

TB Structural Genomics Consortium (<http://www.doe-mbi.ucla.edu/TB/>) & The Integrated Center for Structure and Function Innovation: a PSI-2 Specialized Technology Center (<http://techcenter.mbi.ucla.edu/>).

[1] Cabantous S., Terwilliger T.C., Waldo G.S., & Pédelacq JD, *BIOforum Europe*, 2006, 10, 26-28.

[2] Cabantous S., Pédelacq J.D., Mark B.L., Naranjo C., Terwilliger T.C. & Waldo G.S., *Journal of Structural and Functional Genomics*, 2005, 6 (2-3): 113-119.

[3] Cabantous S., Terwilliger T.C., & Waldo G.S., *Nature Biotechnology*, 2005, 23 (1): 102-7.

MS02 O4

High throughput structural genomics and its general applicability to structural biology Ian A. Wilson, *Joint Center for Structural Genomics (JCSG), La Jolla, CA 92037 USA*. E-mail: wilson@scripps.edu

Keywords: Structural Genomics; Metagenomics; Microbiome

Over the past 7 years, the JCSG (www.jcsg.org) has evaluated, developed, adapted and integrated various methodologies and technologies into a high throughput (HT) production pipeline for all steps from target selection, cloning, expression, crystallization to structure determination. The pipeline which was initially developed using the full proteome screen of *T. maritima* as our main source of targets, now forms the base of our current production pipeline. As one of the 4 PSI production centers, we have focused on large Pfam families without structural coverage, as well as on proposing and validating new protein families, that have recently been added to Pfam. Our biomedical theme project revolves around, The Central Machinery of Life, proteins that are conserved in all kingdoms of life. New and exciting projects in our target portfolio are focused on metagenomes, in particular, the Global Ocean Sampling (GOS) and human microbiomes. To date, we have processed over 11,000 targets, solved over 435 structures and deposited over 405 in the PDB. The *T. maritima* effort has resulted in 163 structures (total unique structures in PDB: 282) that now account for a 15% direct structural coverage of its predicted soluble proteome. Moreover, of its 1877 predicted ORFs, only 95 proteins (5% of the predicted soluble) are currently without fold predictions. As a result of processing such a large number and variety of targets through our HT pipeline, we have been able to test and validate various strategies to increase efficiency, improve yield, and lower cost per structure. This large scale focused studies have resulted in adding a number of routine salvage pathways to our pipeline. Many of these advances in technologies and methodologies can easily be implemented on a smaller scale in individual structural

biology labs. The JCSG, located at The Scripps Research Institute, Genomic Institute of the Novartis Research Foundation, U.C. San Diego, Burnham Institute, and the Stanford Synchrotron Radiation Laboratory/Stanford University, is supported through the NIH Protein Structure Initiative (U54-GM074898), (www.nigms.nih.gov/psi).

MS02 O5

High throughput crystallography in fragment based drug discovery Mladen Vinković, *Astex Therapeutics, Cambridge, UK*.

E-mail: m.vinkovic@astex-therapeutics.com

Keywords: automated crystallization, automated ligand refinement, fragment screening

Fragment based drug discovery is becoming increasingly popular among pharmaceutical and biotech companies. Astex is a pioneer in that field and in the use of protein X-ray crystallography for fragment screening. The advantage of fragments, small molecules with $100 < M_r < 300$, is that they have a better chance of achieving high ligand efficiency than the larger molecules used in traditional HTS. In addition they can be grown to larger and potent molecules while keeping physical, chemical and pharmacological properties in check.

X-ray crystallography has been traditionally used in the lead optimization stage of the drug discovery process to clarify SAR and aid ligand design. Introduction of protein crystallography into the screening phase, where many hundreds of data sets are collected and processed, requires a reduction in the time for solution and refinement of protein-ligand complexes from days to hours. That has been achieved through integration of commercially available crystallization [1] and data collection robotics with our corporate database, as well as by development of an *in house* software platform for automated interpretation of diffraction data (AutoSolve) [2]. A web based GUI enables necessary process flexibility together with model inspection and rebuilding through Astex Viewer [3].

Such a high throughput platform enables the generation of 10s to 100s of hits (protein-fragment crystal structures) and structural protein-ligand elucidation of protein complexes for every relevant compound synthesized in subsequent hit and lead optimization steps, therefore greatly aiding ligand design. Rapid progress from mM hits to nM leads and subsequently drug candidates has been achieved on a number of targets.

[1] Vinković M., *Insights*, 2005, 19, 15.

[2] Mooij W.T.M., Hartshorn M.J., Tickle I.J., Sharff A.J., Verdonk M.L., Jhoti H., *Chem. Med. Chem.*, 2006, 8, 827.

[3] Hartshorn M.J., *J. Comput. Aided Mol. Des.*, 2002, 16, 871