

each case, XAFS would confirm or reject the random character of the solution. Athena and Artemis codes as interfaces for IFEFFIT [2] and FEFF8.4 [3] codes were employed for XAFS spectra interpretation.

$\gamma$ -Fe<sub>2-x</sub>Cr<sub>x</sub>O<sub>3</sub> maghemite for x=0.75, 1 and 1.25 was investigated by XAFS in both Fe and Cr K-edges. Pre-edge decomposition and theoretical modeling of XANES transitions were performed and interatomic distances were determined. Possible distortion of oxygen octahedra around both Fe(III) and Cr(III) cations was checked for potential ferroelectricity explanation, as well as possible vacancy superstructure.

LaFeNiTiO<sub>3</sub> has been confirmed to display magnetic behavior and was studied by its Fe and Ti K-edges. Interatomic distances were modeled for preferential occupation of sites by Fe or Ti atoms and for geometry of oxygen octahedra occupied by different cations.

[1] J. Rodríguez-Carvajal, *Commission on Powder Diffraction Newsletter* **2001**, 26, 12–9. [2] B. Ravel, M.J. Newville, *Synch. Rad.* **2005**, 12, 537–41. [3] J.J. Kas, J.J. Rehr, J.L. Glover, C.T. Chantler *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **2010**, 619, 28–32.

**Keywords:** X-ray diffraction, XANES, EXAFS

## MS56.P18

*Acta Cryst.* (2011) **A67**, C588

### NMR Crystallography applied to dicarboxylic acids

Philippe Ochsenbein,<sup>a</sup> Mohamed El-Hajji,<sup>a</sup> <sup>a</sup>Analytical Science Department, Sanofi-aventis, Montpellier (France). E-mail: philippe.ochsenbein@sanofi-aventis.com

A large number of active pharmaceutical ingredients (API) are developed as salts, generally for solubility reasons. Beside the common inorganic hydrochloride acid (by far the most represented) and sulphuric acid, others carbon-containing acids are widely used as counterions, e.g. fumaric and succinic acids. These organic dicarboxylic acids, having their carbon skeletons that differ only by an inner bond saturated or not, are usually highly difficult to be precisely localized in a <sup>13</sup>C solid state NMR (ssNMR) spectrum. This is e.g. the case by comparison with liquid state <sup>13</sup>C NMR data, since no more than two signals are expected in solution for such dicarboxylic acids; whatever the salt/base stoichiometry may be (for instance, salts of mono-, hemi- or sesqui-fumaric acid could be found in the Cambridge Structural Database). The NMR chemical shifts of the carbons belonging to the base may also vary in the liquid and solid state; which represent an additional difficulty for salt identification. However, using NMR crystallography, we were able to find the peak positions of fumaric and succinic acids mono complexed to a given API, notwithstanding the large number of carbons (35) displayed in their respective cross polarized magic angle spinning (CPMAS) spectrum. Crystallographic studies have been carried out for both salts; they crystallize in triclinic non centrosymmetrical structures, without organic solvent or water. The two crystal structures are remarkably isomorphous; this similarity even somewhat extends to the counterions, where the inner single bond for succinic acid is shorter than expected (both at room and low temperature). In contrast indeed to solution NMR data, four distinct signals are found by ssNMR both with fumaric and succinic salts, a consequence of salt bridges established in each case by only one carboxylic acid. Furthermore, this study allows the attribution without ambiguities of the carbons peak positions of each salt; they could readily be deduced from comparison of the two CPMAS spectra, since the chemical shifts characteristic of the base are found at same positions.

**Keywords:** ssNMR, structure elucidation

## MS57.P01

*Acta Cryst.* (2011) **A67**, C588

### Crystal structures of an enzyme duo involved in bacterial cell wall recycling

Debanu Das,<sup>a,b</sup> Mireille Hervé,<sup>c</sup> Dominique Mengin-Lecreux,<sup>c</sup> Ashley Deacon,<sup>a,b</sup> and Ian A. Wilson,<sup>a,d</sup> <sup>a</sup>Joint Center for Structural Genomics, <sup>b</sup>SSRL, SLAC National Accelerator Laboratory, Menlo Park, (USA). <sup>c</sup>Université Paris-Sud, Laboratoire des Enveloppes Bactériennes et Antibiotiques, IBBMC, and CNRS, Orsay, (France), <sup>d</sup>Department of Molecular Biology, The Scripps Research Institute, La Jolla, (USA).

Peptidoglycan in bacterial cell walls is synthesized by bacteria-specific enzymes by sequential addition of amino acids to UDP-N-acetylmuramic acid to form the peptidoglycan pentapeptide precursor, UDP-N-acetylmuramoyl-L-alanyl- $\gamma$ -D-glutamyl-meso-diaminopimelyl-D-alanyl-D-alanine. In gram-negative bacteria, ~30–60% of the bacterial cell wall is recycled every generation in a complex process that involves ~18 proteins. LdcA catabolizes the bond between L- and D-amino acids in the degradation of the tetrapeptide L-alanyl- $\gamma$ -D-glutamyl-meso-diaminopimelyl-D-alanine to form L-alanyl- $\gamma$ -D-glutamyl-meso-diaminopimelate, which is the substrate for Mpl. Loss of LdcA or Mpl activity results in increased sensitivity to stationary phase lysis or antibiotic susceptibility, respectively. Crystal structures of a novel LdcA from *Novosphingobium aromaticivorans* (NaLdcA) and the Mpl from *Psychrobacter arcticum* (PaMpl) have been determined at 1.89 and 1.65 Å resolution, respectively. NaLdcA and the LdcA from *Pseudomonas aeruginosa* have similar overall structures and a conserved catalytic triad despite only a 20% sequence identity. Modeling a tetrapeptide substrate into the active site of the NaLdcA structure reveals residues that may be important in substrate recognition and how the catalytic triad is positioned for action. An unidentified ligand in the dimer interface might be involved in enzyme dimerization and indicates a possible site for inhibitor design. In PaMpl, ~30 residues are likely to be important for substrate recognition or may be involved in interdomain conformational changes on substrate binding. The PaMpl enzymatic activity has been characterized at different temperatures and with different peptide substrates. At 15 °C, PaMpl has almost twice the activity of Mpl from *E. coli*, consistent with *P. arcticum* adaptation at cold temperatures. These results expand the structural coverage of two important protagonists in bacterial cell wall peptidoglycan recycling and provide novel targets for anti-bacterial drug discovery.

[1] D. Das, M. Hervé, J. Feuerhelm, C.L. Farr, H.J. Chiu, M.A. Elsliger, M.W. Knuth, H.E. Klock, M.D. Miller, A. Godzik, S.A. Lesley, A.M. Deacon, D. Mengin-Lecreux, I.A. Wilson, *PLOS One*, **2011**.

**Keywords:** Peptidoglycan, LdcA, Mpl

## MS57.P02

*Acta Cryst.* (2011) **A67**, C588–C589

### Structure Determination of HP0902, a Putative Secretory Protein from *H. pylori*

Dae-Won Sim,<sup>a</sup> Sung-Hee Lee,<sup>a</sup> Woo Cheol Lee,<sup>b</sup> Jung Hyun Song,<sup>b</sup> Hye-Yeon Kim,<sup>b</sup> Hyung-Sik Won,<sup>a</sup> <sup>a</sup>Department of Applied Life Science, Graduate School, Konkuk University, Chungju. <sup>b</sup>Division of Magnetic Resonance Research, Korea Basic Science Institute, Ochang, Chungbuk (Republic of Korea). E-mail: konapapa@nate.com

As a structural genomics approach to the HP0902, a putative candidate for virulence factor of *Helicobacter pylori*, crystal structures were obtained from two different constructs and validated by nuclear

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 *fdxA* 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- $\beta$  topology forming a symmetric homodimer. The monomer was formed primarily by two entirely antiparallel  $\beta$ -sheets that form a jelly roll  $\beta$ -sandwich. The homodimer is formed by a domain swapping between adjacent edge strands  $\beta 1'$  and  $\beta 8$  from two different subunits in the dimer. The larger  $\beta$ -sheet has a six-stranded 2310581' topology, while the smaller  $\beta$ -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  $\beta$ -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  $\beta 1$  and  $\beta 2$  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) A67, C589

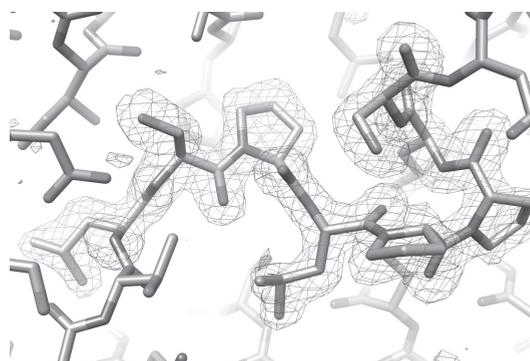
### Primer design for cDNA synthesis based on the crystal structure

Naoki Shibata,<sup>a,b</sup> Tsuyoshi Inoue,<sup>c</sup> Yoshiki Higuchi,<sup>a,b</sup> Yasushi Kai,<sup>c</sup>  
<sup>a</sup>Department of Life Science, Graduate School of Life Science, University of Hyogo, Ako-gun, Hyogo (Japan). <sup>b</sup>RIKEN Harima Institute, SPring-8 Center, Sayo-gun, Hyogo (Japan). <sup>c</sup>Department of Applied Chemistry, Graduate School of Engineering, 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) A67, C589-C590

### Strategies for analysis expression and protein solubility

Fernández-Millán P, Rubio-Cosials A, Jiménez-Menéndez N, Silva-España 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