Pacific Symposium on Biocomputing 5:90-91 (2000)

PROTEIN STRUCTURE PREDICTION IN BIOLOGY AND MEDICINE

Roland L. Dunbrack Jr.
Institute for Cancer Research
Fox Chase Cancer Center
7701 Burholme Avenue
Philadelphia PA 19111
RL_Dunbrack@fccc.edu

Keith Dunker
School of Molecular Biosciences
Washington State University
Pullman, WA 99164-4660
dunker@mail.wsu.edu

Adam Godzik
The Burnham Institute
10901 N. Torrey Pines Rd.
La Jolla, California 92037
adam@burnham-inst.org

Protein structure prediction from sequence has made rapid advances in the last few years, gaining increased attention after the success of the CASP meetings. Specifically, threading predictions were shown to work for many cases previously thought to be outside of the scope of prediction methods. The papers in this session address a wide scope of topics, ranging from techniques for validation of prediction methods and further improvements of threading algorithms, to specific applications of protein structure predictions in biology.

Protein structure prediction is one of the most active fields in computational biology. The interest in structure prediction is fueled by the extraordinary pace of discovery of new protein sequences in genome sequencing experiments. Close to 50% of all protein sequences can be assigned folds with existing fold recognition methods,

Still, many challenges lie ahead – from further improvement of prediction protocols to demonstrating that protein fold and structure prediction can indeed contribute to understanding of important biological problems. The papers presented in this session span a wide spectrum of problems facing the protein structure prediction field.

Analysis of known protein structures in "**Folding nuclei in 3D protein structures**" by Galzitskaya, Skoogarev, Ivankov and Finkelstein addresses a question of finding a possible folding nuclei in proteins. This interesting problem may lead to significant breakthroughs in *ab initio* protein fold prediction. In a similar vein, "**Recognition of protein structure: elucidating the specific roles of β-strands, α-helices and loops**" by Reva and Topiol analyzes protein structures to

determine which of the secondary and supersecondary elements contribute most to the threading-based recognition of structurally analogous proteins.

In the next group, several papers suggest and evaluate possible improvements to the existing fold prediction algorithms. "**Hybrid fold recognition: combining sequence derived properties with evolutionary information**" by Fischer introduces a new hybrid threading algorithm that combines the information coming from analysis of protein structures with information coming from the analysis of families of homologous proteins. This contribution shows that the two sources of information can be successfully combined to arrive at a prediction method better than that from either one alone. "**Eliminating superfluous neighbor pairs while threading fold models**" by Bienkowska, Rogers and Smith evaluates ways of improving energy functions used in threading by adding information about interaction directionality on the atomic level. Backofen, Will, Bond and Clote in "**Algorithmic approach to quantifying the hydrophobic force contribution in protein folding**" analyze the contribution of hydrophobic forces to protein folding by using a simple backbone rebuilding algorithm using only selected elements of energy fields. Finally, a comprehensive analysis of several variants of threading algorithms on a benchmark of hard and medium-hard fold recognition examples is presented in "**How universal are fold recognition parameters. A comprehensive study of alignment and scoring function parameters influence on recognition of distant folds**" by Olszewski.

A final group of contributions describe applications of fold recognition and modeling methods to study specific molecular systems. Samudrala, Xia, Levitt, Cotton, Huang and Davis in "**Probing structure-function relationships of the DNA polymerase alpha-associated zinc-finger protein using computational approaches**" combine several *ab initio* modeling methods to predict and build an atomic model of a 67-residue fragment of a DNA polymerase associated zinc finger protein. The model is used to predict details of the protein function, with additional experimental data used to verify the prediction. Ranganathan, Male, Ormsby, Giannakis and Gordon in "**Pinpointing the putative heparin/sialic acid binding residues in the 'sushi' domain 7 of factor H: a molecular modeling study**" use a different set of threading/modeling tools to study 60-residue heparin binding domains, involved in regulation of the complement system. Finally, algorithms predicting flexibility are used to study a family of cytochromes C by Thorpe, Hespenheide, Yang and Kuhn in "**Flexibility and critical hydrogen bonds in cytochrome C**".

Acknowledgments.

The session organizers thank the session reviewers who helped us in the difficult task of choosing the best contribution from a large number of excellent submissions.