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.

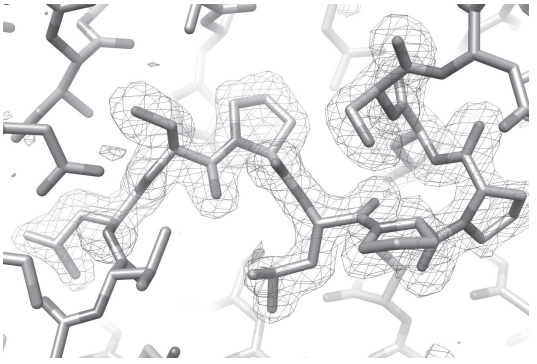

Fig. 1 An Fo-Fc omit map around one of the regions chosen for primer design.

[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