BL32XU
RIKEN Targeted Proteins

1. Introduction

BL32XU is the RIKEN targeted protein beamline dedicated to high-resolution diffraction data collection from protein microcrystals. Since FY2015, we have been developing a fully automated data-collection system dedicated to protein crystallography, which is named ZOO, at the BL32XU [1]. The ZOO system covers all existing experimental schemes in goniometer-based data collection from protein crystals. Furthermore, it has achieved un-attended data collection. Hence, remote users can acquire high-resolution datasets using SPring-8 just by sending crystal samples. BL32XU has supported numerous structure determinations of challenging proteins such as membrane proteins (e.g., GPCR) as part of the BINDS (Basis for Supporting Innovative Drug Discovery and Life Science Research) program since FY2017.

2. Recent activities

The ZOO system was initially developed at BL32XU in combination with a brilliant micro-focus X-ray beam (at minimum, < 1 μm × 1 μm), a pixel array detector, and an automated sample changer. The system consists of several programs such as that for automatic loop centering, a two-dimensional raster scan, defining crystal size and positions, and estimating radiation damage. The automatic data-processing pipeline KAMO immediately processes collected datasets using a computer cluster [2]. The ZOO system secures data quality by automated crystal selection and controlling radiation damage. The enhanced efficiency and quality of data collection accelerate the accumulation of better datasets. Merging many datasets improves the signal-to-noise ratio required for higher-resolution structure determination. Consequently, the ZOO system has accelerated high-resolution structural analysis of challenging proteins [3, 4].

The ZOO system can collect 200–300 full datasets within 24 hours from protein crystals larger than 50 μm (one crystal/loop). This efficiency enables high-throughput protein crystallography, which is a prerequisite for fragment-based drug design (FBDD), which examines many crystals in a complex with small ligands and large-scale small-wedge data collection from membrane protein crystals grown in the lipidic cubic phase (LCP). To analyze many crystal structures, it is critical to efficiently handle the tremendous amount of data produced by the ZOO system. Thus, we developed the automatic data processing pipeline KAMO.

In FY2019, the data-merging function, which consists of hierarchical clustering and merging of small-wedge data, was automated. Numerous data sets can be processed automatically from the diffraction intensities to the reflection file, which can be used for structural analysis with the aid of the automatic data-processing pipeline and merging function. We also developed an automatic data-assessing system to evaluate the results from numerous data sets. This system includes a logging function of experiments and a reporting function to experimenters.

Since the ZOO system can automatically collect data based on the conditions that experimenters

define prior to the experiments, un-attended data collection is now available. To realize routine un-attended data collection and to accelerate high-throughput protein crystallography, we have replaced the function of BL45XU from SAXS to protein crystallography. To switch BL45XU to a fully automated protein crystallography beamline, the ZOO system and the automatic data analysis system were installed with the support of BL32XU. Since May 2019, BL45XU has provided un-attended measurements with the ZOO system, and contributed to high-throughput protein crystallography for academic and industrial users.

The ZOO system has enhanced the efficiency of experiments in goniometer-based data collection. Currently, we are now focusing on developing efficient crystal harvesting to enhance the throughput to harvest protein crystals from crystallization plates. In particular, fishing LCP crystal from the plate is difficult and requires well-experienced skills. To overcome this harvesting problem, AOMUSHI has been developed for LCP samples crystallized by the film-sandwich method using a Diffrax plate. Instead of fishing LCP crystals, AOMUSHI just punches a hole (6-mm diameter) in the Diffrax plate to mount the sample on a special holder. Previously only 20–30% of LCP crystals were available for diffraction measurements, even though a well-experienced operator fished the crystals. Consequently, the development of AOMUSHI allows all crystals to be used by simple and easy handling. Further developments of automation of sample-freezing procedures for the sample holder from AOMUSHI using collaborative robots and software for automatic data collection for large-area holders of AOMUSHI are underway.

Recently we have tried to determine the high-resolution structure from sub-μm crystals by serial synchrotron rotation crystallography (SSROX) at BL32XU. Using a microfocus X-ray beam (1 μm × 1 μm), we achieved structure determination at 2.3-Å resolution from 500–700-nm crystals (Fig. 1) (in preparation). ZOO allows users to automatically acquire datasets from thousands to tens of thousands of tiny crystals. Moreover, to adapt data collection from tiny crystals with weak diffraction, we have developed a new methodology to obtain a low-background dataset. These developments should expand the possibilities of crystal-structure determinations of challenging proteins.

The ZOO system covers all measurement schemes not only a conventional single-point exposure scheme, but also helical data or multiple small-wedge data collection schemes. Automatic data collection by the ZOO system requires a spreadsheet, which describes conditions for data collection. Although users can specify suitable measurement schemes or experimental parameters by themselves, it is sometimes difficult for less experienced users to complete the appropriate parameters. Additionally, challenges may arise when many kinds of crystals are fished in the same loop. Hence, we have further developed a program

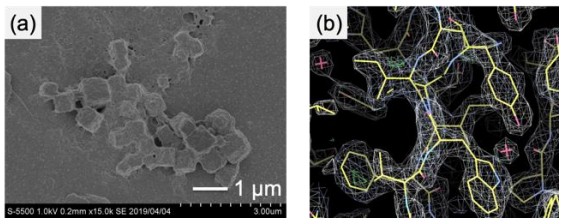

Fig. 1. Sub-μm crystals used for SSROX experiment with the ZOO system. (a) SEM image of sub-μm crystals and (b) electron density map obtained from SSROX experiment.

called HITO to automatically determine the best measurement scheme from the result of a raster scan. The installation of the HITO system in the ZOO system has realized automatic selection of better data collection schemes from spatial information of crystals on each loop (Fig. 2). In the near future, we plan to implement and exploit deep learning in HITO for more complex but automated experiments.

Structure determinations of many difficult and challenging proteins (*e.g.*, membrane proteins) have been achieved by the ZOO system. The parameters for data collection were optimized by deep consideration with numerous test studies. Now HITO is available for users. Hence, researchers, even those unfamiliar with X-ray crystallography, can easily obtain high-quality diffraction data. Highly efficient automatic data collection by the ZOO system yields a vast amount of diffraction datasets.

Our current focus is on analyzing such big data in protein crystallography. To handle such a large amount of data, an automatic structure-analysis pipeline and overview system of such big data are under development. These should contribute to future research on protein structure dynamics.

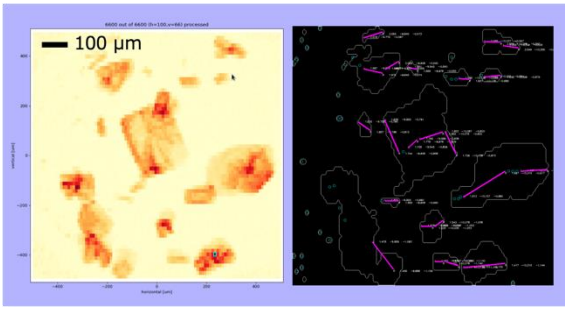

Fig. 2. HITO system for a mixed-scheme measurement. (left) Raster-scan image and (right) result of heat-map analysis by HITO.

Kunio Hirata[*1], Yoshiaki Kawano[*1], Naoki Sakai[*2], and Hiroaki Matsuura[*2]

[*1] SR Life Science Instrumentation Team, Life Science Research Infrastructure Group, Advanced Photon Technology Division, RIKEN SPring-8 Center

[*2] Life Science Research Infrastructure Group, Advanced Photon Technology Division, RIKEN SPring-8 Center

References:

[1] Hirata, K. et al. (2019). *Acta Cryst. SecD,* 75, 138–150.

[2] Yamashita, K. et al. (2018). *Acta Cryst.* D, 74, 441–449.

[3] Miller et al. (2020). *Cell Chem. Biol.,* 27, 1192–1198.

[4] Katoh et al. (2020). *Nat. Chem.* doi: 10.1038/s41557-020-0525.