



Contribution ID : 24

Type : **not specified**

Synchrotrons Enable New Research in Genomics and Proteomics

Thursday, 5 May 2011 09:30 (1:00)

Abstract content

Genome sequencing projects rapidly expand protein sequence space and allow comprehensive approaches to studies of entire cellular systems. The accumulation of sequence data has accelerated significantly and now includes studies of microbiomes and metagenomes. However, many aspects of protein function, including molecular recognition, assembly and catalysis, depend on the 3D atomic structure. Protein structural analysis contributes to an understanding of the evolutionary and functional relationships among protein families that are often not apparent from the genome sequences. In the past ten years, third generation synchrotron sources and dedicated macromolecular crystallography (MX) beamlines have expanded our competence in determining protein structures using X-ray crystallography. MAD or SAD data are collected from cryo-protected crystals and structures are determined with a semi-automatic approach at the synchrotron beamline. Protein models are auto-built and structures refined, verified and analyzed using integrated computational tools and then deposited in public databases. World-wide Structural Genomics efforts took advantage of these resources and contributed a complementary array of the rapid, highly integrated and cost effective methods in molecular biology, proteomics, structure determination and bioinformatics and created efficient structure determination pipelines. For example the semi-automated pipeline of the Midwest Center for Structural Genomics (MCSG), one of the Large-scale Production Centers of the NIH-funded Protein Structure Initiative (PSI), comprises: (1) classifying all available genomic sequences to establish a prioritized target set of proteins from human pathogens and eukaryotes, (2) cloning and expressing proteins of microbial and eukaryotic origin, (3) purifying and crystallizing native and derived proteins for X-ray crystallography, (4) collecting data and determining structures, (5) analyzing structures for fold and function assignment, and homology modeling of related proteins. The pipeline when combined with data collection MX facilities at the third-generation synchrotrons, advanced software and computing resources resulted in significant acceleration of protein structure determination and overall reduction of cost. These new approaches have allowed the MCSG to determine nearly 1300 structures and the entire PSI over 5000 structures. All structures are annotated for function and ligand binding. Homology models are generated for members of protein families and ultimately will provide good structural coverage of major protein families. Moreover the PSI has comprehensively sampled the entire prokaryotic protein sequence space more broadly than ever before. The PSI Centers have cloned nearly 150,000 genes of proteins from a significant fraction of all large protein sequence families and purified nearly 50,000 highly diverse proteins creating a unique resource. Many of these structures represent protein families of high biological and biomedical interest, and many have provided novel functional hypotheses. Results and data are made available to the scientific community. Structural Genomics efforts build on the constantly growing foundation laid by the highly successful genome and metagenome sequencing projects that are enriching and expanding our knowledge of the protein universe. This discovery-based program contributes to studies the co-evolution of protein structure and function and is highly complementary to traditional hypothesis-driven approaches. Structural Genomics provides a wealth of ideas, concepts and understanding of mechanisms for acquisition of novel biological function and the evolution of biological systems.

This work was supported by NIH Grant GM094585 and by the U.S. DOE, OBER contract DE-AC02-06CH11357.

Summary

Primary author(s) : Dr. JOACHIMIAK, Andrzej (Argonne National Laboratory)

Presenter(s) : Dr. JOACHIMIAK, Andrzej (Argonne National Laboratory)