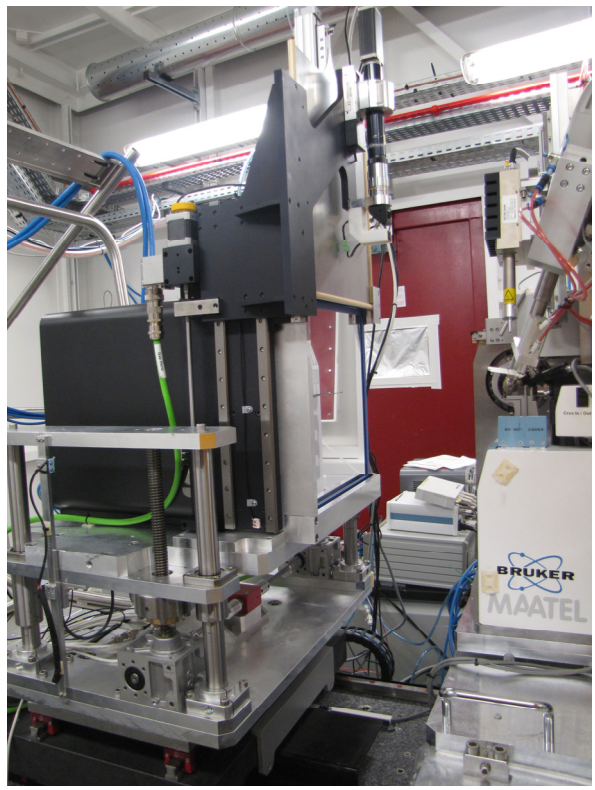

Figure 2: Photo of the vertical translation table for the EIGER X 9M on PROXIMA 2A.

EIGER X 9M Translation Table

With ever increasing frequency, many PROXIMA 2A users grow crystals that diffract to sub-atomic resolution, but until recently the experimental set up could not bring the EIGER X 9M detector close enough to permit the complete data collection to resolutions beyond 1.1Å resolution at the nominal energy of 12.65 keV. By elevating the detector by as much as 120 mm and raising the X-ray energy to 18 keV, the PROXIMA 2A set up can readily collect complete X-ray diffraction data to 0.6Å resolution.

The vertical translation stage of the EIGER X 9M detector is comprised of four stiff columns, which guide the lifting of a 25mm thick base plate. A single motor coupled to a transmission (Phytron, Germany) drives the two lead screws. The two struts connecting the guide columns reinforce the rigidity of the ensemble (see Figure 2). X-ray powder diffraction images of LaB_6 confirm that diffraction rings are observable beyond 0.6Å resolution.

Determining the SOC With a Mini-Kappa

In order to provide extra functionality for the users, especially those who wish to perform sulphur SAD experiments, the PROXIMA 2A beamline recently acquired a motorised mini-kappa arm (MK3, Arinax, France) for the MD2. However, this motorised arm does not convert the system into a proper κ goniometer as the sample needs re-centring for different κ, φ and ω angles. Furthermore, the weight of the MK3 affects the SOC of the ω air-bearing spindle of the MD2. Thus, we are currently investigating methods to automatically characterise the shifts in the

ISBN: 978-3-95450-207-3 doi:10.18429/JACoW-MEDSI2018-WEPH36

position using a ball bearing, a thin fibre or a point, in order to create a geometric displacement map. The site acceptance tests using a point and optical microscope of the MD2 confirmed that the SOC (peak-valley) of the ω-axis degrades slightly of the extra weight of the mini-kappa arm from 2.4 µm at κ = 0° to 4.8 µm at κ = 130°. However our initial attempts to reproduce these measurements with an interferometer and ball bearing yielded much higher SOC values (8.8 µm horizontally and 9.0 µm vertically at κ = 0°), and the measurements were very sensitive to and depended upon the mechanical set up. Consequently, we are continuing our pursuit of more reliable methods of determining the SOC for the definition of a displacement map for the MK3 & MD2.

Figure 3: Result of sample detection from an overhead image of samples stored in the Dewar under liquid nitrogen (LN2): Purple circles: identified lids, red circles: missing pucks, green circles: present pucks, red squares: missing pins, green squares: present pins.

Computer Vision Detection of Samples in LN2

The automated sample-transfer system installed on PROXIMA 2A stores up to 148 samples under a layer of liquid nitrogen (LN2). A limitation of the system is its inability to detect the presence of individual samples. This sometimes leads to confusion on the part of the user and leads to unnecessary interruptions in the smooth operation of the beamline. To address this problem, we have developed a system to determine the exact content of the Dewar by the analysis of optical images acquired by a camera positioned above the storage Dewar.

Given an image of the storage Dewar (Figure 3), we determine the locations of present and missing samples. We considered the problem as that of defining segments of the image corresponding to missing and present samples and solved it as a supervised classification machine-learning task. Using human designated positive and negative examples (the training data set contained ~10,000 examples for each class), we trained a linear classifier to decide whether a particular image patch is likely to correspond to a present or missing sample. Using the results from the classification step, we could then make use of the known geometry of the Dewar to determine the coordinates and indices of all missing and present samples. An image pre-processing and sliding window search were done using scikit-image library; feature computation, classification and clustering using scikit-learn library.

The system is robust with respect to small differences in intensity of lighting across the image, differences in pin types and the solution is also robust with respect to small changes in camera position and scale of the input image.

CONCLUSIONS

The various advances in instrumentation and methodology will render the automation of MX beamlines more robust while also offering increased functionality. *In situ* plate screening will become necessary for the crystalline systems that are particularly fragile and do not resist the cryo-cooling step, such as crystals of membrane proteins, although plate screening can also be useful to verify the diffraction limit of crystals prior to any handling. Extension of the crystallographic resolution limit to 0.6Å and implementation of an easy to use κ goniometry are obvious advantages to the MX users' arsenal. Similarly, knowledge of the incident flux will not only help monitor the stability of the X-rays on the sample, which is important for micro-crystal and anomalous scattering experiments, but it will also allow users to calculate the X-ray dose and avoid unnecessary radiation damage. Finally, the application of artificial intelligence and computer vision technologies will enable the automatic handling of more complex situations and extend the boundaries of automation on the beamline.

ACKNOWLEDGEMENTS

The authors wish to thank the various support groups at Synchrotron SOLEIL, especially Alain Lestrade, Jean-Michel Dubuisson and Bertrand Pillaud.

REFERENCES

[1] L. Jacquamet, J. Ohana, J. Joly, F. Borel, M. Pirocchi, P. Charrault, A. Bertoni, P. Israel-Gouy, P. Carpentier, F. Kozielski, D. Blot, and J. L. Ferrer, "Automated analysis of vapor diffusion crystallization drops with an X-ray beam," *Structure*, vol. 12, no. 7, pp. 1219–1225, 2004.

[2] A. Le Maire, M. Gelin, S. Pochet, F. Hoh, M. Pirocchi, J. F. Guichou, J. L. Ferrer, and G. Labesse, "In-plate protein crystallization, in situ ligand soaking and X-ray diffraction," *Acta Cryst. D*, vol. 67, no. 9, pp. 747–755, 2011.

[3] N. Watanabe, H. Murai, and I. Tanaka, "Semi-automatic protein crystallization system that allows in situ observation of X-ray diffraction from crystals in the drop," *Acta Cryst. D*, vol. 58, no. 10 II, pp. 1527–1530, 2002.

[4] J. H. Kelly, "Engineering Challenges of the VMXi Beamline," in *Proc. MEDSI'16*, Barcelona, Spain, pp. 304–308, 2016. doi:10.18429/JACoW-MEDSI2016-WECA07