Protein Crystallography Initiative

Synchrotron X-ray crystallography has become the dominant method for obtaining atomic resolution information about the structures of biological macromolecules. The high brilliance of synchrotron source overcomes the limitation due to small size of protein crystals. The tunability of synchrotron sources has sparked a rapid growth in the use of anomalous diffraction techniques for solving the phase problem, which in turn has greatly accelerated the speed with which macromolecular structures can be determined. The development of automated robotic operation will soon make high-throughput protein structure determination to become routine at many synchrotron facilities.

Although the wiggler construction was funded (NTD 262 M in 1992) for both materials and protein structure research through a joint proposal of SRRC and Institute of Molecular Biology of Academia Sinica, the push for a protein crystallography facility was only possible until recently. An *ad hoc* Advisory Committee meeting on the Construction of Protein Beamline and End Station was held at SSRL, Stanford, California in April 2001 to help the planning. As happening in other synchrotron facilities world wide, the protein X-ray crystallography with high-throughput capability has taken a high priority at SRRC now. Through the collaboration with the NASDA of Japan, our first modern protein crystallography station with an ADSC Quantum 4R CCD detector was installed at SPring-8 Beamline 12B2 station and open to the bio users in September 2002. Meanwhile, the wiggler 17B2 station was upgraded with a new image plate to improve the data collection speed. Both stations provide MAD capability now. In a short time, numerous protein structures have already been solved.

Our push for protein X-ray crystallography is aligned with NSC's newly initiated National Science and Technology Program for Genomic Medicine (NRPGM). Under the support of this Program, NSRRC has been granted a total of NTD 252M (~ USD 7.3 M) for 3 years to construct dedicated and highly automated beamlines on a new superconducting 28-pole wiggler which will be the X-ray Core Facility of the Program. To meet the anticipated strong demand for beam time from NRPGM, two new dedicated beamlines for protein crystallography are being constructed. One will be a MAD beamline with energy tunability from 6.5 keV to 19 keV and the other a monochromatic beam line with energy fixed at 13 keV for crystal screening and high-resolution structure study (see Beamline Development). A large body of staff with different expertise has been assembled at NSRRC to carry out this project. The facility will serve biology users in year 2005. Major tasks of this project include:

- The design, construction and installation of a 28-pole superconducting wiggle.
- The design and construction of two protein X-ray crystallography beamlines.
- The development and construction of the intelligent robotics system.
- The implementation of robust software for automated crystal characterization, strategy evaluation, data collection and analysis in a sustained production mode.
- The development of new high-throughput pipelines for phasing and structure determination.
- The set up of a massive data archive and storage facilities at the National Center for High-performance Computing.

Acknowledgements

We are grateful to Profs. B. C. Wang, Andrew Wang, and Sunny Chan for their advice in the development of the structural biology program at NSRRC. We are indebted to Dr. T. Ishikawa for his initiation of the NASDA collaboration at SPring-8. Special thanks are due to our STAC member Prof. K. Hodgson for

his help in the development of our new X-ray facility. We thank all bio researchers for their continuous support. The protein crystallography program at NSRRC would not be possible without the strong financial support of NSC.

Information Contact: Yuch-Cheng Jean, ycjean@nsrrc.org.tw

Recent Experimental Results

(1) Crystal Structure of the Hyperthermoplilic Archael DNA-binding Protein Sso10b2 at a Resolution of 1.85 Angstroms by A. H. J. Wang (Academia Sinica)

The families of small, abundant, and basic DNA binding proteins in thermoacidphilic archaea of the genus *Sulfolobus* were first characterized in the 1980s. These proteins can be grouped into three classes according to their molecular sizes (7, 8 and 10 kDa, respectively). Two members of the 7 kDa proteins have been under extensive studies, but little is known about 8 kDa and 10 kDa proteins. *Sulfolobus solfataricus* Sso10b is one of the 10 kDa members. The crystal structure at 2.8 Å resolution of the Sso10b1 has been solved recently. To further explore the function and diversity of the Sso10b1 and Sso10b2 proteins, we present the crystal structure of Sso10b2 from Sulfolobus solfataricus at a substantially higher resolution of 1.85 Å than that of Sso10b1. The Zn-MAD data sets and the high-resolution data were collected at BL12B2 and BL17B2 (NSRRC), respectively. (Publication: J. Bacteriology **185**, 4066 (2003)) (see Fig. 1)

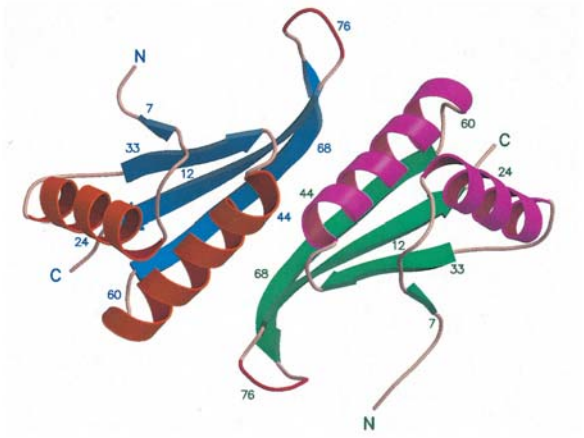

Fig. 1: The ribbon drawing of Sso10b2 dimer. Each Sso10b2 monomer has a β α β α β β topology and comprises two parallel α – helices packed against a four- β - stranded sheet. The Sso10b2 dimer resembles a body with outstretched arms.

(2) Structural determination of the yeast cytosine deaminase by S. H. Liaw (NYMU)

Yeast cytosine deaminase (CD) is an attractive candidate for anticancer gene therapy because of its catalysis of the deamination of the prodrug 5-fluorocytosine into 5-fluorouracil. Recently at BL12B2, Prof. S-H Liaw (NYMU)'s group has solved the crystal structure of the enzyme in complex with the inhibitor 2-hydroxypyrimidine at 1.6 Å resolution by the Se-MAD methods. The current model contains residues 8-158 with clear electron density. The protein structure is composed of a central five-stranded α-sheet, sandwiched by five β-helices. (Publication: J .Biol. Chem **278**, 19111 (2003)) (see Fig. 2).

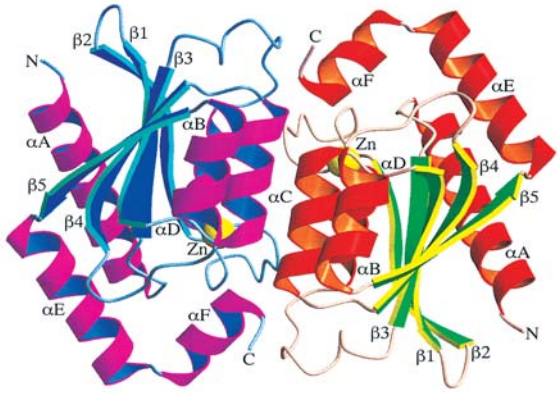

Fig. 2: The dimer structure of yeast cytosine deaminase. One monomer is colored in red (helices) and green (strands), while the other in magenta (helices) and cyan (strands). The zinc ions are shown as yellow spheres.

(3) Structural Basis for the Substrate Specificity of Agrobacterium radiobacter N-Carbamoyl-D-Amino-Acid Amidohydrolase by W.-C. Wang (NTHU)

N-carbamoyl-D-amino acid amidohydrolase (D-NCAase) catalyzes the hydrolysis of N-carbamoyl D-amino acids to produce D-amino acids with release of carbon dioxide and ammonia. Together with a D-specific hydantoinase, the

hydantoinase-NCAase reaction process is currently applied on an industrial scale to convert specific hydantoins into D-carbamoyl derivatives and subsequently into D-amino acids such as D-phenylglycine and D-*p*-hydroxyphenylglycine, the basic building blocks of β-lactam antibiotics. At BL12B2, Prof. W-C Wang (NTHU)'s group has collected several crystal data sets of mutant D-NCAase proteins including N173A, R175A, M5L, and V144C D-NCAases using synchrotron radiation source. Two protein-substrate complex crystals, C172A and C172S complexed with N-carbamoyl-D-phenylglycine were also collected. Both the mutant and complex crystals diffract well to the resolution of about 2.2 Å with the mosaicity about 0.8 and overall Rmerge below 10%. Both structures are solved by molecular replacement method with the wild type D-NCAase crystal structure as the searching model. (Publication: J. Biol. Chem. **278**, 26194-26201 (2003)) (see Fig. 3)

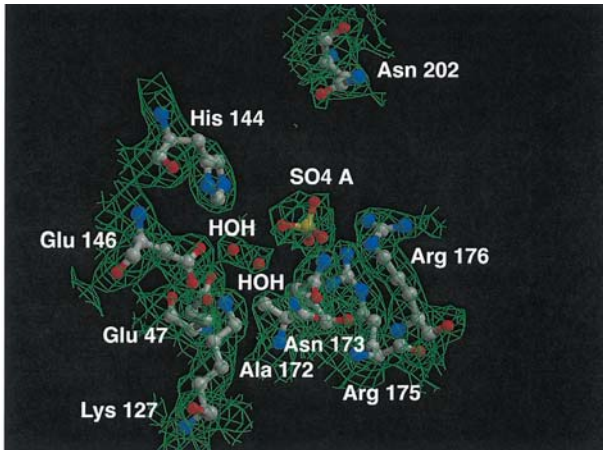

Fig. 3: Electron density map of D-NCAase mutant around residue 172. The map is contoured at the 1.5-s level.

(4) Crystal Structure of a Nucleoside Diphosphate Kinase from Bacillus halodenitrificans: Co-expression of its activity with a Mn-superoxide Dismutase by C. J. Chen (NSRRC)

The moderate halophile, gram-positive bacterium *Bacillus halodenitrificans* (ATCC 49067) when grown in anaerobic condition synthesizes large amounts of a polypeptide complex that contains a heme center capable of reversibly bind nitric oxide. This complex, when exposed to air, dissociates and re-associates into two active components, a Mn-containing super-oxide dismutase (SOD) and a nucleoside diphosphate kinase (*Bh*NDK). The crystal structure of this latter enzyme has been determined at 2.2 Å resolution of which data was collected with Raxis-IV++ detector at NSRRC BL12B2 Beamline. The model contains 149 residues of a total 150 residues and 34 water molecules. *Bh*NDK consists of a four-stranded antiparallel β-sheet, whose surfaces are partially covered by six (-helices, and its overall and active site structures are similar to those of homologous enzymes. However, the hexameric packing of *Bh*NDK shows that this enzyme is different from both eucaryotic and Gram-negative bacteria. The need for the bacterium to pre-synthesize both SOD and NDK precursors that is activated during the anaerobic-aerobic transition is essential. (Publication: J. Struct. Biol. **142**, 247 (2003)) (see Fig. 4).

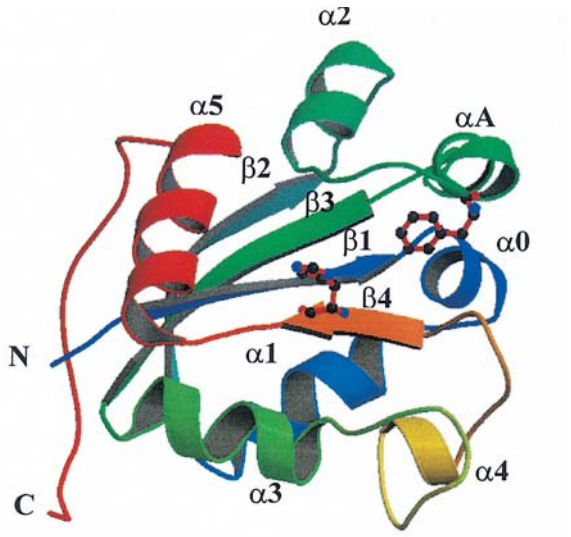

Fig. 4: Ribbon drawing of NDK monomer structure with MOLESCRIPT. The active sites His116 and Phe58 (in ball-and-stick) are located on β 4 and α A, respectively.

(5) Crystal structure of D-Aminoacylase from Alcaligenes faecalis DA1 by S. H. Liaw (NYMU)

D-aminoacylase is an attractive candidate for commercial production of D-amino acids through its catalysis in the hydrolysis of N-acyl-D-amino acids. The crystal structure of the 483-residue D-aminoacylase reveals that the enzyme comprises a small b barrel (residues 7-61, 425-480), and a catalytic TIM barrel (residues 62-414) with a 63-

residue insertion. The small b barrel may be responsible for the structural stabilization, the TIM barrel for catalysis, and the insertion for substrate-mediated conformational switch. In spite of a lack of sequence homology, D-aminoacylase shares significant structural similarity to the TIM barrel metal-dependent hydrolases, in which four conserved histidines and one aspartate are located at the C-terminal ends of strands b1 (HXH), b5 (H), b6 (H), and b8 (D) in the TIM barrel, and define a subtle but sharp sequence signature. (Publication: J. Biol. Chem. **278**, 4957 (2003)) (see Fig. 5).

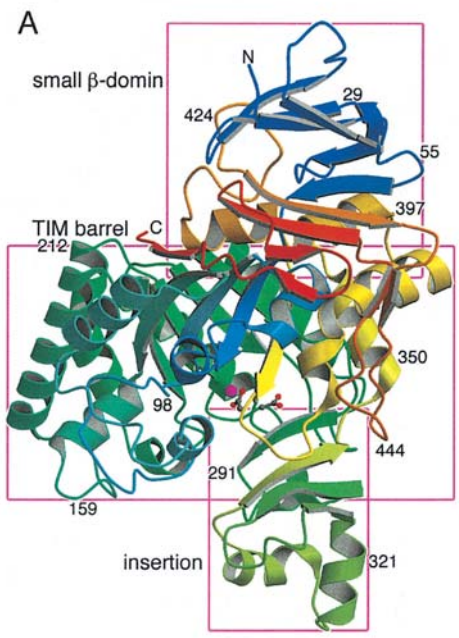

Fig. 5: Ribbon diagram of the C_α backbone of the A. faecalis D-aminoacylase. The tightly bound zinc ion, the metal ligands and two acetate molecules are shown as a pink sphere and ball-and-stick representation.

(6) The Crystal Structure of Rubredoxin from Desulfovibrio Gigas at Ultra-high 0.68 Å Resolution by C. J. Chen (NSRRC)

More and more examples of phasing macromolecular crystal structures based on single-wavelength anomalous dispersion (SAD) show that this method is more powerful and may have more general application in structural biology. Advanced data-collection facilities and cryogenic techniques, coupled with powerful programs for data processing, phasing, density modification and automatic model building, have made the SAD approach gain wider use because of its simplicity and faster data collection and phasing than the multi-wavelength (MAD) method or other methods. It can be performed at any wavelength where anomalous scattering can be observed in synchrotron. For those proteins containing metal ions, metal clusters or heavier atoms with sufficiently large △f", such as Fe, Ni, Cu and S, the SAD experiment can be carried out directly on the native crystals without the need of seleno-derivatives. The SAD method combines the use of SAD data and solvent flattening to resolve phase ambiguity.

We have recently determined the ab initio structure of rubredoxin at ultra-high resolution, 0.68 Å, using the single-wavelength iron anomalous dispersion signal at SPring-8 BL12B2 and NSRRC BL17B2 Taiwan beamline. Rubredoxin from anaerobic sulfide reducing bacteria *Desulfovibrio gigas* is a small redox protein composed of 52 amino acids. It contains with single iron atom bound in a tetrahedral coordination by the sulfur atoms of four cysteinyl residues. The protein is often purified from anaerobic bacteria where it is thought to be involved in electron transfer or exchange processes. The electron density at 0.68 Å ultra-high resolution can help determining very accurate atomic position, especially for the coordination of [Fe-4S] cluster, which reveals detailed information of biological function. (Publication: Biochem. Biophys. Res. Comm. **308**, 684 (2003); The 2003 Annual Meeting of the American Crystallographic Association (ACA), ISSN 0569-4221, **30**, 34 (2003)) (see Fig. 6).

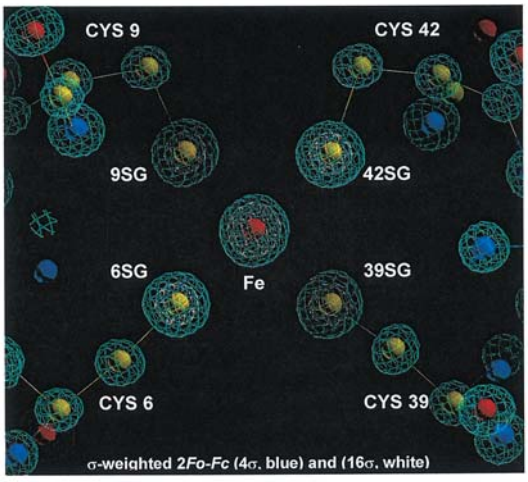

Fig. 6: Sigma-weighted 2Fo-Fc electron density map around [Fe-4S] cluster of rubredoxin (blue: 4 sigma; white: 16 sigma cut-off).