De novo phasing with serial femtosecond crystallography at SACLA

Phase determination has been a major problem in protein crystallography. The protein crystal structure is very informative for understanding the functions of proteins; however, it is difficult to solve the protein structure using μm-size protein crystals, especially membrane protein crystals.

The high brilliance X-ray pulses of the X-ray free electron laser (XFEL) are anticipated to enable crystal structure determination using microcrystals. Serial femtosecond crystallography (SFX) [1] is a new method of structure determination, where single-pulse diffraction patterns are collected from a liquid flow of microcrystals using X-ray pulses of the XFEL. Data collection by SFX will reduce the time spent on the optimization of protein crystallization conditions because large crystals are no longer required. However, structure determination by SFX has mainly been restricted to the molecular replacement method. Thus, the development of *de novo* phasing using protein microcrystals for XFEL has been anticipated.

The single-wavelength anomalous dispersion (SAD) method is the most common experimental phasing method in protein crystallography. Barends *et al.* succeeded in *de novo* phasing by the SAD method with lysozyme crystals using the SFX method at LCLS in 2014 [2]. About 60,000 single-pulse patterns, which were required for the structure determination, were collected using 8.5 keV X-rays in order to measure the anomalous dispersion of Gd ions soaked in lysozyme crystals. Another reason for this X-ray energy was because the maximum available photon energy at LCLS was 9.5 keV. Heavy atoms such as Hg, Se, Pt, and Au are most commonly used in protein phase determination at synchrotron radiation facilities. The absorption edges of these atoms are near 12.4 keV. Such a high photon energy, which is advantageous in high-resolution *de novo* structure determination, is available at SACLA. We are developing an effective *de novo* phasing method using protein microcrystals and the high photon energy XFEL of SACLA [3].

In order to carry out phasing by SFX at SACLA, we prepared microcrystals of native luciferin regenerating enzyme (LRE) and a Hg derivative. The microcrystals were crystallized by a batch method using PEG3350 as a precipitant. Hg-derivative crystals were obtained by soaking native microcrystals in stock solution containing 1 mM HgO. Before SFX data collection, the solution including the microcrystals was mixed with a grease matrix in order to increase the viscosity [4]. The SFX experiment was performed at **BL3** of SACLA. The mixture solution was extruded from the syringe injector system in a DAPHNIS chamber [5] at room temperature. The photon energy was tuned to 12.6 keV, which is near the absorption edge of the Hg atom. We collected 133,958 and 583,291 images for the native and Hg-derivative crystals, respectively. For the native crystals, 26,238 images (20%) were preselected on the basis of the diffraction intensity

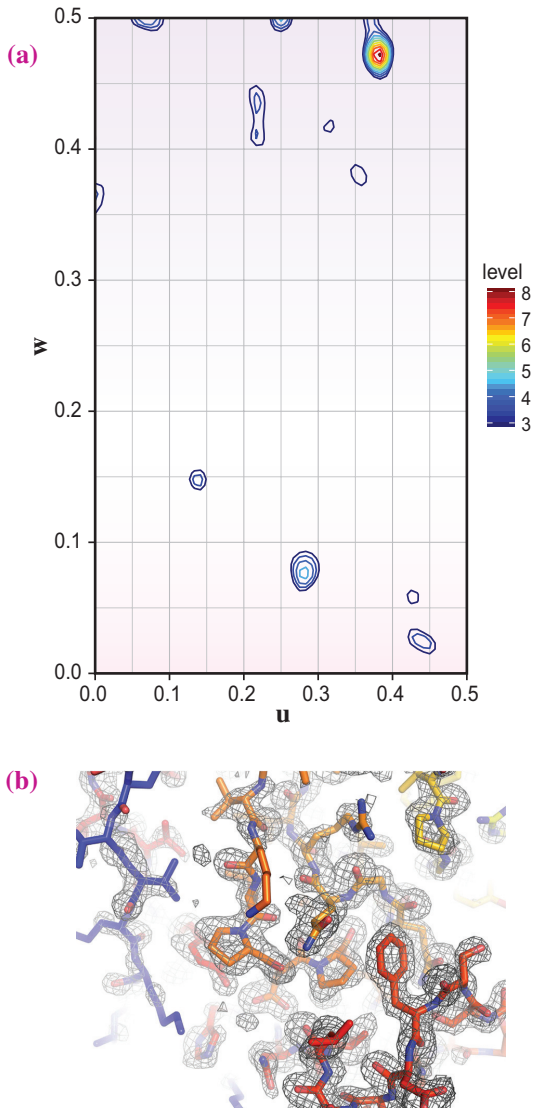

Fig. 1. **(a)** The v = 1/2 Harker section of isomorphous difference Patterson map. **(b)** A refined model of LRE and SIRAS electron density map produced by SHELXE. The electron density map is contoured at 1.0σ. These maps were calculated with 10,792 and 10,000 indexed patterns of native and Hg-derivative LRE crystals, respectively.

and 10,792 patterns of these (41%) were indexed. For the Hg-derivative crystals, we selected 298,061 (51%) images and 85,747 patterns of these (29%) were indexed. Diffraction patterns were processed using the CrystFEL software suite [6]. Monte Carlo integration for the native and Hg-derivative crystals yielded mean SFX multiplicities of 222 at 1.5 Å resolution and 908 at 1.6 Å resolution, respectively. We attempted phasing by the SAD method using all the Hg-derivative data, but this was unsuccessful.

Although SAD phasing was unsuccessful, we succeeded in phase determination by the SIRAS method using 10,792 native and 10,000 Hg-derivative images with a mean multiplicity of 106. The phase calculations and improvements were carried out using the SHELX C, D, and E programs. An isomorphous difference Patterson map showed a significant peak corresponding to the position of the Hg atom at 8.1σ (Fig. 1(a)). When the correct hand was used, an electron density map that was readily interpretable using a mean FOM of 0.615 was obtained and 197 residues were modeled with CC = 27%. Automatic model building was then performed using ARP/wARP with REFMAC5. 304 out of 308 residues were modeled with satisfactory accuracy (R_{work} = 22.3% and R_{free} = 27.6%) (Fig. 1(b)).

In order to examine the quality of the *de novo* phasing, we tested the SIRAS phasing protocol using different numbers of native and derivative patterns. We evaluated the phase quality using the correlation coefficient (CC) between the experimental electron density map given by SHELXE and the Fc map calculated using the final refined structure mentioned above. When the CC value was greater than 0.65, the trial was judged a success. More than 5,000 native patterns and 8,000 derivative patterns were required but in total at least 18,000 patterns were needed for successful phasing (Fig. 2).

In this study, we showed that the combination of Hg-derivative and native crystals enabled *de novo* phasing using SIRAS with a much smaller number of crystals than that required for the SAD method. *De novo* phasing was achieved using data collected in only two hours. The reduction in the amount of data required for structure determination is very important to make the best use of the limited beam time of XFEL because there are only two XFEL facilities in the world, SACLA and LCLS.

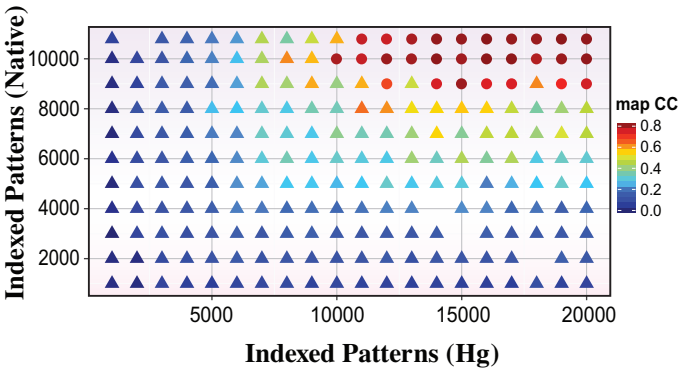

Fig. 2. Phase quality (CC) as a function of the numbers of native and derivative indexed patterns. The success and failure of phasing are represented as circular and triangular symbols, respectively. The CC between the experimental electron density map and the Fc map is indicated by the color. Some data points are missing because SHELXE failed to trace the map.

Toru Nakatsu[a,b,*], Keitaro Yamashita[b] and So Iwata[b,c]

[a] Department of Structural Biology, Kyoto University
[b] RIKEN SPring-8 Center
[c] Department of Cell Biology, Kyoto University

*E-mail: nakatsu@pharm.kyoto-u.ac.jp

References

[1] H.N. Chapman *et al.*: Nature **470** (2011) 73.
[2] T.R.M. Barends *et al.*: Nature **505** (2014) 244.
[3] *K. Yamashita, D. Pan, T. Okuda, M. Sugahara, A. Kodan, T. Yamaguchi, T. Murai, K. Gomi, N. Kajiyama, E. Mizohata, M. Suzuki, E. Nango, K. Tono, Y. Joti, T. Kameshima, J. Park, C. Song, T. Hatsui, M. Yabashi, S. Iwata, H. Kato, H. Ago, M. Yamamoto and T. Nakatsu: Sci. Rep.* ***5*** *(2015) 14017.*
[4] M. Sugahara *et al.*: Nat. Meth. **12** (2015) 61.
[5] K. Too *et al.*: J. Synchrotron Rad. **22** (2015) 532.
[6] T.A. White *et al.*: J. Appl. Cryst. **45** (2012) 335.