MS02 P01

Embedding the ESRF into the High-Throughput Gene-to-Structure Pipeline Joanne McCarthy[a] [a]*European Synchrotron Radiation Facility (ESRF), Grenoble, France.* E-mail: mccarthy@esrf.fr

Keywords: macromolecular crystallography; high-throughput; automation

Macromolecular crystallography (MX) is the most effective method of determining the high resolution structures of biological macromolecules. Each year many thousands of data collections that form the basis for such structure determinations are carried out at the ESRF. The nature of these data collections at synchrotron-based MX facilities is, in routine cases, rather repetitive. Indeed, it has become something of a cliché to describe the process of macromolecular crystal structure determination as a pipeline which can be broken down into a series of discrete steps with the completion of one step triggering the start of the next. At the ESRF these have been automated and linked together to produce a fully automatic data collection pipeline (DCP), a prototype of which is available on all the ESRF MX end-stations. The DCP begins with the scanning of the unique 2D datamatrices on SPINE standard sample holders contained in the SC3 sample changer robot [1]. These unique identifiers provide a means of tracking individual samples through the pipeline. For a set of samples chosen (up to 50), the DCP then carries out automatic screening, characterisation and ranking of crystal quality using the DNA software [2, 3, http://www.dna.ac.uk]. For the best or most appropriate samples, automatic collection, integration and scaling of data can then be carried out via the DCP. Results are stored in, and are accessible via, the ISPyB database (http://ispyb.esrf.fr) which allows the two-way communication of information between users' home labs and the ESRF beam-lines and has paved the way for remote access experiments.

[1] Cipriani, F. *et al*, *Acta Cryst.* 2006, *D***62**, 1251.
[2] Leslie, A. *et al*, *Acta Cryst.* 2002 *D***58**, 1924.
[3] Popov, A. N. and Bourenkov, G. P. *Acta Cryst.* 2003, *D***59**, 1145.

MS02 P02

Gene Composer, a Gene and Expression Construct Design Project Management Tool Peter Nollert, John Walchli, Mark Mixon, Alex Burgin. *Emerald BioSystems, Bainbridge Island, WA, USA*
E-mail: pnollert@emeraldbiosystems.com

Keywords: whole gene synthesis, expression construct design, software

With the goal of improving success rates for eukaryotic protein expression and crystallization, we are developing Gene Composer software to facilitate the information-rich design of protein constructs, their respective nucleic acid coding sequences and expression vectors. The Protein Design Module distills protein structure information from PDB files and comparative sequence information into an interactive alignment viewer. This graphical user interface allows the researcher to simultaneously visualize sequence conservation in the context of known protein secondary structure, ligand contacts, water contacts, crystal contacts, B-factors, solvent accessible area, residue property type and several other property views. The Gene Design Module automates the back-translation of a protein amino acid sequence into a codon-optimized nucleic acid sequence, which can be handed off to be synthesized. The Construct Design Module allows the user to define termini, make insertions or deletions, change residues, add tags, define cloning sites and finally permute those constructs and virtually clone them into one or multiple expression vectors of choice. Using this procedure, Gene Composer will generate all primers and mutagenic oligonucleotides necessary to perform all corresponding wet-lab procedures. We will present each of the Gene Composer modules and a detailed protocol for PCR-based gene synthesis from designed oligonucleotides.

MS02 P03

The Matrix Maker: A Liquid Handling Robot for High-Throughput Protein Crystallization Peter Nollert, Laurelin Ward, Natalie Duncan, Lance Stewart, Mark Mixon. *Emerald BioSystems, Bainbridge Island, WA, USA*
E -mail: pnollert@emeraldbiosystems.com

Keywords: high-throughput optimization, crystallization reagents, formulation

The Matrix Maker liquid handling robot has been developed to produce primary and secondary reagent matrices for protein crystallization. Controlled via customized database software, the Matrix Maker assists in the design of any formulation matrix and dispenses matrices from stock solutions. The system consists of positive displacement pumps capable of dispensing any formulation from up to 60 simultaneous stock solutions. Solutions can be dispensed into virtually any kind of vessel, from tubes to multichamber plates. Interchangeable pumps allow accurate dispensing of volumes from 10 ul to 50 ml. In its original application of making solutions for protein crystallization screening, the Matrix Maker allows a substantial reduction in liquid handling time compared with manual solution preparation. Formulation accuracy is still maintained, as shown by pH measurements and protein crystallization results. The performance and ease-of-use of the Matrix Maker make this instrument an essential component in modern high-throughput protein crystallization laboratories.

MS02 P04

Mail-in data collection at SPring-8 protein crystallography beamlines Nobuo Okazaki[a], Kazuya Hasegawa[a], Go Ueno[b], Hironomi Murakami[b], Masaki Yamamoto[a, b], [a]*SPring-8/JASRI,* [b]*RIKEN SPring-8 Center*
E-mail: okazaki@spring8.or.jp

Keywords: mail-in data collection, remote operation, web based application

The mail-in data collection system at SPring-8 makes it possible that distant users collect diffraction data without visiting SPring-8. The mail-in users only send samples to SPring-8 via home-delivery services as the first step. Then they can request measurement conditions and check results at their laboratory on the Web. The data collection for delivered samples are carried out with the automated beamline operation system using sample auto-changer

SPACE [1] and beamline control software BSS [2]. For smooth communication with distant users via the Internet, we have newly developed the data management system D-Cha (Database for Crystallography with Home-lab Arrangements) which mediates between the users and SPring-8 beamlines. D-Cha provides the GUI for users to deposit the experimental conditions for samples and to browse / download the collected data on web browser.
The mail-in system has been developed and operated for Structural Genomics Project at RIKEN Structural Genomics II (BL26B2) since September 2005. Then the system has been presented for public users at Structural Biology III (BL38B1) since December 2005 and at RIKEN Structural Genomics I (BL26B1) since December 2006. At BL26B1 and BL26B2, mail-in system has been operated on a daily basis. In addition, the commercial mail-in service has just started in July 2006, as the joint project among JASRI, RIKEN and analysis service companies. The mail-in data collection is our first step of remote beamline access at SPring-8. The next step is to achieve the fully remote control data collection based on the mail-in system.

[1] Ueno G, Hirose R, Ida K, Kumasaka T, Yamamoto M, *J Appl. Cryst.* 2004 Dec;37(Pt 6):867-873 "Sample management system for a vast amount of frozen crystals at SPring-8"
[2] Ueno G, Kanda H, Kumasaka T, Yamamoto M, *J Synchrotron Radiat.* 2005 May;12(Pt 3):380-4 "Beamline Scheduling Software: administration software for automatic operation of the RIKEN structural genomics beamlines at SPring-8"

MS02 P05

Recent advances in the Crank automated structure solution suite Navraj S. Pannu, Pavol Skubak, Irakli Sikharulidze, Jan Pieter Abrahams, Rudolf A.G. de Graaff. *Biophysical Structural Chemistry, Leiden University, Leiden, The Netherlands.*
E-mail: raj@chem.leidenuniv.nl

Keywords: automated structure solution, MAD, SAD

For its first release, the CRANK system was shown to effectively detect anomalous scatterers and phase SAD data [1]. Since then, CRANK's speed and robustness has improved, building many structures automatically for SAD, SIRAS, MAD and MAD + native data [2].
One improvement involves using Luzzati parameters refined in the program BP3 to validate the quality and completeness of a substructure obtained. This has proven to be effective and reliable in identifying correct solutions that do not meet the figure of merit levels reported in substructure detection programs necessary to safely assume that a correct solution was found. Thus, this allows for an early termination of substructure detection.
To improve automated model building, an interface to ARP/wARP and REFMAC has been added to also include SAD data directly in model refinement. The multivariate SAD likelihood function, implemented in a modified version of REFMAC has been shown to extend the resolution and phase quality limits required for automated model building with iterative refinement [3]. Recently, the new refinement target was shown to be very effective in combination with the SHELX[C/D/E] pipeline available in CRANK.
The above and other advances are in the latest version of CRANK at http://www.bfsc.leidenuniv.nl/software/crank or the CCP4 pre-release zone.

[1] Ness, S, de Graaff RA, Abrahams, JP, Pannu NS. (2004) *Structure*, 12, 1753.
[2]http://www.bfsc.leidenuniv.nl/software/crank/tests/pipeline1.html
[3] Skubak, P, Ness, S, Pannu, NS. *Acta Cryst D61*, 1626.

MS02 P06

Truncate2 - A Program for Intensity to Amplitude Conversion Norman Stein[a], Charles Ballard[a], [a]*CCP4, Daresbury Laboratory, Warrington, WA4 4AD, UK.*
E-mail: n.d.stein@dl.ac.uk

Keywords: software, protein structure function, structure solution

Conversion of measured intensities to structure factors in protein crystallography is complicated by the fact that background subtraction can result in negative intensities for weak reflections. Truncate2 is a new CCP4 program, designed to replace the original Truncate program, which uses Bayesian statistics to produce positive structure factor values from negative input intensities. Small positive intensities are also boosted by the conversion process. In addition, Truncate2 calculates a number of statistics from the intensity data, such as moments, cumulative intensity distributions and the Wilson plot. When output in graphical form, these can be used to assess data quality and to check for possible twinning. Truncate2 detects significant anisotropy in the data and performs anisotropy correction. A number of quantitative tests for twinning such as the H test and the Britton test have also been introduced. The prior distribution used in Truncate is the Wilson distribution, which is only appropriate in the absence of twinning and translational NCS. Truncate2 is capable of handling the last two cases in a more accurate manner.

MS02 P07

New developments for a full automation of the FIP beamline at the ESRF Jacquamet L., Bertoni A., Borel F., Charrault P., Israel-Gouy P., Iwema T., Kahn R., Joly J., Ohana J., Pirocchi M., Robin A., Serre L., Vernede X. and Ferrer J. L. *Institut de Biologie Structurale Jean Pierre Ebel, CEA; CNRS; Université Joseph Fourier; 41 rue Jules Horowitz,F-38027 Grenoble, France* .
E-mail: lilian.jacquamet@ibs.fr

Keywords: protein data collection, automatic control, robots

FIP (French beamline for the Investigation of Proteins) at the ESRF (European Synchrotron Radiation Facility) pushed developments in automation to reach a fully automated beamline.
- The energy adjustment and beam optimization are completely automated [1].
- The screening of the different protein crystals is ensured by a robotic system: CATS (Cryogenic Automated Transfer System) [2] now commercialized by IRELEC.
- In addition, this robot offers the possibility to analyze crystals directly as they grow in drops inside crystallization plates [3]. FIP has then developed in collaboration with GREINER Bio-One a new crystallization plate devoted to this new application.
- The centering of the protein crystal is improved using the installed UV-laser [4].
- The automation of the data recording and processing with ADP (Automated Data Processing) [5] has also been achieved.