Poster I-57

Targeting Proteins from the Human Cancer Protein Interaction Network (HCPIN) for Structural Genomics

Authors:
Y. J. Huang (*Department of Molecular Biology and Biochemistry, Center for Advanced Biotechnology and Medicine, Ru*)
L. J. Lu (*Department of Molecular Biology and Biochemistry,Yale University,Northeast Structure Genomics Consor*)
M. B. Gerstein (*Department of Molecular Biology and Biochemistry,Yale University,Northeast Structure Genomics Consor*)
G. T. Montelione (*Department of Molecular Biology and Biochemistry, Center for Advanced Biotechnology and Medicine, Ru*)

Short Abstract: The Northeast Structure Genomics Consortium is targeting hundreds of proteins from the Human Cancer Protein Interaction Network. This interaction network has been constructed by combined analysis of seven signal transduction pathways from the KEGG Database, together with data from the Human Protein Reference Database, including >2000 proteins and ~5000 interactions.

Long Abstract:
Structural genomics is an international effort aimed at providing three-dimensional structures, either directly or by modeling, for all proteins in nature. The Northeast Structure Genomics Consortium (NESG) is targeting hundreds of proteins from the Human Cancer Protein Interaction Network (HCPIN). This interaction network has been constructed by combined analysis of seven signal transduction pathways from the KEGG Database, together with data from the Human Protein Reference Database (HPRD). HCPIN includes >2000 proteins and ~5000 interactions. About 42% of residues in this network of cancer-related proteins have 3D structure coverage, as defined by loose homology-modeling criteria (i.e. Blast E_val < 10-6). However, only ~18% of residues are structurally covered using more stringent homology-modeling criteria (i.e. >80% sequence identity). Having high-resolution structures or models for these human proteins associated with cancer is important for understanding enzyme active sites and for drug design. Proteins or domains from the cancer-related network containing segments with no structural coverage based on stringent homology-modeling criteria (i.e. >80% sequence identity) have therefore been targeted by NESG for cloning, expression, purification, and 3D structure determination using X-ray crystallography or NMR methods. Specifically, we have selected an initial set of ~300 human proteins suitable for high throughput structure determinations as our first pass targets. These protein targets are listed online at http://www-nmr.cabm.rutgers.edu/bioinformatics/ZebaView2/index.html. The topologies of the cancer-related interaction network have also been studied, and potential functions of domains within the network have been assessed. Proteins that play central roles as “hubs” or “bottlenecks” in the interaction network are discussed.