MS6-03 **Structural bioinformatics incorporated into ARP/wARP model building.** Tim Wiegels,[a] Marco Biasini,[b] Torsten Schwede,[b] Victor Lamzin,[a] [a]*EMBL Hamburg outstation, Hamburg, Germany,* [b]*Swiss Institute for Bioinformatics, Basel, Switzerland*
E-mail: wiegels@embl-hamburg.de

One of the most challenging tasks in macromolecular crystallography (MX) is the determination of the three-dimensional structure of proteins for which crystals do not diffract to sufficiently high resolution. Computational approaches for automatic model building in MX have traditionally been focused on high-resolution data. Thus their application to data at resolutions worse than ~2.5 Å is limited and typically results in incomplete and highly fragmented models. Therefore, robust methods are urgently needed for automated determination of low-resolution MX structures to high levels of completeness and accuracy.

We address the problem by exploiting the fact that 50% of all crystal structures deposited in the PDB [1] contain multiple copies of subunits or their assemblies in the asymmetric unit - ie. they possess non-crystallographic symmetry (NCS). Often NCS-related parts of the structure are built to different extents. Using these differences helps to advance the model building process and significantly increases the overall completeness of built structures, especially with low-resolution data. The application of the so-called Protein NCS-based Structure (PNS) extender to the classic ARP/wARP model building protocol [2] increases the completeness of built structures by up to 18%. More side chains can be docked and, in many cases, the average length of built chain fragments can be doubled [3].

The FittOFF method (Fitting OF Fragments), utilising the experience accumulated within the ARP/wARP and OpenStructure [4] projects, identifies structural gaps between built protein fragments and fills these gaps with fragments selected from a large database derived from the PDB. In contrast to loop-building approaches commonly used in MX, the processing of structural gaps does not require the built fragments to be sequence-assigned. Gap identification is achieved by docking the built fragments to a secondary structure predicted from the amino-acid sequence. The identified gaps are filtered for false-positives using a knowledge-based approach relating the number of residues contained in a gap to the distances between the anchoring residues. The application of FittOFF to ARP/wARP model building was tested on a set of ten structures each containing up to 300 residues and solved with data at resolutions between 3.0 Å and 3.8 Å. We observed a noticeable increase in model completeness and doubling of the average fragment length.

[1] P.W. Rose, et al., *Nucleic Acids Research* **2011**, *39*, D392-D401
[2] G. Langer, S.X. Cohen, V.S. Lamzin, A. Perrakis, *Nature Protocols* **2008**, *3*, 1171-1179.
[3] T. Wiegels & V.S. Lamzin, *Acta Cryst.* **2012,** D68, 446-453
[4] M.Biasini, V. Mariani, J. Haas, S. Scheuber, A.D. Schenk, T. Schwede, A. Philippsen, *Bioinformatics* **2010**, 26, 2626-2628

Keywords: model building, loop modeling, bioinformatics

MS6-04 **PDB_REDO: from bioinformatics data bank to crystallographic tool.** Robbie P. Joosten, *Netherlands Cancer Institute, the Netherlands*
E-mail: r.joosten@nki.nl

The interaction between MX and bioinformatics is often seen as a unidirectional process in which structure models are passed from creators (crystallographers) to users (bioinformaticians). This view overlooks that bioinformatics also provides algorithms and tools for MX which are used for structure analysis and annotation, sequence and structure alignment, mutant design, model validation, etc. The PDB_REDO project is an example of this exchange between bioinformatics and MX. PDB_REDO started out as a data bank of re-refined X-ray structure models from the Protein Data Bank which allowed bioinformaticians to benefit from the latest developments in structure model refinement methods [1]. Now, PDB_REDO has grown out to a fully automated pipeline of new and existing crystallographic tools, available for structure model optimisation and validation in the MX lab [2]. The current features of PDB_REDO, which include refinement with restraint weight optimization, selection of atomic B-factor and TLS models, (re)building of peptides and side chains, and extensive validation of the results, are discussed.

[1] Joosten, R. P., Vriend, G. (2007). *Science* **317**, 195-196.
[2] Joosten, R. P., Joosten, K., Murshudov, G. N., Perrakis, A. (2012). *Acta Cryst.* **D68**, 484-496.

Keywords: macromolecular crystallography; refinement; validation