4 Structural Biology Group

4-1 Outline

The Structural Biology Group led by Prof. Soichi Wakatsuki has continued to evolve, with activities in structural biology research and synchrotron instrumentation for demanding protein crystallographic experiments during its third year (The latest information is available in the group's web site http://pfweis.kek.jp/index_eg.html). This is the second year of the MEXT (Ministry of Education, Culture, Sports, Science and Technology) grant, the Special Coordination Funds for Promoting Science and Technology (Shinko-Chosei-hi), and we are already well into the maturation phase with a series of new systems that were initiated in 2000 or 2001 for high-throughput X-ray protein crystallography. In addition, the "Protein 3000" project officially started in FY2002 and solicited network research proposals in seven target areas. The KEK Structural Biology Group formed a research consortium with 4 research institutes, 9 university groups, and a technology transfer firm and submitted an application to the project. We were selected as one of eight consortia and started to pursue target-oriented structural genomics of protein modification and transport (Fig. 1). During FY2002, our consortium produced 10 new structures and 48 research papers on the structure and function of the target proteins, a large proportion of which were published in journals with high impact factors. The goal we had set forward was to determine 5 structures in the first year, which we have surpassed already by a factor of two in FY2002, and are advancing even faster in FY2003.

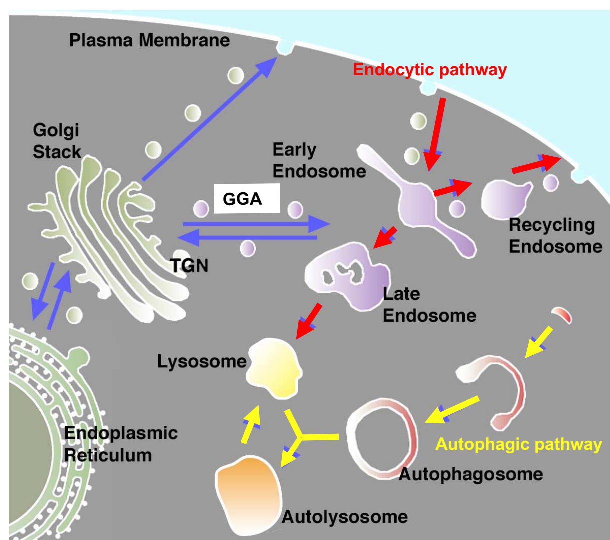

Figure 1
Post-translational modification and transport of proteins.

Increase in the Group Members

FY2002 has seen a dramatic increase in the number of personnel (two postdoctoral fellows, three Ph.D. students and five technicians) and there are now 26 members in the group. Two Japanese Ph.D. students started their Ph.D. thesis projects on protein transport at the Graduate University of Advanced Studies (GUAS) in FY2002. Both of these projects are in collaboration with Prof. Tamotsu Yoshimori of the National Institute of Genetics, Okazaki and are in part supported by the GUAS collaborative research grant between 2002 and 2004. In Autumn 2002, a French Ph.D. student joined the group from Grenoble, France to pursue a Ph.D. course in GUAS on crystallographic studies on vesicle transport and protein modification. As of the end of FY2002, the three students have already accomplished a significant amount of research and are now preparing their first research papers.

Extension of the Structural Biology Building

Using parts of the two grants from the MEXT, we have made substantial improvements of the experimental facilities in the Structural Biology Building. For example, an insect cell based baculovirus protein expression system has been set up for proteins which are difficult to express in active forms using E. coli or yeast expression systems. To cope with the rapid increase in the number of personnel and research activities, the Structural Biology Building, which was built in March 2001, was extended from 429 m^2 to 643 m^2 (an increase of 214 m^2) at the end of March 2003 (Fig. 2). The extension will be used for experiments on protein expression, purification and crystallization. In particular, a high-throughput automated protein crystallization system will be installed in the expanded area (see 4-4).

Figure 2
The Structural Biology Building after the extension in March 2003.

4-2 Structural Biology Program: Intracellular Protein Transport and Protein Glycosylation

GGAs, a New Class of Multi-potent Adaptor Proteins of the Clathrin-Coated Transport Vesicles: Collection, Packaging, and Help in Delivery of Cargos

Adaptor protein (AP) complexes play critical roles in vesicular transport through an intricate network of interactions with cargo receptors, clathrin and various accessory proteins. A novel family of adaptor proteins, called GGAs (Golgi-localizing, γ-adaptin ear homology domain, ARF-binding proteins), has been shown to be important for clathrin-coated vesicle transport between the TGN (trans-Golgi network) and endosomes/lysosomes in lieu of classical AP-1 complexes. GGA is an adapter protein with three domains and a hinge region, each of which interacts with other transport proteins (Fig. 3). International quests for their structures have produced a number of crystal structures of the three domains, VHS, GAT, and GAE and their complexes with signal peptides or target proteins, in just under three years. The N-terminal VHS domain binds TGN sorting receptors, such as mannose 6-phosphate receptors (MPR), sortilin and β-secretase (BACE) by recognizing their acidic cluster dileucine sequences [1]. The GAT domain uses its hydrophobic residues of the N-terminal helix-turn-helix motif to dock onto Switches I and II of GTP-bound ARF (ADP-ribosylation factor), thereby tethering GGA to the TGN membrane for efficient recruitment of cargo receptors by the VHS domain [2] while it employs its C-terminal half to interact with ubiquitin and Rabaptin-5. The proline-rich hinge region recruits the N-terminal β-propeller of clathrin to initiate clathrin cage formation. It also contains sequences capable of binding to the GAE domain and the γ-ear domain of AP-1, as well as an autoinhibitory sequence for the VHS domain. Finally, the γ-ear domain of AP-1 [3] and the GAE domain [4] adopt very similar immunoglobulin folds to interact with Rabaptin-5 and γ-synergin. Taken together these structures shed light on the molecular mechanisms of this new class of adaptor proteins in vesicular transport: docking to the target membrane, recruitment of cargos, regulation through interactions with accessory proteins, and vesicle formation.

References

[1] T. Shiba, H. Takatsu, T. Nogi, N. Matsugaki, M. Kawasaki, N. Igarashi, M. Suzuki, R. Kato, T. Earnest, K. Nakayama and S. Wakatsuki, *Nature* **415** (2002) 937; T. Shiba *et al.*, submitted.

[2] T. Shiba, M. Kawasaki, H. Takatsu, T. Nogi, N. Matsugaki, N. Igarashi, M. Suzuki, R. Kato, K. Nakayama and S. Wakatsuki, *Nature Structural Biology* **10** (2003) 386; "Highlight" 8-1 (p. 30) of this issue for details.

[3] T. Nogi, Y. Shiba, M. Kawasaki, T. Shiba, N. Matsugaki, N. Igarashi, M. Suzuki, R. Kato, H. Takatsu, K. Nakayama and S. Wakatsuki, *Nature Structural Biology* **9** (2002) 527; Y. Yamada *et al.*, in preparation.

[4] M. Inoue *et al.*, in preparation.

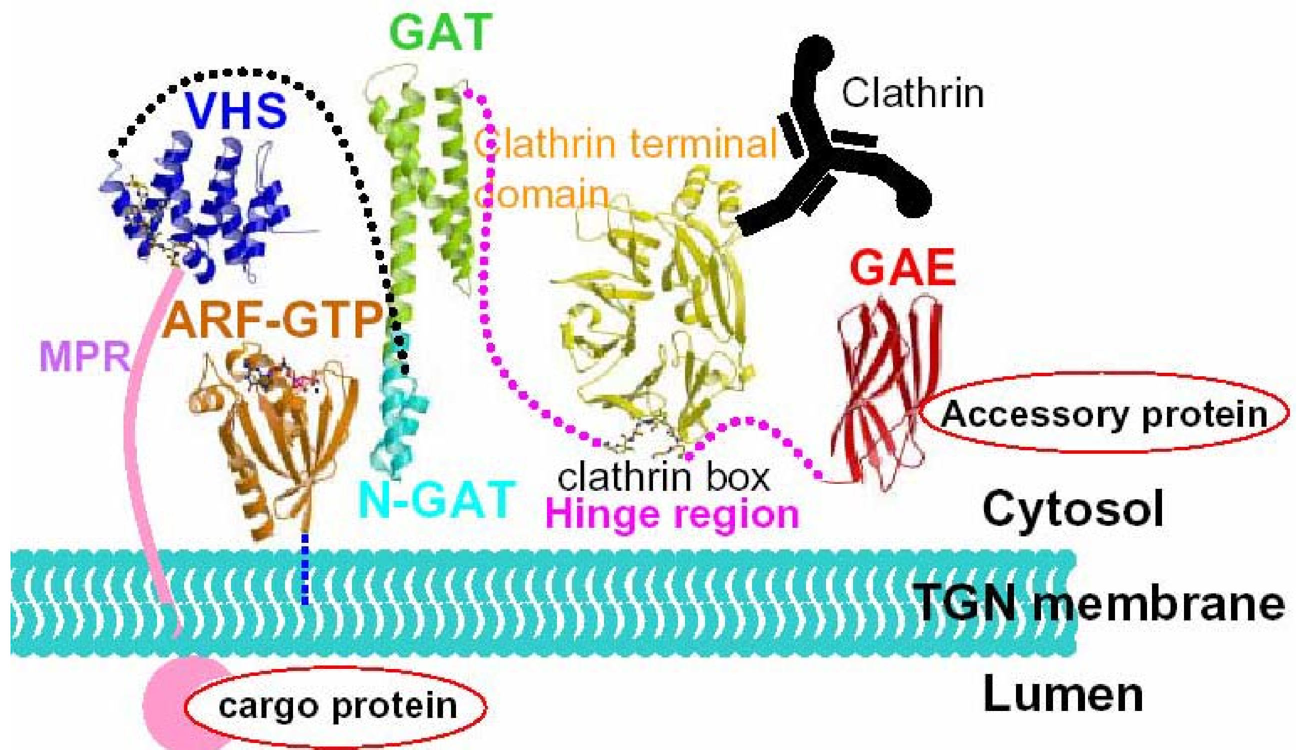

Figure 3
Structures of the individual domains of human GGA1 and their partner proteins.

4-3 Beamline Developments

AR-NW12 X-Ray Undulator Beamline

In February of 2003, we completed the construction and commissioning of a new MAD beamline NW12 (see Sec. 2-3 on p. 64 for details of the insertion device and optics). The beamline features a highly-collimated, high flux, tunable monochromatic X-ray beam optimized for efficient MAD experiments. We have installed a CCD-type X-ray detector, ADSC Q210 on the table in the experimental hutch (Fig. 4). The active area of the detector is 210 mm × 210 mm and the pixel size is 51 µm. An image (4096 × 4096 × 2 byte) can be read out in just under one second. High-speed and high-resolution data collection can be realized with the detector. Combined with the gigabit Ethernet network, the dead time including readout is around 2.5 s per frame. The exposure time with the undulator beam is typically around 5 s. This means that the time to collect one data set (1 degree oscillation and 180 frames) is around 20 minutes. A typical four-wavelength MAD experiment including XAFS measurements takes 1 or 2 hours to complete the data collection, which is around 10 times shorter than at BL-6A or BL-18B.

The experimental table is adjusted so that the deviation of the direct beam position on the detector surface is less than 51 µm when changing the detector distance from the sample. The table has an attenuator, a four-blade slit, a shutter and a one-axis goniometer in front of the detector. In addition, the equipment around the sample, such as the beam stop and the fluorescent detector are all motorized and retractable to allow for the automatic mounting and removal of crystals (Fig. 5). The zoom and focus of the telescope are also motorized for semi-automatic crystal alignment. Later this will be further improved to truly automated crystal centering using an all-in-focus method the group is currently developing. All of the components can be moved together using the beamline control software, allowing for fully automated data collection in the near future.

The shutter can handle oscillation experiments with exposure times as short as 10 ms and with a rotation accuracy of the sample axis of 2.3 µm, which makes measurements with micron size crystals in a short-time exposure possible. Because the synchronization of the shutter timing and the movement of the sample axis is important in such measurements, we have adjusted the timing error to less than 1 ms. However, currently, we limit the shortest exposure time to about one second. Public use of this beamline will be started from May 2003.

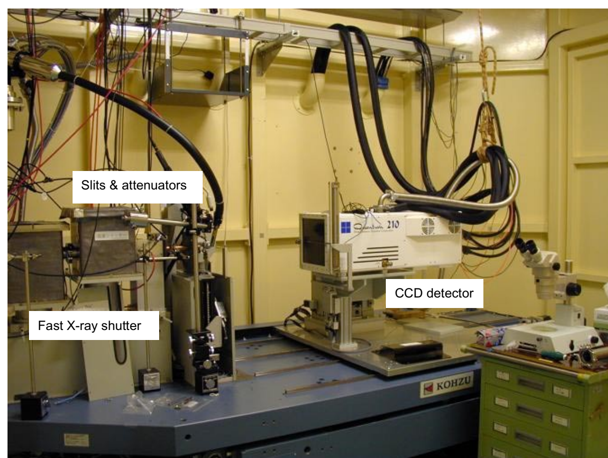

Figure 4
Inside the high-throughput MAD beamline PF-AR NW12.

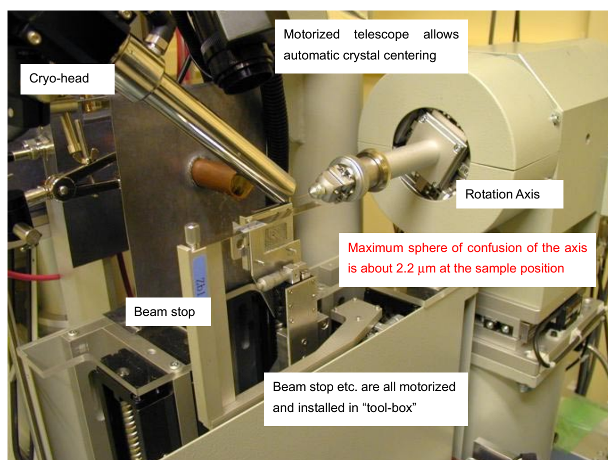

Figure 5
User-friendly experimental set-up of PF-AR NW12.

A Multipole Wiggler MAD Beamline BL-5

We are continuing the construction of the new protein crystallography beamline, BL-5, on the PF ring using a 2.4-m multipole wiggler as a light source for high-throughput MAD experiments (Fig. 6). Two optics hutches, an experimental hutch, and a control cabin adjacent to the experimental hutch have been constructed with a large roof deck (Bridge) covering all the hutches and the control cabin (Fig. 7). In the first optics hutch for the beam shutter, graphite filters, a water-cooled slit, and a beam monitor will be installed. In the main optics hutch, two mirror benders for X-ray collimation and focusing and a water-cooled double crystal monochromator (DCM) have been installed (Fig. 8). Furthermore, an X-Y wire monitor, water-cooled slits (Slit 2), slits for the monochromated X-rays (Slit 3), and a downstream shutter will be installed during FY2003. In the experimental hutch, the combination of a diffractometer with a high precision sample spindle axis (the maximum sphere of confusion: 1 µm), a high

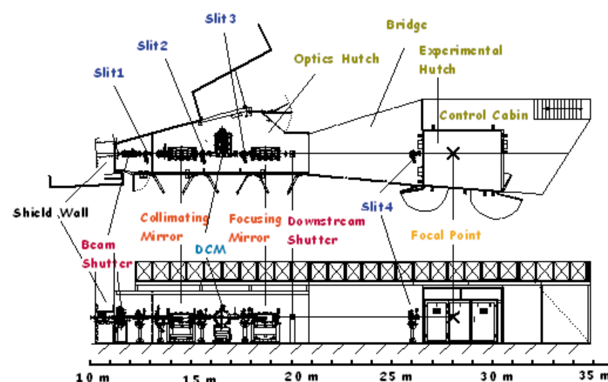

Figure 6
Layout of the multipole wiggler MAD beamline BL-5.

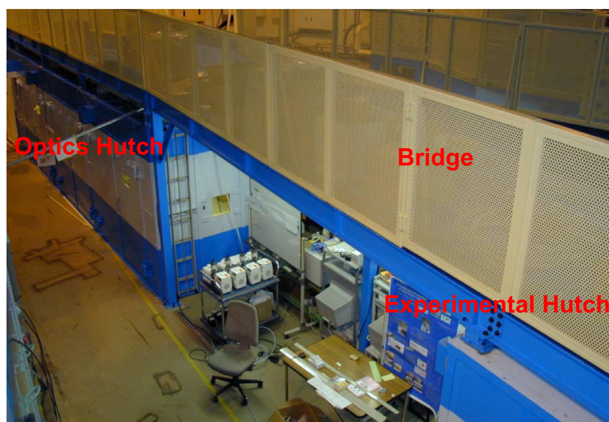

Figure 7
The exterior of BL-5.

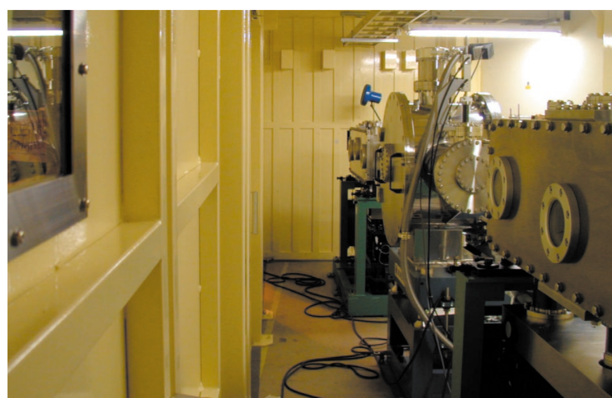
Figure 8
DCM (center) and two mirror benders in the optics hutch.

speed shutter (minimum exposure: 10 ms), and a liquid nitrogen sample cooling system will ensure reliable data collection from extremely small protein crystals. A large size detector (315 mm × 315 mm, the largest CCD detector installed in Japan) with high speed read-out (1 s in full resolution mode) will enable very efficient data collection (10 to 30 minutes per data set) to the highest resolution. Appropriate sample manipulation space with incubators and computational facilities including gigabit network to a collaboratory server will be provided in the control cabin for efficient data collection and data analyses. Commissioning will be started in the Autumn run of 2003 and the beamline will become available for general user experiments in FY2004.

"Protein 3000" Beam Time

The PF Structural Biology Group submitted an S2 beam time proposal 2003S2-002 on target-oriented structural genomics in the "Protein 3000" project on behalf of the 8 university consortia. The proposal was approved with a high score, and preparation was made for its official start in April 2003. This beam time proposal will receive up to 30% of the beam time available from the PF structural biology beamlines, (BL-6A, BL-18B, AR-NW12 and BL-5 when completed). This amounts to 120 - 160 days of beam time per year. A web reservation system is being developed for team members to reserve beam time as late as one day before the beam time.

4-4 High-Throughput Technologies

Large Scale Protein Crystallization Robot

X-ray crystallographic techniques are very well suited for studying complex structures at atomic resolution as long as one can obtain high quality protein crystals. Crystallization of proteins remains a labor-intensive step. One might need to set up thousands of crystallization trials with varied conditions for protein solutions and precipitants before obtaining X-ray quality crystals. The process benefits enormously from automation using robots. We are developing an extremely large-scale protein crystallization robot which can prepare 200,000 crystallization trials per day (Figs. 9 and 10). The robot is based on a parallel approach; dispensing will be carried out for 96 wells simultaneously using disposable chips and crystallization trays which conform to conventional formats but have been designed originally for this system for maximum accuracy and reasonable running costs. We hope that it will crystallize a large number of structures at an unprecedented pace. Based on initial experiences with the crystallization robot, we will also develop much more efficient crystallization kits to overcome the difficulties of crystallizing very large, multi-domain proteins or their complexes and membrane proteins. The system is expected to become operational in Autumn 2003.

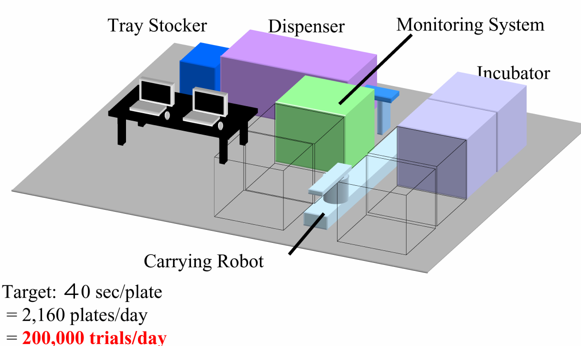

Figure 9
Schematic view of the protein crystallization system.

Figure 10
The protein crystallization robot as installed in the extended Structural Biology Building.

Harvesting of Protein Crystals

We have developed a system for harvesting protein crystals using a micromanipulator for high-throughput protein crystallographic experiments at the beamlines. It is part of a suit of fully automated robotics systems being developed at the PF, including systems for protein crystallization, storage, monitoring of crystal growth, harvesting and freezing crystals, mounting inside a hutch and data acquisition. One of the bottlenecks of these processes is the automatic harvesting of protein crystals from crystallization trays. Here we propose a novel method to pick up crystals using glass chopsticks controlled by a parallel set of piezoelectric materials and a motorized cryo-loop (Fig. 11). A prototype of the protein crystal harvesting system has been developed and the results of the initial experiments to harvest crystals are very encouraging (Fig. 12).

Integrated Controlling System and Unified Database using STARS

STARS (Simple Transmission And Retrieval System), originally developed for use in beamline interlock systems by T. Kosuge of the PF, has been chosen for the control system of the structural biology beamlines (Fig. 13). Towards the realization of fully-automated data collection, we first developed a user-friendly experimental set-up, a "fancy-box", using STARS for beamline AR-NW12. It consists of a beam stopper, the fluorescence detector and the CCD detector shield. A fully automated crystal alignment system is being developed and a sample exchange robot will be installed later in FY2003. These devices will be controlled by a suite of programs which communicate with each other via the STARS server.

To take best advantage of STARS, the computer environments at BL-6A and BL-18B were improved for stability and speed. For example, a Giga-bit network system was introduced and a high-speed and large-capacity RAID hard disk system was installed. The time for data backup was also made substantially shorter: downloading 180 oscillation images takes only several minutes.

As a next step, we have developed an integrated controlling system and a unified database for high throughput protein crystallography experiments. All of the main features of protein crystallography experiments (purification, crystallization, crystal harvesting, data collection, data processing) have been integrated into the software. All

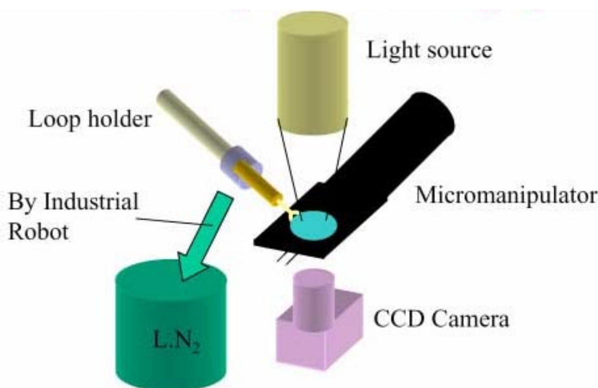

Figure 11
Schematic view of the protein crystal harvesting system.

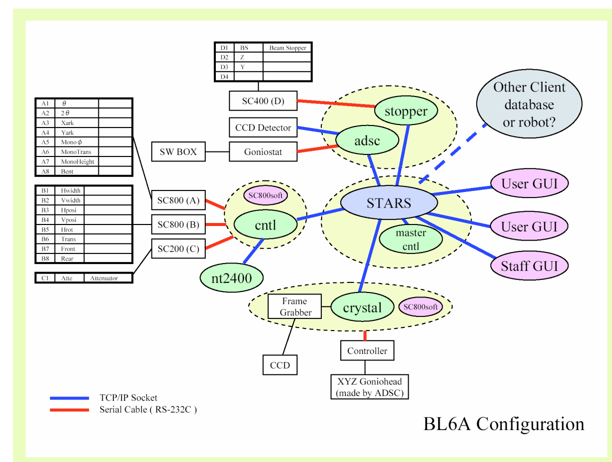

Figure 13
BL-6A control software architecture using STARS.

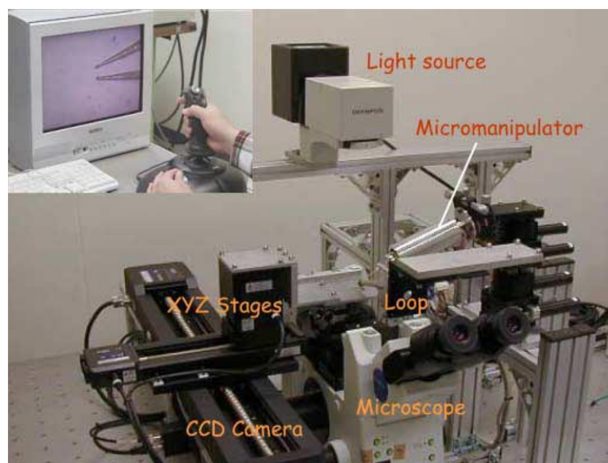

Figure 12
Arrangement of the protein crystal harvesting system and the micromanipulator (inset).

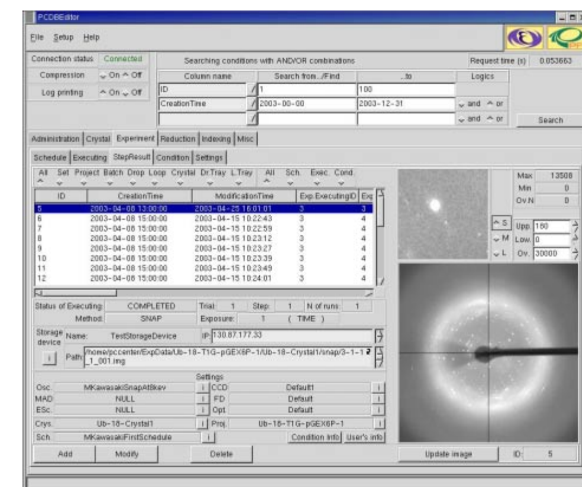

Figure 14
PCCTools GUI.

of the information necessary to perform protein crystallography experiments is stored (except for the raw X-ray data which is stored in a central data server) in a MySQL relational database. The database contains four mutually linked hierarchical trees describing the protein crystals, the data collection and experimental data processing. A database editor was designed and developed. The editor supports basic database functions to view, create, modify and delete user records in the database (Fig. 14). Two search engines were realized: a direct search of necessary information in the database and an object oriented search. The system is based on TCP/IP secure UNIX sockets with four predefined behaviors for sending and receiving commands and replies, which support communications between all connected servers and clients with remote control functions (creating and modifying data for experimental conditions, data acquisition, viewing experimental data, and performing data processing). Two secure login schemes were designed and developed: a direct method (using locally developed Linux clients with secure connection) and an indirect method (using secure SSL connection with secure X11 support from any operating system with X-terminal and SSH support). Part of the system has been implemented on the new MAD beamline, AR-NW12 at the PF-AR for general user experiments.

Development of a New X-Ray Area Detector, X-HARP

With the advent of high brilliance protein crystallography beamlines at synchrotron sources, large, fast and yet affordable X-ray detectors are needed. The next-generation synchrotron sources (XFEL's and ERL's) will demand detectors with even higher performance.

The structural biology group has started R & D study of a new area detector for next-generation structural biology beamlines in collaboration with S. Kishimoto and K. Hyodo (KEK-PF), R. Mochizuki (NHK Engineering Services) and K. Tanioka (NHK Science & Technical Research Laboratories). We have adopted the technology with an amorphous photoconductive layer of selenium operating in the avalanche-mode developed by NHK. In contrast to other amorphous selenium-based solid state area detectors, the NHK group uses an exceptionally high voltage across the selenium membrane, typically 100 V/μm to take advantage of the avalanche effect (Fig. 15). Thus the detector is named "HARP": High-gain Avalanche Rushing amorphous Photoconductor. The final goal is to produce an X-ray area detector with a large area (40 cm × 40 cm or larger), tens of millions of pixels with fast readout in the range of 30 to 90 frames/s.

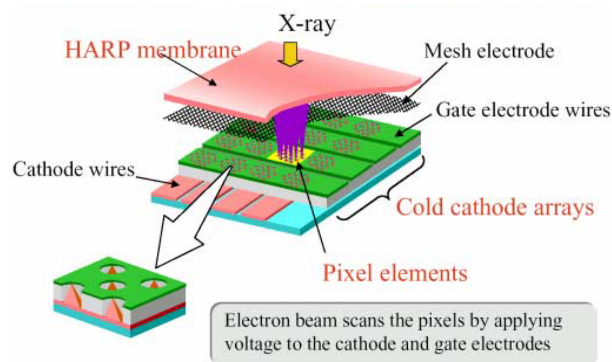

Figure 15
Schematic view of the cold cathode HARP area detector.

Figure 16
X-ray HARP Prototype with a preamplifier (Mochizuki and Tanioka, NHK).

In FY2002, we developed a prototype detector with 10 mm^2 active area with full speed read out (Fig. 16), and tested it on AR-NW2. The spatial resolution was proven to be satisfactory although inactive areas were abundant due to the surface roughness of the HARP membrane. The surface problem is tackled by the NHK group by changing the voltage profile within the HARP membrane. Field emitter array technology will be brought in for the next generation prototype with the intention to make the active area compatible with protein crystallographic experiments. With further development, we hope to produce an area detector useful not only for the synchrotron protein crystallography experiments of today, but also for structural biology and other fields of research using the next-generation synchrotron radiation.