

Structural Genomics combines genomic data and structural biology to advance protein fold and function space

Andrzej Joachimiak, Midwest Center for Structural Genomics & Structural Biology Center, Biosciences, Argonne National Laboratory, Argonne, IL 60439, USA.

The main objective of the Structural Genomic effort is to rapidly determine the structures of strategically selected protein targets to elucidate protein fold space and ultimately provide structural coverage of major protein families with sufficient granularity to allow 3D homology modeling of all proteins using only computational methods. This will provide the foundation for 21st century structural biology when structures of virtually all proteins will be found in the Protein Data Bank or derived by computational methods. The Midwest Center for Structural Genomics (MCSG) is achieving these goals by developing, implementing and refining rapid, highly integrated and cost effective methods for structure determination by x-ray crystallography at 3rd generation synchrotrons. The MCSG has established a structure determination platform that comprises: (1) classifying all available genomic sequences to establish a prioritized target set and includes proteins from human pathogens and higher eukaryotes, (2) cloning, and expressing genes and gene fragments of microbial and eukaryotic origin, (3) purifying and crystallizing native and derivatized 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 MCSG platform contributes large number of unique structures and provides for rapid model validation and deposition in PDB.

This work was supported by the grants from the National Institute of Health (GM62414 and GM074942) and the U.S. Department of Energy, Office of Biological and Environmental Research under Contract W-31-109-ENG-38.