MS07 O1

Is there a Sevres standard for protein structures? Alexander Wlodawer[a], Mariusz Jaskolski[b], Zbigniew Dauter[a], [a]*Macromolecular Crystallography Laboratory, National Cancer Institute, Frederick, MD, USA.* [b]*Department of Crystallography, A. Mickiewicz University, Poznan, Poland.*
E-mail: wlodawer@ncifcrf.gov

Keywords: refinement, restraints, validation

Limited resolution of the diffraction data collected for most of protein crystals do not allow their models to be refined without introduction of stereochemical restraints. Traditionally, the libraries of geometrical targets have been constructed on the basis of well-refined crystal structures of amino acids or small peptides. These standards may not be completely appropriate for large protein structures, where some specific intra- and intermolecular interactions may cause certain stereochemical features to adopt conformations rarely observed in less strained small molecules. The availability in the PDB of a number of protein models refined at ultra-high resolution offers the possibility of creating a library of protein geometry built on most accurately refined protein models. However, even the best refined models display some degree of variability in the analogous structural features, therefore the geometrical targets cannot be treated as a "holy Grail", but should be accompanied by appropriate weights, allowing for a degree of flexibility. The targets of various restraints (very often too tight in the current practice) and the level of corresponding weights appropriate in various circumstances will be discussed, and illustrated with the examples taken from the recent depositions in PDB.

MS07 O2

Improvement of Crystallographic Data near Corners in Mosaic CCD Detectors. A.S. Arvai[1], S. Brockhauser[2], G. Cioci[3], GA. Leonard[3], A. McCarthy[2], S. McSweeney[3], C. Müller-Diekmann[3], C. Nielsen[1], D. Nurizzo[3], R. B.G. Ravelli[2], and X. Nguyen-huu[1], [1] *Area Detector Systems Corp, Poway CA, USA,* [2] *EMBL, 6 rue Jules Horowitz, 38042 Grenoble, France* [3]*ESRF, rue Jules Horowitz, 38042 Grenoble, France.* E-mail: gordon.leonard@esrf.fr

Large area fast-readout CCDs have become the detector of choice for modern Macromolecular Crystallography synchrotron beam lines. These mosaic detectors are flat-field corrected in order to give a uniform response over the whole surface area. Nevertheless, it has been observed that crystalline diffraction intensities are often underestimated when recorded near the corners of a CCD module. We believe that this effect is primarily due to a variation of the point spread function from the center to the corner of each module. The location of these corners on 3x3 and 4x4 CCD detectors often has significant negative impact on structure solution when important medium resolution reflections are measured in these areas. A method to construct and apply a correction for this effect is presented along with some examples showing improvement in R-factors, anomalous signal, and the ability to solve a structure.

MS07 O3

Tandem use of crystallography and mass spectrometry to sequence *ab initio* HPBP Eric Chabriere[a], Mikael Elias[a] Hélène Diemer[b], Frédérique Renault[c], Carlos Contreras-Martel[d], Alain Van Dorsselaer[b] ,[a]*LCM3B, Nancy Universtit.* [b]*LSMBO, Strasbourg.* [c]*CRSSA, France.* [d]*LCCP, IBS, Grenoble.*
E-mail: eric.chabriere@lcm3b.uhp-nancy.fr

Keywords: *ab initio* sequencing, gene missing, atherosclerosis

The Human Phosphate Binding Protein (HPBP) is a serendipitously discovered apolipoprotein from human plasma that binds phosphate [1]. Amino acid sequence relates HPBP to an intriguing protein family that seems ubiquitous in eukaryotes. These proteins, named DING according to the sequence of their four conserved N-terminal residues, are systematically absent from eukaryotic genomes database [2]. As a consequence, HPBP amino acids sequence had to be first assigned from the electronic density map at 1.9 Å. Then, an original approach combining X-ray crystallography and mass spectrometry provides the complete and *a priori* exact sequence of the 38kDa HPBP. This first complete sequence of a eukaryotic DING protein will be helpful to study HPBP and the entire DING protein family which could be involved in various diseases (atherosclerosis, kidney stone, HIV, rheumatoid arthritis)

[1] Morales R., Berna A., Carpentier P., Contreras-Martel C., Renault F., Nicodeme M., Chesne-Seck M.L., Bernier F., Dupuy J., Schaeffer C., Diemer H., Van-Dorsselaer A., Fontecilla-Camps J.C., Masson P., Rochu D.**,** Eric Chabriere**,** *Structure,* 2006, 14, 601-609 (2006)
[2] Berna A., Bernier F., Chabriere E., Perera T., Scott, K., *Intl. J. Biochem. Cell Biol.*, 2007 In press.

MS07 O4

The wwPDB remediation project Ganesh J. Swaminathan[a], Bohdan Schneider[b], [a]*European Bioinformatics Institute, Cambridge, U.K.* [b]*Research Collaboratory for Structural Bioinformatics, Piscataway, USA.* E-mail: jawahar@ebi.ac.uk

Keywords: protein database, structural biology, chemical nomenclature

The worldwide Protein Data Bank (wwPDB) [1], [2] is a group of organizations that act as data centres for deposition, processing and distribution of data in the Protein Data Bank (PDB) [3]. Its mission is to maintain a single PDB archive of macromolecular structural data that is freely and publicly available to the global community. The members are RCSB PDB (USA), BMRB (USA), MSD-EBI (Europe) and PDBj (Japan).
The evolution of experimental methods, functional knowledge of proteins, and methods used to process PDB data have introduced inconsistencies into the archive since its inception in 1971. The wwPDB (http://www.wwpdb.org/) has collaborated on a project to remediate the PDB archive and create a new set of corrected files. All existing entries have been reviewed and errors have been corrected where possible in order to ensure the uniformity of archived entries and allow scope for future developments in structural biology. Some of

areas concentrated upon in the wwPDB remediation efforts are

a) Improving the detailed chemical description of non-polymer and monomer chemical components.
b) Standardizing atom nomenclature.
c) Updating sequence database references and taxonomies.
d) Resolving any remaining differences between chemical and macromolecular sequences.
e) Improving the representation of viruses.
f) Verifying primary citation assignments.
g) Ensuring consistency between beamline and synchrotron facility names with BioSync.

A complete release of remediated data files is available from a dedicated ftp site. This area is separated from the original archive of deposited data files, which will continue to be available as previously. The details of the remediation process as well as its impact on end-users and developers of software will be discussed along with new features in PDB format 3.1.

[1] Berman,H.M., Henrick,K., Nakamura,H., *Nat. Struct. Biol.*, 2005, 10, 980.
[2] Berman,H., Henrick,K., Nakamura,H., Markley,J.L., *Nucleic Acids Res.*, 2007, 35, D310.
[3] Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N., Bourne,P.E., *Nucleic Acids Res.*, 2000, 28, 235.

MS07 O5

The determination of protonation states in proteins J R Helliwell[a,b], H U Ahmed[a], M Blakeley[c], M Cianci[b,@], J A Hubbard[d], S Fisher [a], L Govada[e], N E Chayen[e], E Carpenter[e] and J Squire[e] [a]*School of Chemistry, Brunswick street, The University of Manchester, M13 9PL. UK,* [b]*CCLRC Daresbury Laboratory, Warrington, Cheshire WA4 4AD, UK,* [c]*EMBL Grenoble Outstation, BP 156, 38042, Grenoble, Cedex 9, France* [d]*Computational, Analytical & Structural Sciences, GlaxoSmithKline, Gunnels Wood Road, Stevenage, Hertfordshire, SG1 2NY. UK,* [e]*Sir Alexander Fleming Building, Imperial College of Science, Medicine and Technology, South Kensington, London SW7 2BZ, UK, @ Current address:PETRA III Project, EMBL Outstation, DESY Laboratory, Hamburg, Germany.* E-mail: john.helliwell@manchester.ac.uk

Keywords: Protonation states in proteins; X-rays; neutrons.

Biology rests on chemical reactions. Many chemical reactions involve hydrogens or protons. The determination of hydrogens, as is well known, is particularly challenging for protein crystallography. For X-rays as the probe, even with the best possible brilliance synchrotron radiation beamline, mobility of a hydrogen can render the diffraction signal invisible and that of a proton is by definition not visible to X-rays. For neutrons as the probe the issues are technical to do with converting hydrogens to deuteriums, by soaking in D_2O or perdeuteration biological expression, and/or the need for big crystals to be used with weak neutron fluxes. The protonation states of aspartic acids and glutamic acids as well as histidine are investigated in three X-ray cases; Ni, Ca concanavalin A at 0.94Å, a thrombin-hirugen binary complex at 1.26Å resolution and a cardiac myosin binding protein 'C1 domain' at tested resolutions of both 1.2 and 1.3 Å. The truncation of the Ni, Ca concanavalin A data at various test resolutions between 0.94 and 1.50Å provided a test comparator for the thrombin-hirugen and 'mycC0', 'unknown', carboxylate bond lengths. The complementary technique of using neutron protein crystallography has provided evidence of the protonation states of histidine and acidic residues in concanavalin A, and also the correct orientations of asparagine and glutamine side chains. Again the truncation of the neutron data at various test resolutions between 2.5 and 3.0Å, even 3.25 and 3.75Å resolution, examines the limits of the neutron probe. These studies indicate a widening of the scope of both X-ray and neutron probes in certain circumstances to elucidate the protonation states in proteins, and which will have basic and applied (ie Pharma-industrial) research potential.