FA1-MS12-O1

Recycling PDB: Automatic Structure Solution System-BALBES. Fei Long[a], A. Alexei Vagin[a], Paul Young[a], Garib N. Murshudov[a]. [a]*Chemistry Department, University of York, York, YO10 5YW, UK.*
E-mail: garib@ysbl.york.ac.uk

Automated pipelines for structure solutions show great potentials in the field of macromolecular crystallography but present daunting challenges in terms of their effectiveness and efficiency. This talk will describe BALBES[1], a molecular-replacement pipeline developed in York. The pipeline integrates three components. (1) A specifically designed internal database (DB). DB comes from tailoring the ever rich source of Protein Data Bank (PDB), in order to achieve efficient model search in molecular replacement. All entries in the PDB have been analysed according to sequence identities and three dimensional similarity and only non-redundant sets of protein structure were stored. The entries in DB are clustered into hierarchical trees based on sequence alignment. The relevant domains and tertiary information are presented in DB. (2) A manager system that employs collection of protocols and algorithms. (3) Core engine programs such as MOLREP, REFMAC and SFCHECK. BALBES requires as input only experimental data – sequence and the reflection data, requires no users' intervention. Systematic tests of BALBES were carried out using the newly released structure factor files. Overall success rate of tests is more than 75% and detailed analysis of the tests will be presented.
Tests show that combination of automatic molecular replacement with automatic model building procedures (e.g. ARP/wARP) shows potential of complete automation of structure solution.
Some of the newly implemented features such as automatic search for ensembles of search models and their use in molecular replacement will also be presented.

[1] Long F, Vagin A, Young P, Murshudov GN. "BALBES: a molecular replacement pipeline" *Acta Cryst* **2008**:D64;125-134

Keywords: molecular replacement; automation; webserver

FA1-MS12-O2

Challenges in Structure Validation - Going beyond the Protein. Robbie P. Joosten. *CMBI/NCMLS, Radboud University Nijmegen Medical Centre, The Netherlands.*
E-mail: rjoosten@cmbi.ru.nl

With the advent of high-throughput methods in macromolecular X-ray crystallography a wealth of structural data is becoming available. Because these methods rely highly on automation, this data is also increasingly uniform. This leads to new possibilities for statistical analysis of protein structures which should result in improved or novel means of protein structure validation.

Besides protein, a macromolecular crystal may contain other compounds such as nucleic acids, carbohydrates, ligands, chemical agents used in crystallization or cryoprotection, ions, and structured water. With many powerful validation tools available for protein (e.g. MolProbity [1] and WHAT_CHECK [2]), the challenge of X-ray structure validation is now to look beyond the protein. Some validation tools for non-protein entities exist (e.g. pdb-care [3] and ValLigURL [4]), but there are many examples in the current PDB that show there is still much work to be done.

[1] Davis I.W., Leaver-Fay A., Chen V.B., Block J.N., Kapral G.J., Wang X., Murray L.W., Arendall W.B. 3rd, Snoeyink J., Richardson J.S., Richardson D.C., *Nuc. Ac. Res*, **2007**, 35, 375-383. [2] Hooft R.W.W., Vriend G., Sander C., Abola E.E., *Nature*, **1996**, 381, 272. [3] Lütteke, T., von der Lieth, C.W., *BMC Bioinf.*, **2004**, 5, 69. [4] Kleywegt G.J., Harris M.R.., *Acta Cryst. D*, **2007**, 63, 935-938.

Keywords: structure validation; ions; protein structure

FA1-MS12-O3

DEA: The Combination of the DEDM-EDM Procedure with Automatic Model Building Packages to Solve Difficult Protein Phasing Cases. Dritan Siliqi[a], Rocco Caliandro[a], Benedetta Carrozzini[a], Giovanni Luca Cascarano[a], Carmelo Giacovazzo[a], Annamaria Mazzone[a]. [a]*Institute of Crystallography-CNR. Via G. Amendola, 122/O 70126 Bari, Italy.*
Email: dritan.siliqi@ic.cnr.it

Modern crystallography uses various techniques for phasing proteins (e.g., ab initio, SAD-MAD, SIR-MIR, MR or ab initio modelling), several tools for extending and improving phases [i.e. Electron Density Modification (EDM), Difference Electron Density Modification (DEDM [1]), hybrid direct methods] and automatic procedures to build and refine structural models (AMB). The final aim of the methodologists is to reduce as much as possible the manual efforts in all the steps of the phasing process. This presentation indicates how EDM, DEDM and AMB programs may be combined into a single procedure (DEA) [2] to automatize phase refinement and model building steps: it has been shown that DEA succeeds when different combinations of the single programs fail. In particular, it is shown that the iterated use of particular combinations of available tools may greatly increase the efficiency of the structure solution process. From the point of view of the execution time, DEA shows a supplementary practical advantage. Most of the computing time is spent for the iterated application of the AMB programs: the use of the DEDM- EDM cycles reduces the number of AMB applications, and thus dramatically reduces the total computing time. DEA has been included in the package IL MILIONE [3], which may be integrated by external AMB programs via suitable scripts. As outlook, we foresee that embedding DEDM-EDM cycles within AMB procedures can greatly increase their efficiency.

[1] Caliandro, R., Carrozzini, B., Cascarano, G.L., Giacovazzo, C., Mazzone, A.M. & Siliqi, D. **2009**. *Acta Cryst.* D65, 249-256.

[2]Caliandro, R., Carrozzini, B., Cascarano, G.L., Giacovazzo, C., Mazzone, A.M. & Siliqi, D. **2009**. *Acta Cryst*. D65, 000-000. [3] Burla, M.C., Caliandro, R., Camalli, M., Carrozzini, B., Cascarano, G.L., De Caro, L., Giacovazzo, C., Polidori, G. Siliqi, D. & Spagna, R. **2007**. *J. Appl. Cryst*. 40, 609-613.

Keywords: phase refinement; methods development; automatic structure solution

FA1-MS12-O4

APLx–Automated Protein-Ligand Crystallography Workflow. Romeu Pieritz[a], Leonard Leonard[a], Sean McSweeney[a]. [a]*European Synchrotron Radiation Facility, Grenoble, France.*
E-mail: pieritz@esrf.eu

Fragment-based approaches is a new paradigm for small-molecule drug discovery. The methodology is complementary to high-throughput screening experiments performed on modern X-ray facilities. The APLx project aims to develop an automated system to be used during the MX experiment for online data analysis. The system is designed to control and analyse the workflow to promote structure determination and screening inhibitors from a library of synthetic compounds. The first prototype executes in parallel for different data sets molecular replacement, structure refinement and ranks the structures obtained based on the "Rfree" index. The prototype is used to analysis the data during the MX experiment and can help the scientist to decide what it is the best data set for further studies. The automated workflow is used on the ESRF Macromolecular crystallography beamlines and results presented show the overall performance of the system. When completed the systematic use of this parallel workflow during the MX experiment will decrease the overall time to obtain a valid molecular structure including bound ligands. This research is funded by the SOUTH Consortium - 6th Framework Programme of the European Commission (LSH-2005-2.1.1-4).

Keywords: computational analysis of crystallographic data; automation; protein ligands

FA1-MS12-O5

Proteopedia: Scientific Wiki Bridging 3D Structure-Function. Joel L Sussman[a,b], Eran Hodis[c], Israel Silman[a,d], John Moult[f], Eric Martz[g], Jaime Prilusky[a,e]. [a]*The Israel Structural Proteomics Center.* [b]*Depts of Struct Biol.* [c]*Comp Sci & Applied Math,* [d]*Neurobiol.* [e]*Bioinformatics Unit, Weizmann Inst, Rehovot, Israel.* [f]*Center for Adv Res in Biotech, U MD Biotech Inst, Rockville, MD.* [g]*Dept of Microbiol, U MA, Amherst, MA.*
E-mail: Joel.Sussman@weizmann.ac.il

Rather than relying on printed text to provide the understanding of biomacromolecular structures, a collaborative website called *Proteopedia* provides a new resource by linking written information and 3D structural information [1]. *Proteopedia* displays protein structures and other biomacromolecules interactively. These 3D interactive images can be rotated and zoomed, and are surrounded by text with hyperlinks that change the appearance of the 3D structure to reflect the concept explained in the text. This makes the complex structural information readily accessible and comprehensible, even to non-structural biologists. Using *Proteopedia*, anyone can easily create descriptions of biomacromolecules linked to their 3D structures, *e.g.*:
(a) Proton Channels:
http://proteopedia.org/wiki/index.php/Proton_Channels
(b) HIV-1 protease:
http://proteopedia.org/wiki/index.php/HIV-1_protease (c) *Beta*-adrenergic receptor:
http://www.proteopedia.org/wiki/index.php/A_Physical_Model_of_the_β2-Adrenergic_Receptor
Aside from content added by the hundreds of registered users of *Proteopedia*, pages on each of the more than 56,000 entries in the PDB have been automatically created, and are primed for expansion by users. Members of the scientific community are invited to request a user account to edit existing pages and to create new ones. An account can be obtained from the homepage at http://www.proteopedia.org

[1] Hodis, E., Prilusky, J., Martz, E., Silman, I., Moult, J. & Sussman, J. L., *Genome Biol*, **2008**, 9, R121.

Keywords: molecular computer graphics; journal publication; computer-aided education