5.7 Biological Science I

Part-A: Protein X-ray Crystallography Beamlines and Users' Results

1. Overview of the protein X-ray crystallography beamlines

We completed the construction of two insertion device (ID) high-throughput MAD beamlines for protein crystallography, AR-NW12A and BL-5A, in FY2003 and FY2004, respectively. In addition, we continued operating two bending magnet (BM) beamlines, BL-6A and BL-18B. Owing to the high efficiency of data collection and high demand for beam time, we began to assign beam time in half-day (9-hr daytime, 15-hr overnight) units on the ID beamlines in the 3rd period of FY2004. This allowed for more efficient beam time assignment and timely access to the synchrotron for users. At the end of February 2005, we stopped operation of BL-18B and closed it in order to construct a new ID beamline BL-17A, specifically designed for measurements of small size crystals and low-energy experiments. The newly developed ID beamlines (AR-NW12A, BL-5A and the under-construction BL-17A) have the following common features; (1) high-speed data acquisition using CCD detectors, (2) fast and reliable X-ray energy tunability using double-crystal monochromators (DCMs), (3) extremely precise sample rotation axes, and (4) motorized stages in the experimental stations. The updated specifications of the beamlines are summarized in Table 1.

To operate all the beamlines efficiently, a network-based beamline control system has been developed, providing not only a common user interface but also a function to enable experiments to be carried out remotely using secure TCP/IP communication. As part of the system, software using a relational database is being developed for the storage of all the necessary information related to the experiments. Together with the sample exchange robots installed on the ID beamlines, fully automated experiments will become possible. Detailed descriptions of the beamlines and software are provided in the following sections.

Experimental proposals from researchers mainly at universities and research institutes are reviewed irrespective of nationality by the PF Program Advisory Committee (PF-PAC) and approved by the Advisory Councils for Scientific Policy and Managements. The numbers of accepted proposals for the period FY2001-2005 and the first half of FY2006 are shown

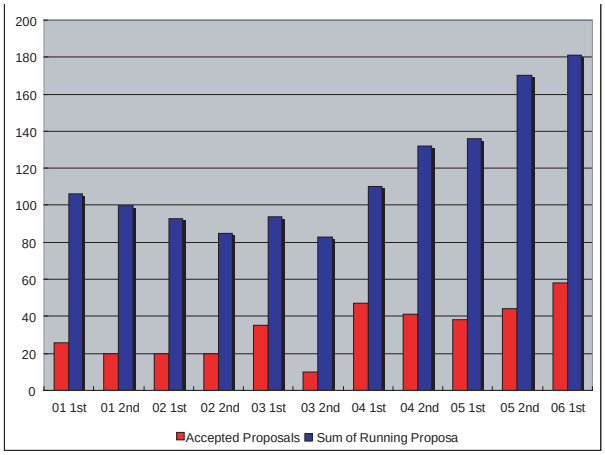

Figure 1
Statistics of experimental proposals.

Table 1 Structural Biology Beamlines at the Photon Factory.

	BL-6A	AR-NW12A	BL-5A	BL-17A
Start of operations	1987	2003	2004	2006(planned)
Synchrotron ring	PF	PF-AR	PF	PF
Injection	once a day (9:00)	twice a day (10:00, 21;00)	once a day (9:00)	once a day (9:00)
X-ray source	Bending Magnet	Undulator	Multi Pole Wiggler	Short Gap Undulator
Wavelength range (Å)	0.91-1.33	0.7-1.9	0.7-1.9	0.95-1.1 1.4-2.0
Energy resolution (ΔE/E)	1×10^{-3}	2×10^{-4}	2×10^{-4}	3×10^{-4}
Photon flux (photons/sec @ 1.0 Å)	1.0×10^{10}	2×10^{11}	2×10^{11}	$> 10^{10}$(@1.0 Å), $> 10^{11}$(@2.0 Å)
Slit size (mm)	0.1	0.2	0.2	0.02
Detector	Quantum 4R	Quantum 210	Quantum 315	not decided
Type	CCD	CCD	CCD	-
Active area (mm^2)	188 × 188	210 × 210	315 × 315	-
Pixel size (μm^2)	81.6 × 81.6	51 × 51	51 × 51	(The specifications of BL-17A are estimated values.)
Pixel number	2304 × 2304	4096 × 4096	6144 × 6144	
Frame data size (MB)	11	34	75	
Readout time (sec)	8	1	1	
Typical exposure time (1.0° oscillation)	30 sec	5 sec	5 sec	-
Typical data collection time (180 frames)	2.0 hr	20 min	20 min	-
Backup time for 1 data set (180 frames) (using the IEEE1394 interface)	3 min	6 min	6 min	-
Camera distance (mm)	60 ~ 400	60 ~ 1000	60 ~ 1000	-
Detector vertical offset	0 ~ 25 deg	0 ~ 100 mm	0 ~ 180 mm	-
Software for image processing	HKL2000, DPS/mosflm			

as red bars in Fig. 1. A running sum of the number of current (effective) experimental proposals is also shown in Fig. 1 as blue bars. The number of effective proposals decreased gradually until FY2003, after which an upward trend continued through to FY2006. Proposals rejected by PF-PAC are excluded from the numbers shown in Fig. 1.

2. High-throughput ID beamlines, BL-5A and AR-NW12A

Beamlines BL-5A and AR-NW12A have been stable and productive since their opening for user operation, and the beamlines are being improved day by day. For example we have installed one more set of four-blade slits immediately upstream of the sample position at AR-NW12A, forming part of a double-slit system along with the existing four-blade slits. This allows users to easily change the beam divergence, useful for studies of crystals with large cell dimensions. Another important improvement at AR-NW12A is the use of the 1st harmonic of the undulator beam for low-energy experiments (wavelengths from 1.5 Å to 2.0 Å). This decreases the heat load on the optical elements drastically compared to using the 3rd harmonic with a short undulator gap.

The micro-channel crystal in the DCM at BL-5A caused problems several times as the cooling efficiency gradually decreased during operation, with the beam intensity at the sample position eventually decreasing to about half of the normal intensity. It was neccessary to replace the crystal after each operation period, resulting in the destruction loss of two precious crystals. In addition, there was trouble with the AR-NW12A X-ray CCD detector (ADSC Quantum 210) at the beginning of June 2004. Several stripe-like noise lines were frequently observed in one quadrant of the read-out images, although the level of the noise was not serious. This problem was fixed during the summer shutdown and has not reoccurred.

3. Bending magnet beamlines, BL-6A and BL-18B

BL-6A is a conventional beamline for protein crystallography using a BM as a light source. It is the oldest protein crystallography beamline at the PF, and has been operational since 1987. In FY1999, we began a refurbishment program for BL-6A, which was completed at the beginning of FY2004. Currently, BL-6A functions as a modern beamline with a CCD detector data acquisition system and a high-precision optical bench. In 2005, the following further upgrades took place. (1) Installation of the "FancyBox" interface already in use at beamlines AR-NW12A, BL-5A and the under-construction BL-17A. (2) Development of a compact high-precision sample rotation axis with a spherical confusion of a few microns. (3) Development of a compact high-speed X-ray shutter. (4) Installation of the new graphical user interface already in use at the ID beamlines (see below). With these improvements, all four beamlines will have the same architecture and look-and-feel.

BL-18B was closed on the morning of February 28, 2005. Initially built in 1993 as a beamline for anomalous dispersion work using a rapidly tunable DCM and the millisecond time-resolved Laue method with a bent-cylindrical focusing mirror, it has sincecontributed significantly to the progress of structural biology as a pioneer beamline for MAD experiments in Japan. After over 10 years of operation, its mission has been completed, producing many experimental results leading to numerous publications.

4. Construction of a new ID beamline, BL-17A

At the end of February 2005, we started to construct a new protein crystallography beamline, BL-17A. The existing BL-17A, B and C will be moved to a new BL-18B. The light source of BL-17A is a mini-pole (mini-gap) undulator which will be installed in the 2.5 GeV PF ring after the improvement of the straight sections. The new beamline is designed for measurements of very small protein crystals and for low-energy experiments. In FY2004, the construction of two optics hutches and the deck were completed. The first beam was delivered in December 2005, and operation for general users will begin in FY2006.

5. Development and improvement of the beamline control software, graphical user interface and experimental environment

1. Beamline control software

A unified graphical user interface (GUI) for beamline control has been newly developed and made available for public use in FY2004 (Fig. 2). Previously, the GUI for each experimental operation (for example XAFS and sample alignment) had been developed separately. Although such concurrent development is one of the advantages of our STARS system, we have unified the three GUIs for crystal centering, XAFS measurements and diffraction image collection in order to prevent users from making mistakes during beamline operations. The new interface provides users with a smooth and intuitive operating environment. For instance, it is possible to extract the wavelength values for MAD data collection automatically from the XAFS measurement. The unified GUI will be installed at all the structural biology beamlines, so that users can carry out their experiments in the same manner at each beamline. In addition, automatic loop centering software and a control module for the crystal exchange robot are under development. In the near future, these functions will be also be installed in the system, to achieve a higher level of user-friendliness and semi-automated data collection.

We are also developing an integrated control system based on the unified database PCCS (Protein Crystallography experiment Control System, Fig. 2). This system is designed to manage all experimental information, including structural biology research activities such as over-expression, purification and crystallization. In addition, the system allows fully automated measurements

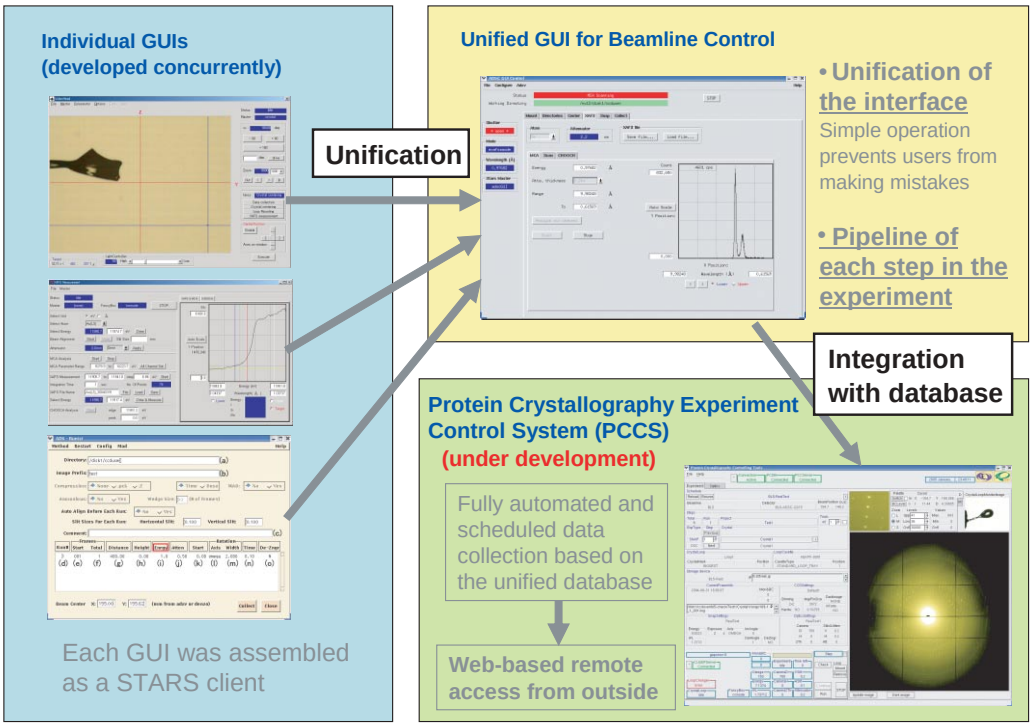

Figure 2
Schematic representation of the development of the beamline control software, graphical user interface, unified database system and overall integration.

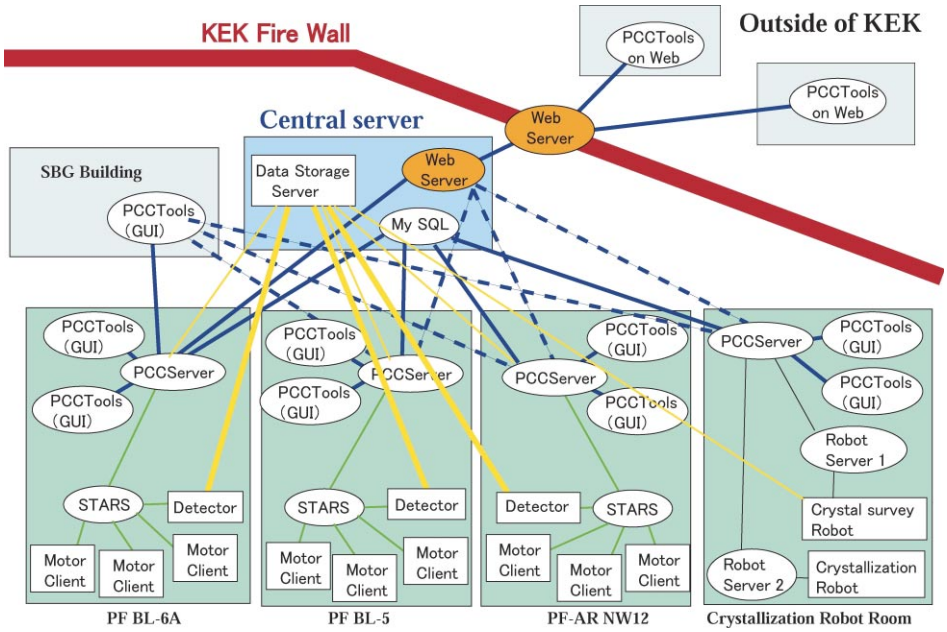

Figure 3
Schematic diagram of the information network.

and multiple access from various places at the PF. We began to commission and debug the beamline part of the software at the end of FY2004. In the future, PCCS will be accessible from outside the PF via a web–server, allowing users to collect and process data from their laboratories. The system is developed on a central server connected with the beamlines and the experimental facilities with a high-speed optical fiber network.

2. Central server and high-speed network

A new central server system with high-speed network was installed in FY2004 (Fig. 3). It connects the PF structural biology beamlines and the various experimental equipment in the associated laboratories via a Gbit network. The central server is a multi-CPU machine with redundant interfaces. The server machine is an SGI Altix 3700, which has 16 Itanium2 processors. Each processor has a clock speed of 1.3 GHz, a 3-MB L3 cache and a peak performance of 5.2 Gflops. A total of 32-GB of globally shared memory is available. A single SGI Linux operating system controls the machine. This presents an environment resembling a very

large Linux workstation for users. An SGI TP9100 RAID storage system provides a 9-TB RAID5 hard disk, which is connected by two 2-Gbit fibre-channel storage area network to the server. This system is easily expandable; the architecture supports as many as 512 processors in the same system. We can increase the CPU, the capacity of RAID hard disks and the interface without any change in the operating system itself. A large amount of image data can be sent with very high speed (max 77 MB/s) to the server from the beamlines. This very large bandwidth is indispensable for the delay- and trouble-free transmission of such a large amount of data.

An intelligent multilayer modular switch (Cisco Catalyst 6503) was installed as the gateway for the LAN, improving the network performance to a level ready for future remote access from sites outside the PF. LDAP/NFS was also installed to consolidate user accounts and experimental data. Users can access their data under the same environment from any PCs connected to the LAN.

3. Improvement of experimental environment

A new room for data processing and analysis was prepared in the PF-AR northwest building, financially supported by the Protein 3000 project (described later). Users who have finished their beam time can use the room to process, analyze and backup their data efficiently, under the same environment as at the beamlines. A similar workspace will be installed on the deck of the new beamline, BL-17A.

6. Research highlights from the users

1. Structural Basis of the Allosteric Activation of HutP-mediated Anti-termination of Transcription

Bacteria exploit a variety of mechanisms to regulate transcription elongation, in order to control gene expression in response to changes in their environment. Among these mechanisms, a common path is the modulation of mRNA secondary structures by RNA-binding proteins, either to pause the transcription near the terminator region or to allow synthesis of the full-length transcript. One such anti-terminator protein is HutP, Histidine utilizing Protein, of *Bacillus subtilis*, which is responsible for the regulation of the expression of the hut structural genes of the organism in response to changes in intracellular levels of L-histidine. In the *hut* operon, *hutP* is located just downstream from the promoter, while the five subsequent structural genes *hutH*, *hutU*, *hutI*, *hutG* and *hutM* are positioned further downstream from the promoter. In the presence of L-histidine, HutP binds to the nascent *hut* mRNA leader transcript. This allows the anti-terminator to form, thereby preventing the formation of the terminator and permitting transcriptional read through into the *hut* structural genes. In the absence of L-histidine, HutP does not bind to the *hut* mRNA, thus allowing the formation of a stem loop terminator structure within the nucleotide sequence located in between the *hutP* and structural genes.

In order to understand the structure of HutP, we solved the crystal structure of the HutP, revealing a novel fold where three dimers are arranged in a 3-fold axis to form a hexamer [1]. We also identified a minimal RNA binding element sufficient for HutP binding: three UAG trinucleotide motifs each separated by 4 nucleotides and located just upstream of the terminator, spanning the region between +496 and +515 nucleotides [1]. On the basis of in vitro selections and site-specifi c mutational analyses, we also identified UAGNNNNUAGNNNNUAG as the recognition motif, in which N indicates any base and the important chemical groups of the bases for HutP recognition within the core region of the UAG motif [2]. At present, the attenuation/anti-termination protein-RNA complex structures are available for only two proteins, including the full-length TRAP and the amino-terminal peptide fragment of the LicT. However, the activation of TRAP and the conformational changes before binding to the cognate RNA is remained unknown.

Our recently published study on HutP and their complexes, using data obtained at PF AR-NW12A, revealed not only the anti-termination complex structure of HutP but also the intermediate structures of HutP to explain allosteric activation of HutP before binding to the cognate RNA [3]. The quaternary complex structure shows how HutP specifically recognizes the conserved sequences within the *hut* mRNA, and reveals the unexpected direct role of the Mg^{2+} ion for mediating the L-histidine-dependent structural rearrangement in the protein (Fig. 4). The overall HutP protein forms a hexamer structure. The bound 21-mer RNA was recognized on both the top and bottom surfaces of the cylinder of the HutP in a previously unreported triangular conformation.

To unravel these structural changes in HutP, we have solved two additional crystal structures (uncomplexed HutP and HutP-L-histidine-Mg^{2+}) and found that

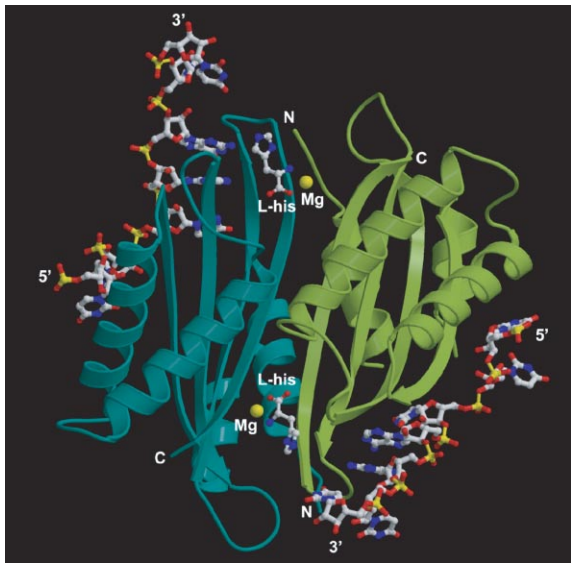

Figure 4
HutP dimer (blue and green) viewed along the two pseudo-two-fold axis. The L-histidine and RNA are shown as ball-and-stick models and the Mg^{2+} ions are represented by yellow spheres.

the Mg^{2+} ion coordinates with the L-histidine to facilitate an appropriate structural rearrangement before the recognition of its cognate RNA. In considering the various structural and biochemical studies on HutP, we have proposed a model for the structural rearrangement of HutP at each stage. From this evidence, it appears that once HutP has undergone this structural rearrangement it binds specifically to the RNA sequences within the terminator regions and wraps the terminator/anti-terminator region of hut mRNA around the protein; the RNA structure may subsequently reorganize and destabilize the terminator structure [3].

References

[1] T. Kumarevel, Z. Fujimoto, P. Karthe, M. Oda, H. Mizuno and P. K. R. Kumar, *Structure,* **12** (2004) 1269.

[2] T. Kumarevel, S. C. B. Gopinath, S. Nishikawa, H. Mizuno and P. K. R. Kumar, *Nucleic Acids Res.,* **32** (2004) 3904.

[3] T. Kumarevel, H. Mizuno and P. K. R. Kumar, *Nature,* **434** (2005) 183.

2. Structural basis for Ca^{2+}-induced activation of human PAD4

PAD (protein-arginine deiminase, protein L-arginine iminohydrolase) is a Ca^{2+}-dependent enzyme that catalyzes the conversion of protein arginine residues to citrulline residues. To date, five types of human PADs, designated 1, 2, 3, 4, and 6 have been characterized by cDNA cloning. PAD4 is the only type of PAD present in the cell nucleus and functions in the citrullination of histones H2A, H3, H4, and nucleophosmin/B23. PADs and citrullinated proteins are associated with human diseases such as rheumatoid arthritis (RA). RA is characterized by large numbers of antibodies directed against citrullinated proteins and produced by the recognition of protein citrulline residues as a major epitope of autoantigens. Recently, a significant association was reported between RA and functional variants of the gene that encodes PAD4 in the Japanese population. To gain insight into the Ca^{2+}-dependent molecular mechanism of protein citrullination by PADs, we determined the crystal structures of Ca^{2+}-free PAD4 and of a Ca2+-bound inactive mutant with and without an artificial substrate for PAD, benzoyl-L-arginine amide (BA), at resolutions of 2.8 Å, 2.3 Å, and 2.6 Å using diffraction data recorded at PF AR-NW12A, PF BL38B1 and SPring-8's BL45PX [4].

The PAD4 molecule has an elongated shape similar to a rubber boot (Fig. 5a), an approximate size of 125 × 45 × 50 Å, and is in close contact with another PAD molecule related by a crystallographic 2-fold axis to form a functional dimer (Fig. 5b). The polypeptide chain of the molecule is folded into two domains (N- and C-terminal domains). The N-terminal domain consists of the amino acid residues from Met1 to Pro300 and is further divided into two immunoglobulin-like sub-domains (sub-domains 1 and 2). A nuclear localization signal (^{56}PPAKKKST^{63}) is positioned on the molecular surface in sub-domain 1 and is disordered in all three structures. Functional mutations of the gene encoding PAD4 in the Japanese

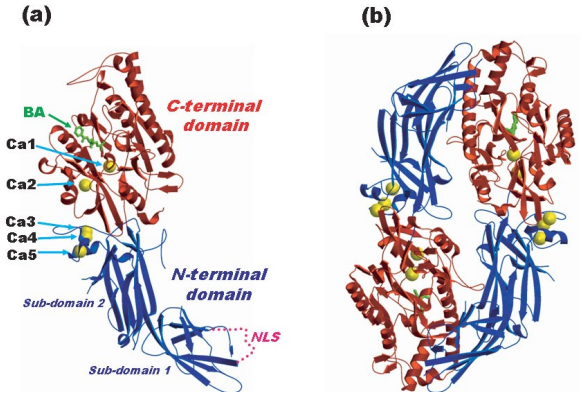

Figure 5
Overall structure of PAD4. (a) Ribbon representation of monomeric form of Ca^{2+}-bound inactive mutant with an artificial substrate, benzoyl-L-arginine amide (BA). (b) Dimeric form of the substrate complex. A crystallographic two fold axis runs vertically at the center of the dimer.

population for RA are also positioned in sub-domain 1, far from the active site in the C-terminal domain, and are unlikely to affect catalytic function directly.

Asn301 to Pro663 form the C-terminal domain with a structure of five ββαβ modules and contain the catalytic residues Asp350, His471, Asp473, and Cys645. The C-terminal domain of Ca^{2+}-bound mutant PAD4 is almost the same as that of the Ca^{2+}-bound mutant PAD4 with BA, indicating that substrate binding has no effect on the formation of the active site cleft. In contrast, large conformational differences are observed between the C-terminal domains of Ca^{2+}-free PAD4 and Ca^{2+}-bound mutant PAD4. Many disordered portions that form the acidic concave surface at the molecular surface of Ca^{2+}-free PAD4 are well ordered in the Ca^{2+}-bound mutant PAD4 and create an active site cleft upon Ca1 and Ca2 binding. The conformational changes that occur around the active site strongly suggest that binding of Ca1 and Ca2 to the acidic concave surface is crucial for substrate recognition. So far, such characteristic Ca^{2+}-induced generation of the active site cleft has not been observed in other Ca^{2+}-dependent enzymes such as calpain and transglutaminase 3, showing that human PAD4 is activated with a novel mechanism by Ca^{2+}.

References

[4] K. Arita, H. Hashimoto, T. Shimizu, K. Nakashima, M. Yamada and M. Sato, *Nature Struct. Mol. Biol.* **11** (2004) 777.

3. X-ray Crystallography of Kinesin Motors

Molecular motors are enzymes that convert energy liberated by chemical reactions into mechanical work. Kinesin superfamily proteins are a family of molecular motors most of which transport several organelles along their cellular track microtubules by using the energy of ATP hydrolysis [5]. Recently, some kinesins have been shown to destabilize the microtubule from both ends for reorganizing the cellular track or for mitosis [6]. To elucidate the molecular mechanisms of the translational movement along the microtubules and of the microtu-

bule destabilization we have solved the crystal structures of two types of kinesins: a transporter K1F1A [7] and a microtubule destabilizer KIF2C [8].

Crystal structures of the transporter KIF1A were solved in the ATP state (with ATP analog, AMPPNP) and in two ADP-phosphate states (transition states of ATP hydrolysis with two classical ADP-phosphate analogs, ADP-AlFx and ADP-vanadate), and were compared with the known structures in the ADP state [9]. These structures computationally docked with microtubules showed that KIF1A uses two microtubule-binding loops L11 and L12 in an alternating manner during ATP hydrolysis. In the ATP state, KIF1A extends loop L11 and is strongly bound to the microtubule. During ATP hydrolysis (ADP-phosphate states), KIF1A transiently raises both microtubule-binding loops (L11 and L12), actively detaching itself from the microtubules. After hydrolysis (ADP state), KIF1A extends another loop L12 to the microtubules. This L12 loop interacts with the flexible C-terminal region E-hook of the microtubule, and this flexible binding allows the diffusive movement of KIF1A along the microtubules. These structural results are consistent with the KIF1A movement directly visualized by single molecule analysis [10, 11]

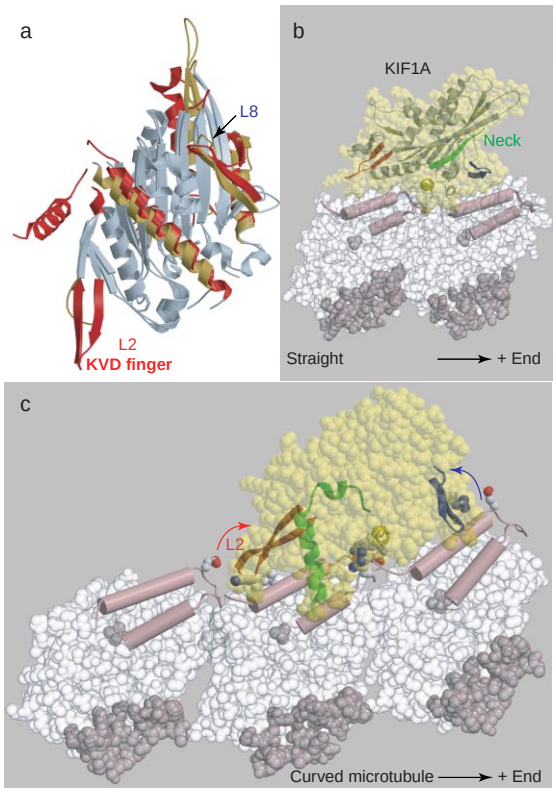

Figure 6
In silico modeling to dock the atomic model of KIF2C to that of the microtubule protofilament. (a) Comparison of the structure of KIF2C with that of KIF1A seen from the microtubule binding side. The overlapping structures are colored with gray. The highly divergent structural elements of KIF2C and KIF1A are colored with red and yellow, respectively. (b) Docking of KIF1A (yellow) to the straight microtubule protofilament (white). H11 and H12 of tubulin are shown as brown cylinders, and the M-loop is shown in dark-brown. (c) Docking of KIF2C (yellow) to the curved protofilament. The interface of KIF2C fits very well with curved protofilament and the loops L2 and L8 cooperatively stabilize the curved protofilament (red and blue allows).

The overall architecture of the microtubule destabilizer KIF2C was very similar to that of the transporter KIF1A (Fig. 6a). However, large structural differences were found in the following three class specific regions: the class-specific N-terminal neck, loop L2, and loop L8. Since all these regions directly face the microtubule, the microtubule-interface of KIF2C is very different from that of KIF1A; KIF1A fits very well with the straight protofilament of microtubule, whereas KIF2C fits with the curved protofilament of the microtubule (Fig. 6b,c). The microtubule protofilament naturally takes a straight conformation at its side wall and curved conformations at both ends. Thus, when KIF2C reaches the end of microtubule it makes full contact with the curved protofilament. The N-terminal neck region then inserts deeply into the interprotofilament groove, destabilizing the lateral interaction of the protofilament. Also, the L2 (KVD finger) and L8 loops cooperatively stabilize the curved protofilament at both ends of the microtubule. These effects are enough to shift the microtubule dynamics to depolymerization without further active processes.

Finally, we note that both KIF2C and KIF1A use the energy of ATP hydrolysis for detachment from the microtubule, enabling KIF2C to continuously destabilize the microtubule. This design principle of the active detachment from its effecter molecule is similar to that of G-proteins, protein kinases, and other nucleotidases. In other words, the energy of hydrolysis is used for the active detachment necessary to cycle reactions such as signal transduction, movement along the microtubule, or microtubule destabilization. This conserved strategy might reflect the evolutionary pathway of these classes of proteins.

References

[5] N. Hirokawa, *Science* **279** (1998) 519.
[6] N. Homma, Y. Takei, Y. Tanaka, T. Nakata, S. Terada, M. Kikkawa, Y. Noda and N. Hirokawa, *Cell* **114** (2003) 229.
[7] R. Nitta, M. Kikkawa, Y. Okada and N. Hirokawa, *Science* **305** (2004) 678.
[8] T. Ogawa, R. Nitta, Y. Okada and N. Hirokawa, *Cell* **116** (2004) 485.
[9] M. Kikkawa, E.P. Sablin, Y. Okada, H. Yajima, R.J. Fletterick and N. Hirokawa, *Nature* **411** (2001) 439.
[10] Y. Okada and N. Hirokawa, *Science* **283** (1999) 1152.
[11] Y. Okada, H. Higuchi and N. Hirokawa, *Nature* **424** (2003) 574.

4. Crystal Structure and Molecular Mechanism of Transcription Factor DksA

Stringent control, a complex of regulatory events in bacterial cells starved of amino acids, is triggered by elevated concentrations of guanosine-tetraphosphate (ppGpp), also as the "magic spot". Complexed with RNA polymerase (RNAP), this nucleotide selectively regulates the transcription of genes involved in amino acid metabolism. ppGpp both inhibits the transcription of rRNA and tRNA genes, and also stimulates the expression of proteins required for amino acid biosynthesis and transport. The overall effect of ppGpp action is thus to increase amino acid pools in the cell.

An intriguing and unresolved discrepancy exists between the small but reproducible ppGpp effects observed in highly purified *in vitro* systems and the dramatic range of regulation observed *in vivo*. This apparent discrepancy could be due to a requirement for cellular factor(s) that modulate ppGpp activity *in vivo* Ein fact, the existence of such an auxiliary factor was proposed nearly 30 years ago. Recently, the DksA protein was shown to greatly amplify the inhibition of rRNA transcription by ppGpp in vitro, and thus DksA may play the role of the missing *in vivo* modulator of ppGpp activities.

Recently, we solved the structure of the bacterial (*Escherichia coli*) DksA protein with 2.0 Å resolution using X-rays from the PF AR-NW12A beamline [12]. We show the surprising result that the DksA protein, while lacking any sequence similarity closely resembles in structure another well known transcription factor, GreA. Both structures contain a long α-helical coiled-coil domain with invariant acidic residues at the tip. It was proposed recently that upon binding to RNAP, GreA protrudes its coiled-coil domain deeply into the substrate entry (secondary) channel towards the RNAP active site, where its invariant acidic residues coordinate a catalytic Mg^{2+} ion.

We have already determined the RNAP/ppGpp complex structure [13], revealing that ppGpp binds in the RNAP secondary channel in close vicinity to the RNAP active site. The structure also identifies two Mg^{2+} ions bound to each di-phosphate in ppGpp. Whereas one of the ppGpp-bound Mg^{2+} ions is buried within the protein and is well fixed by the protein residues, the second Mg^{2+} ion is accessible from the outside through the RNAP secondary channel and is loosely bound by only ppGpp phosphates.

Given the previously demonstrated modulation of ppGpp activity by DksA and the structural similarity with GreA, the DksA structure implies a molecular mechanism in which DksA, like GreA, binds to RNAP, protrudes its coiled-coil domain through the secondary channel towards the ppGpp biding site, and stabilizes the RNAP/ppGpp complex through coordination of the loosely ppGpp-bound Mg^{2+} by invariant acidic residues (Fig. 7). To verify the proposed mechanism we have carried out a limited set of focused biochemical experiments that showed that DksA indeed directly binds to RNAP positioning the tip of its coiled-coil domain near the RNAP active site, and that invariant acidic residues are crucial to the DksA function - mutations of these residues resulted in loss of the DksA effect on the ppGpp activity. Thus the secondary channel emerges as a common regulatory entrance for transcription factors.

References

[12] A. Perederina, V. Svetlov, M. N. Vassylyeva, T. H. Tahirov, S. Yokoyama, I. Artsimovitcha and D. G. Vassylyev, *Cell*, **118**(2004) 297.

[13] I. Artsimovitch, V. Patlan, S. Sekine, M. N. Vassylyeva, T. Hosaka, K. Ochi, S. Yokoyama and D. G. Vassylyev, *Cell*, **117** (2004) 299

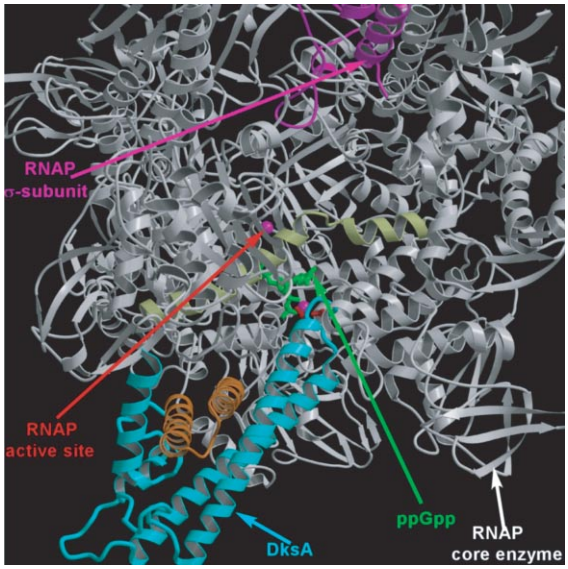

Figure 7
Overview of the model structure of the RNAP/DksA/ppGpp complex.

5. The Atomic Structure of Rice dwarf virus

Rice dwarf virus (RDV), a causal agent of rice dwarf disease, is a member of the genus *Phytoreovirus* in the family *Reoviridae*. It is transmitted to rice, wheat, barley, and other gramineae plants by insect vectors, with the main vectors being leafhoppers (*Nephotettix* species), after multiplication of the virus in the insect. Infection by RDV results in chlorotic specks on leaves and the stunting of plant growth. RDV is prevalent and it is one of the viruses that cause the most economical damage in China, Japan and other Asian countries. Each viral particle has an icosahedral shape approximately 700 Å in diameter [14], consisting of two concentric layers of proteins that encapsidate a genome of 12 discrete double-stranded (ds) RNAs. The core particle is composed of P1, a putative RNA polymerase; P5, a putative guanylyltransferase; and P7, a non-specific nucleic acid-binding protein. The core is encapsidated with a thin layer of P3 core capsid proteins. The outer layer of the virus is composed of mainly P8 proteins and a small number of P2 proteins which are required for vector transmission. Particles with only P8 in their outer layer can infect insect vectors via needle injection, while core particles with both P2 and P8 proteins can infect insect vectors by both needle infection and membrane feeding. The total molecular mass of a particle is about 70 million Dalton.

We have solved the atomic structure of RDV, determined at 3.5 Å resolution by X-ray crystallography [15]. The model consists of P3 inner capsid proteins, P8 outer capsid proteins, and fragments of P7, the nucleic acid binding protein (Fig. 8j). The atomic structure suggests a self-assembly mechanism for both homologous and heterologous capsid proteins. Insertion of the amino-terminal arm of one P3 protein into another P3 protein was found to be essential for the dimeric association of this core capsid protein. The interaction between two P3 proteins alters the conformation of the amino-terminal

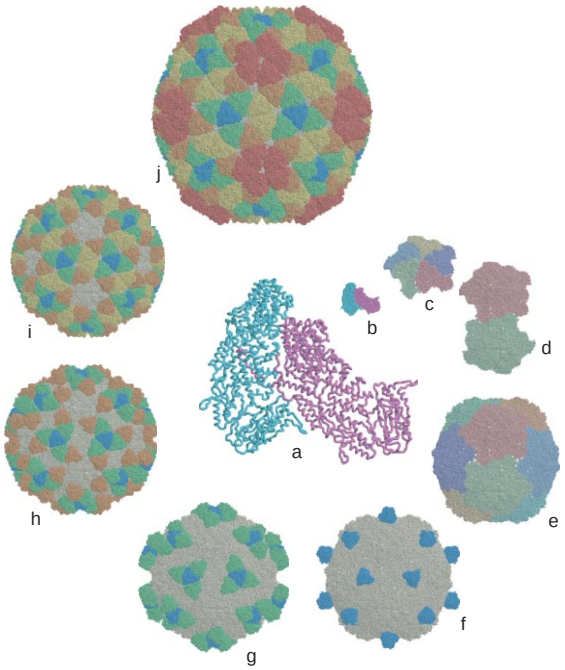

Figure 8
On the basis of the apparent interactions between the various proteins, the following sequence of RDV particle assembly is proposed. (a) Insertion of the amino-terminal arm of P3B into P3A initiates the assembly of a P3 dimer. (b) This P3A-P3B dimer acts as a unit piece in the jigsaw puzzle. (c) A pentameric structure of dimers of P3 protein forms around an icosahedral 5-fold axis, and then (d) this pentameric structure assembles (e) to form the core structure of an RDV particle. (f) The trimeric aggregate of P8 proteins acts as a unit and these trimers attach to the icosahedral three-fold axis at the T-site first. Orientation of the T-trimer on the surface of the core at the icosahedral 3-fold axis is defined by electrostatic complementarities. (g) R-trimers then attach via interactions with the inner shell and with the T-trimers. (h) Q-trimers and (i) S-trimers attach to the core surface and, at the final stage of viral assembly, (j) P-trimers attach at the icosahedral 5-fold axes to form the complete virus particle.

region of one of the other P3 proteins and a mere ten residues in one amino-terminal region block the extension of the amino-terminal loop of the other protein (Fig. 8a). This interaction triggers an overall structural change in the second P3 protein to show tight coupling of the two P3 molecules, and it appears that this dimeric structure initiates the assembly of the inner core of the virus. The interactions among P8 trimers involve side-by-side contacts among trimers. The electrostatic potential on the surface of each P8 trimer has obvious patches of positive and negative charge, allowing clear electrostatic complementarity with adjacent subunits. The interactions between the P3 core capsid protein and the P8 outer capsid protein involve predominantly hydrogen bonds and electrostatic interactions, with surface-charge complementarity. The electrostatic potential of the surface of P8 trimers has clear positively and negatively charged patches over the surface of each trimer, which exhibit complementarity in terms of electrostatic charges with adjacent P8 trimers, which can, thus, associate to form the outer capsid layer of the virus (Fig. 8f). A hierarchy of structural organization of this double-shelled virus is proposed on the basis of inter-subunit inter-atomic distances and the electrostatic surface potentials of the various subunits (Fig. 8).

References

[14] H. Mizuno, H. Kano, T. Omura, M. Koizumi, M. Kondoh, T. Tsukihara, *J. Mol. Biol.*, **219** (1991) 665-669.
[15] A. Nakagawa, N. Miyazaki, J. Taka, H. Naitow, A. Ogawa, Z. Fujimoto, H. Mizuno, T. Higashi, Y. Watanabe, T. Omura, R. H. Cheng, T. Tsukihara, *Structure*, **11** (2003) 1227-1238

6. Crystal Structure of the First Natural Diels-Alderase, Macrophomate Synthase

The Diels-Alder reaction is a cycloaddition whose mechanism involves the overlap of the π-orbitals of the two unsaturated systems in which an alkene (dienophile) adds to a 1,3-diene to form a 6-membered ring. The reaction is synthetically very useful because it forms cyclic products with high regio- and stereoselectivity under mild conditions. It has been applied to the synthesis of complex pharmaceutical and biologically active compounds. Catalytic methods with biomolecules such as RNA and protein antibody have also been developed. The reactions catalyzed by these biomolecules show remarkable enantio- and diastereoselectivity. Recently, natural Diels-Alderases such as solanapyrone synthase, lovastatin nonaketide synthase and macrophomate synthase (MPS) have been reported in the biosynthesis of secondary natural products. The function and catalytic mechanism of the natural Diels-Alderase are of great interest due to the diversity of molecular skeletons in natural Diels-Alder adducts. However, the details of catalysis of natural Diels-Alderases are still poorly understood.

The phytopathogenic fungus *Macrophoma commelinae*, isolated from spots on the leaves of *Commelina communis* has the ability to transform 2-pyrone derivatives into the corresponding benzoate analogues. This complex aromatic conversion is catalyzed by only one enzyme, macrophomate synthase (MPS), with oxalacetate as a substrate for the C3-unit precursor. MPS is a Mg^{2+}-dependent enzyme with 339 amino acid residues (Mw=36,244Da), the sequence of which showed no significant similarity with known proteins in a homology search. The catalytic mechanism of the whole pathway was investigated extensively, and it was shown that it proceeds through three separate steps including decarboxylation, two carbon-carbon bond formations, and decarboxylation with concomitant dehydration. In the absence of 2-pyrone, MPS simply acts as a decarboxylase with high catalytic efficiency. Furthermore, the involvement of a Diels-Alder reaction at the second step is proposed, based on the previously reported reaction type and the stereospecificity of the reaction. We present the first atomic resolution structure of natural Diels-Alderase [16] (Fig. 9).

The molecule is hexameric with point group symmetry 32. The protomer core region consists of 8 stranded β-barrel surrounded by 8+3 α-helices with a $(\beta/\alpha)_8$ barrel fold. The C-terminal α8-helix (residues 275-298) of each protomer protrudes from the core and joins to the β-bar-

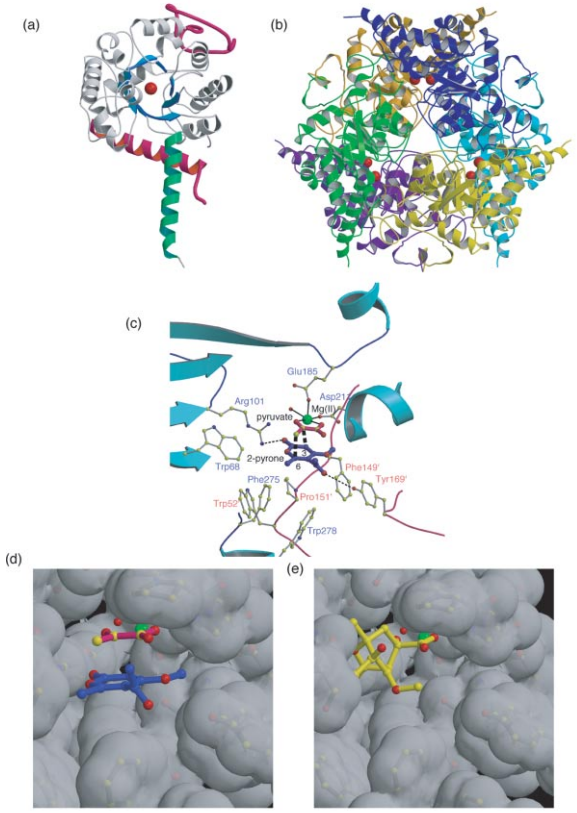

Figure 9
(a) Promoter structure of MPS showing the α-helix swapped $(\beta/\alpha)_8$ barrel fold. (b) The functional unit of MPS with point group symmetry 32. (c) The residues in the active site and proposed model for the very early transition state of the Diels-Alderreaction. (d) The space-filling model of the active site with transition state of substrates pyruvate enolate (red) and 2-pyrone 2 (blue) and (e) reaction intermediate (yellow).

rel of the 2-fold-related protomer. With these swapped helices two protomers are closely associated to form an extensively hydrophobic dimer interface.

In the second step of the reaction, the cycloaddition of the enolate and the 2-pyrone 2 takes place (Fig. 9). The steric congestion of the peptide backbone allows the 2-pyrone access only from one side of the enolate plane where the catalytic pocket is open. The proposed model for the very early transition state of the Diels-Alder reaction is as follows. In this binding model, two planes (2-pyrone and pyruvate enolate) are placed in parallel at π-orbital-overlapping distance. Several features are worth noticing in this model. First, the 2-pyrone molecule is likely to be fixed in place through two hydrogen bonds between the carbonyl oxygen of 2-pyrone and Arg101, and the C5-acyl oxygen and Tyr169. Tyr169 is in turn placed in proper orientation via stacking with Phe149. The flexible loop (residues 139-170) with hydrophobic side-chains (Phe149, Pro151 and Trp152) from the 3-fold-related protomer shields this transition state from the solvent. The stacking direction of 2-pyrone to pyruvate enolate is exactly the one expected from the product.

Both the R101S and Y169F mutants dramatically disturbed MPS activity while retaining decarboxylase activity, suggesting the importance of these hydrogen bonds in the carbon-carbon bond-forming reaction. The experimental result that 2-pyrones lacking a C5-acyl group are not converted into normal aromatic products gives further support for this binding structure. Generally speaking, the hydrogen-bonds between LUMO-energied substrate and some moieties in the reaction medium accelerate the Diels-Alder reaction. The intermediate is substantially reoriented from the early transition state with respect to the enzyme because of the conformational constraints imposed upon the adduct. The rather large hydrophobic cavity of this enzyme enables this rotation (reorientation) to occur without any steric congestion. The enzyme also has substantial van der Waals contacts to this intermediate. The first natural Diels-Alderase is found to adopt such several ingenious strategies.

References

[16] T. Ose, K. Watanabe, T. Mie, M. Honma, H. Watanabe, M. Yao, H. Oikawa, and I. Tanaka, *Nature* **422** (2003) 185.

7. Structure of Toc34, a Novel GTPase of the Chloroplast Translocon Complex

Protein transport is vital to the well-being of all living organisms. Newly synthesized proteins must be modified and transported to their correct destinations - such as outside of the cell, the nucleus, plasma membranes or various compartments knows as organelles. If proteins are mislocalised, they cannot function properly, leading to organelle malfunction and disease. Mitochondria and chloroplasts are examples of essential organelles; mitochondria produce ATPs which are used as the chemical fuel of the cell, and chloroplasts are the venue for photosynthesis. Both have their own DNA to synthesize core proteins, but also require many important proteins that are coded by the DNA of the cell's nucleus. They are synthesized in cytosol and incorporated into the organelles. The protein import machineries of the two organelles have many aspects in common, although those of mitochondria have been better characterized.

Toc34, Toc159, and Toc75 are the most well-characterized components of the Toc (translocon at the outer membrane of chloroplast) complex that associates with precursor proteins during protein transport across the chloroplast outer membrane. Toc34 and Toc159 are two homologous GTPases. Substantial biochemical data suggest that there is a direct link between the function of these GTPases and the import of proteins into the main channel of the Toc complex, Toc75. The key receptor of the complex, Toc159, is believed to interact via its highly acidic N-terminal domain with the predominantly positively-charged transit region of a protein destined to be transported into chloroplasts. The role of Toc34, however, in the protein import is unclear. It may either regulate the function of Toc75 by its GTPase molecular switching or function by interacting with the G-domain of Toc159.

The crystal structure of Toc34 [17] (Fig. 10) has

revealed a set of unique GTP binding (G) motifs that have not been observed in other GTPase structures reported to date. Excluding G1 (P-loop) of Toc34, which can be easily superimposed with that of the canonical GTPase, oncogenic p21 (Ras), the Toc34 and Ras G motifs do not share structural homology. For instance, the two strictly conserved residues of Ras, Thr35 and Asp119, which have counterparts which have been observed in all other GTPases, are absent in Toc34. Although Glu210 of Toc34 plays a similar role to that of Ras Asp119 by making two hydrogen bonds with the exocyclic guanine ring, no obvious candidate is found to act as Thr35 of Ras in Toc34, mainly due to the unavailability of the GTP-bound structure of Toc34. More strikingly, the presumed catalytic Gln61 of Ras, which directs a water molecule for nucleophilic attack on the oxygen of g-phosphate, is replaced by Leu97 in Toc34, suggesting a different mechanism for GTP hydrolysis in Toc34 to that of Ras.

More importantly, the 28 kDa of Toc34 molecules which lack the C-terminal 52 residues including the trans-membrane domain (residues 266-284) were found as dimers, having a large buried surface area at the interface of the crystal. Gel-filtration experiments also revealed that a third of the recombinant protein population existed as dimers in solution at the physiological condition, implying a specific role for the dimerization. A close inspection of the dimer interface demonstrates that Arg133 of each interacting monomer hydrogen-bonds to the b-phosphate oxygen of the GDP residing on the other monomer. When Arg128, one of the key residues involved in the dimerization was mutated to Ala, the population of the Toc34 dimer decreased to undetectable levels. The subsequent GTPase assay of the Arg128 mutant also showed a significantly reduced GTPase activity compared to that of the wild type. This suggests that the dimerization stimulates the two interacting GTPases. A similar result was obtained from the mutation of Arg133 to Lys, which resulted in a negligible reduction of the dimer population, but the mutant showed a pronounced reduction of the GTPase activity, comparable to that of the Arg128 mutant. The mutation of two arginines thus suggests that each interacting Toc34 monomer acts as a GTPase activating protein on the other one; and Arg133 is the "arginine finger".

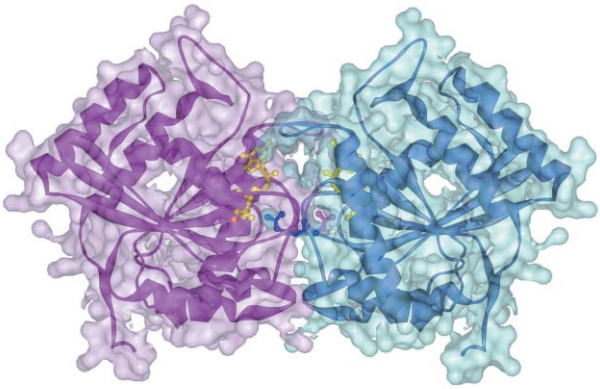

Figure 10
Molecular diagram of the Toc34 dimer. The two GDP molecules are depicted as yellow ball-and-stick models. Side chains of the two Arg133s are shown in ball-and-stick models.

References

[17] Y.-J. Sun, F. Forouhar, H.-M. Li, S.-L. Tu, Y.-H. Yeh, S. Kao, H.-L. Shr, C.-C. Chou, C. Chen, and C.-D. Hsiao, *Nature Struct. Biol.,* **9** (2002) 95.

8. Allosteric modulation of the RNA polymerase

Rifamycins are the clinically important antibiotics which target bacterial RNA polymerase (RNAP). The proposed mechanism in which rifamycins sterically block the extension of nascent RNA beyond three nucleotides does not alone explain why certain RNAP mutations confer resistance to some but not other rifamycins. Here we show that unlike rifampicin and rifapentin, and contradictory to the steric model, rifabutin inhibits formation of the first and second phosphodiester bonds. We report 2.5 Å resolution structures of rifabutin and rifapentin complexed with the *Thermus thermophilus* RNAP holoenzyme (Fig. 11). The structures reveal functionally important distinct interactions of antibiotics with the initiation sigma factor. Strikingly, both complexes lack the catalytic Mg^{2+} ion observed in the apo-holoenzyme, whereas an increase in Mg^{2+} concentration confers resistance to rifamycins. We propose that a rifamycin-induced signal is transmitted over approximately 19 Å to the RNAP active site to slow down catalysis. Based on structural predictions, we have designed enzyme substitutions that apparently interrupt this allosteric signal [18].

Streptolydigin (Stl) is a potent inhibitor of bacterial RNAPs. The 2.4 Å resolution structure of the *Thermus thermophilus* RNAP-Stl complex showed that, in full agreement with the available genetic data, the inhibitor binding site is located 20 Å away from the RNAP active site and encompasses the bridge helix and the trigger loop, two elements that are considered to be crucial for RNAP catalytic center function (Fig. 11). Structure-based biochemical experiments revealed additional determinants of Stl binding and demonstrated that Stl does not affect NTP substrate binding, DNA translocation, and phosphodiester bond formation. The RNAP-Stl complex structure, its comparison with the closely related substrate bound eukaryotic transcription elongation complexes, and biochemical analysis suggest an inhibitory mechanism in which Stl stabilizes catalytically inactive (preinsertion) substrate bound transcription intermediate, thereby blocking structural isomerization of RNAP to an active configuration. The results provide a basis for a design of new antibiotics utilizing the Stl-like mechanism [19].

Tagetitoxin (Tgt) inhibits transcription by an unknown mechanism. A structure at a resolution of 2.4 Å of the *Thermus thermophilus* RNAP–Tgt complex revealed that the Tgt-binding site within the RNAP secondary channel overlaps that of the stringent control effector ppGpp, which partially protects RNAP from Tgt inhibi-

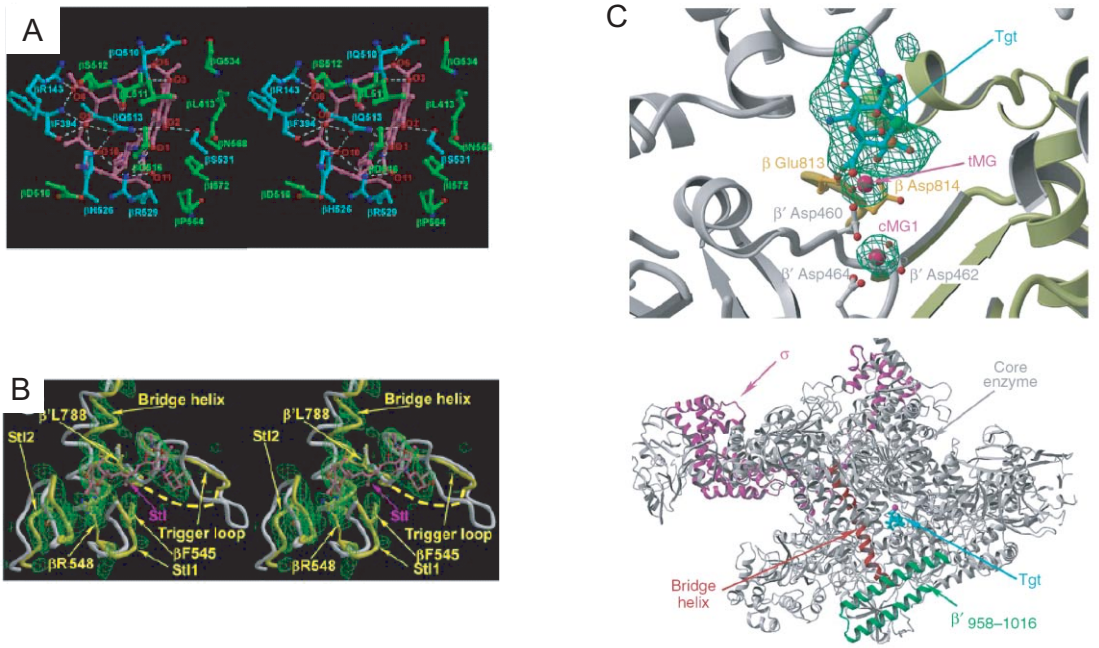

Figure 11
(A) RNAP/Rifs binding residues making polar (cyan) and hydrophobic (green) interactions with the Rifs' ansa ring are shown.
(B) The RNAP-Stl complex structure. The initial $|F_{stl} - F_{nati}|$ difference ED map (green, 3.5 Å resolution, contoured at 2.3 σ level), where F_{stl} and F_{nati} are the structure factor amplitudes of the RNAP-Stl complex and the apo-holoenzyme, respectively, is superimposed on the ribbon diagram of the RNAP structure (yellow) and Stl (pink balls-and-sticks model) in the complex; the structure of the apo-holoenzyme is shown in white. The protein side chains which moved upon complex formation in the Stl binding site, as revealed by the substantial ED in the difference ED map, are modeled as balls and sticks.
(C) Upper panel: experimental 3.5-Å resolution and $|F_{Tgt} - F_{nat}|$ (both 3.0-σ-level) omit electron density map (green) superimposed on the RNAP1–Tgt structure. The residues (balls and sticks) and protein backbone (ribbon diagram) of the β and β' subunits are shown in yellow and gray, respectively. Lower panel: overall view of the RNAP–Tgt complex structure showing that Tgt binds within the secondary channel in close vicinity to the active site (marked by cMG1, magenta sphere).

tion. Tgt binding is mediated exclusively through polar interactions with the β and β' residues whose substitutions confer resistance to Tgt *in vitro*. Importantly, a Tgt phosphate, together with two active site acidic residues, coordinates the third Mg^{2+} ion, which is distinct from the two catalytic metal ions. We show that Tgt inhibits all RNAP catalytic reactions and propose a mechanism in which the Tgt-bound Mg^{2+} ion has a key role in stabilization of an inactive transcription intermediate. Remodeling of the active site by metal ions could be a common theme in the regulation of catalysis by nucleic acid enzymes(Fig. 11) [20].

References

[18] Artsimovitch I, Vassylyeva MN, Svetlov D, Svetlov V, Perederina A, Igarashi N, Matsugaki N, Wakatsuki S, Tahirov TH, Vassylyev DG.. *Cell* **122** (2005) 351.
[19] Temiakov D, Zenkin N, Vassylyeva MN, Perederina A, Tahirov TH, Kashkina E, Savkina M, Zorov S, Nikiforov V, Igarashi N, Matsugaki N, Wakatsuki S, Severinov K, Vassylyev DG., *Mol Cell* **19** (2005) 655.
[20] Vassylyev DG, Svetlov V, Vassylyeva MN, Perederina A, Igarashi N, Matsugaki N, Wakatsuki S, Artsimovitch I., *Nature Struct. Mol. Biol.* **12** (2005) 1086.

Part-B: Research Activity of Structural Biology Research Center

1. Overview

The Structural Biology Research Group was formed in May 2000. The aims of the research group are user support of synchrotron radiation X-ray crystallography studies of macromolecules, highly advanced technical development and in-house structural biology research. The group has grown steadily during the last five years; the structural biology building was extended from 429 m^2 to 643 m^2, and the number of staff members increased from four (one professor and three research assistants) to about thirty, including the group leader Professor Soichi Wakatsuki, an associate professor (Dr. Ryuichi Kato) and five research associates (Drs. Noriyuki Igarashi, Naohiro Matsugaki, Masato Kawasaki, Masahiko Hiraki and Yusuke Yamada) as the core staff members. While about half of the members are primarily engaged in beamline operation and development and the remaining half in biological research, the synergy between the two activities is a unique aspect of this group. In conjunction with the recent change of status of KEK from a government institute to an agency in April 2004, the Structural Biology Research Group became the Structural Biology Research Center in May 2003. Three graduate students (attached to the Graduate University for Advanced Studies, SOKENDAI) including one from France are studying and carrying out their own

research projects relevant to the group's research field under the guidance of the staff members.

During FY2001-FY2003, the activities of the group were supported by a fund called "Special Coordination Funds for Promoting Science and Technology" awarded to the group together with the universities of Hokkaido, Tokyo, Kyoto and Osaka and the research institute of NHK (the Japanese national broadcasting corporation) by MEXT (Ministry of Education, Culture, Sports, Science and Technology). We built and commissioned a new high-throughput beamline BL-5A, developed an assortment of technologies for the automated handling of protein crystals, built a prototype of a next-generation two-dimensional X-ray "HARP" detector, and developed software which facilitates rapid and accurate structure determination. We also made improvements to the experimental environment of the beamlines and sample preparation laboratories using the same research fund. Subsequently a five year national project "Protein 3000" was begun by MEXT in FY2002. The project consists of two programs; a "Comprehensive Program" carried out by RIKEN, and "Individual Analysis Programs" carried out by eight consortia of universities and institutes including the KEK-PF Structural Biology Research Center. In addition, in FY2004 a new research and development program, "Development of Systems and Technology for Advanced Measurement and Analysis" was launched by JST (Japan Science and Technology Agency). We proposed a project to develop a next-generation detector coupled with a micro-focus beamline, which was selected and commenced in the same year for an initial 3 years. During this project we will develop a new beamline BL-17A, optimized for data collection from small crystals, and develop an advanced prototype of the next-generation two-dimensional X-ray "HARP" detector in collaboration with the research institute of NHK and associated companies.

Highlights of the R&D projects and biological research are described below to illustrate the synergistic approach of the group in pursuing in-house structural biology research, developing and improving fundamental research tools for synchrotron based protein crystallography, and operating the user facilities.

2. Protein 3000 Project - Individual Analysis Program

There have been many national and international post-genome projects started in the last few years. The Life Sciences Division of MEXT began a five-year project for determining 3,000 protein structures or unique folds in FY2002 (Fig. 12), and the structural genomics project at RIKEN will carry out the major part of this Protein 3000 project. Their primary goal is to contribute to the worldwide effort of determining all the representative structures in the structure space by determining 2,500 structures in the five years. On the other hand, structural biologists at Japanese universities and KEK-PF have been proposing a network of structural genomics consortia to pursue target-oriented structural genomics projects aimed at specific biological or medical targets. Each consortium consists of X-ray protein crystallography, NMR, and bioinformatics groups tightly coupled with those specialized in medical, pharmaceutical and biological sciences that share the same biological interests in their pursuit of structure-function relationships. By the end of October 2005 a total of 1,650 structures had been determined, with a total of 2,149 related publications. In total 2,200 structures will be determined during FY2005, with the remaining 800 completed by the end of the project. The Structural Biology Research Center has two missions in the Protein 3000 project. One is to promote our own original structural biology

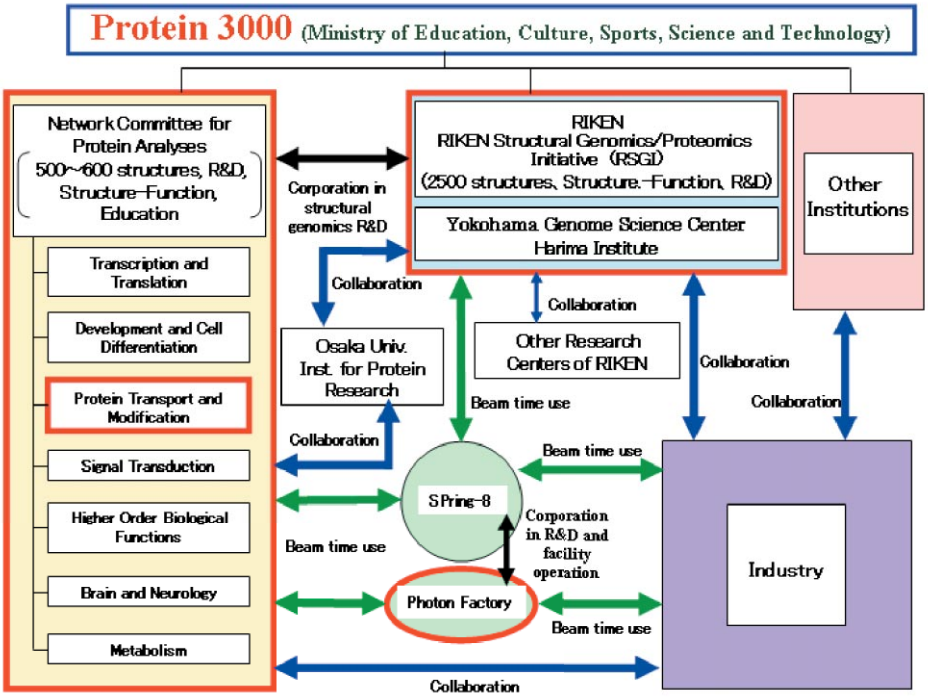

Figure 12
Schematic diagram of Protein 3000 project.

Table 2 Numbers of structures solved using PF beamlines by target-oriented structural genomics committee of the Protein 3000 project.

Team leader and organization	Targets	Number of structures solved
Masaru Tanokura, The Univ. of Tokyo, Graduate School of Agricultural and Life Sciences	Development and Cell Differentiation	78
Isao Tanaka, Hokkaido Univ., Graduate School of Science	Transcription and Translation	26
Yoshifumi Nishimura, Yokohama City Univ., Graduate School of Integrated Science	Transcription and Translation	30
Soichi Wakatsuki, Institute of Materials Structure Science, KEK	Protein Transport and Modification	114
Kunio Miki, Kyoto Univ., Graduate School of Science	Higher Order Biological Functions	29
Fuyuhiko Inagaki, Hokkaido Univ., Graduate School of Pharmaceutical Sciences	Signal Transduction	10
Atsushi Nakagawa, Prof. of Osaka Univ., Institute for Protein Research	Brain and Neurology	33
Seiki Kuramitsu, Osaka Univ., Graduate School of Sciences	Metabolism	67
Total		387

As of November 13, 2005

Table 3 Members of the "Posttranslational Modification and Transport" network of the Protein 3000.

Functional Analyses	Intracellular trafficking	Akihiko Nakano (RIKEN, Univ. of Tokyo), Kazuhisa Nakayama (Kyoto Univ. Pharmaceutical), Hiroshi Ohno (RIKEN Laboratory of Epithelial Immunobiology), Hiroaki Kato (Kyoto Univ. Pharmaceutical), Masayuki Murata (Univ. of Tokyo, Arts and Sciences), Syuya Fukai (Tokyo Inst. of Technology), Soichi Wakatsuki (KEK-PF)
	Post-translational modification	Toshisuke Kawasaki (Kyoto Univ. Pharmaceutical), Naoyuki Taniguchi (Osaka Univ. Medicine), Yoshifumi Jigami (AIST), Koichi Kato (Nagoya City Univ. Pharmaceutical), Sumihiro Hase (Osaka Univ. Science), Soichi Wakatsuki (KEK-PF)
Structural Analyses	X-ray crystallography	Takamasa Nonaka (Nagaoka Univ. of Technology), Nobutada Tanaka (Showa Univ. Pharmaceutical), Hiroaki Kato (Kyoto Univ. Pharmaceutical), Shuya Fukai (Tokyo Inst. of Technology), Soichi Wakatsuki (KEK-PF),
	NMR, Small angle X-ray scattering, Bioinformatics	Koichi Kato (Nagoya City Univ. Pharmaceutical), Mikio Kataoka (Nara Inst. of Science and Technology), Kei Yura (JAERI CCSE)

research, and the other is to support other structural researchers in the project when they use the PF protein crystallography beamlines. To this end, we have established an operation scheme to reserve about 30% of the beam time available at our beamlines for users of the eight university consortia of the project and developed a web-based beam time reservation system to facilitate this. As summarized in Table 2, many structures have been reported under this scheme.

The Structural Biology Research Center serves as the leading institute of one of the eight consortia of the Protein 3000 project, pursuing structural and functional analyses of protein transport and modification. Our consortium consists of nine universities and four research institutes (Table 3). Cell signaling and intracellular traf-

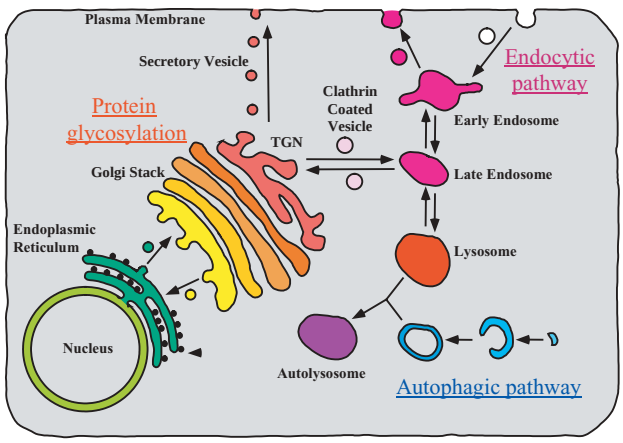

Figure 13
Network of protein glycosylation and transport in cells.

ficking are the means by which eukaryotic cells deliver cargo proteins to various organelles, cell membranes, and extracellular destinations (Fig. 13). The accurate distribution of these proteins is crucial for a range of cellular functions and activities. Mutations in the genes encoding protein transport regulators underlie a number of genetic diseases. Hence, an understanding of the biological and biomedical function of transport proteins is indispensable for making progress in treating human diseases. Furthermore, more than half of all human proteins undergo post-translational processing and modification such as glycosylation, acetylation, phosphorylation, geranyl-genylation, and farnesylation. In particular, glycobiology has recently gained prominence as a post-genomic science for its role in modulating protein function and transducing cellular signals. The recent interest in proteomics, the study of structure-function relationships of proteins, further increases the relevance of structural exploration in glycobiology and protein transport.

As part of the project, the Structural Biological Research Center has sought to maximize the efficiency of the large-scale expression, purification, and crystallization of proteins by trying to eliminate bottlenecks in each process. X-ray crystallographic studies using synchrotron X-ray radiation are conducted at KEK-PF, Kyoto University, Nagaoka University of Technology, Tokyo Institute of Technology and Showa University. NMR experiments are conducted at Nagoya City University, and small angle X-ray scattering studies at Nara Institute of Science and Technology (NAIST). From these institutes, Prof. Mikio Kataoka of NAIST and Dr. Shuya Fukai of the Tokyo Institute of Technology newly joined our consortium in FY2004. In FY2005, an additional three laboratories (Drs. Kenji Yamamoto, Tamao Endo and Akira Kurosaka) also joined. Each structural analysis project maintains close contacts with the groups responsible for functional analyses. Our initial research plan was to accomplish the structural and functional analyses of more than 70 proteins during the five years of the project. To date, we have completed structural analyses of 155 proteins and complexes and are currently working on about 50 additional targets. Furthermore, we have submitted eleven domestic and three international patent applications, based on our R&D and structural and functional studies of the target proteins.

As mentioned above, the post-translational modification of proteins is closely associated with intracellular transport systems, and new molecular interactions are being identified and characterized, thus providing new targets for our project. A logical extension of the current proteomics studies is the incorporation of the post-translational modification and transport machineries, in particular membrane complexes, which are key players in membrane trafficking. Our future research plan thus includes structural analyses of these complexes. With this in mind we will further develop the integrated and synergistic approaches within our consortia, and establish new research collaborations with groups in other research networks.

3. Development of Systems and Technology for Advanced Measurement and Analysis

Synchrotron X-ray crystallography is the most powerful technique available for determining the three-dimensional structures of biological molecules. The advancement of synchrotron facilities and experimental techniques over the past twenty years has had major impacts in biology and other life sciences. However, as scientists try to solve significantly more difficult and complex structures, it becomes harder to obtain crystals of sufficient sizes. Even with the latest 3rd generation synchrotron facilities such as SPring-8 (Japan), APS (US) and ESRF (EU), the structural analysis of protein crystals of a few microns or smaller is currently out of reach. While every effort has been made and more advanced techniques are being developed to push the limits of ring-based synchrotron facilities, the next (4th) generation synchrotron light sources such as the X-ray free electron laser (X-FEL) and energy recovery linac (ERL), are expected to produce X-ray beams with unprecedented qualities. The higher brilliance X-rays with shorter pulse-lengths available from such light sources may be useful for the structural analysis of single- or nanometer size crystals and various theoretical studies have been reported on the feasibility of such experiments. It will be crucial to develop next-generation two-dimensional X-ray detectors with high-speed readout and high sensitivity to match the characteristics of the next-generation structural analyses. Such detectors will be also useful for collecting weak diffractions from small crystals using micro-focus beamlines.

During FY2001-FY2003, we developed a prototype high-speed X-ray area detector based on a visible light HARP (High-gain Avalanche Rushing amorphous Photoconductor) camera in collaboration with NHK-ES (NHK Engineering Service) under the funding of MEXT. Following this program we are now improving the X-ray HARP detector, funded by a JST "Development of Systems and Technology for Advanced Measurement and Analysis" program (FY2004-FY2009, Fig. 14). The purpose of the new program is to develop a new measurement system for biological macromolecules based on a high-speed and high-sensitivity X-ray detector system coupled with a micro-focus beamline optimized for small crystals. The micro-focus beamline for the project, BL-17A, is being constructed at the 2.5 GeV PF-ring. The detector system is also currently being developed, and is described below. When a new prototype of the X-ray HARP detector is developed, it will be immediately be installed at BL-17A for a series of test experiments. Based on the experience gained from these studies the high-speed, high-resolution structural analysis system will be further improved. Upon completion, the system is expected to provide a solid base for developing X-ray detectors for the next generation of light sources. The

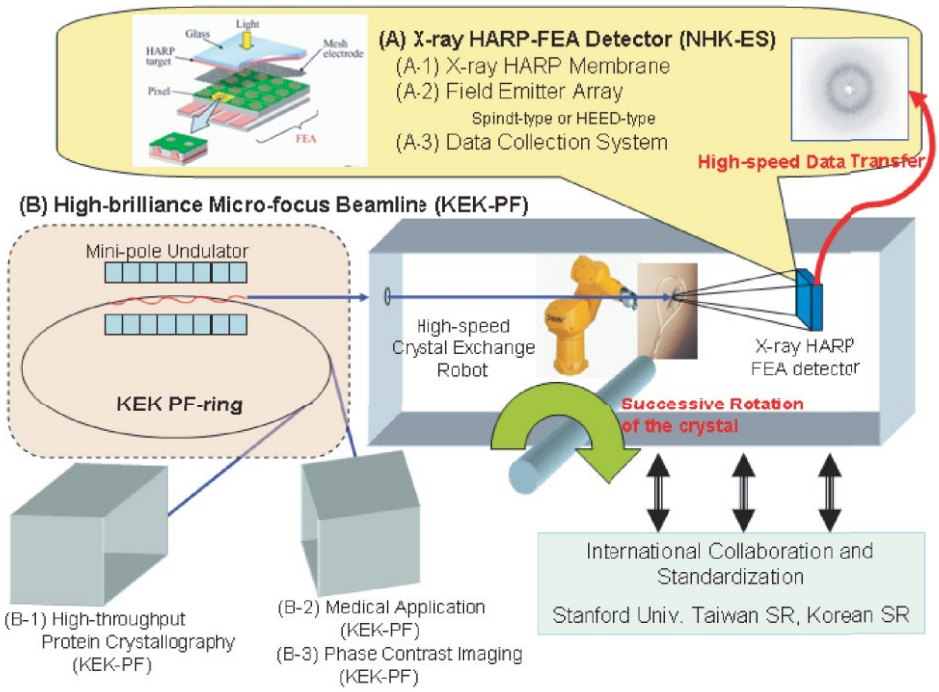

Figure 14
An outline of the "Development of Systems and Technology for Advanced Measurement and Analysis." program. In the project, we will mainly develop (A) an X-ray HARP-FEA detector in collaboration with NHK-ES and (B) a high-brilliance micro-focus beamline.

new detector will also be used for phase contrast imaging, medical imaging diagnosis, and many other applications which require high-speed and high-sensitivity X-ray imaging systems.

4. Structural Biology Research Highlights

1. The molecular mechanism of ubiquitinated-cargo recognition by GGA

Lysosome is an acidic compartment found in all animal cells, and contains digestive enzymes for degrading unnecessary proteins and glycolipids. Mannose-6-phosphate receptors (M6PRs) transport newly synthesized lysosomal enzymes from the *trans*-Golgi network (TGN) and the plasma membranes to endosomes. GGA (Golgi-localizing, γ-adaptin ear domain homology, ARF-binding) proteins mediate the sorting of these lysosomal cargo receptors at the TGN. The GGA protein is composed of four functional regions; VHS, GAT, hinge and GAE. Since the discovery of GGA in 2001, we have been pursuing structure-function analyses of each domain of GGA in complex with its partners (Fig. 15). Lysosomal cargo receptors interact with the VHS domain of GGA. We have reported the structure of GGA1-VHS in complex with the cytoplasmic tail of M6PR [21]. We have also reported the interaction mode between GGA1-VHS and β-secretase which cleaves the Alzheimer's amyloid precursor protein [22].

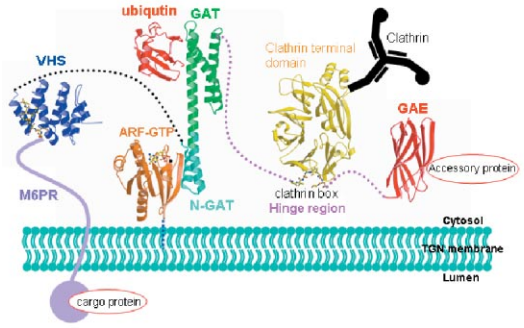

Figure 15
The interaction network between GGA and other proteins which are involved in intra-cellular protein transport.

The GAT domain of GGA consists of two subdomains; the amino (N) terminal (N-GAT) and the carboxyl terminal (C-GAT) domains. The N-GAT domain has a helix-loop-helix structure that is responsible for ADP-ribosylation factor (ARF) binding. In 2003, we reported the complex structure between N-GAT and ARF, explaining how GGA is recruited to the TGN membrane by ARF [23]. On the other hand, the binding partners of the C-GAT subdomain remained unknown until 2004 when ubiquitin was found to bind C-GAT. The ubiquitination of cell surface receptors serves as a signal for lysosome degradation after the internalization of receptors *via* endocytosis. Thus knowing that ubiquitin interacts with C-GAT uncovered a novel role for GGA in the ubiquitin-dependent sorting of cargo proteins in the endocytosis pathway. Quite recently, we solved the crystal structure of the complex between GGA3 C-GAT and ubiquitin [24].

In the GGA3 C-GAT/ubiquitin complex crystal, a hydrophobic patch of helices α1 and α2 of C-GAT interacts with the hydrophobic ubiquitin Ile44 surface, which is commonly used as a binding site by a variety of ubiquitin-binding modules. The ubiquitin Ile44 surface can be divided into three hydrophobic pockets. We compared the known structures of ubiquitin in complex with various ubiquitin-binding modules, and found that the three pockets of the Ile44 surface generally accommodate hydrophobic residues in all cases. We

also found that Arg42 of ubiquitin is a key residue for determining the shape and charge distribution of the Ile44 surface, facilitating interaction with the structurally divergent ubiquitin binding modules. In addition to the ubiquitin-binding site mentioned above, biochemical and NMR data on the C-GAT/ubiquitin interaction suggested that another hydrophobic patch of helices α2 and α3 of C-GAT also interacts with the ubiquitin Ile44 surface. This double-sided ubiquitin binding allows the efficient recognition of ubiquitinated cargos by GGA.

References

[21] 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.

[22] T. Shiba, S. Kametaka, M. Kawasaki, M. Shibata, S. Waguri, Y. Uchiyama and S. Wakatsuki, *Traffic,* **5** (2004) 437.

[23] T. Shiba, M. Kawasaki, H. Takatsu, T. Nogi, N. Matsugaki, N. Igarashi, M. Suzuki, R. Kato, K. Nakayama and S. Wakatsuki, *Nature Struct. Biol.,* **10** (2003) 386.

[24] M. Kawasaki, T. Shiba, Y. Shiba, Y. Yamaguchi, N. Matsugaki, N. Igarashi, M. Suzuki, R. Kato, K. Kato, K. Nakayama and S. Wakatsuki, *Genes Cells,* **10** (2005) 639.

2. Substrate recognition mechanism of a glucuronyltransferase, GlcAT-P, which is required for HNK-1 carbohydrate biosynthesis

Carbohydrate molecules on the cell surface modulate a variety of cellular functions, including cell-to-cell interactions. The HNK-1 carbohydrate epitope, which is recognized by HNK-1 monoclonal antibodies, is found on many neural cell adhesion molecules and is thought to be required for development of brain and nerve systems. It has been reported that there are two mammalian glucuronyltransferases, GlcAT-P and GlcAT-S, which are associated with the biosynthesis of the HNK-1 carbohydrate epitope. Another glucuronyltransferase, GlcAT-I which has the acceptor substrate proteoglycan has been characterized. These enzymes catalyze the transfer of glucuronic acid (GlcA) from a donor substrate, uridine diphosphoglucuronic acid (UDP-GlcA), to a reducing terminal residue of oligosaccharide chains? in the presence of manganese. There are many differences in substrate specificity of the acceptor sugars. To elucidate the acceptor substrate specificity of GlcAT-P, we solved the crystal structures of GlcAT-P with and without substrates using the PF-AR NW12A beamline [25].

The overall structure of GlcAT-P consists of twelve β -strands and seven α-helices (Fig. 16a). The N-terminal region (β1-α1-β2-α2-β3-α3'-α3-β4), referred to as an UDP-GlcA binding region, contains an α/β Rossmann-like fold. The C-terminal region (β5-α4-β6-β7-β8-β9-α5-α6-β10-β11-β12), referred to as an acceptor substrate binding region, contains two β-sheets. The C-terminal long loop, which extends to another molecule, is related by the non-crystallographic two-fold axis (Fig. 16b). A UDP moiety of UDP-GlcA which is a donor substrate is recognized with tyrosine (Y93), arginine (R165), aspartic acid (D197), histidine (H311) and arginine (R313). All

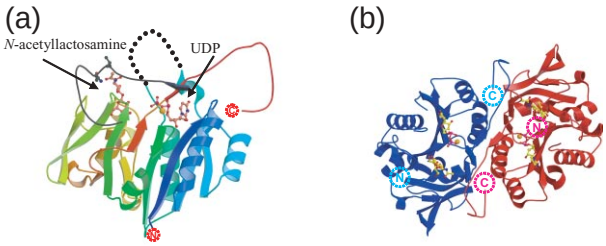

Figure 16
The structure of GlcAT-P in complex with Mn^{2+}, UDP and *N*-acetyllactosamine. (a) Monomer structure of the GlcAT-P quaternary complex. The C-terminal region of the neighbor molecule is shown in gray. (b) Dimer structure of the complex.

these residues are conserved among GlcAT-P, GlcAT-S and GlcAT-I not only on their primary sequences but also on their stereochemical dispositions.

Many amino acid residues (especially the acidic residues; glutamic acid (E228), aspartic acid (D254) and glutamic acid (E284) in GlcAT-P) interact with the acceptor sugar substrate analogue N-acetyllactosamine (Galβ1-4GlcNAc) in GlcAT-P. They are also conserved among GlcAT-P, GlcAT-S and GlcAT-I. However, there are some differences among them. Firstly, the crystal structure of the GlcAT-P complex shows that the C-terminal long loop from the neighboring molecule serves as the key in the recognition of *N*-acetyllactosamine. Two amino acid residues, valine (V320) and asparagine (N321) of the C-terminal long loop from the neighboring molecule, interact with the GlcNAc residue through a hydrophobic interaction and a hydrogen bond, respectively. V320 is found only in GlcAT-P, and not in GlcAT-S or GlcAT-I. N321 is conserved in GlcAT-P and GlcAT-S, but not in GlcAT-I. We propose that this C-terminal long loop is critical for the acceptor specificity. Secondly, phenylalanine (F245) is important for the interaction with the acceptor substrate *N*-acetyllactosamine. GlcAT-S and GlcAT-I have tryptophan at the position corresponding to F245 of GlcAT-P. Although a similar aromatic interaction was found between tryptophan (W243) of GlcAT-I and the galactose ring of the acceptor substrate, both the aromatic ring of W243 and the galactose ring are tilted by 30 degrees keeping their parallelism as compared to those in GlcAT-P. These differences in the C-terminal residues of the neighboring molecule and in the stacking between an aromatic residue and galactose ring account for the differences in the specificity of the substrate acceptor recognition.

References

[25] S. Kakuda, T. Shiba, M. Ishiguro, H. Tagawa, S. Oka, Y. Kajiwara, T. Kawasaki, S. Wakatsuki and R. Kato, *J. Biol. Chem.* **279** (2004) 22693.

3. First structure of human sialidase, Neu2, and potential drug design for influenza viruses

Sialidases are glycohydrolytic enzymes widely distributed among species ranging from viruses to mammals. The removal of sialic acid from the non-reducing

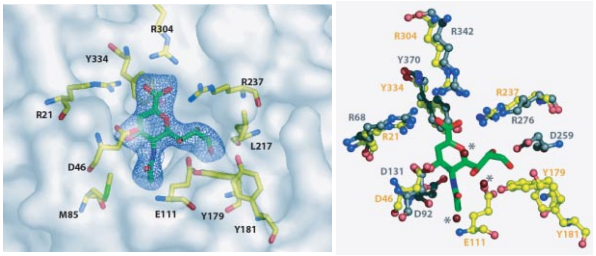

Figure 17
Left panel: molecular surface representation of Neu2 active site showing residues involved in its inhibitor, DANA coordination. DANA is represented as a green stick model with its omit map contoured at 2σ (blue). Right panel: ball-and stick representation of the active site residues of Neu2 (yellow) and viral sialidase (gray) around the DANA molecule (green).

terminus of complex carbohydrates is considered to be the initial step for various biological events such as protein degradation, infection processes, antigenic expression, differentiation, signal transduction and intercellular interactions. Neu2 has been recently identified as a human cytosolic sialidase, one of the four known human sialidases. Although its precise function in living cells is unfortunately as yet unknown, many observations suggest that it may act on ganglioside GM3 leading to the alteration of the cytoskeleton functions, consistent with the decreased invasiveness of transfected melanoma cells. We have determined the first crystal structures of a mammalian sialidase from human, both in its free form and also in complex with an inhibitor, 2-deoxy-2,3-dehydro-N-acetyl neuraminic acid (DANA) [26]. The core of Neu2 folds as a canonical six-bladed β-propeller, displaying its active site in a shallow crevice. The structure of the Neu2 and DANA complex shows the inhibitor lying in the catalytic cleft surrounded by ten conserved amino acids. The arginine triad, conserved among sialidases, stabilizes the carboxylate group of DANA. The catalytic aspartate Asp46 becomes visible in the Neu2-DANA complex structure while it is disordered in the apo-form of the protein, illustrating the dynamic nature of substrate recognition (Fig. 17, left panel). These observations suggest that the molecular mechanism of the catalytic reaction is common among all sialidases, from virus to human. Furthermore, the interaction between Neu2 and DANA shows similarities with bacterial and viral counterparts, but also exhibits significant differences in active site arrangement and the dynamic nature of the loops which contain residues responsible for catalysis and substrate recognition (Fig. 17, right panel). Tamiflu® (oseltamivir) and RELENZA® (zanamivir) are recently-developed drugs which target influenza viruses. They bind to neuraminidase, an influenza virus sialidase, and inhibit the release of virus particles from the host cell. There have been reports of some side-effects from the use of these drugs, and a detailed structural comparison between the virus and human sialidases is expected to provide clues for developing more effective drugs with fewer side-effects.

References

[26] L.M. Chavas, C. Tringali, P. Fusi, B. Venerando, G. Tettamanti, R. Kato, E. Monti and S. Wakatsuki, *J. Biol. Chem.*, **280** (2005) 469.

4. Double sided ubiquitin binding of Hrs-UIM in endosomal protein sorting

Hrs plays an essential role in sorting of mono-ubiquitinated receptors to multivesicular bodies for lysosomal degradation, through recognition of ubiquitinated receptors by its ubiquitin interacting motif (UIM). Here, we present a 1.7 Å structure of a complex of Hrs-UIM and ubiquitin (Fig. 18). Hrs-UIM forms a single α-helix which binds two ubiquitin molecules on either side. These two ubiquitin molecules are related by pseudo two-fold screw symmetry along the helical axis of the UIM, corresponding to a shift by two residues on the UIM helix. Both ubiquitin molecules interact with the UIM in the same manner using the Ile44 surface with equal binding affinities. Mutational experiments show that both binding sites of Hrs-UIM are required for efficient degradative protein sorting. Hrs-UIM belongs to a new subclass of

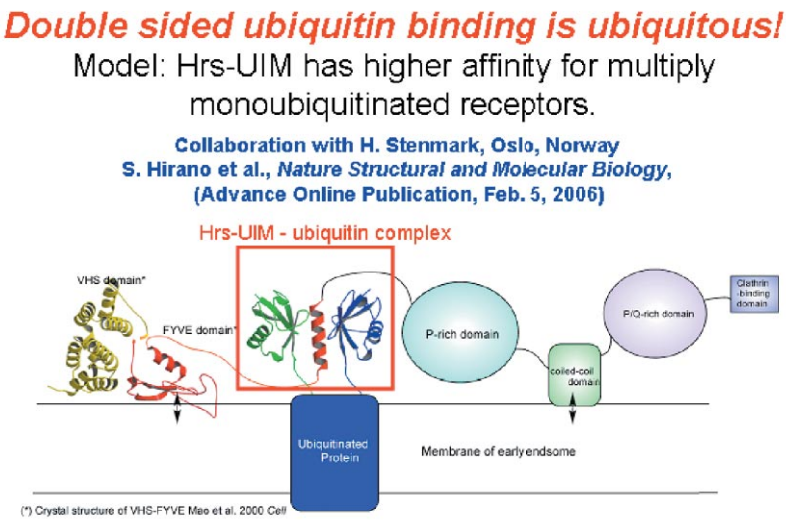

Figure 18
Interaction between Hrs and ubiquitin molecules. Hrs is a multi domain protein composed of VHS, FYVE, UIM, P-rich, coiled-coil, P/Q-rich and clathrin binding domains. Hrs-UIM interacts with two ubiquitin molecules to improve higher affinity for ubiqutinated receptor proteins.

double-sided UIMs, in contrast to its yeast homologue Vps27p which has two tandem single-sided UIMs [27].

References

[27] S. Hirano, M. Kawasaki, H. Ura, R. Kato, C. Raiborg, H. Stenmark and S. Wakatsuki, *Nature Struct. Mol. Biol.,* (2006) in press.

5. Beamlines

The activities of beamline development, operation and management are already described in PART-A of this chapter.

6. Robotics for High-throughput and Automated Protein Structural Research

1. Large-scale protein crystallization and monitoring system

Protein crystallization remains one of the bottlenecks in the crystallographic analysis of macromolecules. We have developed a large-scale protein crystallization and monitoring system (Fig. 19) that is unrivalled in terms of speed by any other system in the world. The system has shortened the time required for crystallization setup more than 100-fold; the time required to setup 480 crystallization conditions using the new system is only 10 minutes, whereas the same task performed manually would take 20 hours. Users can view images of the crystallization drops remotely using web-based browsing software and quickly search for the appropriate crystallization conditions. About 900 crystallization plates (each plate contains 96 crystallization drops) were set up for the initial screening of crystallization conditions at the beginning of operations in the autumn of 2003, and we have succeeded in making a large number of crystals using this system. In some cases, we could determine the protein structures directly from the crystals grown by the system without further optimization (Fig. 20). In other cases, crystals of size and quality sufficient for X-ray diffraction experiments were obtained by further optimization (manual screening) based on the conditions found using the crystallization system.

Figure19
Exterior view of the large-scale protein crystallization system in operation.

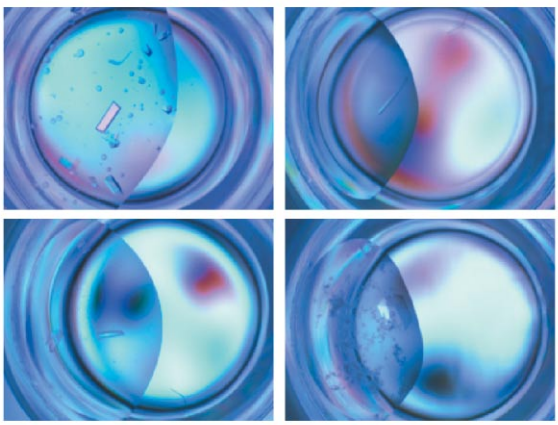

Figure 20
Some examples of protein crystals obtained using the large-scale crystallization system.

To improve the system, we are developing a high-speed observation system with 96 cameras operating in parallel to capture images of all the drops on the crystallization plates on a timescale of tenths of a second. Moreover, we plan to develop a nanolitre dispensing system to reduce sample volume and dispensing time, and an automated crystal scoring system based on various image processing techniques. As a longer-term development project, we also plan to develop a fully automated system which encompasses the whole rage of crystal preparations: protein crystallization, crystal observation, crystal scoring, picking up the crystal, soaking the crystal in a cryoprotectant, and freezing and storing the crystal into a cassette. To this end, we have already developed a seal cutting system which can cut out the seal of the specified wells in which prospective crystals are found, a cryoprotectant exchange system and a subsidiary system for supplying consumables.

2. Automated sample-exchange robots

The high-flux ID beamlines BL-5A and AR-NW12A need significantly shorter exposure time for X-ray protein crystallographic experiments compared to the (BM) beamlines. Combined with fast-readout CCD detectors and shutters, a high quality dataset can be collected within 10 to 30 minutes for 180 degree oscillations. This means that each user (group) can collect 20 to 30 high quality X-ray diffraction datasets during 24 hours of beam time. To further improve the overall throughput, the time required for the offline experimental steps must be reduced as much as possible, in particular the time taken for manual crystal exchange in the experimental hutches. To this end, we have adopted an automated system developed by Stanford Synchrotron Radiation Laboratory (SSRL) for exchanging samples from a liquid nitrogen Dewar, mounting it on the diffractometer and centering the crystal. We selected this system based on its reliability, its track record of about three years of user operation, its large capacity (continuous operation of 288 crystals using three cassettes with 96 crystals each), and its compatibility with commercially available cryo-loops. In addition, we modified the SSRL system from a single cryo-tong to a double cryo-tong in order to reduce the number of trips between the Dewar and the

Figure 21
Overview of the sample exchange robot installed at AR-NW12A beamline. The inset shows an inside view of the Dewar in which three SSRL cassettes are held in place.

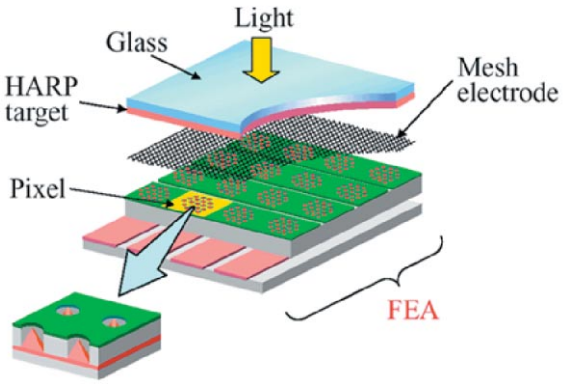

Figure 22
Schematic representation of the HARP area detector with FEA. The glass plate is replaced by beryllium, silicon or other material for an X-ray detector.

diffractometer by a factor of two. We have developed two exchange robots with double tongs, and installed them on BL-5A and AR-NW12A (Fig. 21) during the summer shutdown of 2004. In order to improve the stability of the system, we have been modifying the position calibration and the orientation of the cryo-loops in the cassettes in the Dewar. These robots will increase the throughput of the beamlines from 20 to 30 data sets per day to over 100. Easily-operated tools for inserting the protein crystals into the cassette are available for users.

7. Development of a Next-Generation X-ray Area Detector

X-ray detectors are one of the key components of synchrotron X-ray protein crystallography beamlines. In recent years, tapered fiber optics coupled CCD detectors have become widely used throughout the world. They match most of the current requirements but fall short for the expected characteristics of future light sources and ever increasing demands of structural analyses of complex structures. To prepare for such high demands in the near future, we have been working on a project to develop a next-generation X-ray area detector in collaboration with NHK-ES. The core of the detector system consists of a matrix field emitter array (FEA) and an X-ray HARP. A prototype FEA-HARP has been developed for HDTV applications by the research institute of NHK (Fig. 22). In the FEA-HARP, the FEA is faced to the photoconductor plate. After exposure to light, the photoconductor stores the image in the form of a surface charge distribution on the back of the photoconductor plate, which is scanned by electron beams generated by the FEA and readout individually. The photoconductor plate is comprised of a HARP target layer formed on a glass face plate. The HARP target layer, consisting mainly of amorphous selenium, uses an avalanche multiplication effect across the high-electric field to amplify the signals. These characteristics show the following advantages over the currently available CCD and other area detectors; (1) higher sensitivity, owing

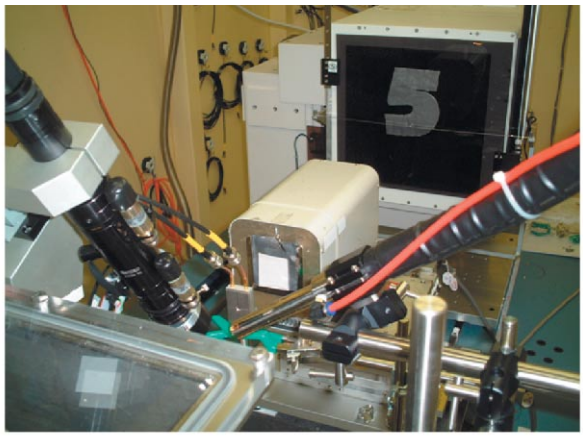

Figure 23
View of the test experiment of the prototype X-ray HARP tube camera at BL-5A. A protein crystal diffraction experiment with 180 degree rotation was completed within 12 seconds.

to the amplification effect by the HARP membrane, (2) higher spatial resolution, achieved by the single driven FEA, (3) higher framing rate of 10-120 images/second, and (4) other characteristics such as a higher dynamic range, lower noise and radiation hardness.

The maximum framing rate of 120 images/second matches the proposed frequency of LCLS, an X-FEL being developed at SLAC (Stanford, California, USA). The X-ray FEA-HARP detector system will be used for experiments at next generation light sources such as LCLS. Fig. 23 shows a view of a test experiment using the prototype X-ray HARP detector with an electron tube to collect diffraction images from a typical protein crystal at BL-5A. Diffraction images covering a 180 degree range could be recorded in only 12 seconds. The new detector system has shown tremendous possibilities which could bring a breakthrough for protein crystallographic analyses of macromolecules.

The development of this detector system is a six-year project starting from FY2004 and is divided into two terms with specific milestones (Table 4). The detector is also planned to be used for other applications such as low-dose medical imaging and real-time phase-contrast imaging methods, which will benefit from the high sensi-

Table 4 Milestone of the X-ray HARP-FEA detector in the "Development of Systems and Technology for Advanced Measurement and Analysis." program.

	First term (1-3 year)	Second term (4-6 year)
Pixel number	1000 × 1000	6000 × 6000
Active area (mm)^2	22.6 × 22.6	300.0 × 300.0
Framing rate (images/sec)	30 – 90	10 – 120
Dynamic range (bits)	14	16

tivity and high framing rate of the detector.

8. Industrial Use and Collaboration with Companies

Industrial users can use PF beamlines under various schemes. The TARA/SBSP project, a consortium between KEK-PF, TARA/SBSP and several companies, has been promoting industrial use since FY1995. Accompanying the change of status of KEK from a government institute to an agency in 2004, this scheme will come to a close at the end of FY2005. However, many industrial users have requested to continue using the PF protein crystallography beamlines. Furthermore, they request to use the high-throughput ID beamlines, BL-5A, NW12A and BL-17A. To realize this, we are now constructing a new scheme between KEK and the industrial companies (Fig. 24). From April 2006, this new scheme of industrial use will start under the new consortium.

In addition to the industrial use of beam time as descried above, we promote tighter collaborations with some companies. The number of these collaborations has gradually increased and stands at eight in FY2005. The contents of the collaborations are different depending on the companies involved. For example, one contract includes protein expression, purification, crystallization, diffraction data collection and structure determination wheras another contract includes only beam time assistance. Throughout these collaborations with companies, we receive not only scientific output, technical experiments and financial support, but also return our social capital to a public society.

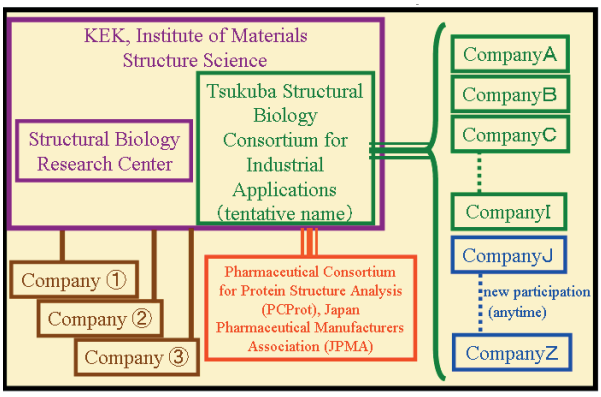

Figure 24
Future Plan of the Industrial Use and Collaborations between KEK and Industry.

9. International Collaboration

We are promoting international collaborations aggressively. Now, there are six international biological collaborations with researchers in foreign countries (USA, UK, Italy, Australia, India, and Norway). In addition, we have accepted many staff members from foreign countries: a graduate student from France, a technician from Australia, and many short-stay post doctoral fellows and researchers from China, Taiwan, Korea, Australia, USA and the UK. As one example of international collaboration, we have adopted an automated sample exchange system which was originally developed by Stanford Synchrotron Radiation Laboratory (SSRL). We were also successful in obtaining a grant (FY2005-FY2007) from JST for collaboration with Professor So Iwata in the UK on a membrane protein crystallization project.

Part-C: Short-term strategy of the Structural Biology Research Center

There are six areas on which we will focus our structural biology activities at the PF in the next five years: (1) beam line operation and technology development, (2) structural biology of post-translational modification and transport of proteins, (3) Post Protein3000 project, (4) University of Tokyo Outstation, (5) industrial partnership, and (6) preparation for the next PF lightsource project.

(1) Beam line operation and technology development

We will continue our user operation and further develop the high-throughput technologies. The current development projects, the micro focus beam line BL-17A, super-fine diffractometers and the X-ray area detector, will continue. The results of these R&D projects will be spread across all the PF beamlines in due course in order to maintain a common user environment.

In this regard it will be vital for the PF SBRC to keep good relations with other synchrotrons. In particular, we are discussing with our counterparts at SPring-8 for future exchanges of personnel through short visits and joint applications for government funding.

Another aspect is the exchange of ideas with other synchrotrons outside of Japan. S. Wakatsuki serves on several science advisory boards including Vizier (an EU-funded structural proteomics project), SSRL, and JCSG-2 (Joint Center of Structural Genomics-2), the SESAME beam line committee, and is also an external reviewer of BIOXHIT and 3D-Repertoire (both EU funded network projects). We have also started a 4-year exchange program with UK groups funded by JST, one with Prof. So Iwata of Imperial College London on membrane protein crystallography and another with Prof. Louise N. Johnson and the Diamond synchrotron on compatibility issues between SSRL-type exchange robots and EU standards.

(2) Structural biology of post-translational modification and transport of proteins

Over the last 5 years we have developed a number of strong collaborations with both domestic and overseas cell and molecular biology groups on structural analyses of post-translational modification and transport of proteins. The staff, post-doctoral fellows, and Ph.D. students now have an in-depth understanding of the field and in many cases compete well with structural biologists in the same fields. We often publish structure papers within a few months or back to back in the same journal with those by other groups in USA, Europe and Asia. We also collaborate with other crystallographers, for example Prof. J. Deisenhofer of University of Texas, in solving crystal structures of importance. These collaborations are extremely important for the SBRC to stay abreast of the field and we will try to expand such collaborative work in the coming years.

The emphasis on structural biology projects will shift from individual components and their complexes with their partners to systems approach as well as medically relevant targets (see below) in the near future. The systems approach will be particularly critical for understanding the higher level regulations of molecular events in the cell and on the organism level. We will make systematic approaches on a number of protein networks including protein glycosylation, complex formation and intracellular transport. Projects will be chosen both for fundamental understanding of molecular cell biology as well as for applied biology, such as the development of efficient drugs.

(3) The next MEXT protein project

Over the last three years the SBRC has provided about 30% of the user beam time on the structural biology beam lines (BL-6A, BL-5A and AR-NW12A) to the eight consortia of the Protein 3000 Project. While we will continue this activity with additional beamtime from the new in-vacuum short gap undulator BL-17A until the end of the project in March 2007, we are keen to propose a new micro-focus MAD beamline as part of the next national protein project. The beamline will take advantage of low energy SAD phasing in the energy range 4 to 6 keV using an in-vacuum short gap undulator based on the experience of BL-17A.

Although much of the next national project is still under discussion at MEXT, the Life Science Department is starting a pilot project on medically relevant target proteins in FY 2006. In this project, they are trying to set up about 20 well defined projects on targets prepared by a committee of leading medical scientists in Japan. We have applied for two new projects and are already in discussion with the medical groups to start structural and functional studies on two new proteins, one involved in liver cancer and another in HIV infection. In addition, we have already started other projects on medically relevant protein targets, for example POMT1 (protein O-mannosyltransferase 1, an integral membrane protein) which is responsible for Walker-Warburg syndrome, a muscle dystrophy, and a-glucosidase (95 kDa with complex N glycosylation) which is the target of the enzyme replacement therapy for Pombe disease. We will further expand this line of research over the next five years.

(4) University of Tokyo Outstation

The University of Tokyo is planning to expand SR activities in the university by establishing two outstations, one at the Photon Factory and another at SPring-8. Several beam lines are planned for VUV-SX spectroscopy, hard X-ray materials science, and structural biology. S. Wakatsuki has been asked by the University to head the structural biology division and is in the process of preparing beam line projects for both PF and SPring-8. Success of the outstations will strongly depend on cooperation with the two synchrotron facilities. To this end, a working group has been formed at the university to discuss strategies for fundraising and design of the beam lines. The structural biology division is planning a total of 4 beam lines; two undulator protein crystallography beam lines at the Photon Factory and a high brilliance microfocus beam line and EXAFS/SAXS beam line at SPring-8. Higher priorities are given to one protein crystallography beam line at each synchrotron. The outstation beamline(s) at the PF will be integrated into the SBRC for design, construction, and operation. In addition, the structural biology division will establish a core crystallography group whose responsibility will be to act as liaison with a large number of top level biology groups within the University of Tokyo.

As part of the University of Tokyo Outstation, S. Wakatsuki will assume an adjunct professorship at the Department of Medical Genomics, Frontier Science Division, University of Tokyo from April 1, 2006. With an additional adjunct associate professor position which will be prepared in due course, we will establish a System Structural Biology course in the University of Tokyo and will accept graduate students at both M.S. and Ph.D. levels.

(5) Industrial partnership

The most important point in which we concentrate our industrial collaboration is establishing a consortium for the industrial use of the structural biology beam lines. This is expected to commence on April 1st 2006, initially with 9 companies from the pharmaceutical and food industries. This consortium will be open to companies who commit themselves to using the PF structural biology beam lines for more than 4 days per year. The consortium will share and test beta versions of the latest technologies that we develop for high throughput structural analyses. An example is the double-tong crystal exchange robot based on the single-tong SSRL sample exchange robot.

In parallel, we will continue and expand our collaborative program with several pharmaceutical companies

in the area of protein production using baculo virus expression systems and E. coli as well as structural analyses optimized for structure-based drug design.

We will also continue the training program for industrial users in which we either accept young inexperienced crystallographers for on-site training or help them as crystallographic problems arise. This has been highly successful in that we have trained a young protein crystallographer in synchrotron experiments and phase determination, and also gave crystallographic advice to another company which collected several hundred data sets with solved structures last year alone.

Finally, we are in the process of negotiating with a pharmaceutical company to build a PRT (program research team) beamline for high throughput protein crystallography. Since the number of straight sections in the PF and PF-AR rings is rather limited, construction of a new beam line requires careful discussion and preparation on the facility side. New schemes for industrial PRT beam lines (general contract schemes including tax, depreciation rate, beam time access and maintenance contract) are being discussed between the relevant parties.

(6) Towards the next Photon Factory Lightsource project.

The Photon Factory is proposing an ERL-based next-generation lightsource project in which structural biology will play a major role. We will be pushing the limit of crystal size and large unit cell in the next PF project, aiming for sub-micron protein crystals and extremely large biological samples, i.e. near diffraction limit protein crystallography. A combination of K-B mirrors and capillary optics as well as X-ray beam monitors for active feed back will be investigated to achieve the optical properties required for these experiments. In addition, we will further develop the sample manipulation project by moving from piezo-based micro manipulators to laser tweezers and improving the diffractometers from 200 nm rotation error to sub 100 nm. There will be synergy between these activities and those being developed for X-FEL based single molecule structural analyses.