Development of High-Throughput Protein Crystallography

Huge amount of genetic information on the entire genomes has been produced from the human and other genome sequencing projects. This information has been used extensively in a new branch of science, called the post-genome science, and the result is a better human life. Diseases, an arch enemy of human beings, could now be better understood from the basic research in genomics and proteomics. A whole new pattern for the diagnosis and treatment of human ailment, which will bring radical changes to the way health care is delivered, is also introduced. These, if properly used, will generate a new and powerful economy to help flourish the existing computer and electronic industries in Taiwan. To achieve this goal, the National Science Council (NSC) has launched the "National Research Program for Genomic Medicine (NRPGM)" in 2002 as the first phase of Taiwan's Biotechnology Initiative. The purpose of NRPGM is to capitalize on the knowledge incorporated in the human genome sequence to develop Taiwan's competitiveness in medical research.

Genomic information is a powerful tool for studying gene function, since it introduces a complete set of structural families into which all proteins can be classified. This is also one of the many goals of structural genomics. To understand gene function, we need to know its three-dimensional structure. The protein structure could be determined by X-ray crystallography or nuclear magnetic resonance (NMR) techniques. The development of modern synchrotron radiation facilities in the past ten years allows the protein crystallographers to collect complete data sets in tens of minutes. With tunable energy and highly collimated X-ray photons, the time required for accurate structure determination has been reduced by one to two orders of magnitude. Structure determination of large unit cell protein and extreme small crystals also has become possible. It could be expected that thousands of protein structures will be solved and their functions analyzed in a coordinated way worldwide in the near future.

With strong synchrotron sources and modern detectors, the data collection time is null compared to the time required for routine processes, such as sample mounting, crystal centering, and the determination of data collection parameters. To make the best use of beam time, these routine processes should be fully automated. By automation we mean that the software and hardware should complete the assigned jobs automatically without human intervention. Automatic screening and scoring of a large number of crystals followed by data collection from the best crystals would let the crystallographers concentrate on other important issues. Users will also benefit from the automated data processing during an experiment. They can monitor the experiment and make decision to collect more data or to stop the current experiment. Automation is an essential part of the high-throughput structural genomics programs.

To obtain a complete collection of protein folds and representative members of all protein families, reproducible and cost-efficient high-throughput methods must be used. Our approach to achieve high-throughput protein crystallography at NSRRC is to implement advanced instrumentation, automation, and an intelligent control system. For beamline instrumentation, a unique superconducting wiggler is installed in the storage ring of the Taiwan Light Source to provide high intensity X-ray beams at photon energy up to 19 keV. The brilliance of this insertion device is almost equal to the third-generation synchrotron X-ray source, as shown in Fig. 1. Three carefully designed beamlines shown in Fig. 2 are built with the following features: (1) full remote control, (2) full system monitoring and diagnostics, (3) stable and reproducible optics producing high-intensity beams over long range of

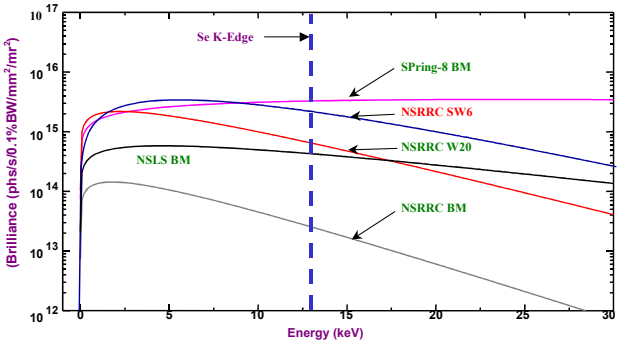

Fig. 1: Source spectrum of the specially-designed Superconducting Wiggler (SW6) in NSRRC.

wavelengths. For experimental instrumentation, the state-of-the-art fast and large area charge coupled device (CCD) detector, high-precision diffractometer, robotic system for automatic sample mounting and centering, high speed data network with large data storage, and user-friendly beamline control system, are used to collect single crystal diffraction data in high-throughput mode, as shown in Fig. 3.

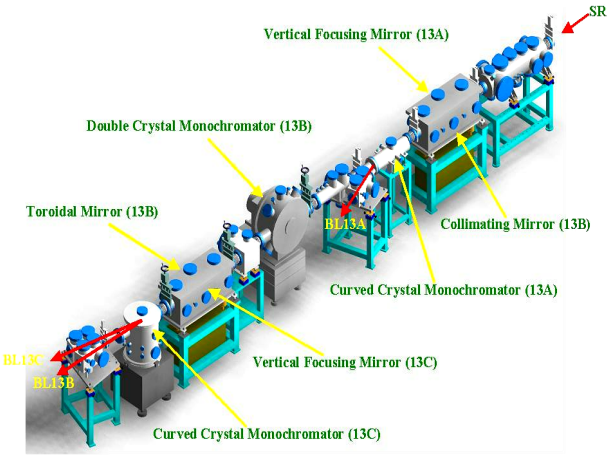

Fig. 2: Layout of the three high-throughput protein crystallography beamlines.

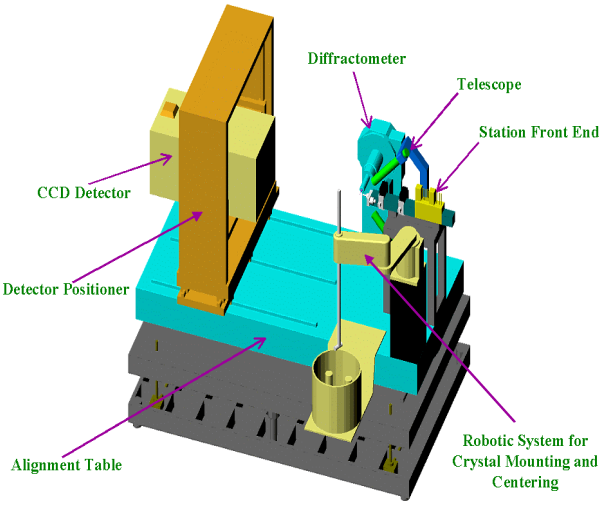

Fig. 3: 3D CAD design of the experimental station.

With a user-friendly beamline control system that incorporates automated sample changing, mounting, viewing, centering, and data acquisition, users can concentrate on data collection and analysis without worrying about the beamline operation. A large area and fast readout CCD detector is used in the current application, which substantially reduces the cycling dead time, and hence minimizes the time-dependent effect of radiation damage on the sample and improves the quality of data collected. A powerful computer system equipped with all necessary crystallographic software will allow users to process their data *in situ* to enhance the success of the project.

This high-throughput protein crystallography facility is supported by the National Research Program for Genomic Medicine (NRPGM) launched by National Science Council (NSC). The facility comprises a Multi-wavelengths Anomalous Dispersion (MAD) beamline with energies from 6.5 keV to 19 keV, and a monochromatic beamline with fixed energy tunable from 12 keV to 14 keV for crystal screening and high-resolution structure study. A large team with members of different expertise has been assembled at NSRRC to carry out this project. Great thanks are given to all staffs at NSRRC, especially those in the Beamline Group and Biology Group, who are directly involved in this project. This facility will be open to the biology users in the year 2005.

AUTHOR
Y.-C. Jean
National Synchrotron Radiation Research Center, Hsinchu, Taiwan

PUBLICATION

· P. C. Tseng, Y. C. Jean, C. H. Chang, L. J. Huang, H. S. Fung, Y. F. Song, C. I. Ma, Y. S. Huang, F. Chao, C. J. Chen, and K. L. Tsang, *"Conceptual design report for the high-throughput biological crystallography beamlines at the SRRC"*, SRRC Report, SRRC/RBM/IM/02-04 (2002).

CONTACT E-MAIL
ycjean@nsrrc.org.tw