

a more focused approach, the JCSG is leveraging its HTSB platform to address more challenging targets and capitalize on our extensive experience to develop the best strategies to enhance chances of success. In parallel, we process our internal biomedical-theme targets in a HT mode. This project is centered the microbial communities that inhabit specific niches and environments of the human body to investigate the structural basis for host/commensal-microbe interactions. We are initially focusing on secreted proteins from commensal bacteria in the human gut to explore their symbiotic relationship with their human host. The gut microbiota is dominated by poorly characterized bacterial phyla, which contain an unusually high number of uncharacterized proteins and remain largely unstudied. Their influence upon human development, physiology, immunity, and nutrition are only starting to surface and, thus, represents an exciting new frontier for HTSB where we can investigate the contributions of these microorganisms to human health, as well as to disease. Supported by NIGMS: U54-GM094586

Keywords: structural genomics, high-throughput structural biology, human microbiome

MS58.P01

Acta Cryst. (2011) A67, C591

ARCIMBOLDO goes super: *ab Initio* phasing on the supercomputer Calendula FCSC

Isabel Usón,^{a,b} Dayté Rodríguez Martínez,^a Massimo Sammito,^a Kathrin Meindl,^a Iñaki Martínez de Ilarduya^a. ^a*Departament of Structural Biology, Instituto de Biologia Molecular de Barcelona (IBMB-CSIC) Barcelona Science Park, (Spain).* ^b*ICREA (Spain).* E-mail: uson@ibmb.csic.es

A supercomputer (FCSC: www.fcsc.es) provides the ideal environment for ARCIMBOLDO, as it opens new dimensions to the incorporation of prior knowledge, allowing to tackle increasingly difficult structures. Extensions in the method and its successes will be reported. It also makes the method accessible to users who do not have a grid.

Ab Initio phasing of macromolecular structures with no heavy atoms has been limited to cases with up to around 1000 atoms in the asymmetric unit, diffracting to atomic resolution [1].

Both the size and resolution barriers have been overcome in the case of several test and previously unknown structures. Thus, cases with a few thousand atoms, diffracting to 2Å have been solved through a combination of location of model fragments such as polyalanine alpha-helices with the program PHASER [2] and density modification with the program SHELXE [3]. Given the difficulties in discriminating correctly positioned fragments, the method has to test many putative groups of fragments in parallel, thus calculations are performed in a grid. The method has been called after the Italian painter Arcimboldo [4], who used to compose portraits out of fruits and vegetables. In the case of our program, most collections of fragments remain a “still-life”, but some are correct enough for density modification to reveal the protein’s portrait.

[1] G.M. Sheldrick, H.A. Hauptman, C.M. Weeks, R. Miller, I. Usón, *International Tables for Macromolecular Crystallography vol. F*, (eds., M.G. Rossmann and E. Arnold), Boston, **2001**, 333–345. [2] A.J. McCoy et al. *J. Appl. Crystallogr* **2007**, *40*, 658–674. [3] G.M. Sheldrick *Z. Kristallogr* **2002**, *217*, 644–650. [4] T. Tannenbaum, D. Wright, K. Miller, M. Livny, in *Beowulf Cluster Computing with Linux* (ed., T. Sterling) (MIT Press, Cambridge, Massachusetts, USA), **2002** 307–350 [4] D. Rodríguez, C. Grosse, S. Himmel, C. González, I. M de Ilarduya, S. Becker, G.M. Sheldrick, I. Usón, *Nature Meth.* **2009**, *6*, 651–654.

Keywords: *ab initio* phasing, macromolecule, supercomputing

MS58.P02

Acta Cryst. (2011) A67, C591

Accelerating *ab initio* phasing with *de novo* models

Kam Y. J. Zhang, Rojan Shrestha, Francois Berenger, Zhang Initiative Unit, Advanced Science Institute, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, (Japan). E-mail: kamzhang@riken.jp

The *ab initio* phasing is one of remaining challenges in protein crystallography. Recent progress in computational structure prediction has enabled the generation of *de novo* models with high enough accuracy to solve the phase problem *ab initio*. This “*ab initio* phasing with *de novo* models” method first generates a huge number of *de novo* models and then selects some lowest energy models to solve the phase problem using molecular replacement. The amount of CPU time required is huge even for small proteins and this has limited the utility of this method. Here, we describe an approach that significantly reduces the computing time required to perform the “*ab initio* phasing with *de novo* models”. Instead of performing molecular replacement after the completion of all models, we initiate molecular replacement during the course of each simulation. Our approach principally focuses on avoiding the refinement of the best and the worst models and terminating the entire simulation early once suitable models for phasing have been obtained. In a benchmark dataset of 20 proteins, our method is over two orders of magnitude faster than the conventional approach. We have observed that in most cases molecular replacement solutions were determined soon after the coarse-grained models were turned into full atom representations. We have also found that all-atom refinement could hardly change the models sufficiently to enable successful molecular replacement if the coarse-grained models were not very close to the native structure. Therefore, it remains critical to generate good quality coarse-grained models to enable subsequent all-atom refinement for successful *ab initio* phasing by molecular replacement.

Keywords: phasing, prediction, computation

MS58.P03

Acta Cryst. (2011) A67, C591-C592

Structure analysis of 2D membrane proteins using X-ray powder diffraction data

R. A. Dilanian,^a C. Darmanin,^b J.N. Varghese,^b S.W. Wilkins,^c T. Oka,^d N. Yagi,^e H. M. Quiney,^a K.A. Nugent^a ^a*School of Physics, The University of Melbourne, VIC 3010, (Australia).* ^b*CSIRO Preventative Health Flagship, CSIRO Molecular Health and Technologies, Melbourne, VIC 3052, (Australia).* ^c*CSIRO Materials Science and Engineering, PB33 Clayton Sth MDC, Melbourne, VIC 3169, (Australia).* ^d*Department of Physics, Faculty of Science and Technology, Shizuoka University, Shizuoka, 422-8529, (Japan).* ^e*Research and Utilization Division, SPring-8/JASRI, Hyogo 679-5198, (Japan).* E-mail: roubend@unimelb.edu.au

The majority of known integral membrane proteins (IMPs) have a natural propensity to form two-dimensional (2D) crystals during the crystallization processes [1]. This limits the possibility that their molecular structures may be obtained using the standard methods of protein crystallography. Powder diffraction methods are, in contrast, not critically sensitive to the quality and dimensions of crystals, which suggests their use in the structure analysis of protein crystals [2]. The application of powder diffraction methods for the structure analysis of proteins, however, is still regarded as intractable because of the large number of unresolved (overlapping) reflections. The development of new methodologies for powder diffraction structure analysis is, therefore, timely and desirable and could significantly expand the list

of biological molecules whose structures may be determined from x-ray diffraction data.

Here, we describe a novel approach for the structure analysis of 2D IMP crystals using x-ray powder diffraction data [3]. We apply our method to the recovery of the structure of the bacteriorhodopsin molecule to a resolution of 7 Å.

We use a priori information about unit cell lattice parameters, space group transformations and chemical composition in a bootstrap process that resolves the ambiguities associated with overlapping reflections. The measured ratios of reflections that can be resolved experimentally are used to refine the position, shape and orientation of low-resolution molecular structures within the unit cell, leading to the resolution of the remaining overlapping reflections. The molecular model is then made progressively more sophisticated as additional diffraction information is included in the analysis. Our approach can be used to provide reliable low-resolution phase information that can be further refined by the conventional methods of protein crystallography.

Keywords: integral membrane proteins, 2D crystals, powder diffraction

[1] M. Baker, *Nature Methods* **2010**, 7, 429-433. [2] R.B. Von Dreele, *Methods Enzymol* **2003**, 368, 254-267. [3] R.A. Dilanian, *et al.*, *Protein Science* **2011**, 20, 457-464.

MS58.P04

Acta Cryst. (2011) **A67**, C592

Towards more complete models in macromolecular crystal structure determination

Tim Wiegels,^a Marco Biasini,^b Victor Lamzin,^a ^a*European Molecular Biology Laboratory (EMBL) - Hamburg Unit, Hamburg (Germany).*

^b*Swiss Institute of Bioinformatics, Biozentrum, University of Basel, Basel (Switzerland).* E-mail: wiegels@embl-hamburg.de

Macromolecular machines play central roles in life processes and present important targets in biomedical and pharmaceutical research. However, determining the three-dimensional structures of large molecular assemblies is a challenging task in macromolecular crystallography (MX). Crystals of such structures rarely diffract to high resolution and often only noisy and inaccurate electron density maps are obtained. Computational approaches for model building in MX have historically been focused on high-resolution data, thus their application at lower than 3.0 Å resolution data is limited and typically results in incomplete and highly fragmented models. Hence, robust methods that would improve the completeness and the accuracy of models are urgently needed for the automated determination of low-resolution MX structures.

To address this aim within the ARP/wARP software project [1], we exploit the fact that 50% of all crystal structures deposited in the PDB [2] contain multiple copies of subunits or their assemblies in the asymmetric unit "non-crystallographic symmetry or NCS". We noticed that during automated model building with ARP/wARP, particularly in its initial steps, NCS-related parts of the structure are rarely built in exactly the same way. The reasons for that are manifold - including limited resolution of the data and poor initial phases. However, a beneficial side effect of differently built NCS-related copies is that each provides information that is not present in another copy; combining this (which we call *intrinsic*) information improves the model building process and increases the overall completeness of built structures at low resolution. The use of NCS during model building with ARP/wARP provides a significant improvement in many cases, and often requires less model-building cycles. In the best case, at 3.2 Å resolution, the model completeness improves from 55% to 73%, more side chains

can be docked in sequence, and the length of the built fragments increases.

Density that cannot be easily interpreted as part of a protein chain can be regarded as a poorly defined connection between two built chain fragments. Such connection generally contains not only loops but also helices or strands. In an approach using *complementary* information, we use structural fragments from the PDB for the interpretation of such density. Tests are currently underway, with very promising preliminary results. Rebuilding of 10-residue long gaps in high-resolution structures can already be achieved with an r.m.s.d. of under 0.5 Å. Further results and application to model building will be presented at the conference.

[1] G. Langer, S.X. Cohen, V.S. Lamzin, A. Perrakis, *Nature Protocols* **2008**, 3, 1171-1179. [2] P.W. Rose, *et al.*, *Nucleic Acids Research* **2011**, 39, D392-D401

Keywords: model building, automation, macromolecular crystallography

MS58.P05

Acta Cryst. (2011) **A67**, C592

ShelXle – A Qt GUI for SHELXL

Christian B. Huebschle,^a George M. Sheldrick,^b Birger Dittrich,^a ^a*Institute of Inorganic Chemistry, Georg-August-University of Goettingen (Germany).* ^b*Department of Structural Chemistry, Georg-August-University of Goettingen (Germany).* E-mail: chuebsch@moliso.de

ShelXle is a graphical user interface (GUI) for small-molecule least-squares refinements of single-crystal diffraction data with SHELXL[1]. It is designed like a integrated development environment and combines an editor with syntax highlighting and auto completer with a graphical representation of the three dimensional structure. ShelXle is a tool for expert users of SHELXL giving full control over the *.res/*.ins input file. Non expert users can rapidly learn how to appreciate the full capability of SHELXL by exploring the functionality of ShelXle. The electron density and difference electron density maps (F_o and $F_o - F_c$) can be visualized as wire framed isosurfaces. A 'rename mode' provides the ability to re label atoms including residues and/or parts and assigning free variables for occupation constraints. Molecules can be moved so that their centers of gravity lie inside the unit cell by just one click. If there is more than one chemical identical molecule present in the asymmetric unit then one can inherit labels semi-automatically from a previously labeled molecule. The 'auto HFIX' function uses electron density ($F_o - F_c$) for the placement of Hydrogen atoms with suitable constraints/restraints. For convenience functions to update the number of atoms in the cell (UNIT) and the weighting scheme (WGHT) are build in. A refinement history and a save history allow to go back to previous file versions.

The three dimensional representation of the molecule is drawn using OpenGL. Several stereoscopic projection modes are available including one for Zalman Monitors. The GUI is written entirely in C++ using Qt. Some tools are written in Fortran using Intel's MKL library. The program has been thoroughly tested prior to this launch. ShelXle is available for Windows (XP/Vista/7),

MacOS X (10.5/10.6), Linux (SuSE [11.1-11.4] / Debian) and as source. On all systems it is easy to install. ShelXle is licensed under LGPL(2.1) and can be downloaded free of charge at <http://ewald.ac.chemie.uni-goettingen.de/shelx/>.

[1] G.M. Sheldrick *Acta Cryst.* **2008**, *A64*, 112–122.

Keywords: small molecule refinement, graphics, interface