magnetic resonance (NMR) spectroscopy. We suggest that this kind of combination of structural information could be useful for clever inspection of structure by structural genomics.

Infection of the gastric pathogen, *H. pylori*, induces severe gastric disorders including peptic ulcer and stomach cancer. HP0902, identified as one of secretory proteins from *H. pylori*, is predicted to interact with VacA, a representative virulence factor secreted from the bacterium. In addition, HP0902 is over-expressed in a mutant strain lacking the *fdx*A gene, which regulates the resistance of *H. pylori* to an antibiotic. It would be reasonable to consider the secreted proteins from *H. pylori* as candidates for virulence factor of the bacterium, as they can contribute to gastric inflammation. One such protein is HP0902, of which function is unknown. In this respect, crystal structure of HP0902 was approached, in terms of structural genomics.

We first succeeded in solving the crystal structure HP0902 using a construct with an N-terminal His-tag, at 1.4Å resolution. Although the His-tag was not seen in the electron density map, the N-terminal residues were located at dimeric interface and contributed significantly to dimer contact. Thus, we additionally solved the structure using a different construct with a C-terminal His-tag, to ensure the N-terminal conformation in the absence of His-tag attached. Unfortunately, the protein without tagging was not crystallized. Thus, structural inspection for untagged HP0902 was further complemented by NMR spectroscopy in solution, via backbone NMR assignments and chemical shift analysis.

The determined structure of HP0902 showed an all-β topology forming a symmetric homodimer. The monomer was formed primarily by two entirely antiparallel β-sheets that form a jelly roll β-sandwich. The homodimer is formed by a domain swapping between adjacent edge strands β1' and β8 from two different subunits in the dimer. The larger β-sheet has a six-stranded 2310581' topology, while the smaller β-sheet has a four-stranded 4967 topology, respectively. All those are conserved features in cupin superfamily proteins. Cupins are ubiquitous proteins sharing a highly conserved topology of β-barrel, but are classified into 35 protein families, with greatly diversified functions and sequences. In addition, most proteins with the highest score of structural homology to HP0902 are functionally uncharacterized. Thus, unfortunately, a structural fold and homolog search could not be successful in suggesting function of HP0902. However, HP0902 is folded without bound metal ion and possesses additionally extended stretch between β1 and β2 strands. Its dimeric interface is formed by frequent hydrogen bonding, instead of hydrophobic clustering. Those structural properties distinct from other cupin family proteins might provide functional specificity to HP0902. Thus, the present results constitute fundamental, critical data for progressing studies to identify function and/or virulence and to elucidate its structural mechanism.

Keywords: Helicobacter pylori, HP0902, structural genomics

MS57.P03

Acta Cryst. (2011) A**67,** C589

Primer design for cDNA synthesis based on the crystal structure

Naoki Shibata,[a,b] Tsuyoshi Inoue,[c] Yoshiki Higuchi,[a,b] Yasushi Kai,[c] [a]*Department of Life Science, Graduate School of Life Science, University of Hyogo, Ako-gun, Hyogo (Japan).* [b]*RIKEN Harima Institute, SPring-8 Center, Sayo-gun, Hyogo (Japan). Department of Applied Chemistry, Graduate School of Engineering,* [c]*Osaka University, Suita, Osaka (Japan).* E-mail: shibach@sci.u-hyogo.ac.jp

Novel proteins isolated from natural source are often functionally and genetically unknown. In such case, structural analysis may provide not only a clue as to the function but also a partial amino acid sequence. Once the overall fold has been determined, structural similarity can be examined using structure comparison services, which may suggest the functional aspect of the protein. Partial amino acid sequence can be determined based on the electron density shape, which allows design of the target-specific primers for polymerase chain reaction (PCR). Possible DNA sequences that encode the corresponding peptide sequence are diversified depending on the codon degeneracy. However, the DNA polymerases being used for PCR can elongate a new strand even from a primer that has several mismatches around the 5'-end and the middle. The target-specific primers can be used to amplify a double-stranded DNA fragment of the target gene from the cDNA fragment library, which can be then used as a probe for hybridization to obtain the full-length gene of interest.

A functionally and genetically unknown protein from Pleurotus mushroom has been isolated and crystallized. The crystal structure of the protein was solved by MIRAS at 1.79 Å resolution. The amino acid types were tentatively determined based on the electron density shapes, with which several refinement cycles were performed. Two regions in which the electron densities unambiguously indicated amino acid types were chosen for primer design (Fig. 1). The PCR product amplified using the primers designed from the amino acid sequences of the two regions had the size of ~330 bp, consistent with the number of residues (110 aa) corresponding to the amplified part. The product was used as a probe for isolation of the full-length cDNA of interest. The full-length cDNA was cloned by the RARGIP method [1] based on the lone linker PCR technique [2]. 145 out of 226 residues were correctly assigned by X-ray sequence. Further refinement cycles were performed with the gene-derived sequence.

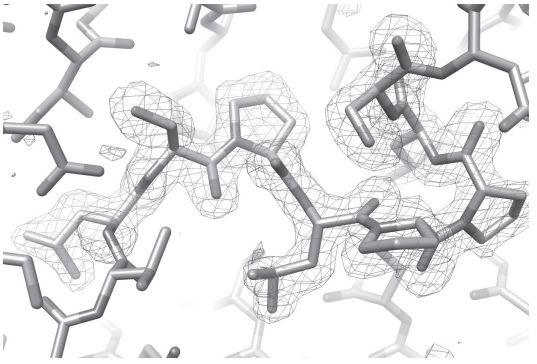

Fig. 1 An Fo-Fc omit map around one of the regions chosen for primer design.

[1] K. Abe *Mamm. Genome* **1992**, *2*, 252-259. [2] M.S. Ko, S.B. Ko, N. Takahashi, K. Nishiguchi, K. Abe *Nucleic Acids Res*. **1990**, *18*, 4293-4294.

Keywords: protein, sequencing, structure

MS57.P04

Acta Cryst. (2011) A**67,** C589-C590

Strategies for analysis expression and protein solubility

Fernández-Millán P, Rubio-Cosials A, Jiménez-Menéndez N, Silva-Espiña C, & Solà M., *IBMB-CSIC. Barcelona (Spain).*

To get soluble protein is one of the major bottlenecks that precede crystallographic studies. During the last years several techniques and strategies have been developed to address this problem. However, many of them imply an economical cost and technologies that are not always available.

We will describe a general plan for protein solubility analysis by using a combination of four different but complementary strategies. In this plan, different constructs of a protein of interest are designed and

amplified by PCR introducing in all of them the same double restriction sites present in a library of expression vectors, which will add different tags (MBP, GB1, His-tag at the N- or C-terminus) to each construct. All constructs are subsequently tested for bacterial overexpression using different strains and at different culture conditions. The last step consists of a solubility screening for each construct tested at a specific strain and expression condition, by using different buffers designed according to a semi-rational approach. The results from different analysis will be presented.

Keywords: methods, expression, solubility

MS57.P05

Acta Cryst. (2011) A**67,** C590

Community nominated projects at the Joint Center for Structural Genomics

Qingping Xu,[a,b] Ashley Deacon,[a,b] Ian A. Wilson,[a,c] [a]*Joint Center for Structural Genomics,* [b]*Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, (USA).* [c]*Department of Molecular Biology, The Scripps Research Institute, La Jolla, California 92037, (USA).* E-mail: qxu@slac.stanford.edu

The Joint Center for Structural Genomics (JCSG), funded through the NIH NIGMS Protein Structure Initiative (PSI), has implemented a high-throughput structural biology pipeline, which has delivered more than 1100 structures over the last 10 years. Numerous technological advances have been incorporated to allow efficient, parallel processing of a large numbers of targets from both bacteria and eukaryotes. Until recently, the primary focus of the PSI was to increase structural coverage of protein sequence space, by determining structures from large protein families that lacked any structural representative. However, in 2010, the PSI entered its third phase (PSI-Biology), where the emphasis shifted to make high-throughput structure determination more broadly available to the biological research community. NIGMS has funded several High-throughput Enabled Structural Biology Centers to focus on specific biological topics, as part of the PSI-Biology Network. In addition, individual researchers are able to propose "community-nominated" targets to the PSI through the PSI Structural Biology Knowledgebase (http://cnt.sbkb.org/CNT/targetlogin.jsp).

Here we will present specific examples of "community-nominated" projects at the JCSG, highlighting the diversity and scope of these biological projects. These projects are sometimes initiated as a follow-up to structures previously solved by the JCSG, where we team up with biologists to perform biochemical studies aimed at elucidating the functional role of these novel proteins. Frequently, the scope of the project is then expanded to target other related proteins. In other cases, biologists come to us with exciting new projects, targeting structures of novel and biologically significant proteins or protein families, proteins involved in interesting pathways, essential proteins but still of unknown function identified through microarray studies, or protein-protein/protein-DNA complexes. Results from some of these projects will be outlined.

The JCSG is funded by NIGMS/PSI, U54 GM094586 and U54 GM074898. SSRL operations are funded by DOE BES, and the SSRL SMB program by DOE BER, NIH NCRR BTP and NIH NIGMS.

Keywords: high-throughput structural biology, collaborative biological research, community targets

MS57.P06

Acta Cryst. (2011) A**67,** C590

X-ray fluosescence tool for rapid in-house evaluation of heavy atom derivatives

Takashi Matsumoto and Akihito Yamano, *Rigaku Corporation, Tokyo (Japan).* E-mail: t-matumo@rigaku.co.jp

The MIR phasing method still plays a crucial role in the structure determination of a novel protein. However, to acquire adequate heavy atom derivatives, one must repeatedly go through the tedious procedure, transferring crystals to various heavy atom solutions with different concentration for different soaking time ranging from minutes to days [1]. Moreover, the result of soaking can only be obtained by checking a difference Patterson map that requires a full data collection. If one can assess the presence of heavy atoms in a protein crystal prior to a data collection, it will greatly reduce the amount of work required to obtain heavy atom derivatives.

To meet this need, we developed an X-ray fluorescence (XRF) tool that can be used to detect heavy atoms introduced by soaking. This XRF tool is composed of an SDD, an XYZ stage, counting circuits and software, and can be added to an existing in-house single crystal system.

We performed some experiments to evaluate this tool. The tool was placed on a Rigaku MicroMax 007HF rotating anode X-ray generator equipped with a VariMax Mo multilayer optic. Some lysozyme crystals were soaked in K_2PtCl_4 solution and subjected to both XRF and X-ray diffraction (XRD) measurements. We found that those crystals showing clear XRF spectra of Pt tend to have discrete peaks corresponding to Pt's on Harker sections and have higher occupancy when refined against ΔF's derived from the XRD measurement.

This XRF tool can be used not only to eliminate unwanted "Native" crystals in-house prior to beam time, but also to estimate amount of heavy atoms introduced in a protein crystal by soaking.

[1] T.L. Blundell, L.N. Johnson, *Protein Crystallography, London: Academic Press.* **1976**, 183-239.

Keywords: fluorescence, tooling, derivative

MS57.P07

Acta Cryst. (2011) A**67,** C590-C591

The Joint Center for Structural Genomics

M.-A. Elsliger,[a,b] A.M. Deacon,[a,c] A. Godzik,[a,d] S.A. Lesley,[a,e] K.O. Hodgson,[a,c] J. Wooley,[a,f] K. Wüthrich,[a,b] I.A. Wilson,[a,b] [a]*JCSG.* [b]*The Scripps Research Inst., La Jolla, CA.* [c]*SSRL, Stanford University, Menlo Park, CA.* [d]*Stanford-Burnham Institute, La Jolla, CA.* [e]*Genomics Inst. of the Novartis Research Foundation, San Diego, CA.* [f]*UC San Diego, La Jolla, CA.* elsliger@scripps.edu

The Joint Center for Structural Genomics (JCSG; www.jcsg.org) is one of the four PSI:Biology Center for High-Throughput (HT) Structure Determination. Our mission in this 3rd phase of the Protein Structure Initiative (PSI) is to extend HT Structural Biology (HTSB) to a broader biological community. To meet these goals, each of the HT centers has been partnered with Biological Centers that present challenging new opportunities for HTSB, in addition to our internal biomedical-theme projects. The JCSG is in its 12th year of operation and provides a robust and flexible HTSB platform that is applicable to a large variety of targets from both bacteria and eukaryotes. Our main approach involves processing large numbers of targets via an extensive combination of bioinformatics and biophysical analyses to efficiently characterize each target in order to optimize its path through our pipeline. In close collaboration with our PSI:Biology Partnerships and