obtained and 6 structures were determined. Among them, Bs139 protein functions as phosphoribosylglycinamide formyltransferase (GART), an important enzyme in the de novo pathway of purine biosynthesis. Bs139 crystal diffracted to 2.5 Å resolution at home X-ray source and the structure was determined by molecular replacement (MR). Bs154 protein is a putative deoxyuridine 5'-triphosphate nucleotidehydrolase (dUTPase), which plays important role in DNA replication. Se-YosS crystal diffraction datasets were collected at Beijing Synchrotron Radiation Facility (BSRF) and the structure was determined by multi-wavelength anomalous diffraction (MAD) method.

Keywords: structural genomics, *Bacillus subtilis*, nucleotide metabolism

P.04.22.7
Acta Cryst. (2005). A**61**, C257

ACTOR Gets an AGENT: Automation for Multiple Instruments

Angela Criswell, Kris Tesh, Charlie Stence, Wenjeng Li, John Ziegler, Terry Hahn, Keith Crane, Russ Athay. *Rigaku/MSC, Inc., The Woodlands, TX, USA*. E-mail: acriswell@RigakuMSC.com

Laboratory automation has improved the efficiency and capabilities of the modern crystallographer. To date, robotic methods are used for all steps of the crystallographic pipeline, from preparing solutions for crystal growth all the way to mounting crystals on goniometers and collecting diffraction data. Additionally, data processing has been automated to provide nearly immediate results to the scientist. With the need for crystal transport and data collection maturing, there has arisen the need to evaluate crystals and then choose to move collection worthy crystals to the most appropriate diffraction setup for full data set collection. ACTOR Gantry Enabling Numerous Targets, AGENT, has been added to the crystallographer's tool chest to fill this gap. AGENT allows ACTOR to mount crystals on multiple adjacent diffractometers. Crystals can then be evaluated, ranked, and, if they exceed a quality threshold, data can be collected. Based on the crystal diffraction properties, the best instrument and experimental parameters can be selected for data collection. Not limited to a single detector or generator, AGENT can outsource samples to multiple detectors on multiple generator platforms. Thus, AGENT provides the ultimate in high-throughput technology, while still retaining the scheduling flexibility required for optimal crystallographic data collection.

Keywords: automation, robots, structural genomics

P.04.22.8
Acta Cryst. (2005). A**61**, C257

Crystal Structures of Sortase B from *Staphylococcus aureus* and *Bacillus anthracis* Reveal Catalytic Amino Acid Triad in the Active site

Rongguang Zhang[1], Ruiying Wu[1], Grazyna Joachimiak[1], Sarkis K. Mazmanian[2,3], Dominique M. Missiakas[2,4], Piotr Gornicki[3], Olaf Schneewind[2,3], Andrzej Joachimiak[1], [1]*Structural Biology Center, Midwest Center for Structural Genomics, Argonne National Laboratory, 9700 South Cass Ave., Bldg 202, Argonne, IL 60439, USA*. [2]*Committee on Microbiology, University of Chicago, 920 E. 58th St., Chicago, IL 60637, USA*. [3]*Department of Molecular Genetics and Cell Biology, University of Chicago, 920 E. 58th St., Chicago, IL 60637, USA*. E-mail: rgzhang@anl.gov

Surface proteins of Gram-positive bacteria perform critical biological functions during the pathogenesis of human infections. These functions are only carried out when they are attached to the cell wall envelope. The anchoring process of the surface protein is accomplished by sortases via a transpeptidation reaction involving a C-terminal sorting signal containing a conserved five-amino acid motif. Sortase B recognizes NPQ**T**N in *S. aureus*, and NPKTG in *B. anthracis*, cleaves the polypeptide after the Thr residue and attaches the mature protein to the cell wall peptidoglycan. The catalytic mechanism for similar reaction has been proposed. Questions whether a thiol ion pair intermediate plays a key role in the sortase-catalyzed reaction and which residues constitute the active site remain unsolved. In this paper, we report 1.6 and 2.0 Å resolution crystal structures of SrtB from *B. anthracis and S. aureus,* respectively, provide a first detailed view of the active site and enables the design of new experiments with a goal to target the protein for new class of drugs that would inhibit cell wall anchoring in gram-positive bacteria.

Keywords: sortase, *Bacillus anthracis*, structural genomics

P.04.22.9
Acta Cryst. (2005). A**61**, C257

The Joint Center for Structural Genomics: A Multi-tiered Approach to Structural Genomics

Marc-A. Elsliger[1,6], A.M. Deacon[1,2], A. Godzik[1,3], P. Kuhn[1,6], S.A. Lesley[1,4], R.C. Stevens[1,6], K.O. Hodgson[1,2], J. Wooley[1,3,5], I.A. Wilson[1,6], [1]*The JCSG*. [2]*Stanford Synchrotron Radiation Laboratory*. [3]*The San Diego Supercomputer Center*. [4]*The Genomics Institute of the Novartis Research Foundation, California*. [5]*The University of California, San Diego*. [6]*The Scripps Research Institute*. E-mail: elsliger@scripps.edu

The JCSG (www.jcsg.org) has made substantial progress in the ultimate goal of HT structure determination by truly automated means. We have implemented a 3-tiered pipeline strategy where targets are characterized, categorized by behavior, and then processed in parallel by appropriate methods. Tier 1 is focused primarily on gathering experimental data on the proposed targets and is heavily reliant on complete automation and the processing of a large number of targets through initial crystallization trials. Targets are then either advanced towards structure determination by MAD or MR (Tier 2), or enter an appropriate salvage pathway (Tier 3). Salvage pathways have been developed which attempt to customize processing of smaller sunsets of targets through parallel processing methods. Such pathways include NMR and DXMS screening, mutagenesis, protein refolding, protein co-expression, baculovirus protein expression and orthologs. These strategies have been successfully applied to a prokaryote (*T. maritima*) and eukaryote (mouse) proteomes.

Collaborations with the scientific community are an important part of the development, production and dissemination aspects of the JCSG. These collaborations are created, supported, and dynamically managed to match the programmatic needs while maximizing the leverage of available resources. (NIGMS/PSI (P50-GM 62411).

Keywords: structural genomics, automation, high-throughput

P.04.22.10
Acta Cryst. (2005). A**61**, C257-C258

When Structures of Unknown Proteins are Determined, What is next?

Ludmilla Shuvalova[a], George Minasov[a], Shyamala Rajan[a], Joseph S. Brunzelle[b], Xiaojing Yang[a], Wayne F. Anderson[a], [a]*Dept. MPBC, Northwestern Univ. Feinberg School of Medicine, Chicgao,USA*. [b]*LS-CAT, Argonne National Laboratory, USA*. E-mail: l-shuvalova@northwestern.edu

As part of Midwest Center for Structural Genomics (MCSG), our current aim is to solve high-resolution protein structures with less than 30% sequence identity to known structures. This approach unavoidably brings the realization that a large fraction of protein targets will be functionally uncharacterized. Three-dimensional structures of such proteins may furnish insight into their function. In the following case study we present recently determined x-ray crystallographic structures of proteins representing this category.

The RBSTP1166 protein from *Bacillus stearothermophilus* consists of 216 amino acids and related sequences appear to occur in a very small range of species. Preliminary structural comparisons suggest the protein may be a glycoside hydrolase.

YfiT, a hypothetical protein from *Bacillus subtilis* is found to have a divalent cation bound by three conserved histidines.The localization of the metal, its coordination geometry, the surrounding residues and the ligands involved suggest that YfiT might function as a peptidase or hydrolase.

An outer surface protein from *Bacillus cereus* has a two-domain

structure. The large domain shows (βα)8 barrel motif and the small domain suggests structural similarity to cyclophilin A.

Keywords: structural genomics, high-resolution protein structures, domain structure

P.04.22.11

Acta Cryst. (2005). A**61**, C258

A High-efficiency, Low-cost Platform for Structural Genomics Studies at Peking University

Yu-He Liang[a], Lanfen Li[a], Xiaoyan Zhang[a], Dan Li[a], Xiaofeng Zheng[a], Yuhui Dong[c], Peng Liu[c], Dingchang Xian[c], Jie Nan[a], Jia-Wei Wang[b], Hai-Fu Fan[a,b], Xiao-Dong Su[a], [a]*Department of Biochemistry and Molecular Biology, College of Life Sciences, Peking University, Beijing 100871, China*. [b]*Institute of Physics*. [c]*Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China*. E-mail: su-xd@pku.edu.cn; liangyh@pku.edu.cn

A high-throughput, high-efficiency and low-cost platform based on Beckman-Coulter robotic system Biomek FX for structural genomics has been set up. Several projects of structural genomics are in processing. Now, the platform has a capacity to process more than 1000 genes/year for structural and functional analyses. *B. Subtilis,* a model ogranism for Gram-positive bacteria and *S. Mutans,* the primary pathogen of dental caries were selected as our main target sources. So far, more than 450 *B. subtilis* and 250 *S. mutans* proteins and some proteins from other sources were selected as targets for this platform, the selected targets are mainly related to important metabolism pathways, and/or of potential for drug design. Up to 2005 Jan., 20 protein structures from the selected targets were determined, among them, eight structures were determined ab-initio. The application of beamline at **BSRF** (Beijing Synchrotron Radiation Facilities) and the ***OASIS-2004*** program have been crucial components for the operation of our platform. The use of SAD (single-wavelength anomalous diffraction) phasing methods combined with direct methods in ***OASIS-2004*** has increased the efficiency significantly, 5 out of 8 ab-initio determined structures have been solved this way.

Keywords: structural genomics, BSRF, OASIS-2004

P.04.22.12

Acta Cryst. (2005). A**61**, C258

Progress in the Whole Cell Project of a Model Organism, *Thermus thermophilus* HB8

Akio Ebihara[a], Noriko Nakagawa[a,b], Mayumi Kanagawa[a], Shinya Satoh[a], Yoshihiro Agari[a], Nobuko Maoka[a], Hitoshi Iino[a], Aiko Kashihara[a], Chizu Kuroishi[a], Ryoji Masui[a,b], Mikako Shirouzu[a,c], Takaho Terada[a,c], Kunio Miki[a,d], Shigeyuki Yokoyama[a,c,e], Seiki Kuramitsu[a,b,c], [a]*RIKEN Harima Institute at SPring-8*. [b]*Department of Biology, Graduate School of Science, Osaka University*. [c]*RIKEN Genomic Sciences Center*. [d]*Department of Chemistry, Graduate School of Science, Kyoto University*. [e]*Department of Biophysics and Biochemistry, Graduate School of Science, University of Tokyo, Japan*. E-mail: ebihara@spring8.or.jp

One of the long-term goals of structural and functional genomics is the interpretation of all fundamental biological phenomena at atomic resolution. An extremely thermophilic bacterium, *Thermus thermophilus* HB8, is a promising model organism for structural and functional studies, because of the small genome size, the availability of genetic tools for functional analysis, and the thermostability of its proteins. Toward this aim, the "Whole Cell Project" of this bacterium is currently in progress (http://www.thermus.org/). The complete genome sequence identifies approximately 2,200 ORFs, and about 2,000 expression plasmids have been constructed. The target proteins were overproduced in *E. coli*, purified, crystallized, and characterized by X-ray crystallography, through which about 200 protein structures have been solved. As part of functional studies, we have constructed the gene disruption plasmids using the thermostable selective marker (kanamycin resistance) and analyzed mRNA by the DNA microarray system.

Keywords: structural genomics, functional genomics, *Thermus thermophilus*

P.04.22.13

Acta Cryst. (2005). A**61**, C258

Structural Proteomics of Proteins Coded by the *cag* PAI of *Helicobacter pylori*

Giuseppe Zanotti, Laura Cendron, Anke Seydel, Alessandro Angelini, Roberto Battistutta, *Department of Chemistry and VIMM, University of Padua, Italy*. E-mail: giuseppe.zanotti@unipd.it

H. pilory is a Gram-negative bacterium that colonizes the stomach of probably half of the human population. It is associated with gastritis, peptic ulcers and mucosa-associated lymphoid tissue lymphomas. Many factors contribute to the virulence of *H. pylori* [1]. Among them, the enzyme urease, the Neutrophil Activating Protein, NAP [2] and the secreted protein toxin VacA. However, the major genetic difference in HP isolates is the presence or absence of a specific pathogenicity island, named *cag*-PAI. It is a 40-kb locus that contains about 30 ORFs, whose function is unknown, with few exceptions.

We have cloned, expressed, and purified several proteins of the *cag* pathogenicity island of *H. pylori*. They all have been expressed in *E. coli.* We have already solved the structure of CagZ, using the Se-Met method [3] and the structure will be described in detail. We have also obtained crystals of a second protein, CagS, and its structure determination is in progress, along with crystallization tests on other *cag* proteins. Our final goal is to determine, in collaboration with other groups [4], most of the proteins coded by the *cag*-PAI island.

[1] Covacci, et al., *Science*, 1999, **284**, 1328. [2] Zanotti, et al., *J. Mol. Biol*., 2002, 323, 125-130. [3] Cendron L., Seydel A., Angelini A., Battistutta R., Zanotti G., *J. Mol. Biol*., 2004, **340**, 881. [4] *The Helicopter Structural and Molecular Biology Consortium.*

Keywords: structural genomics, bacterial pathogenesis, MAD phasing

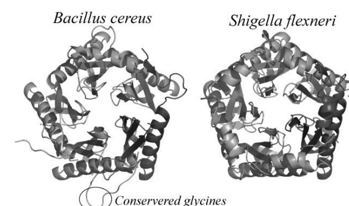
P.04.22.14

Acta Cryst. (2005). A**61**, C258

Trying To Deduce Function From Structural Variability And Conservation

Joseph S. Brunzelle[a], Xaiojing Yang[b], George Minasov[b], Ludmilla Shuvalova[b], Wayne F. Anderson[b], [a]*LS-CAT, APS, Argonne National Laboratory*. [b]*Dept of Mol Pharm and Bio Chem, Northwestern University*. E-mail: j-brunzelle@northwestern.edu

A comparative case is presented for two similar proteins from *Shigella flexneri* and *Bacillus cereus* which are homologs of the *E. coli* protein ybjQ. These two proteins are members of the COG0393, a widely conserved family of proteins in bacteria and archaea that are functionally uncharacterized. All members of the sequence family are about 100 residues. The two examples presented are both homopentamers and have 54% sequence identity. Despite the high sequence identity, the *B. cereus* protein contains regions of structural variation. A sequence alignment of the protein family reveals a pair of conserved glycines at residues 44 and 45. These conserved glycines are located in a loop that it is a region of structural variation in the *B. cereus* protein. This area of structural variation has been predicted as a region of disorder from the DisEMBL server which may be important to the function of these proteins.



Keywords: structural genomics, conformational change, macromolecular structure

P.04.22.15

Acta Cryst. (2005). A**61**, C258-C259

Comprehensive Structure-based Functional Analysis on Transcription Factors

Hiroshi Itou[a], Ui Okada[b], Min Yao[b], Nobuhisa Watanabe[b], Isao Tanaka[b], [a]*Biomolecular Structure Laboratory, National Institute of*