PLENARY LECTURES

Raymond Stevens
Professor
Department of Molecular Biology and Chemistry
The Scripps Research Institute
La Jolla, Ca 92037 USA

Structure and Function of the Human GPCR Superfamily

G protein-coupled receptors comprise the largest family of human proteins that communicate signals across the membrane and recognize millions of diverse molecules such as adrenaline, opioids, caffeine, dopamine, chemokines to name only a few. Over the past 10 years we have built a process pipeline to determine representative members of the G protein-coupled receptor phylogentic tree in order to understand the similarities and differences within this protein superfamily. In 2007, we solved the crystal structures of the human β_2-adrenergic receptor bound to the partial inverse agonist carazolol and timolol at 2.4 Å and 2.8 Å resolution and are now conducting NMR and HDX studies to understand receptor dynamics. In 2008, we determined the structure of the human adenosine A_{2A} receptor bound to the antagonist ZM241385 at 2.6 Å resolution. As part of a biotechnology start-up to use the technologies for specific drug discovery, Receptos has determined the structure of the human S1P1 receptor and now has a Phase I clinical trial underway. More recently, we have determined the structures of the human CXCR4 chemokine, human dopamine D3 receptor structures, and agonist structure of the human adenosine A2a receptor. The collective structures provide a high-resolution view of a human G protein-coupled receptor bound to diffusible ligands. Ligand-binding site accessibility is enabled by the extracellular loops which are held out of the binding cavity by a set of disulfide bridges and unique structural motifs. An exciting discovery is the role of cholesterol in receptor stability and potential function. Future studies include the determination of representative members from the different branches of the G protein-coupled receptor phylogenetic tree including class A, B, and C G protein-coupled receptors, as well as the receptors bound to agonists and G-proteins in an activated state.

This work was supported by NIGMS PSI:Biology for GPCR structure processing (U54GM094618) and the NIH Roadmap Initiative (JCIMPT) for technology development (GM073197).

Natalie Strynadka
Professor
Department of Biochemistry and Molecular Biology
Faculty of Medicine
University of British Columbia
Life Sciences Centre
2350 Health Sciences Mall
Vancouver, British Columbia
Canada V6T 1Z3

Structure-based analysis of protein transport across bacterial membranes

Bacteria have evolved several dedicated and sophisticated assemblies to transport proteins across their biological membranes. Recent advances in our understanding of the molecular details governing the specific actions of these protein secretion systems has come from an integrated approach of x-ray crystallography, NMR, mass spectroscopy, electron microscopy and in vitro reconstitution/ in vivo phenotypic analysis. Highlights of recent advances will be presented with an emphasis on that of the Type III Secretion system, the so-called bacterial “injectisome” encoded exclusively by pathogenic Gram negative strains including Salmonella typhimurium, Yersinia pestis, Pseudomonas aeruginosa and enteropathogenic Escherichia coli. A structure-based and genetic piecing together of the Type III Secretion System indicates that more than two dozen proteins assemble into a large needle shaped complex spanning the inner and outer bacterial membranes as well as that of the infected host cell, providing a direct conduit for the transport of essential bacterial virulence effectors from bacterial to host cytosol. A molecular understanding of the Type III systems being garnered from these studies provides the foundation for the development of new classes of vaccines and antimicrobials to combat these pathogens in the clinic and community.

SESSION 1: ADVANCES IN PROTEIN EXPRESSION, PURIFICATION AND CRYSTALLIZATION

Anthony Kossiakoff
Professor
Biochemistry & Molecular Biophysics
University of Chicago
929 East 57th Street
Chicago, IL 60637
CIS W210

Chaperone-Enabled Biology-Structure Technology Platform

M. Paduch, A. Koide, S. Rizk, J.Piccirilli, S. Sidhu, S. Koide, A. Kossiakoff
Department of Biochemistry and Molecular Biology
University of Chicago.

The Chaperone-Enabled Biology-Structure (CEBS) technology platform is built around generating and utilizing customized “Synthetic Affinity Binders” (sABs) as crystallization chaperones and high performance cell biology reagents. sABs are antibody-like molecules that are produced using phage display mutagenesis to bind tightly and specifically to a variety of protein and DNA target classes. Highly controlled phage display selection strategies allow for targeting sABs that can: bind to a predetermined region on a protein’s surface, trap a desired conformational state of a protein and to capture and stabilize a transient protein-protein complex. The power of the approach has been demonstrated through the determination of a number of structures of recalcitrant protein systems. sABs have also been shown to be superior cell biology reagents compared to traditional monoclonal antibodies for applications in imaging, pull downs and as perturbers of protein-protein interactions. A uniquely powerful feature of our approach is that by studying both structure and function using the same identical set of sABs, we can determine the functional importance of a specific region/conformation/complex that will provide researchers with a direct link between structure and function. These endeavors may stimulate new research initiatives into complex biological systems that have not been undertaken because of daunting technical challenges, but now the CEBS approach can clearly overcome.

Pär Nordlund
Structural Genomics Consortium and
Department of Medical Biochemistry and Biophysics
Karolinska Institutet
Stockholm SE-171 77

School of Biological Sciences
Nanyang Technological University
50 Nanyang Avenue
Singapore 639798

Understanding specificity and mechanism in human nucleotide metabolism and lipid signaling pathways

Human nucleotide metabolism and lipid signaling pathways are of outstanding importance for therapeutic developments target disease such as cancer, inflammation and the metabolic syndrome. We have pursued systematic structural and biochemical studies of human nucleotide metabolism and lipid signaling pathways to derive detailed understanding of the structural basis for regulation, catalysis and substrate recognition of these proteins, as well as their interaction mode with ligands and drugs. We have in the last 6 years solved 33 novel structures of enzymes in the human nucleotide metabolism and added mechanistic understanding to many of these proteins. This has lead to that the structural coverage of enzymes in human nucleotide metabolism has been dramatically improved and is approaching completion. Interactions studies of these proteins with the metabolome of intermediates of the nucleotide metabolism has identified novel potential regulatory feedback loops in this pathway with therapeutic implications.

Lipid based signalling is driven by enzyme networks that mediate the dynamic generation of lipophilic compounds such as ecoisanoids, phosphoinositides, and sphingolipids. Structural studies of these proteins are, however, often challenging when they are integral or interfacial to cellular membranes. We have recently solved 11 structures of enzymes in ecosanoid and phosphoinositide signalling pathways revelling modes for lipid and membrane recognition as well as regulation and catalysis.

As for all structural biology projects, the availability of suitable proteins for structural studies is the determining factor for success and strategies for production, stabilization and crystallisation of proteins in these pathways will be summarized. Also, some recent developments of directed evolution strategies for stabilizing challenging proteins will be discussed

Jennifer L. Martin
Associate Professor
Institute for Molecular Bioscience
University of Queensland
Brisbane, Queensland, Australia 4072

Structure and function studies of macrophage proteins using the UQSG pipeline

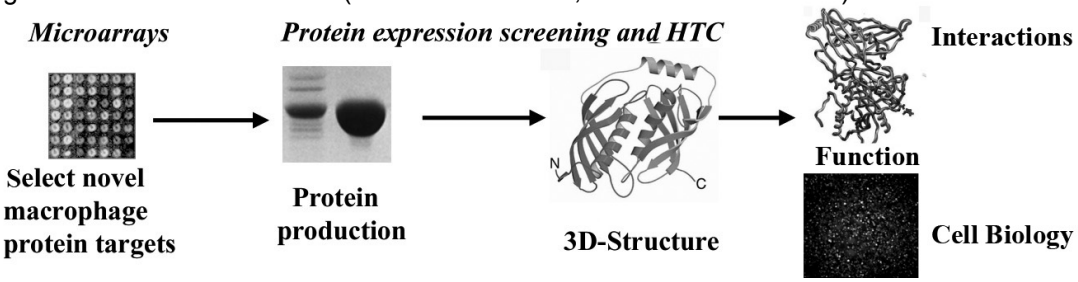
Kai-En Chen[1], Felicia Goh[1], Juliana Arrifin[1], Matthew Sweet[1], Stuart Kellie[2], Parimala Vajjhala[2], Justine Hill[2], Bostjan Kobe[2,1] and Jennifer L Martin[1,2]

[1] Institute for Molecular Bioscience, Division of Chemistry and Structural Biology,
[2] School of Molecular and Microbial Sciences, University of Queensland, Brisbane, Queensland, Australia

At the University of Queensland (UQ) in Australia, we established a small-scale focused pipeline that allows the parallel processing of dozens of protein targets (Cowieson et al 2005, J Struct Funct Genomics 6:13-20). The UQ pipeline intentionally focuses on protein targets that are likely to be biomedically significant. Thus, the resulting structural data informs on function for uncharacterized proteins and, in the long-term, will provide the starting point for the structure-based design of drugs.

Our approach is to link gene expression analysis by cDNA microarrays (to guide target selection) with the parallel processing of proteins for structural studies. The combination of expression profiling and structural analysis provides a powerful way to identify protein function, as the former suggests a cellular role (eg involvement in a pathway), while the latter suggests a molecular or biochemical function (eg an enzymatic activity). At the same time, this strategy retains the cost-effective nature of pursuing the technically more tractable protein structures first.

Our targets are proteins from macrophages, because these cells are abundant and play a central role in immunity and inflammation. Through array profiling, hundreds of proteins have been identified, including many of unknown function, that are expressed in a restricted manner in the macrophage lineage or are induced by pro-inflammatory stimuli. These molecules are considered to be candidate regulators of innate immunity (Wells et al 2003, Genome Res 13:1360-5). Full length cDNAs encoding proteins expressed in mouse macrophages are available because of the extensive polling of the macrophage transcriptome through the FANTOM consortium (Carninci et al 2005, Science 309:1559-63).



Importantly, we have substantial in-house macrophage expertise at UQ so that new targets can be evaluated for regulation of expression, subcellular localization and effect on macrophage function. The functional data adds substantial value to the derived crystal structure information (Aagaard et al 2005, Structure 13:309-17). We have also initiated alternative approaches to crystallography to gain structural information, including SAXS, Synchrotron CD, NMR, chemical cross-linking and mass spectrometric analysis (Mouradov et al 2006, Protein Eng Des Sel 19:9-16; Forwood et al 2007, Proc Natl Acad Sci USA 104:10382-87; Cowieson et al 2008, Proteins 70:1142-6; Cowieson et al 2008, J Biol Chem 283: 16187-93; Cowieson et al 2008, Curr Opin Struct Biol 18:617-22).

This is an academic program of research, where a researcher chooses their list of targets using set selection criteria and works with these targets through cloning, protein expression screening and protein purification and applying the UQSG parallel processing pipeline methods. Work on 12 macrophage protein targets by one PhD student, Kai-En Chen, will be presented, from target selection through to structure determination and functional analysis.

Jan Steyaert
Vrije Universiteit Brussel
pleinlaan 2, 1050
Brussel, Belgium

Conformation selective nanobodies for GPCR research.

Juan José Fung1,4, Els Pardon2,3,4, Søren G F Rasmussen1,4, Brian K Kobilka1 & Jan Steyaert2,3

1. Department of Molecular and Cellular Physiology, Stanford University School of Medicine, 279 Campus Drive, Stanford, Palo Alto, California 94305, USA.
2. Department of Molecular and Cellular Interactions, Vlaams Instituut voor Biotechnologie (VIB), Vrije Universiteit Brussel, B-1050 Brussels, Belgium.
3. Structural Biology Brussels, Vrije Universiteit Brussel, B-1050 Brussels, Belgium.
4. These authors contributed equally to this work.

The complex behavior of GPCRs in respons to natural or synthetic ligands and proteins can be attributed to the receptor's structural plasticity manifested in multiple functionally distinct conformational states. Obtaining crystals of GPCRs in an agonist bound, active-state has proven to be challenging due to the instability of this state in the absence of G protein. Here we describe the generation of camelid single-domain antibody fragments (nanobodies) that have G protein like properties towards the human β 2 adrenergic receptor (β 2AR). To generate receptor specific nanobodies, a llama was immunized with purified agonist bound β2AR reconstituted at high density into phospholipid vesicles. A library of nanobody clones was generated and screened against agonist bound β2AR. Seven clones were identified that recognize the agonist bound but not the inverse agonist bound receptor. One of these nanobodies was found to increase the affinity of β2AR for agonists 30-fold and to induce conformational changes at TM6, two effects that are indistinguishable from those observed upon adding G protein to the receptor. This nanobody that faithfully mimics the effects of Gs binding was used to obtain diffraction quality crystals and to solve the first structure of an active agonist-bound state of the human β 2 adrenergic receptor.

References

Fung, J. J., Pardon, E. , Rasmussen, S.G.F, Kobilka, B.K., & Steyaert, J. (2011) Camelid Single-chain Antibody Fragments Stabilize a G Protein-Coupled Receptor Active State, Submitted. Rasmussen, S.G.F., Choi, H.-J., Fung, J.-J., Pardon, E., Casarosa, P., Seok Chae, P., DeVree, B.T. , Rosenbaum, D.M., Thian, F.S., Kobilka, T.S., Schnapp, A., Konetzki, I., Sunahara, R.K., Gellman, S.H., Pautsch, A. , Steyaert, J., Weis, W.I., Kobilka, B.K. (2011) Structure of a nanobody-stabilized active state of the β2 adrenoceptor. Nature 363, 446–448.

Cristina Viola
EMBL Grenoble
6, rue Jules Horowitz
BP 181
38042 Grenoble Cedex 9, France

Expression strategies for protein complex production

Lakshmi Sumitra Vijayachandran, Cristina Viola, Frederic Garzoni, Simon Trowitzsch, Christoph Bieniossek, Maxime Chaillet, Christiane Schaffitzel, Didier Busso, Christophe Romier, Arnaud Poterszman, Timothy J. Richmond and Imre Berger

It is well known that many vital biological functions carried out by the cell are more often performed by multiprotein complexes rather than by single proteins. A major challenge in the field of structural biology is to obtain atomic resolution of proteins in the context of the multiprotein assemblies whereby they carry out their task. Protein complexes, especially those found in eukaryotic hosts, tend to be present only in very low quantities in the cell. Since relatively large quantities of homogeneous material are required for analysis by biophysical methods such as X-ray crystallography, heterologous overproduction is often necessary. In our group, we have developed the ACEMBL and MultiBac baculovirus expression systems to recombinantly produce multiprotein complexes in either E. coli or insect cells. We also addressed parallelization and automation of gene assembly for multiprotein complex expression by developing robotic routines for multigene vector generation. We introduced several improvements of baculovirus expression system performance including modifications of transfer plasmids, methods for generating composite multigene baculoviral DNA and simplified and standardized expression procedures. In order to address the problem of stoichiometrically imbalanced subunits, we developed a polyprotein expression strategy for producing multiprotein complexes. Together, these developments provide a useful toolbox for tackling challenging protein complexes.

SESSION 2: STRUCTURAL GENOMICS OF INTEGRAL MEMBRANE PROTEINS

Wayne A. Hendrickson
Professor
Howard Hughes Medical Institute
Biochemistry & Molecular Biophysics
Columbia University
New York, NY 10032

New York Structural Biology Center
89 Convent Avenue
New York, NY 10027

Membrane Protein Structures from a Structural Genomics Approach

The New York Consortium on Membrane Protein Structure (NYCOMPS) employs a structural genomics approach for studies on membrane proteins. During Phase 2 of the Protein Structure Initiative (PSI-2), NYCOMPS developed an automated system at the New York Structural Biology Center (NYSBC) for the cloning, expression and purification of membrane proteins, and we produced a number of crystal structures from that work. These structures derived from bioinformatic expansions of seeds from *Escherichia coli* membrane proteins into a battery of genomes from which recombinant cDNAs could be cloned for expression testing in *E. coli*. Screening and initial purification is centralized at NYSBC; structure determination efforts are distributed to the participating laboratories. Now as a PSI-Biology Center for Membrane Protein Structure Determination, NYCOMPS is focusing its efforts on expansions from human membrane proteins and on membrane proteins implicated in human metabolic disorders, and proteins are expressed in eukaryotic systems as well as in *E. coli*. I will report on the experiences of NYCOMPS investigators in structural analyses of membrane proteins and on structure-inspired tests of biochemical and cellular function.

Robert Stroud
Professor
Molecular Structure Group (MSG)
UCSF
Genentech Hall, Rm. S412C
600 16th. Street
San Francisco
California 94158-2517

Approaches to the Structures of Eukaryotic Membrane Proteins

Robert Stroud, John Pak, William Harries, Avner Schlessinger, Louis Metzger, Andrew Waight, Bjørn Panyella Pedersen, Zygy Roe-Zurz, Sarika Chaudhary, Larry Miercke, Joseph O'Connell III, Yaneth Robles, Rebecca Robbins, Meseret Tessema, James Holton, Andrej Sali.

Guided by methods of proteomics (1,2), new technology for expression that we have sought to develop (3-5), novel crystallization schemes, and protein engineering, we seek to understand mechanisms aimed at the solute transporter families, channels and transporters. In one case the structure of a human Rh factor (6) shows it to be a channel that conducts ammonia as NH3, defining a paradigm for how these channels remain insulating to water, protons or any ions. The structure of a human water channel from the brain (7) shows how it develops specificity for water, and remains insulating against all ions, even hydronium ions or protons. Current goals include attempts to determine the structure and mechanisms of drug resistance transporters. These include sodium or proton driven secondary transporters

References

1. Kelly L, Pieper U, Eswar N, Hays FA, Li M, Roe-Zurz Z, Kroetz DL, Giacomini KM, Stroud RM, Sali (2009) A survey of integral alpha-helical membrane proteins *J Struct Funct Genomics.* **10,** 269-280 (NIHMS150748)

2.Schlessinger A, Matsson P, Shima JE, Pieper U, Yee Sw, Kelly L, Stroud RM, Ferrin TTE, Giacomini KM, Sali A. (2010) Comparison of human solute carriers. *Protein Sci.* **19**, 412-28. (PMID: 20052679)

3.Lee JK, Stroud RM. (2010) Unlocking the eukaryotic membrane protein structural proteome. *Curr Opin Struct Biol* **20,** 464-70.

4. Hays FA, Roe-Zurz Z, Li M, Kelly L, Gruswitz F, Sali A, Stroud RM. (2009). Ratiocinative screen of eukaryotic integral membrane protein expression and solubilization for structure determination. *J Struct Funct Genomics* **10:** 9

5. Hays F, Roe-Zurz Z, Stroud R. M. Overexpression and Purification of Integral Membrane Proteins in Yeast (2010) *Methods in Enzymology* Chapter **29** p 697- 709

6. Gruswitz, F., Chaudhary, S., Ho, J.D., Schlessinger, A., Pezeshki, B., Ho, C-M, Holton, J., Sali, A., Westhoff, C.M., Stroud, R.M. (2010) Function of human Rh based on structure of RhCG at 2.1 A *Proc Natl Acad Sci* U S A **107**, 9638-9643

7. Ho, J. D., Yeh, R., Sandstrom, A., Chorny, I., Harries, W. E., Robbins, R. A., Miercke, L. J., & Stroud, R. M. (2009) Crystal structure of human aquaporin 4 at 1.8 A and its mechanism of conductance. *Proc Natl Acad Sci U S A* **106,** 7437- 7442.

Elisabeth Carpenter
Principal Investigator
Integral Membrane Protein Group
Structural Genomics Consortium
University of Oxford
Old Road Campus Research Building
Old Road Campus
Roosevelt Drive
Headington
Oxford OX3 7DQ
UK

Human integral membrane protein structural biology at the SGC

Integral membrane proteins represent approximately 25% of the genes in the human genome and are the targets for more than 50% of all therapeutic drugs. The number of structures available for this important class of proteins is however still very limited due to the difficulty of expressing, purifying and crystallising these hydrophobic proteins. Human membrane proteins are particularly under-represented with only 21 structures solved. At the SGC we are tackling this challenging area by putting together a pipeline for production of human membrane proteins using the baculovirus expression system. To date we have screened over 200 human integral membrane proteins for expression and purification in the Baculovirus system. For each protein we start with a series of constructs, screen for proteins that can be purified on a small scale, test a variety of detergents to see which gives stable, monodisperse protein and then we purify the protein on a 3L scale. To date we have obtained over 25 proteins that can be purified on a scale of at least 50 ug per litre and we have crystals of three targets. We have recently obtained diffraction to beyond 3.4A from crystals of a human ABC transporter and the structure has been phased using a mercury derivative, giving interpretable maps. The initial structure and interpretation will be presented and discussed.

Oliver P. Ernst
University of Toronto
Department of Biochemistry
1 King's College Circle
Toronto, ON

Crystal structure of metarhodopsin II

Hui-Woog Choe, Yong Ju Kim, Jung Hee Park, Takefumi Morizumi, Emil F. Pai, Norbert Krauß, Klaus Peter Hofmann, Patrick Scheerer & Oliver P. Ernst

Charité - Universitätsmedizin Berlin, Germany
Chonbuk National University, Korea
Queen Mary University of London, UK
University of Toronto, Canada

Rhodopsin is a G-protein-coupled receptor (GPCR), in which the inactivating ligand 11-cis-retinal is covalently bound in the seven-transmembrane helix bundle of the apoprotein opsin. After light-induced cis/trans-isomerization of the ligand, the Schiff base linking retinal and Lys296 deprotonates to form the active Meta II state. To overcome the problem of subsequent rapid Schiff base hydrolysis and all-trans-retinal dissociation, we formed Meta II for structural analysis by soaking opsin crystals with all-trans-retinal. Here we present the 3.0 Å and 2.85 Å crystal structures, respectively, of Meta II alone or in complex with a C-terminal fragment derived from the α-subunit of the heterotrimeric G protein. The Gα fragment binds in a large crevice at the cytoplasmic side and identifies the reversibly formed Meta II as the active, G-protein-interacting state. By comparison with rhodopsin and early photoproducts we propose how retinal translocation and rotation induces the gross-conformational changes characteristic for Meta II. The structures can now serve as models for the large GPCR family.

Isabel Moraes
Membrane Protein Laboratory
Imperial College
South Kensington Campus
London, UK SW7 2AZ

Challenges and Opportunities in Structure Determination of Membrane Proteins

I. Moraes, S. Iwata, G. Evans

Diamond Light Source Ltd, Imperial College London, UK

Membrane proteins are the supreme example where more effort is needed in structural biology. In spite of their abundance and importance, of the more than 60,000 protein structures in the Protein Data Bank, only around 200 of these structures represent unique membrane proteins. To facilitate structural studies on membrane proteins, The Membrane Protein Laboratory at Diamond Light Source Ltd (Diamond-MPL) was created in a collaboration between Imperial College London and Diamond, funded by the Wellcome Trust. The MPL is a research and training state-of-the-art user facility open to scientists from laboratories anywhere in the world interested in solving the 3-dimensional structures of membrane proteins by X-ray crystallography. Because membrane proteins are unstable, hard to crystallise and crystals difficult to handle. More systematic approaches and technical developments are needed to improve the success rate of the structure determination of membrane proteins. The MPL has a formal collaboration with I24 microfocus beamline at Diamond to develop new techniques for crystallisation and structural determination of membrane proteins, including prototype systems for high throughput methods, improving handling of small and delicate crystals, and methods for collecting and merging data from a large number of small crystals. The beamline is unique in its ability to deliver a tunable X-ray beam (6 – 25keV) of variable size between 5 μm and 50 μm on to crystal samples. This versatility is coupled with a state-of-the-art pixel array detector.

SESSION 3: INFECTIOUS DISEASE RESEARCH I: VIRAL AND PARASITIC DISEASES

Raymond Hui
Principal Investigator
Enabling Technology & Parasitology
Structural Genomics Consortium
University of Toronto
MaRS Centre, South Tower
101 College St., Suite 700
Toronto, ON, M5G 1L7
Canada

Using Structural Genomics to Study Druggable Targets in Protozoan Parasites

The genomes of protozoan parasites contain many genes conserved in humans, as well as those distinctive from all known genes. Amongst the conserved genes are those that are essential and potential targets of novel drugs desperately needed. Studying such targets comes with challenges and advantages. On one hand, finding compounds that selectively inhibit parasite proteins and not their human homologues can be a difficult problem. On the other hand, available compounds can be used to compare the chemical sensitivity of the parasite and human proteins. The latter is particularly true for proteins that are targets for well studied human diseases, including heat shock proteins and protein kinases. We expressed a number of HSP90 paralogues and protein kinases from kinetoplastid and apicomplexan parasites, screened them against family-specific inhibitor libraries and, where possible, determined their 3D structures by means of crystallography. The resulting chemical profiles and structures enable us to assess the druggability of specific HSPs and protein kinase, and also demonstrate structural genomics is a highly efficient approach for characterization of potential drug targets.

David Stuart, Oxford, UK
MRC Professor of Structural Biology
Division of Structural Biology
Nuffield Department of Clinical Medicine
University of Oxford
Henry Wellcome Building for Genomic Medicine
Oxford, OX3 7BN, UK

Methods for highish throughput crystallography applied to infectious and non-infectious virus particles

Bruno Coutard
Architecture et Fonction des Macromolécules Biologiques
AFMB UMR 6098
CNRS, Universités Aix-Marseille I & II
Case 932
163 Avenue de Luminy
13288 Marseille cedex 9 France

Structural Genomics on viral replicative machinery: A tool for antiviral discovery.

RNA viruses infecting human and animal can have a devastating impact on health and economy, as exemplified by the SARS-CoV or H5N1 viruses. New potential drug targets against these RNA viruses can be identified through comprehensive structural characterization of the replicative machinery. Since these viral proteins are naturally produced in the host cell as large polyproteins and since virus production can require the use of high bio-safety level facilities, the structural genomics strategy relies on the production of recombinant proteic domains through easy-to-handle expression systems. This presentation will describe tools that were devised within the VIZIER consortium (http://www.vizier-europe.org/) to produce viral proteins for crystallographic and functional studies, and that could be, for most of them, applied to any kind of soluble proteins. The resulting crystal structures often represent the first structural data obtained on neglected virus families. Since the corresponding proteins are usually essential for the viral replication, these structures are valuable information for antiviral design. As a consequence, the initial structural genomics project significantly shifted towards and paved the way to antiviral drug development.

Johannes Bauer
University of Tuebingen
Interfakultaeres Institut fuer Biochemie (AK Stehle)
Hoppe-Seyler-Str. 4,
D-72076 Tuebingen
Germany

The GD1A Glycan is a Cellular Receptor for Adenoviruses Causing Epidemic Keratoconjunctivitis

Johannes Bauer1, Emma C. Nilsson2, Rickard J. Storm2, Sara Spjut3, Susanne M. C. Johansson2, Aivar Lookene4, Mikael Elofsson3, Niklas Arnberg2,5 and Thilo Stehle1,6

1. Interfaculty Institute of Biochemistry, University of Tuebingen, Germany
2. Division of Virology, Department of Clinical Microbiology, Umeå University, Sweden
3. Department of Chemistry, Umeå University, Sweden
4. Department of Chemistry, Tallin University of Technology, Estonia, 6 Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, USA
5. Laboratory for Molecular Infection Medicine in Sweden (MIMS), Umeå University, Sweden

Adenovirus type 37 (Ad37) is a leading cause of epidemic keratoconjunctivitis (EKC), a severe and highly contagious ocular disease. Whereas most other adenoviruses infect cells by engaging CD46 or the coxsackie and adenovirus receptor (CAR), Ad37 binds to sialic acid-containing cell surface molecules. However, the identity of the Ad37 receptor has remained unknown. Using glycan array screening, we found that the receptor-recognizing knob domain of the Ad37 fiber protein specifically binds to a branched hexa-saccharide (the carbohydrate portion of the GD1a ganglioside) that features two terminal sialic acids. Soluble GD1a glycans and GD1a-binding antibodies efficiently prevented Ad37 virions from binding to and infecting human corneal cells. X-ray crystallographic studies reveal that the two terminal sialic acids dock into two of three previously established sialic acid-binding sites in the trimeric Ad37 knob. Based on the fact that the fiber knob protein contains three sialic acid binding sites a trivalent inhibitor was designed. This glycoconjugate efficiently inhibited binding of Ad37 to human corneal cells as well as Ad37 infection. The co-crystal structure of the Ad37-inhibitor complex demonstrates that the inhibitor engages now all three binding sites of the Ad37 fiber protein. Thus, biological competition experiments and structure-function analysis of the virus-inhibitor interaction show that such molecules may be suitable for topical treatment of EKC.

SESSION 4: AUTOMATED AND ADVANCED TECHNIQUES FOR PROTEIN STRUCTURE DETERMINATION

Thomas Terwilliger
Los Alamos National Laboratory
Mailstop M888
Los Alamos, NM 87545, USA

University of Washington
Department of Biochemistry
Seattle, WA, 98195, USA

X-ray crystal structure determination with PHENIX and Rosetta

Terwilliger, T. C., DiMaio, F., Read, R.J., Baker, D.

We have combined the power of Rosetta structure modeling with Phenix automated molecular replacement, model-building, density modification, and refinement to yield a new general-purpose and easy-to-use tool for crystallographic structure determination. Molecular replacement (MR) solutions are obtained with phenix.automr, rebuilt with Rosetta including electron density map information, and then are further rebuilt with phenix.autobuild. The combination of Rosetta rebuilding and phenix rebuilding is the key part of this method. This combination merges the benefits of structure-modeling, in which homology models can now be created that are more accurate than the templates used to create them, with crystallographic structure determination and refinement, in which models are built that are consistent with measured crystallographic structure factors. These methods are now all accessible through the Phenix tool "phenix.mr_rosetta". Phenix.mr_rosetta can be very useful for cases where the search model used in molecular replacement is slightly too distant to rebuild successfully with phenix.autobuild. It can also be useful in cases where the model is too distant to even find a molecular replacement solution, and pre-refinement with Rosetta can yield an improved search model.

The PHENIX software is available at http://www.phenix-online.org and Rosetta is available at http://depts.washington.edu/uwc4c/express-licenses/assets/rosetta

Thomas Szyperski
UB Distinguished Professor
Director, UB High-Field NMR Facility
Adjunct Senior Researcher at the Hauptman-Woodward Medical Research Institute
State University of New York at Buffalo
816 Natural Sciences Complex
Buffalo, New York 14260-3000

Development and Application of NMR Methodology for HTP Protein Structure Determination

The Northeast Structural Genomics Consortium (NESG; http://www.neg.org) has contributed more than 400 NMR structures to the Protein Data Base (http://www.pdb.org). For such HTP structure determination, the development of methodology focuses on increasing (i) the data acquisition speed, along with (ii) spectral resolution and precision of chemical shift measurements, (iii) the reliability of semi-automated data processing and analysis, including structure calculations, and (iv) the precision and accuracy of NMR structures. Recent developments and the impact on the structure determination of selected targets are presented, and future perspectives are discussed.

Wladek Minor
Professor of Molecular Physiology and Biological Physics
University of Virginia
PO Box 800736
Charlottesville, VA, 22908

HKL-3000: Toward the Future of Structural Biology

HKL-3000 integrates data collection, data reduction, phasing, and model building to significantly accelerate the process of structure determination, and on average, minimize the number of data sets and crystals required for structure solution. Execution of the package merges several modules and software applications into the structure determination pipeline. There are modules for experimental control of some beamlines and home instruments, data reduction, phasing by SAD/MAD or molecular replacement, fast model building, and initial structure refinement. The system is being developed and tested in the high-throughput environment of the Midwest Center for Structural Genomics (MCSG), the Center for Structural Genomics of Infectious Diseases (CSGID) and the New York Structural Genomics Research Consortium (NYSGRC). The robustness of HKL-3000 has improved considerably over time and currently over 1500 structures have been determined with it.

The continuous advancements in the decision-making procedures within HKL-3000 have made it the system of choice for MCSG and CSGID projects. Transforming raw images into a solved structure (with 70% of the model built) in 10-15 minutes is no longer a surprise, but a routine operation, for crystals that diffract to 2.5 Á or better. Our experience with the determination of hundreds of structures by experimental phasing methods helped us to establish rules for the best approaches when the available data fall into three categories: unsolvable with current data, borderline and easy. Our current work concentrates on improving the approach to borderline cases of structure determination rather than optimizing intermediate calculations for easy cases, thus shifting borderline cases into the easy category, and unsolvable ones into borderline.

An important observation is that simple experimental protocols are sufficient in most cases and may even be optimal for the most challenging ones. Feedback from fast preliminary structure solution has proved to be one of the critical components of success.

References

Minor W, Cymborowski M, Otwinowski Z, Chruszcz M (2006) Acta Crystallographica Section D: Biological Crystallography 62: 859-66.

Kirillova O, Chruszcz M, Shumilin IA, Skarina T, Gorodichtchenskaia E, Cymborowski M, Savchenko A, Edwards A, Minor W (2007) Acta Crystallographica Section D: Biological Crystallography 63: 348-54.
Otwinowski Z, Borek D, Majewski W, Minor W (2003) Acta Crystallographica. Section A: Foundations of Crystallography 59: 228-34.

Zheng H, Chruszcz M, Lasota P, Lebioda L, Minor W (2008) Journal of Inorganic Biochemistry 102(9): 1765-76.

Chruszcz M, Wlodawer A, Minor W (2008) Biophysical Journal 95(1): 1-9.

Wlodawer A, Minor W, Dauter Z, Jaskolski M (2008) Febs Journal 275: 1-21.

Otwinowski Z, Minor W (1997) Methods in Enzymology 275: 307-326.

Patrick Shaw Stewart
Douglas Instruments Ltd
Douglas House, Westfield Farm
East Garston
near Hungerford, Berkshire
RG177HD, UK

Random microseeding: a theoretical and practical exploration of the Microseed Matrix Screening (MMS) method, with new recommendations for achieving crystallization success.

Patrick D. Shaw Stewart, Richard A. Briggs, Stefan A. Kolek, Peter Baldock

Douglas Instruments Ltd

Douglas Instruments has completed theoretical and practical studies of Microseed Matrix Screening (MMS), a powerful approach to protein crystallization, introduced by Ireton and Stoddard in 2004 [1]. The method was automated by D'Arcy et al. [2], who first used seeding with random screening kits. Experience shows that MMS with random screens generates more hits or better-diffracting crystals in about 75% of cases where at least one crystal can be obtained. Our studies included analysis of the stability of seed crystals suspended in unconventional stock solutions and the effectiveness of seed crystals from novel sources. This has given rise to new recommendations for maximizing the number of hits, reducing the prevalence of salt crystals, increasing the diversity of crystal forms, and harvesting seed crystals from microfluidic devices. We also investigated nucleation with microporous glass, zeolites, precipitates and mixtures of crushed crystals of proteins that were unrelated to the target protein. These were less effective than conventional MMS, but may still be useful - they can of course be used before the first hits have been found. Throughout the project, the statistical significance of experiments was increased by focusing on "pregnant" conditions - defined as conditions that reliably gave crystals when seeds were present, but which otherwise gave no crystallization.

References:

[1] Ireton and Stoddard, Acta Crystallographica section D60 (2004), 601–605;
[2] D'Arcy et al. Acta Cryst (2007). D63

SESSION 5: STRUCTURE-ENABLED BIOLOGY

Ian A. Wilson
Department of Molecular Biology & Skaggs Institute for Chemical Biology
The Scripps Research Institute
La Jolla, CA 92037 USA.

High Throughput Structural Biology Applied to Challenging Biological Systems

The Joint Center for Structural Genomics (JCSG) high-throughput (HTP) platform, assembled over the past eleven years, delivers large numbers of protein structures to the community by both x-ray crystallography and NMR on a wide range of targets from bacteria to human, including challenging targets, such as eukaryotic proteins, protein-protein and other macromolecular complexes.

Our primary mission as a PSI:Biology Center for High-Throughput Structure Determination is to provide a robust, flexible HTP structure determination pipeline that meets the challenges and embraces the new opportunities that arise from PSI:Biology Partnerships projects, future PARS, and the community. The JCSG is leveraging its HTP platform to promote the biological and biomedical impact of our structures through extensive functional and bioinformatics analyses both internally and via collaborations within the PSI:Biology network as well with the general scientific community.

Using a more focused approach, in close collaboration with our assigned Partnership Centers, we are pursuing extremely challenging targets, such as human protein-protein and protein-nucleic acid complexes, where we capitalize on our extensive experience over the past decade to develop the best strategies to enhance chances of success. In parallel, we continue to process our biological/biomedical theme and structure coverage targets in a HTP mode. The microbial communities that inhabit specific niches and environments of the human body influence human development, physiology, immunity, and nutrition represent an exciting frontier for structural genomics where we can investigate the contributions of these microorganisms to human health and well being, as well as to disease. We are also continuing to contribute to the original PSI mission of advancing structural coverage of the expanding protein universe.

As part of the PSI:Biology Network, we strive to promote widespread use of PSI resources, materials, methodologies and data to the general scientific community. The JCSG also continues to develop new technologies and methodologies, both experimental and computational, to address a wide spectrum of targets, while keeping cost per structure to a minimum and quality to the highest standard. Supported by NIGMS: U54-GM094586.

Dino Moras
Research Director, CNRS
Institut de Génétique et de Biologie Moléculaire et Cellulaire
1, rue Laurent Fries
ILLKIRCH, FRANCE 67400

The Role of DNA in Transcription Regulation by Nuclear Receptors: An Integrated Structural Biology Approach

Transcription regulation requires a coordinated action of nuclear receptors (NRs) and a variety of cofactors complexes. Physiologically, NRs bind to hexanucleotide DNA motifs called hormone response elements (HREs). While detailed structural information is available on the isolated DNA and ligand binding domains little is known on the overall topology of the full length receptors. The structure of the full-length NRs in different functional states to unravel the relevant architectures and the role of DNA in the organization is a still missing crucial information. The question has been addressed using integrated structural biology approaches.

Shigeyuki Yokoyama
Director
RIKEN Systems and Structural Biology Center
1-7-22 Suehiro-cho
Tsurumi-ku, Yokohama
Kanagawa 230-0045
Japan

Structural Basis for Specific Aminoacyl-tRNA Synthesis

In protein biosynthesis, each codon is translated into its specific amino acid by aminoacyl-tRNA. For most of the amino acids used in translation, their cognate aminoacyl-tRNA synthetases (aaRSs) synthesize aminoacyl-adenylate from the amino acid and ATP, and then transfer the aminoacyl moiety to the 3′-terminal adenosine of the cognate tRNA, depending on strict recognition of both amino acid and tRNA. In most cases, the anticodon of tRNA serves as the major determinant for the recognition by the corresponding aaRS. However, several aaRSs recognize other part of tRNA for specific aminoacylation. Alanyl-tRNA synthetase (AlaRS) and histidyl-tRNA synthetase (HisRS) recognize the G3:U80 wobble base pair and the –1 guanosine, respectively. We have determined crystal structures of AlaRS and HisRS in complex with their cognate alanine and histidine tRNAs, respectively, which reveal the unique mechanisms of tRNA recognition. On the other hand, for some amino acids, the aminoacyl-tRNA is not synthesized in the canonical, direct manner, but indirectly synthesized by conversion of the aminoacyl moiety of another aminoacyl-tRNA. For example, glutaminyl-tRNA is synthesized from glutamyl-tRNA by “transamidosome” in most of microbes. The selenocysteine, the twenty-first amino acid in protein synthesis, is synthesized from the serylated form of its specific tRNA. We will present structural basis of specific tRNA recognition in these indirect aminoacyl-tRNA synthesis mechanisms.

Yunyu Shi
Hefei National Laboratory for Physical Sciences at Microscale and School of Life Science
University of Science and Technology of China
Hefei, Anhui 230026
People's Republic of China

Structural basis of epigenetic regulation

Yunyu Shi, Jihui Wu

We started structural genomics (SG) project ten years ago, which supported by Chinese academy of science. This project is biological-functional driven. Initially, targets are mainly selected from human hematopoietic stem/progenitor cells (HSPC). Since epigenetic regulations play a major role in cellular differentiation, we focus our research to proteins related with chromatin remodeling, histone modification, histone chaperon as well as small RNA regulation. We also interested in the molecular mechanism and structural basis of epigenetic regulation for stem cell programming and reprogramming. Both X-ray crystallography and NMR techniques were used. Examples will be given in the talk. Rtt106p is a Saccharomyces cerevisiae histone chaperone with roles in heterochromatin silencing and nucleosome assembly. Structural analysis of Rtt106p reveals a DNA binding role required for heterochromatin silencing. MOZ (monocytic leukemic zinc-finger protein) is histone acetyltransferases. It is important for HOX genes expression and embryo development. BRPF2 is a scaffold protein with multiple domains and form complex with MOZ. We studied domain structures and histone tails binding specificities for BRPF2 and reveal the biological significance of these bindings. Hfq is a bacterial Sm/Lsm post-transcriptional regulator. It can alter RNA structure to help RNA paring and thus called RNA chaperon. Structural analysis by both X-ray crystallography and NMR reveal E. Coli. Small RNA, DsrA binding and ATPase activity of Hfq requires cooperation of distal and proximal binding faces from adjacent hexamer.

Fumiaki Yumoto
University of California, San Francisco
600 16th GH S416A
San Francisco, CA 94158

Structural insight into transcriptional coactivation of LRH-1 by beta-catenin

Fumiaki Yumoto1, Phuong Nguyen1, Elena P. Sablin1, John D. Baxter2, Paul Webb2, Robert J. Fletterick1

1 University of California San Francisco
2 The Methodist Hospital Research Institute

The crystal structure of a complex of the armadillo repeat region of beta-catenin and the nuclear receptor liver receptor homolog 1 (LRH-1 or NR5A2) is solved at 2.8 Å resolution. The surface of beta-catenin for interaction of LRH-1 partly overlaps defined contact sites for peptide segments of beta-catenin partners, including T-cell specific transcription factor 4 (TCF-4). The surface of LRH-1 that engages beta-catenin is comprised of helices 1, 9 and 10 and is distinct from known interaction surfaces of LRH-1 and proteins, including the corepressor and coactivator binding site associated with activation function 2 (AF-2). Targeted mutagenesis of amino acid residues forming both sides of the LRH-1/beta-catenin interface reveals that it is essential for forming stable interactions between these proteins in mammalian cells and GST-pull down assay in vitro. Mutagenesis of the analogous binding site on the androgen receptor (AR) affects its association with beta-catenin, providing evidence that the observed LRH-1/beta-catenin interaction may be prototypic. The structure of LRH-1/beta-catenin opened new insight into transcriptional regulation of nuclear receptor by beta-catenin.

Session 6: Structure Enabled Understanding of Signaling Networks & Systems Biology

Titia Sixma
Groupleader
Division of Biochemistry
Netherlands Cancer Institute, Plesmanlaan 121,
1066 CX Amsterdam,The Netherlands

Regulation of ubiquitin specific proteases

Alex Faesen[*], Mark Luna-Vargas[*], Pim van Dijk[*], Andrea Alfieri[*], Reggie Ekkebus[#], Farid el Oalid[#], Annette Dirac[&],, Huib Ovaa[#], Titia Sixma[*]

[*]Divisions of Biochemistry, [#]Cell Biology and [&]Molecular Carcinogenesis, Netherlands Cancer Institute, Amsterdam, The Netherlands, t.sixma@nki.nl

Ubiquitin conjugation is a central regulatory system in the cell. It consists of target protein modification with one or more ubiquitins which lead to a change in fate for the target protein. A cascade of E1,E2,E3 proteins is responsible for the modification but this is balanced by active deubiquitinated enzymes (DUBs). Since both the modification and the readout of ubiquitination takes place in a a combinatorial fashion through weak interactions analysis of questions regarding specificity and regulation are very challenging. We have made use of a number of structural genomics approaches to characterise a series of DUBs of the Ubiquitin specific protease family. We will compare enzymatic activities, chain specificity and the regulation by ubiquitin-like domains.

Frank Sicheri
Professor
Department of Molecular and Medical Genetics
University of Toronto

Senior Investigator
Samuel Lunenfeld Research Institute
Mount Sinai Hospital
600 University Ave.
Toronto, Ontario, M5G 1X5
Canada

Structural principles of signal transduction

Lucia Banci
Professor of Chemistry
CERM and Department of Chemistry
University of Florence
Via L. Sacconi 6
50019 Sesto Fiorentino (Florence)
Italy

Mechanistic Systems Biology: from structures to functional pathways

Functional processes require the concerted action of a number of proteins and biomolecules which must have the suitable conformation, be located in the proper cellular compartment, and be able to interact with each other in the correct mode. In order to achieve these conditions, proteins, after they are synthesised, must undergo a number of processes to reach their mature, functional state. The process of folding requires a number of steps, often involving interaction with other proteins which facilitate the process and guide it in the correct direction. Furthermore, the folding process should occur in the cellular compartment(s) where the proteins function, which then defines the nature of the interacting proteins and the mode of interaction [1]. Some folding processes are tightly linked to redox reactions where disulfide bond formation, occurring upon interaction with specific chaperones and through a redox relay system, induce secondary and tertiary structure formation.

Most of the interactions along a functional process must be transient in nature as the process must proceed along various steps. Among these processes, metal transfer from metal transporters to the final recipient proteins occurs through a series of protein-protein interactions so that the metal ion is transferred from one protein to the next. [2,3] These transfer processes are determined by metal affinity gradients among the various proteins, with kinetic factors contributing to the selectivity of the processes [4].

A number of pathways responsible for copper trafficking in the cell and for the folding of the involved proteins will be presented with a particular focus on mitochondria. The features of the protein-protein complexes and the fact that their interaction is only metal-mediated [5], as well as the structural properties which modulate protein-protein recognition and metal specificity will be presented and discussed within an integrated approach where, from single structures to protein complexes, entire processes are described in their cellular context.

1] Banci, L., Bertini, I., Cefaro, C., Cenacchi, L., Ciofi-Baffoni, S., Felli, I. C., Gallo, A., Gonnelli, L., Luchinat, E., Sideris, D. P., and Tokatlidis, K., Molecular chaperone function of Mia40 triggers consecutive induced folding steps of the substrate in mitochondrial protein import, Proc.Natl.Acad.Sci.USA, 107, 20190-20195, 2010

2] Banci L, Bertini I, Cantini F and Ciofi-Baffoni S. Cellular copper distribution: a mechanistic systems biology approach. Cell Mol Life Sci: 67, 2563-2589, 2010.

3] Banci L, Bertini I, McGreevy KS and Rosato A. Molecular recognition in copper trafficking. Nat Prod Rep 27: 695-710, 2010

4] Banci L, Bertini I, Ciofi-Baffoni S, Kozyreva T, Zovo K and Palumaa P. Affinity gradients drive copper to cellular destinations. Nature 465: 645-648, 2010

5] Banci L, Bertini I, Cantini F, Felli IC, Gonnelli L, Hadjiliadis N, Pierattelli R, Rosato A and Voulgaris P. The Atx1-Ccc2 complex is a metal-mediated protein-protein interaction. Nat Chem Biol 2: 367-368, 2006

Fuyuhiko Inagaki
Biomedical and Pharmaceutical Science Course
Laboratory of Structural Biology
Hokkaido University
N21, W11, Kita-ku, Sapporo, Japan

Structural Biology of Selective Autophagy

Fuyuhiko Inagaki, Yoshinori Ohsumi, Nobuo N Noda

Faculty of Advanced Life Science

Autophagy is an intracellular bulk degradation system in which isolation membranes enclose a portion of cytoplasm to form a double membraned vesicle called an autophagosome and deliver its inner constituents to the lytic compartments(1). Atg8 and its mammalian homolog LC3 are ubiquitin-like modifiers that are localized on isolation membranes as a conjugate with phosphatidylethanolamine and play crucial roles in the formation of autophagosomes. Although autophagy is in principle a non-selective, starvation-induced response, recent studies have also shed light on another mode of autophagy that selectively and constitutively degrades aggregated proteins, surplus or damaged organelles, and even invasive bacterial cells. Defects in autophagy cause the accumulation of ubiquitin-positive protein inclusions, leading to severe liver injury and neurodegeneration. p62 functions as a receptor to link such protein inclusions to LC3, thus playing a pivotal role in their efficient engulfment into autophagosomes. This process is similar to the cytoplasm-to-vacuole targeting (Cvt) pathway in yeast; a model of selective autophagy in which Atg19 functions as a receptor to link a huge assembly of vacuolar enzymes, aminopeptidase I and ƒÑ-mannosidase to Atg8, thus playing a pivotal role in their efficient engulfment into Cvt vesicles. Using NMR and X-ray crystallography, we report here the structural basis for LC3-p62 and Atg8-Atg19 interactions. Remarkably, LC3 and Atg8 were shown to interact with p62 and Atg19, respectively, in a quite similar manner; they recognized the side-chains of Trp and Leu in a four-amino acid motif, WXXL, in p62 and Atg19 using hydrophobic pockets conserved among Atg8 homologs. Together with mutational analyses, our results show the fundamental mechanism that allows Atg8 homologs, in association with WXXL-containing proteins, to capture specific cargo molecules, which endows isolation membranes and/or their forming machineries with target selectivity. We also showed that WXXL motif is used to recruit Atg8/LC3 to facilitate efficient conjugation reaction with phosphatidylethanolamine.

(1) Noda, N., Ohsumi, Y. and Inagaki, F. (2009) Chemical Review, 109, 1587-1598

John R. Cort
Pacific Northwest National Laboratory
PO Box 999
MS K8-98
Richland WA
99352 USA

Structural and Functional Characterization of Pathogen Effector—Host Protein Interactions

John R. Cort(1), Joshua N. Adkins(1), Fred Heffron(2), Roslyn N. Brown(1), Eric D. Merkley(1), Ernesto S. Nakayusu(1), Kristen M. Meyer(1), Sayed Imtiaz(1), and PSI:Biology partners in the Midwest Center for Structural Genomics

(1)Pacific Northwest National Laboratory
(2)Oregon Health & Science University

The interface between pathogenic bacteria and the host cells they infect is defined by numerous secreted effector proteins that interact with host-cell biomolecules to modulate cell function along a trajectory that varies in time and space. Understanding the molecular details of these interactions will enhance understanding of infectious bacterial diseases. Previous research in our laboratory has identified many novel putative secreted effector proteins in Salmonella Typhimurium, and we are now identifying their host cell protein targets and characterizing these interactions and their functional consequences in the PSI:Biology project Program for the Characterization of Secreted Effector Proteins. Our approach combines high resolution mass spectrometry with chemical cross-linking and affinity pulldown experiments to discover and validate host target proteins of effectors in Salmonella as well as other bacterial pathogens. We are also conducting in vivo transfections and transductions to express effector proteins in macrophage and epithelial cell lines, together with proteomic analysis to identify host phenotypes elicited by the pathogen. These experiments will allow identification of potential interacting host proteins for uncharacterized effectors, and will enable development and testing of functional hypotheses for these effectors. Complementing these studies are high resolution structures of pathogen effectors, host targets, and in ideal cases, the resulting protein complexes determined by our PSI:Biology partners, the Midwest Center for Structural Genomics. In this presentation we will describe our approaches to target identification and selection, identification of interacting proteins, and functional characterization of these interactions, and we will present results from the initial set of targets from Salmonella Typhimurium to be tested with our protocols.

SESSION 7: STRUCTURAL GENOMICS AND DRUG DISCOVERY I

James C Sacchettini
Professor of Biochemistry, Biophysics and Chemistry
Department of Chemistry
Texas A&M University
College Station, TX 77843-3255

Antimicrobial Drug Discovery and Structural Genomics

Despite over sixty years of drug treatments available, mortality from bacterial infections continues to increase at a significant rate. The failure to control many bacterial infections infections is due to a number of factors that complicate treatment, including the expense of delivering treatment to infected individuals and complex treatment regimens, the emergence of drug-resistant and multi-drug resistant strains, and importantly the lack of novel drugs. Several groups have come to the conclusion that target-based approaches have not worked well for antibacterial drug discovery. Given the scope of the problem, it is critical to develop new approaches to improve our ability to discover novel and safe antibacterials.

Steve Almo
Professor
Departments of Biochemistry and Physiology & Biophysics
Albert Einstein College of Medicine
Jack and Pearl Resnick Campus
1300 Morris Park Avenue
Ullmann Building, Room 409
Bronx, NY 10461

Sequence, Structure, Function, Immunity

Cell surface receptors and adhesion molecules are gatekeepers of cellular function, as they are responsible for the detection and integration of signals arising from the extracellular milieu. In adaptive and innate immunity, these molecules underlie the initial recognition and ultimate destruction of foreign pathogens and malignancies, and at the same time are critical components of the tolerance mechanisms that protect the host from harmful autoimmune responses. Soluble versions of these receptors and their cognate ligands, as well as monoclonal antibodies (mAbs) targeted against these proteins represent a major class of protein therapeutics for the manipulation of immune responses to treat a wide range of infectious diseases, autoimmune diseases and malignancies.

We describe our growing program focusing on the comprehensive structural and functional analysis of immune function. This program has three major underlying premises: **1)** simple strategies can be developed that identify those candidate molecules for which structure determination would be particularly informative in terms of mechanism and biology; **2)** structures of receptors, ligands and their complexes can be used to develop mutant molecules with altered affinities, selectivities and oligomeric properties and **3)** this structural and biochemical information can be leveraged to develop new murine model systems to generate *in vivo* structure-function relationships and new therapeutic opportunities. We highlight several examples from the Immunoglobulin and Tumor Necrosis Factor Superfamilies.

Ekaterina Kouznetsova
SGC Stockholm, MBB
Karolinska Institute
Scheeles väg 2
17177 Stockholm
Sweden

Crystal structures of human tankyrase in complex with novel inhibitors

Ekaterina Kouznetsova (1), Elisabet Wahlberg(1), Tobias Karlberg(1), Natalia Markova(1,2), Ann-Gerd Thorsell(1), Åsa Frostell-Karlsson(2), Delal Öncü(3), Björn Kull(3), Herwig Schüler(1) and Johan Weigelt(1)

(1)SGC, Karolinska Institutet, Scheeles väg 2, 17177 Stockholm, Sweden.
(2)GE Healthcare, Bio-Sciences AB, 75184 Uppsala, Sweden.
(3)ACTAR AB, 17717 Stockholm, Sweden.

Tankyrases (TNKS1/PARP5a/ARTD5 and TNKS2/PARP5b/ARTD6) belong to a family of ADP-ribosyltransferases (PARPs), enzymes that catalyze ADP-ribosylation of a number of nuclear and cytosolic proteins. ADP-ribosylation is a reversible post-translational modification that affects a variety of fundamental cellular processes, such as transcriptional regulation and maintenance of genome stability. Tankyrases are better known for their role as positive regulators of telomere elongation. Tankyrases also act as positive regulators of Wnt signaling through regulation of axin levels. Inhibitors of tankyrase catalytic activity have a potential for use as therapeutic agents in cancers and also as research tools in the study of Wnt-signaling and DNA repair pathways. We assembled a ~200 compound collection of previously known PARP inhibitors and chemically related molecules. We used surface plasmon resonance and thermal shift assays to profile these compounds against human PARP catalytic domains in general and the tankyrases in particular. A number of verified hit compounds were assayed for their effect on autocatalytic activity of tankyrases. The result is a selectivity profile of some widely used research tools. In addition, we solved crystal structures of TNKS2 in complex with six compounds from this collection. These structures point to important structural features of tankyrase-ligand interactions, and they can serve as a basis for rational design of more potent and selective inhibitors.

Guido Hansen
Institute of Biochemistry
University of Lübeck
Ratzeburger Allee 160
23538 Lübeck
Germany

Structural biology on Legionella

G. Hansen, R. Wrase, V. Ramachandran, N. Petermann, M. Hoffmann & R. Hilgenfeld

Institute of Biochemistry, University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany

Legionnaires' disease is a severe type of pneumonia with a high fatality rate caused by Gram-negative bacteria of the genus Legionella. With over 600 outbreaks and more than 32.000 cases of Legionnaires' disease between 1995 and 2005 worldwide, the pathogen has become a significant public health risk. Although a limited number of virulence proteins have been identified and partially characterized, structural information on proteins from Legionella is still scarce. Especially the lack of structures of macromolecular assemblies limits our understanding of the intracellular life cycle of Legionella, which is regulated through a complex network of protein-protein interactions of pathogen and host-cell components. At the Institute of Biochemistry in Lübeck we are using X-ray crystallography to determine structures of virulence proteins of Legionella. Additional unknown proteins important for pathogenicity are identified by employing a differential proteomics approach. Resulting 3D-protein structures are used for rational design of mutant proteins to correlate structure and function and rational structure-guided development of new antibiotics. We are especially interested in large protein assemblies and complexes with host-cell proteins. In an effort to identify new factors contributing to virulence in a proteomics approach the protein expression level of L. pneumophila wildtype and a FlaR-deficient strain was analysed. Five proteins with significantly changed expression profiles could be identified: HtpB, PilN, trigger factor and two additional proteins of unknown function. As a first target for structural characterization the type-IV pilus biogenesis protein PilN essential for pilus assembly was chosen, purified and crystallized. The full structure determination of PilN will help clarify the role of the pilus during infection of host cells. Furthermore, we have determined the 3D-structures of three important virulence proteins from Legionella: Mip (macrophage infectivity potentiator protein), a dimeric protein expressed on the cell surface, exhibits a peptidyl prolyl cis-trans isomerase activity and is essential for the initiation of the infection. Based on the structure using in-silico screening we recently identified promising small molecule inhibitor hits. HtrA (heat shock requirement A) is a periplasmatic protein with protease and chaperone function, its activity is regulated by oligomerization. We have determined the structure of the HrtA homolog in Legionella, DegQ, which forms a 580 kDa dodecameric complex that encapsulates a large internal cavity. The structure led us to propose a model for regulation of the catalytic activity of DegQ. FeoB, a polytopic transmembrane GTPase involved in bacterial ferrous iron acquisition, contributes to virulence in a number of intracellular pathogens. We determined the structure of the cytosolic domain of FeoB in free and substrate-bound form.

SESSION 8: HYBRID METHODS AND MACROMOLECULAR COMPLEXES

Andrej Sali
Professor, Department of Bioengineering and Therapeutic Sciences;
Director, California Institute for Quantitative Biosciences at UCSF
University of California San Francisco
UCSF MC 2552
Mission Bay, Byers Hall
1700 4th Street, Suite 503B
San Francisco, CA

Determining architectures of macromolecular assemblies by aligning interaction networks to electron microscopy density maps

Comprehensive lists of interacting proteins can be generated by proteomic techniques. In addition, density maps of protein complexes can be determined by electron microscopy. Here, we describe an integrative method for assembling proteins of known or unknown structures into a complex by fitting their interaction network into a density map of their assembly. Assembly models are optimized with respect to a scoring function that includes the quality-of-fit of components in the map, proteomics-derived restraints, as well as the shape complementarity between pairs of components. The utility of the method is demonstrated using a benchmark of 10 complexes. We analyze the effect of the quality and quantity of the proteomics data and the resolution of the density map on the accuracy of the final model. Given sufficient proteomics data and a map at resolution of at least 20 Å, the positions and orientations of the subunits can be determined, even if some of their structures are not known and the proteomics data contain false positives.

John A. Tainer
Professor, Department of Molecular Biology
and The Skaggs Institute for Chemical Biology
The Scripps Research Institute
10550 North Torrey Pines Road (MB26)
La Jolla, CA 92037

Accurate structures, conformations, and assemblies of macromolecules in solution by high throughput X-ray scattering (SAXS) combined with crystallography

Robert P. Rambo[a], Greg L. Hura[a,], Michal Hammel[a], John A. Tainer[a,b]

[a]Lawrence Berkeley National Laboratory, Berkeley, CA
[b]Department of Molecular Biology & The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA

Conformations and assemblies of proteins, DNA, and RNA, plus their detailed structural chemistry, encode key information needed to define the biological outcomes in cell biology. We are developing SAXS combined with crystallography as a premiere tool for defining macromolecular conformations and connections suitable to join proteins to pathways and at the proteomic scale[1-3]. Crystallography supplies unsurpassed structural detail for mechanistic analyses. Yet, crystallography is restricted to describing conformations of macromolecules within crystal lattices. Advances in SAXS are making this technique increasingly powerful and robust for efficiently examining complexes in solution, as aided by interfaces allowing biologists to do these experiments[4]. Structures of flexible filaments of the XLF with XRCC4·Ligase IV[5], the extended and flexible Nbs1 link to Mre11[6], as well as of DNA-PKcs in complex with Ku and DNA[7] support the promise of SAXS for examining the assemblies and conformations of dynamic complexes in solution[8]. In principle, SAXS can provide reliable solution data on small and large macromolecules[2]. In practice, SAXS can be limited by problems in samples and analyses, which can be reduced or avoided[3]. Emerging results on dynamic complexes show that SAXS has great potential to provide accurate shapes, conformations, and assembly states in solution and inform biological functions in fundamental ways[3-9].

1. Putnam, C.D., Hammel, M., Hura, G. L., and Tainer, J. A. (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution, *Quarterly Reviews in Biophysics* **40**, 191-285.
2. Hura, G.L., Menon, A.L., Hammel, M., Rambo, R.P., Poole, F.L., Tsutakawa, S.E., Jenney, F.E., Frankel, K.A., Hopkins, R.C., Scott, J.O., Dillard, B.D., Classen, S., Adams, M.W.W. and Tainer, J.A. (2009) Rapid and robust proteomics-scale solution structural analyses determined efficiently by x-ray scattering (SAXS) *Nature Methods* **6**, 606-612.
3. Rambo RP, Tainer JA. (2010) Improving small-angle X-ray scattering data for structural analyses of the RNA world. *RNA* **16**, 638-646.
4. Classen S, Rodic I, Holton J, Hura GL, Hammel M, Tainer JA. (2010) Software for the high-throughput collection of SAXS data using an enhanced Blu-Ice/DCS control system. *J Synchrotron Radiat.* **17**, 774-81.
5. Hammel M, Yu Y, Fang S, Lees-Miller SP, Tainer JA. (2010) XLF regulates filament architecture of the XRCC4·ligase IV complex. *Structure* **18**,1431-42.
6. Williams RS, Dodson GE, Limbo O, Yamada Y, Williams JS, Guenther G, Classen S, Glover JN, Iwasaki H, Russell P, Tainer JA (2009). Nbs1 flexibly tethers Ctp1 and Mre11-Rad50 to coordinate DNA double-strand break processing and repair. *Cell* **139**, 87-99
7. Hammel M, Yu Y, Mahaney BL, Cai B, Ye R, Phipps BM, Rambo RP, Hura GL, Pelikan M, So S, Abolfath RM, Chen DJ, Lees-Miller SP, & Tainer JA. (2010) Ku and DNA-dependent protein kinase (DNA-PK) dynamic conformations and assembly regulate DNA binding and the initial non-homollogous end joining complex. *J Biol Chem.* **285**, 1414-1423.
8. Perry JJ, Cotner-Gohara E, Ellenberger T, Tainer JA. (2010) Structural dynamics in DNA damage signaling and repair *Curr Opin Struct Biol*. **20**, 283-94.
9. Rambo RP, Tainer JA. (2010) Bridging the solution divide: comprehensive structural analyses of dynamic RNA, DNA, and protein assemblies by small-angle X-ray scattering. *Current Opinion Structural Biol*. **20**, 128-1376.

Pedro Serrano
Joint Center for Structural Genomics. TSRI
10550 North Torrey Pines Rd.
La Jolla, CA, 92122

New strategy for the use of NMR spectroscopy in the crystallography-centric Joint Center for Structural Genomics

Pedro Serrano,1,2 Michael Geralt,1,2 Biswaranjan Mohanty,1,2 Lukas Susac,1,2 Reto Horst,1,2 Torsten, Herrmann,5 Lukasz Jaroszewski,1,4 Ian A. Wilson,1,2,3 and Kurt Wüthrich 1,2,3

1. Joint Center for Structural Genomics, The Scripps Research Institute, La Jolla, CA
2. Department of Molecular Biology , The Scripps Research Institute, La Jolla, CA
3. Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA
4. The Burnham Institute for Medical Research, La Jolla, CA
5. Centre Europen de RMN Ã TrÃ¨s Hauts Champs de Lyon, Villeurbanne, France

The Joint Center for Structural Genomics (JCSG) has implemented an approach to use solution NMR spectroscopy for expanding structure coverage of the protein universe by solving first structure representatives of large protein families that could not yet be successfully tackled by X-ray crystallography. This new strategy is based on previously established orthogonality between protein sequence and structure determination success rates either by NMR spectroscopy or by X-ray crystallography, and uses the following four-step protocol: (i) Sequence-based bioinformatics to prioritize protein orthologues for high NMR structure determination success rate. (ii) Preparation of structure-quality protein solutions guided by 1D 1H-NMR and 2D [15N,1H]-correlation NMR-profiling. (iii) Extensively automated NMR structure determination. (iv) Interactive structure refinement and validation. So far, this approach has provided structural representatives for more than 3000 protein sequences, which are members of 10 distinct pfam families. The resulting repertoire of three-dimensional architectures shares low sequence identity with proteins deposited in the PDB, and includes both new folds and well-characterized structure families.

Michael Plevin
Institut de Biologie Structurale
41, rue Jules Horowitz
Grenoble, 38027, France

SeSEAM: a Systematic Mutagenesis-Driven Strategy for Site-Resolved NMR Studies of Supramolecular Assemblies

Michael J. Plevin, Carlos D. Amero, M. Asunción Durá, Marjolaine Noirclerc-Savoye, Arnaud Perollier, Benoit Gallet, Thierry Vernet, Bruno Franzetti, and Jérôme Boisbouvier

Institut de Biologie Structurale

It is becoming increasingly feasible to use solution NMR spectroscopy to analyse the function, local structure, intermolecular interactions and dynamics of large protein assemblies of up to 1 MDa. However, due to the size-limitations of standard NMR assignment strategies, the process of obtaining sequence-specific assignments of such systems remains a considerable challenge. We have fully assigned isoleucine (δ1,γ2), alanine, and leucine and valine (proR and proS) methyl groups of a 468 kD dodecmeric aminopeptidase complex using an approach based on site-directed mutagenesis. We produced a library of single-site mutants each of which is designed to “turn off” a resonance in 2D methyl correlation spectra. The library was expressed using appropriate methyl-labelling protocols and each mutant was analysed using 2D methyl TROSY experiments. The missing peak in each spectrum can accordingly be assigned to a specific methyl group in the mutated residue. A key attraction of this approach is the simplicity of the assignment process; complicated NMR experiments and time-consuming specialized data analysis is not necessary. The concept of resonance assignment by mutagenesis is not new. However, modern automated laboratory techniques (mutagenesis, cloning, expression, purification, etc) and advances in NMR spectrometer technology (increased sensitivity, sample changers, small volume probes, etc) make it possible to implement this simple approach in a substantially more comprehensive, systematic and straightforward way. Each step in the protocol (library construction, protein expression and purification, data acquisition and analysis) can be conveniently automated which makes this methodology ideally suited to laboratories familiar with higher-throughput techniques. The SeSAM approach, therefore, offers an alternative, general and convenient method for overcoming the hurdle of NMR assignment in larger protein systems.

SESSION 9: INFECTIOUS DISEASE RESEARCH II: BACTERIAL DISEASES

Wayne Anderson
Professor
Molecular Pharmacology & Biological Chemistry
Northwestern University
303 E. Chicago,
Chicago, IL 60611-3008

Structural Genomics to aid drug discovery for infectious diseases

W. F. Anderson[1*], A. Edwards[2*], D. Fremont[3*], A. Joachimiak[4*], W. Minor[5*], C. Orengo[6*], Z. Otwinowski[7*], S.N. Peterson[8*] and A. Savchenko[2*]

[*]Center for Structural Genomics of Infectious Diseases, Chicago, IL
[1]Northwestern University, Chicago, IL
[2]University of Toronto, Toronto, ON
[3]Washington University, St Louis, MO
[4]University of Chicago, Chicago, IL
[5]University of Virginia, Charlottesville, VA
[6]University College London, London, GB
[7]University of Texas Southwestern Medical Center, Dallas, TX
[8]JC Venter Institute, Rockville, MD

Structural genomics approaches and high throughput structure determination applied to proteins from organisms that cause infectious diseases provides the three-dimensional structures for many proteins that are potential drug targets. The resulting array of structures is a starting point for structure aided drug discovery efforts. The Center for Structural Genomics of Infectious Diseases (CSGID) is applying state-of-the-art structural biology technologies to the characterization of proteins from the National Institute for Allergy and Infectious Diseases (NIAID) category A-C pathogens and organisms causing emerging, or re-emerging infectious diseases. CSGID target selection emphasizes potential biomedical relevance. Selected proteins include known drug targets and their homologs, essential enzymes, virulence factors and vaccine candidates. Suggestions for target proteins are solicited from the broader scientific community and 25% of CSGID target proteins arise from community requests. The ultimate goal is to generate a library of structures for proteins and their complexes with small molecules that are available to the scientific community and can serve as a starting point for further research and structure aided drug discovery for infectious diseases. The CSGID has already deposited over 300 structures from bacterial pathogens into the Protein Data Bank. Additionally, expression clones for over 1500 proteins are available through BEI Resources. These structures are providing important structural insights on metabolic pathways that are potential drug targets and on proteins that are potential sites of interaction with the host, such as surface exposed proteins.

This project has been funded in whole or in part with Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under Contract No. HHSN272200700058C.

Edward N. Baker
Distinguished Professor
University of Auckland
School of Biological Sciences
3A Symonds St.
Auckland, New Zealand

Structure and assembly of bacterial pili, from analysis of virulence-related gene clusters

Hae Joo Kang, Christian Linke, Paul Young, Neil Paterson, Thomas Proft and Edward Baker

The genomes of many Gram-positive bacterial pathogens contain pathogenicity islands that include clusters of 3-5 genes that are predicted to be located on the bacterial surface. Most were annotated as of unknown function, but have proved to be involved in the formation of pili - long, thin, hairlike protein assemblies that function in attachment to host cells. We have undertaken structural analyses of proteins encoded by one such gene cluster, from Group A Streptococcus. Structures are now available for all the components of the pilus, together with a sortase enzyme involved in its assembly and a putative chaperone resembling a signal peptidase. We also discovered previously-unsuspected internal crosslinks in the form of isopeptide bonds. These pilin structures, together with those from several other Gram-positive pathogens, have established an entirely new paradigm in which pili are assembled as novel covalent polymers.

Peter J. Myler
Seattle Biomedical Research Institute
307 Westlake Ave N,
Seattle, WA 98109-5219

Three Hundred New Protein Structures from Microbial Pathogens

Peter J. Myler[1,3,4], Robin Stacy[1,], Lance Stewart[2], Wesley C. Van Voorhis[3,5], Gabriele Varani[6,7], Garry W. Buchko[8]

[1]Seattle Biomedical Research Institute, Seattle, WA, USA
[2]Emerald BioStructures, Bainbridge Island, WA, USA
[3] Department of Global Health, University of Washington, Seattle WA, USA
[4] Department of Medical Education & Biomedical Informatics, University of Washington, Seattle WA, USA
[5] Department of Medicine, University of Washington, Seattle WA, USA
[6] Department of Chemistry, University of Washington, Seattle WA, USA
[7] Department of Biochemistry, University of Washington, Seattle WA, USA
[8]Pacific Northwest National Laboratory, Richland, WA USA

The Seattle Structural Genomics Center for Infectious Disease (SSGCID) is funded by NIAID to solve protein structures from biodefense organisms and emerging infectious diseases. SSGCID target selection focuses on eight bacterial and twelve protozoan genera, as well as ssDNA and negative-strand ssRNA viruses. Community input is actively solicited to identify essential enzymes, virulence factors, drug targets and vaccine candidates of biomedical relevance for our structure determination pipeline (http://ssgcid.org). Target genes are PCR amplified, cloned and screened for expression in *E. coli*. Soluble proteins are purified in milligram quantity, screened for crystallization, and crystals analyzed by X-ray diffraction using in-house source or off-site synchrotron beam-lines. Small proteins which fail to crystallize are queued for structure determination by NMR. Community requests that fail at any stage are recycled back through the pipeline using synthetic gene design, including terminal deletions and other surface mutations, while also testing orthologues from related species. Rescue of a limited number of targets are attempted in wheat germ cell-free and baculovirus expression systems. A small number of high-value targets are also selected for fragment screening to obtain multiple ligand-bound structures. Since project inception in late 2007, over 6000 targets have been selected, with more than 1600 unique proteins expressed and purified for crystallization trials. To date, SSGCID has deposited over 300 protein structures in the Protein Data Bank (PDB), representing the majority of available structures for several organisms. Target-to-structure success rates vary considerably between genera, ranging from 20% (*Brucella*) to 1% (*Toxoplasma*), and about one third of all SSGCID structures contain bound ligands. By the end of 2012, we expect to have solved more than 500 protein structures from pathogenic microbes, which will serve as starting points for structure-based drug design projects. We make all expression clones, purified proteins, and structures publicly available, and engage the scientific community for selection of high-impact targets from infectious diseases for structural characterization.

Bridget Mabbutt
Macquarie University, Sydney
Herring Rd
Sydney
Australia, 2109

Plasticity of Vibrio and A. baumanii genomes: structural genomics for mobile gene functions

Chandrika Deshpande, Bhumika Shah, Vani Sureshan, Steve Harrop, Ian Paulsen, Yan Boucher, Hatch Stokes and Bridget Mabbutt

Department of Chemistry & Biomolecular Sciences, Macquarie University

The wide extent of lateral gene transfer (LGT) presents challenges to genomics of micro-organisms: no one cell line is representative of the gene content of a species as a whole. Instead, prokaryotic genomes harbour features such as pathogenicity islands, integrons or transposons in which are arrayed highly novel genes for niche adaptation and host defence. In many cases, specific function cannot be readily assigned due to lack of sequence homologues within annotated genomes. These bacterial gene targets have proved to be highly amenable to the cloning and screening procedures of structural genomics. I present our output of protein crystal structures from genes derived from Vibrio integron arrays and genomic islands of new Acinetobacter baumanii strains. Our structures reveal new enzyme variants (glyoxalase, acetyltransferase), drug transport/binding proteins and antibiotic resistance factors which all clearly confer fitness advantage. A significant component of our structural output also contains entirely novel protein folds.

Alexei Savchenko
Department of Chemical Engineering and Applied Chemistry
University of Toronto
112 College street room 23, Toronto, Ontario, Canada M5G 1L6

Structural insight into the function of effector proteins – bacterial pathogenic factors orchestrating eukaryotic cell biology

Alexei Savchenko, Alexander U. Singer, Craig Daniels, Tatiana Skarina, Xiaohui Xu

Department of Chemical Engineering and Applied Chemistry, University of Toronto

Many Gram-negative bacteria interact with human, animal, or plant hosts and inject a specific set of proteins called effectors into the cytosol of host cells through various (type III, IV and VI) secretion systems. While the specific function of the majority of effector proteins remains unknown collectively they share the ability to alter host cellular pathways/systems to promote the bacterial life-style and contribute to virulence. Distantly related bacterial pathogens may harbour closely related effectors, demonstrating that one pathogenicity mechanism can give rise to a multitude of diseases that range from bubonic plague in humans to fire blight in fruit trees. Thus, characterization of the specific functions of effectors represents a major task for research into bacterial pathogenesis. We have structurally and functionally characterised representatives of several common families of effectors, including IpaH, NleG and HopA1. Our results demonstrate that these bacterial proteins are able to mimic cellular functions otherwise specific to eukaryotic proteins and interfere with critical cell processes such as the ubiquitination signalling system.

Session 10: Advanced and Automated Methods in Structural Biology and Structural Genomics

Soichi Wakatsuki
Structural Biology Research Center
Photon Factory, Institute of Materials Structure Science
High Energy Accelerator Research Organization (KEK)
Tsukuba, Ibaraki 305-0801
Japan

Beamline developments for targeted structural proteomics

Soichi Wakatsuki[1], Naohiro Matsugaki[1], Yusuke Yamada[1], Leonard Chavas[1], Masato Kawasaki[1], Ryuichi Kato[1], Masahiko Hiraki[1], Noriyuki Igarashi[1], Kunio Hirata[2], Takashi Kumasaka[2,3] & Masaki Yamamoto[2]

1 Structural Biology Research Center, Photon Factory, Institute of Materials Structure Science, High Energy Accelerator Research Organization (KEK), Tsukuba, Ibaraki 305-0801, Japan
2 RIKEN SPring-8 Center, Harima Institute, Sayo, Hyogo 679-5198, Japan
3 JASRI/SPring-8, Sayo, Hyogo 679-5198, Japan

"Targeted Proteins Research Program (TPRP)" is a nationwide structural biology effort which includes 35 target-oriented structural biology projects and 10 R&D projects in protein production, chemical library, structural analysis and bioinformatics. Two synchrotron radiation facilities, SPring-8 & Photon Factory, have developed two complementary micro-focus beam lines and, together with Hokkaido Univ., Osaka Univ., and Kyoto Univ., developing techniques to facilitate user access and experiments at the two synchrotron sites. At SPring-8 a micro-beam beamline, BL32XU, has been constructed to provide highly intense 1 micron by 1 micron beam while the Photon Factory has built a micro-focus beam line, BL1A, optimized for low-energy SAD phasing experiments at around 4 keV. Each of the 35 target-oriented structural proteomics projects aims to solve structures of challenging targets in close collaboration with groups in cell biology, biochemistry, bioengineering, pharmacology, or medicine. The above mentioned newly developed beamlines as well as several already operational beamlines of the two facilities provide a unified platform for protein structure analyses with one-stop beamtime request and allocation scheme.

The micro-beam beamline BL32XU of SPring-8 now produces an extremely intense and stable micro focused beam, 6.2×10^{10} photons/sec with 0.9 micron by 0.9 micron focus. It has been shown that this beam can enhance qualities of diffraction patterns from large complexes such as nucleosome core particles and membrane proteins. On PF BL1A beamline, an intense low energy beam at around 4 keV is provided to enhance anomalous signals from light atoms for SAD phasing of macromolecules which are difficult to prepare heavy atom or selenomethionine derivative crystals. A cryo freezing device designed by Isao Tanaka's group of Hokkaido University is available for loop-less mounting of cryocooled crystals on site. A new crystal cassette, compatible with both SPring-8 (SPACE) and PF (SAM/PAM) type robots, has been developed for cryocooled crystals. Results from the two new beamlines as well as the overall analysis core of the TPRP will be reported.

Lewis E. Kay
Departments of Medical Genetics, Biochemistry and Chemistry
University of Toronto
Medical Sciences Building
1 King's College Circle
Toronto, ON, M5S 1A8
Canada

Seeing the Invisible by Solution NMR Spectroscopy

Many biochemical processes such as ligand binding, enzyme catalysis, molecular recognition and protein folding proceed through the formation of functionally important intermediates that escape detection using traditional structural biology methods. A Nuclear Magnetic Resonance (NMR) approach for seeing such 'invisible' states will be described. A related problem is one where 'near invisible' systems, such as supra-molecular machines, are studied by solution NMR. New NMR approaches and applications that bring such systems into focus will be presented.

Andrzej Joachimiak
Midwest Center for Structural Genomics and Structural Biology Center
Biosciences Division
Argonne National Laboratory
Argonne, Illinois 60439
USA

Human Microbiome - the Next Structural Frontier

Metagenomics aims to analyze, reconstruct and characterize how microbial communities organize and evolve their collective genomes to thrive in the environments they co-habit. Human microbiomes harbor a large number of microbial species, many of which are uncharacterized or unculturable. The human gut microbiome represents a highly complex microbial community that has a significant impact on human physiology. The gut microbiome performs a diverse set of functions supplementing those performed by the host and is believed to significantly enhance the metabolism of amino acids, carbohydrates, xenobiotics, methanogenesis, biosynthesis of vitamins and other compounds. Therefore, microbiomes represent a potential pool of genes coding for novel proteins and functions. The Midwest Center for Structural Genomics (MCSG) (a component of Protein Structure Initiative:Biology project) targets proteins from microbes found in human gut microbiome in healthy and diseased individuals. The MCSG structure determination pipeline is being applied to the microbial proteins that emanate from these projects. We have analyzed gut metagenomic data to identify novel, previously uncharacterized, human microbiome-specific protein families. The analysis revealed an abundance of carbohydrate metabolizing enzymes. A set of reagent genomes was used to extract genes from gut microbiome and produce proteins for structural studies. A number of structures from normal and enteric microflora have been determined. The structural and functional results will be discussed.

This work was supported by National Institutes of Health Grant GM074942, GM094585 and by the U.S. Department of Energy, OBER, under contract DE-AC02-06CH11357.

James Fraser
University of California San Francisco
301E Byers Hall
1700 4th Street, UCSF
San Francisco, CA 94158

Crystal freezing quenches protein motions and biases hydrophobic packing

James S. Fraser(1), Henry van den Bedem(2), Avi Samelson(3), P. Therese Lang(3), James M Holton(4), Nathaniel Echols(5), Tom Alber(3)

1 - California Institute of Quantitative Biosciences (QB3) and Department of Cellular and Molecular Pharmacology, UCSF
2 - Joint Center for Structural Genomics, Stanford Synchrotron Radiation Lightsource
3 - Department of Molecular and Cell Biology, UC Berkeley
4 - Advanced Light Source, Lawrence Berkeley National Laboratory
5 - Lawrence Berkeley National Laboratory

Over the last 30 years, data collection for macromolecular X-ray crystallography has almost completely transitioned from at room temperature (273-298K) at low-power sources to cryogenic temperatures (90-100K) at brilliant synchrotron beamlines. Although cryopreservation is an invaluable tool for limiting radiation damage when determining structures, few studies have examined how the freezing process might bias the functional interpretation of structural results. Previously, we found that Ringer sampling of room temperature, but not cryogenic, electron density revealed a network of alternative side chain conformations essential for catalysis the proline isomerase CypA. Motivated by this temperature-dependent difference, we compared 30 proteins with both high-resolution room and cryogenic temperature datasets using newly developed tools to evaluate protein core packing, refine multiple conformations, and sample electron density in real space. Our results confirm that freezing remodels the crystal lattice and reduces protein volume. Surprisingly, freezing also artificially drives improved packing and changes the distribution of side chain conformations in both exposed and buried regions. As with CypA, we find that temperature-sensitive networks of alternative side chain conformations are correlated with chemical exchange detected by NMR experiments. For example, in H-Ras electron density sampling discovers an intrinsic allosteric side chain network, supported by NMR data, in the room temperature, but not cryogenic, electron density maps. Accounting for how cryogenic X-ray data collection compresses proteins and quenches internal fluctuations may lead to improvements in protein structure prediction, functional assignment, and design.

Giuseppe Melacini
McMaster University
1280 Main St W.
Hamilton, ON L8S 4M1
Canada

Understanding Signaling Networks through the Covariance Analysis of NMR Chemical Shifts

Rajeevan Selvaratnam, Somenath Chowdhury, Bryan VanSchouwen and Giuseppe Melacini*

Department of Chemistry and Chemical Biology and Department of Biochemistry and Biomedical Sciences

Allostery is a fundamental mechanism of regulation in biology and it is critical for the control of signaling networks. The residues at the end points of long-range allosteric perturbations are commonly identified by the comparative analyses of structures and dynamics in apo and effector-bound states. However, the networks of interactions mediating the propagation of allosteric signals between the end points often remain elusive. We will show that the covariance analysis of NMR chemical shift changes caused by a set of covalently modified analogs of the allosteric effector (i.e. agonists and antagonists) reveals extended networks of coupled residues. Unexpectedly, such networks reach not only sites subject to effector-dependent structural variations, but also regions that are controlled by dynamically driven allostery. In these regions the allosteric signal is propagated mainly by dynamic rather than structural modulations, which result in subtle but highly correlated chemical shift variations. The proposed chemical shift covariance analysis (CHESCA) [1] identifies inter-residue correlations based on the combination of agglomerative clustering (AC) and singular value decomposition (SVD) [2]. AC results in dendrograms that define functional clusters of coupled residues, while SVD generates score plots that provide a residue-specific dissection of the contributions to binding and allostery. The CHESCA approach was validated by applying it to the cAMP-binding domain of the exchange protein activated by cAMP (EPAC) and the CHESCA results are in full agreement with independent mutational data on EPAC activation. Overall, CHESCA is a generally applicable method that utilizes a selected chemical library of effector analogs to quantitatively decode the binding and allosteric information content embedded in chemical shift changes.

References

[1] Selvaratnam R, Chowdhury S, VanSchouwen B, Melacini G (2011) Proc Natl Acad Sci USA, In press.

[2] Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Proc Natl Acad Sci USA 95: 14863-14868.

SESSION 11: COMPUTATIONAL APPROACHES TO STRUCTURE AND FUNCTION

Shoshana J. Wodak
Molecular Structure and Function Program
Hospital for Sick Children
555 University Av.
Toronto, Ontario M5G 1X8

Department of Biochemistry
University of Toronto
Toronto, Ontario

Department of Molecular Genetics
University of Toronto
Toronto, Ontario
CANADA

Protein-protein interactions: predicting the structure of complexes and their interfaces

Shoshana J. Wodak[1,2,3] and Marc F. Lensink[4,5]

[1]Molecula Structure and Function Program, Hospital for Sick Children, 555 University Av., Toronto, Ontario M5G 1X8 CANADA
[2]Department of Biochemistry, University of Toronto, Toronto, Ontario, CANADA
[3]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, CANADA
[4]Genome and Network Bioinformatics, CP 263, BC6, Université Libre de Bruxelles, Blvd du Triomphe, 1050 Brussels, BELGIUM,
[5]Structure and Function of Biological Membranes, Université Libre de Bruxelles, Brussels, BELGIUM

Recent large-scale studies of protein-protein interactions and protein complexes in several model organisms and human have revealed that the majority of the proteins in the cell function as part of multi-component assemblies. Unraveling the structural and functional details of these assemblies has therefore become a priority for structural and cellular biologists and biochemists.
Unfortunately, determining the three-dimensional structure of large protein assemblies remains a significant challenge for structural biology, whereas atomic resolution structures of individual proteins are being determined with increasing speed and reliability. This has created a gap that computational methods are well positioned to fill.

We will discuss two types of computational methods, which have become increasingly popular in recent years. One involves the so-called docking procedures, which address the challenge of deriving the 3D structure of a protein complex starting from the atomic resolution structures of the individual components. The other concerns computational procedures for predicting protein interfaces. The latter procedures aim at identifying residues on the surface of a single protein that make-up the recognition interface with another protein component.

We will review recent progress in these computational approaches, with examples from the work of groups participating in CAPRI (Critical Assessment of Predicted Interactions), a community-wide experiment, which is playing an important role in fostering the development of improved docking algorithms and closely monitoring their performance (http://www.ebi.ac.uk/msd-srv/capri/) (1,2).

1- Docking and scoring protein interactions: CAPRI 2009.
Lensink MF, Wodak SJ.
Proteins. 2010 Nov 15;78(15):3073-84.

2- Blind predictions of protein interfaces by docking calculations in CAPRI.
Lensink MF, Wodak SJ.
Proteins. 2010 Nov 15;78(15):3085-95.

Adam Godzik
Program Director, Professor
Bioinformatics and Systems Biology
Sanford-Burnham Medical Research Institute
10901 North Torrey Pines Road
La Jolla, CA 92037

Joel L. Sussman
The Israel Structural Proteomics Center
Weizmann Institute of Science
Rehovot 76100 ISRAEL

Proteopedia - a Scientific 'Wiki' Bridging the Rift Between 3D Structure and Function of Biomacromolecules

Jaime Prilusky1, Eran Hodis2,*, David Canner3,5 ,Wayne Decatur6, Karl Oberholser7, Eric Martz8 & Joel L. Sussman3,4

1. Dept. of Biological Services, Bioinformatics Unit, Weizmann Institute of Science, Rehovot 76100 ISRAEL
2. Dept. of Computer Science & Applied Mathematics, Weizmann Institute of Science, Rehovot 76100 ISRAEL
3. Dept. of Structural Biology, Weizmann Institute of Science, Rehovot 76100 ISRAEL
4. The Israel Structural Proteomics Center, Weizmann Institute of Science, Rehovot 76100 ISRAEL
5. Dept. of Chemistry, Princeton Univ., Princeton, NJ 08544-1009
6. Dept. of Molecular, Cellular & Biomedical Sciences, Univ. of New Hampshire, Durham, NH 03824
7. Dept. of Chemistry and Biochemistry, Messiah College, Grantham, PA 17055
8. Dept. of Microbiology, Univ. of Massachusetts, Amherst, MA 01003-5720,

*Present address: Cancer Program, Broad Institute of MIT and Harvard, Cambridge, MA 02142

Biologists and biochemists without a background in structural biology have traditionally found it difficult to explore 3D models of biomacromolecules underlying biological functions and disease. Most communications of new structures rely on text and 2D images, since molecular visualization programs may be foreign to the non-expert. A new collaborative web-resource called Proteopedia (http://www.proteopedia.org) intuitively links written information and 3D structural information in a manner accessible to a broad scientific audience and allows for easy addition of structural annotations. Protein and other biomacromolecule structures are displayed in interactive format (see Fig. 1). These interactive images are surrounded by descriptive text containing links that change the appearance (such as view, representations, colors or labels) of the adjacent 3D structure to reflect the concept explained in the text. Further, scene changes occur smoothly, allowing the reader to maintain perspective within the overall structure. 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 structure, e.g. see the articles on 'HIV-1 protease' at: http://proteopedia.org/wiki/index.php/HIV-1_protease, 'Lac repressor' at: http://proteopedia.org/wiki/index.php/Lac_repressor, and 'HMG-CoA Reductase' at: http://www.proteopedia.org/wiki/index.php/HMGR. Aside from content added by Proteopedia's over 1,100 registered users, articles on each of the more than 71,000 entries in the Protein Data Bank (PDB) have been automatically generated with information aggregated from various resources (including evolutionary conservation), providing a useful reference article for each PDB entry that includes an interactive 3D model. Members of the scientific community are invited to request a user account to edit existing articles and to create new ones. Fig. 1 Proteopedia's Homepage at: http://www.proteopedia.org Refence: Hodis, Prilusky, Martz, Silman, Moult and Sussman (2008) Genome Biology 9, R121.

Alexey G. Murzin
MRC Laboratory of Molecular Biology
Hills Road
Cambridge CB0 2QH
UK

Structure shortcuts to function prediction

Knowledge of protein structures is essential to our understanding of known functions and can assist prediction of unknown activities. Like sequence-structure relationships, structure-function relationships are also very complex, hindering the direct predictions of protein function from structure. This is arguably the main reason of why functional annotation of new proteins is trailing several years behind their structural characterisation. The progress in function discovery for genomic proteins of unknown function so far has been achieved mainly through bioinformatics inference enhanced by protein structure data. Protein structure is a hub for the integration of all kinds of data on the protein superfamily. It facilitates multiple structural and functional predictions for its known and probable members, including specified activities of uncharacterised protein families. One of the most successful integrated approaches to function prediction is the matching of structurally characterised proteins of unknown function and "hypothetical functions of unknown protein". There is circumstantial biochemical evidence for the existence of many yet unexplored metabolic pathways. Some of these pathways are deduced from known chemical structures of the end products, others provide probable routes between identified intermediates or suggest new variations of established pathways. There are associated enzymatic activities, many of which were partly purified in the past but not sequenced at the time. These "hypothetical" enzymes probably account for a substantial fraction of genomic proteins of unknown function. Many of them may have already been characterised by Structural Genomics, although their identities remained unknown. Selected examples of functional annotations of various Structural Genomics targets by this and similar integrated approaches, including our confirmed and new functional predictions, will be presented.

SESSION 12: FUNCTIONAL GENOMICS

Gaetano T. Montelione
Rutgers University
Dept. of Molecular Biology & Biochemistry
CABM Room 014A
679 Hoes Lane
Piscataway. N.J. 08854

Targeting the Human Cancer Pathway Protein Interaction Network (HCPIN) by Structural Genomics

G. T. Montelione[1], T. A. Acton[1], A.Ertekin[1], R. Guan[1], G. Liu[1], P. Rossi[1], R. Shastry[1], A. Kuzin[2], F. Forouhar[2], L Ma[1], J. Seetharaman[2], S. Vorobieva[2], B. Wu[3], A. Yee[3], J. Mills[4], T. Szyperski[4], J. Hunt[2], L. Tong[2], C. Arrowsmith[3], and Y.J. Huang[1,]

Northeast Structural Genomics Consortium, [1]Center for Advanced Biotechnology and Medicine, Rutgers University, Piscataway, NJ; [2]Columbia University, NY, NY; [3]Ontario Cancer Institute and Structural Genomics Consortium, Toronto, CANADA.; [4]SUNY at Buffalo, Buffalo, NY.

Structural genomics provides unique opportunities for characterizing and understanding systems biology. As a step towards better integrating protein 3D structural information in cancer systems biology, we have constructed a Human Cancer Pathway Protein Interaction Network (HCPIN) by analysis of several classical cancer-associated signaling pathways and their proposed physical protein-protein interactions (Huang et al. 2008). The current version of HCPIN was constructed by identifying “core proteins” associated with classical cancer-associated cellular processes using the KEGG database, and then expanded to include binding partners of these “core proteins” based on physical protein-protein interaction data (e.g. co-IP, Y2H data) obtained from the Human Protein Reference Database (HPRD). The seven core biological processes used to provide the “core proteins” include (i) cell cycle progression, (ii) apoptosis, (iii) MAP kinase (MAPK) pathway, (iv) innate immune response, (v) TGF-□ signaling pathway (vi) phosphotidyl inositol kinase (PI3K) signaling pathway, and (vii) JAK-STAT pathways. The HCPIN database includes some 2,800 human proteins, with some 9,500 putative pair-wise physical interactions. Many well-known cancer-associated proteins (e.g. p53, NF-kB, EGF receptor, etc) play central roles as “hubs” or “bottlenecks” in the HCPIN. While some 50% of residues in these proteins are in sequence segments that meet criteria sufficient for approximate homology modeling (Blast E-val < 10^{-6}), < 30% of residues in these proteins are structurally covered using high-accuracy homology modeling criteria (i.e. Blast E-val < 10^{-6} and at least 80% sequence identity) or by actual experimental structures. Since these human protein structures will be used for many different kinds of biophysical studies, we have defined our goal as structural coverage of HCPIN at the 80% sequence identity level, rather than the 30% level used in our other PSI work. The NESG HCPIN website (available at www.nesg.org) provides a comprehensive description of this biomedically important multi-pathway network. It is a useful tool for cancer biology research, providing experimental and homology models of HCPIN proteins, information about their protein partners, and access to extensive expression and sample production data generated by the NESG.

The long-range goal of this effort is to provide a comprehensive 3D structure-function database for human cancer-associated proteins and their interaction network. The NESG is targeting > 1,000 human proteins (> 3,000 domains) from the HCPIN for sample production and 3D structure determination [1]. About 60 human protein structures from the HCPIN, including three protein-peptide and protein-protein complexes, have been determined and deposited in the PDB to date, with several more in final stages of refinement. Information provided on the NESG HCPIN website also includes the experimentally-identified binding partners of each HCPIN protein, as well as information about disorder/order transitions that may occur upon complex formation. These data will drive our efforts to determine structures of HCPIN proteins and their complexes. We will present the NESG progress on the HCPIN protein production and some HCPIN protein structures solved by NESG, including Myd88, ARID3A, Unc119, DOC-1, and others.

Huang, Y.J., Hang, D., Lu, L.J., Tong, L., Gerstein, M.B., and Montelione, G.T., *Targeting the human*

cancer pathway protein interaction network by structural genomics. Mol Cell Proteomics, 2008 7: 2048 - 2060.

John L Markley
Steenbock Professor of Biomolecular Structure,
Department of Biochemistry
University of Wisconsin–Madison
171A Biochemistry Addition
433 Babcock Drive
Madison, WI 53706-1544
USA

Automated approaches to protein NMR assignment and structure determination

Arash Bahrami, Roger A. Chylla, Gabriel Cornilescu, Hamid R. Eghbalnia,‡ Woonghee Lee, Marco Tonelli, William M. Westler, John L. Markley

National Magnetic Resonance Facility at Madison and Center for Eukaryotic Structural Genomics, Biochemistry Department, University of Wisconsin-Madison
‡Department of Molecular and Cellular Physiology, University of Cincinnati

We present novel approaches to streamlining protein NMR data collection and analysis. “Newton” is a fast maximum-likelihood reconstruction (FMLR) approach developed for use in automated signal recognition (input of peak lists to automated assignment programs) and rapid evaluation of amino acid constructs and solution conditions for their suitability for NMR structure-function studies. “ADAPT-NMR” (Assignment-directed Data collection Algorithm utilizing a Probabilistic Toolkit in NMR) represents a paradigm for the automated integration of protein NMR data collection and analysis. With a [^{15}N,^{13}C]-protein in the NMR probe and input data consisting of the protein sequence and orthogonal planes from ^{1}H-^{15}N and ^{1}H^{13}C correlation experiments, ADAPT-NMR commences data collection by directing the NMR spectrometer to select a particular 3D NMR experiment from a panel of possibilities according to an objective function designed for goal-directed optimality. Analysis, carried out on-the-fly, while a data set is collected, serves as the basis for the selection of the next data set to be collected (3D experiment type and angle(s) of tilted planes). Data collection ceases when the desired level of analysis is reached. ADAPT-NMR supports comprehensive peak identification, backbone chemical shift assignment, and protein secondary structure determination without manual intervention in less than a day for proteins containing up to 110 residues. “PINE” (Probabilistic Interaction Network of Evidence) is then used for automated side chain assignments from additional NMR data sets. “PONDEROSA” (Peak-picking Of Noe Data Enabled by Restriction of Shift Assignments) accepts input information consisting of a protein sequence, backbone and side chain resonance assignments, and 3D-NOESY spectra, and returns assignments of NOESY cross peaks, distance and angle constraints, and a reliable NMR structure represented by a family of conformers. “PINE-SPARKY” and “ADAPT-NMR Enhancer” are visualization packages that support the comparison of automated assignment results with experimental spectral data for verification and refinement.

Supported by NIH grants P41 RR02301, 1 U54 GM094584, U01 GM094622.

Akeo Shinkai
RIKEN SPring-8 Center
1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan

Structural and functional genomics of a model organism Thermus thermophilus HB8: toward functional discovery of functionally unknown proteins

Akeo Shinkai*, Yoshitaka Bessho*, Yoshihiro Agari*, Hitoshi Iino*, Kenji Fukui*, Aya Kitamura*, Keiko Sakamoto*, Kwang Kim‡, Ryoji Masui*‡, Shigeyuki Yokoyama*‡≠, Seiki Kuramitsu*‡

*RIKEN SPring-8 Center, Hyogo, Japan
‡Osaka University, Osaka, Japan
≠RIKEN Systems and Structural Biology Center, Tokyo, Japan

The ultimate goal of our research is to understand all of the biological phenomena in the cell at an atomic-resolution, on the basis of the structures and the functions of all of the molecules, and to predict all of the biological phenomena in the cell. To achieve this goal, we chose the extremely thermophilic organism, Thermus thermophilus HB8, as a model organism, because the essential 2,200 genes (proteins) encoded in its genome have been selected during evolution and are common to many organisms, including human. Furthermore, as this strain is thermostable, this model organism is a good candidate for X-ray structural analyses of its macromolecules. The total number of protein structures from this organism is now more than 490, including 100 or so solved by other groups. This number is about 22% of the total number of proteins of this organism, which is the highest among all organisms. Surprisingly, our genome analysis revealed that as many as 1/3 of the 2,200 total proteins (genes) are functionally uncharacterized, although many of them are essential and common to many organisms, including human. In order to interpret the whole-cell phenomena, the identification of the functionally-uncharacterized (especially hypothetical) proteins is essential. Therefore, we identified their functions (1) by using the structural data obtained from structural genomics, and (2) by genome-wide functomics data obtained from mRNA analysis (transcriptomics), from protein analysis (proteomics), and from metabolite analysis (metabolomics), and then (3) we confirmed the function of each isolated protein in vitro. The plasmids for protein expression and for gene disruption are available to the public through the RIKEN BioResource Center (see http://www.thermus.org). Some of the information about protein production (protein expression and purification), crystallization, structure, and functomics are also available from our database (http://www.thermus.org) and public databases. In this presentation, we will show examples of classification of several functionally uncharacterized proteins focusing on their characteristics of transcriptional regulation. For instance, cyclic AMP receptor protein (CRP), which is one of the global transcription factors that regulate multiple genes, positively regulates 22 genes, including three hypothetical genes. These genes may be involved in DNA/RNA metabolism including clustered regularly interspaced short palindromic repeat (CRISPR)-associated (Cas) proteins that are components of a novel host defense system against invading foreign replicons, such as phages. In total, we have elucidated the functions of seven transcriptional regulators out of about 70 regulators most of which are functionally unknown, and found that 84 genes, including 28 functionally unknown [clusters of orthologous groups of proteins (COG) code S or non-categorized in COG] genes could be categorized based on their transcriptional regulatory mechanism.

References:

Shinkai A. et al. (2007) J Bacteriol 189, 3891-3901
Shinkai A .et al. (2007) J Bacteriol 189, 8758-8764
Sakamoto K. et al. (2008) Gene 423, 153-159
Agari Y. et al. (2008) Mol Microbiol 70, 60-75
Agari Y. et al. (2010) J Mol Biol 395, 270-281
Sakamoto K. et al. (2010) Microbiology 156, 1993-2005
Agari Y. et al. (2010) FEMS Microbiol Lett 313, 127-134
Agari Y. et al. Microbiology, in press.

Allan Matte
Biotechnology Research Institute
National Research Council Canada
6100 Royalmount Ave
Montreal, QC
Canada H4P 2R2

Structural Proteomics of Bacterial Protein Complexes at the BSGI

Allan Matte1, Rong Shi2, Irena Ekiel1,3, Laura MacDonald3, Ariane Proteau2, Eunice Ajamian2, Christine Munger2, Stephane Raymond2, Linhua Zhang2, Zongchao Jia4, Kalle Gehring2, Guennadi Kozlov2, Albert M. Berghuis2 and Miroslaw Cygler1,2 (Bacterial Structural Genomics Initiative)

1. NRC-BRI, Montreal QC Canada,
2. Department of Biochemistry, McGill University, Montreal, QC Canada,
3. Department of Chemistry and Biochemistry, Concordia University, Montreal QC Canada
4. Department of Biochemistry, Queen's University, Kingston ON Canada

Bacterial proteins participate in numerous transient and more stable associations (complexes) in order to perform a variety of biochemical processes. Both high-throughput interactome and individual investigator-initiated studies have provided a wealth of data on protein complexes in bacteria, particularly for E. coli. For the past few years the BSGI has focused efforts on the preparation, crystallization, structural analysis and functional characterization of a number of protein complexes, mainly from E. coli. Various approaches for protein expression, complex formation and stabilization were optimized for each of the systems under study. The structures determined include those involved in Fe-S cluster assembly, dihydroxyacetone metabolism and type II cohesin-scaffoldin interactions. In some instances, NMR titration experiments and small angle X-ray scattering have also been utilized to gain further structural insights. Using the structures as a starting point, mutagenesis combined with in vitro and in vivo functional assays have provided new biological insights into these processes. Research supported by CIHR.

SESSION 13: STRUCTURAL GENOMICS AND DRUG DISCOVERY II

Stephen Burley
Distinguished Lilly Research Scholar
Lilly Biotechnology Center
Eli Lilly and Company
10300 Campus Point Drive, Suite 200
San Diego, CA 92121

Challenges in Fragment Based Structure Guided Drug Discovery

Stefan Knapp
Principal Investigator
Phosphorylation Dependent Signalling Group
Nuffield Department of Clinical Medicine
Structural Genomics Consortium (SGC)
University of Oxford
UK

Targeting Phosphorylation Signalling Networks

Despite large research efforts in the kinase area, selective targeting of kinases and rational target selection remains an enormous challenge. As a consequence the overall effort in kinase drug discovery is currently declining and the remaining drug discovery efforts focuses on only a small subset of already clinically validated targets in the oncology area. This development suggests that the phosphorylation signaling drug development area requires novel strategies and ideas in order to spawn new discovery efforts. Our laboratory is interested in identifying new targeting strategies for protein kinases and in understanding the structural mechanisms of kinase activation. As a basis for such studies we solved more than 50 novel catalytic domain structures, established a panel of close to 200 recombinant enzymes for inhibitor cross screening and a library of kinase targeted inhibitors. In this talk I will present data on large scale kinase active site comparisons and inhibitor parallel cross-screening that led to the identification of unique structural features and the development of highly specific inhibitors. I will also present data that highlight the influence of structural changes induced by kinase regulators and/or kinase activation states on inhibitor binding and present examples how kinases can be targeted outside the ATP binding site.

William Zuercher
Department of Chemical Biology
Platform Technology and Science
GlaxoSmithKline
Research Triangle Park
North Carolina
USA

The Published Kinase Inhibitor Set: A resource to develop probes for the untargeted kinome

Protein kinases are chemically tractable drug targets, yet <10% of the human kinome has been thoroughly explored with selective small molecule inhibitors. It is likely that pharmacological evaluation of the whole kinome will uncover new opportunities for drug discovery. We will describe our strategy to probe the untargeted kinome via compound sharing and pre-competitive collaborations.

CLOSING LECTURE

Aled Edwards
Driector/CEO
Structural Genomics Consortium
University of Toronto
MaRS South tower, Suite 711
101 College St.
Toronto, ON M5G 1L7

Structural genomics: Opportunities to drive biology

The dramatic advances in genome biology are not being rapidly translated into biological understanding or into cures for diseases. A survey of the human protein literature reveals that most papers are published on proteins that were known prior to the sequencing of the human genome. As one example, >65% of the papers published in 2009 on human protein kinases focused on the 10% of protein kinases that were "hot" in the early 1990's.

Several features of biomedical research and research funding conspire against the investigation of the genome's "dark matter", but one is of particular relevance to structural biologists. On many occasions, exciting genetic observations stall because the next critical step in gaining understanding – the functional analysis of the encoded protein – requires techniques, equipment and experience that many geneticists do not have available to them. Although many of these genetically "validated" proteins are "uncharacterized", they likely represent the richest source of truly novel science. This remarkable opportunity should serve as a call to arms to structural biologists – the discipline with the deepest knowledge of protein biochemistry, and the discipline best poised to make a real difference in the global effort to understand human biology.