Amino Acids (2008) 35:513–515
DOI 10.1007/s00726-008-0078-x

EDITORIAL

Editorial

Guo-Ping Zhou

Published online: 12 April 2008

The effects of protein structures as well as protein–protein interactions on the functions of cells are always one of the most important research subjects in life science due to their close relevancy to the health of human beings. For instance, some kinds of cancers are induced by the mutants in cellular genes, resulting in "ill-proteins" leading to terrible consequences; Parkinson's disease and prion diseases are caused by the miss-folding of proteins, leading to the propagation of infectious diseases. Accordingly, in order to inderstand in-depth the molecular mechanisms of various disease forms, it is indispensable to acquire the structural information of the proteins concerned and their interactions in the cellular metabolism pathways. This kind of knowledge can provide useful insights for therapeutic treatments.

This special issue of Amino Acids contains 12 review and research articles reporting the recent progresses in the areas of protein structure, protein–protein interaction, structural bioinformatics, network biology, and system biology.

The review article entitled "Protein function prediction with high-throughput data" by Dr. Xing-Ming Zhao (Institute of Systems Biology, Shanghai University, Shanghai, China), Dr. Kazuyuki Aihara (ERATO Aihara Complexity Modelling Project, JST, Tokyo, Japan), and Dr. Luonan Chen (professor, ERATO Aihara Complexity Modelling Project, JST, Tokyo, Japan) presents a comprehensive framework of utilising machine learning techniques to predict protein function based on high-throughput data. Furthermore, the utilisation of various kinds of data sources, e.g., protein–protein interaction and microarray, and the integration of these data sources for protein function prediction were also presented. Since the functions of proteins are closely relevant to their subcellular localization, introduced in their review are also some very powerful web-servers for identifying protein subcellular localization developed recently, such as the "Cell-PLoc" package.

To provide insight into the mechanism of phosphoryl transfer in the nitrogen pathway and its regulation of bacterial K^+ uptake, one important approach is to determine the structures of the relevant proteins and their complexes, including membrane proteins, many of which are key targets for drug discovery. To report their findings on the NPr protein by the NMR studies, Dr. Xia Li (Eppley Institute for Research in Cancer and Allied Diseases, University of Nebraska Medical Center, Nebraska Medical Center, Omaha, NE), Dr. Alan Peterkofsky (Laboratory of Cell Biology, National Heart, Lung and Blood Institute, National Institute of Health, Bethesda, MD), and Dr. Guangshun Wang (Assistant Professor, Eppley Institute for Research in Cancer and Allied Diseases, University of Nebraska Medical Center, Nebraska Medical Center, Omaha, NE) contributed a research article entitled "Solution structure of NPr, a bacterial signal-transducing protein that controls the phosphorylation state of the potassium transporter-regulating protein IIA^{Ntr}". They observed that both the 3D structural and dynamic information of the protein obtained by their NMR spectroscopy studies are the foundation for understanding the specificity of protein-mediated signal transduction and its regulation of bacterial K^+ utilisation, a discovery that is essential for understanding the protein function as well as its regulation, but is not readily observed by X-ray diffraction.

G.-P. Zhou (✉)
Protein Structure in Chemical Sciences,
University of Alberta, T6G2E8 Edmonton, AB, Canada
e-mail: gzhou@crystal.harvard.edu

The accurate extraction of 3D structural information of a protein from its amino acid sequence is one of the unsolved problems in bioinformatics. The research article by Dr. Takeshi Kikuchi (Professor, Department of Bioscience and Bioinformatics, College of Information Science and Engineering, Ritsumeikan University, Kusatsu, Shiga, Japan) in "Analysis of 3D structural difference in the IgG binding domains based on the interresidue average distance statistics" addresses this problem. The techniques developed in his group can successfully predict several folding properties of proteins, and the information thus obtained can be used for detailed protein 3D structure predictions.

Knowledge of structural classes, which categorise proteins based on the amount and arrangement of the constituent secondary structures, is applied in numerous important predictive tasks that address structural and functional features of proteins. To this end, Drs. Lukasz A. Kurgan (professor), Tuo Zhang and Hua Zhang (Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada), and Drs. Shiyi Shen and Jishou Ruan (College of Mathematical Science and LPMC, Nankai University, Tianjin, China) propose two novel structural class assignment methods that use either true or predicted 1D secondary structure as the input in their paper. Using a large set of low identity sequences and the structure-based assignment performed in SCOP as the gold standard, the accuracies of their methods equal 76 and 75%, respectively. Their method can be used to perform automated assignment of structural classes from the protein sequences.

Recently, a collective effort from multiple research areas has been conducted to understand biological process at the system level. These researches require the ability to simulate particular biological systems, such as cells, organs, organisms, and communities. The paper by Drs. Li-Hong Ren, Yi-Zhen Shen, and Xiang-Feng Zhang (College of Information Sciences and Technology, Donghua University, Shanghai, China) and Yong-Sheng Ding (professor, Engineering Research Center of Digitized Textile and Fashion Technology, Ministry of Education, Donghua University, Shanghai, China) proposes a novel bio-network simulation platform for system biology studies by combining agent approaches. The bio-network platform can be used to simulate the behaviours of biological systems and model them in terms of bio-entities and society-entities. As a demonstration, they discuss how a protein–protein interaction (PPI) network with remarkable ability emerges from a society of autonomous interactive components. The proposed approach can be used to develop various simulators, from design to development, and complexity validation of biological systems.

The paper by Drs. Loris Nanni and Alessandra Lumini (DEIS, IEIIT—CNR, Università di Bologna, Viale Risorgimento, Bologna, Italy) proposes a system for predicting the membrane proteins type directly from the amino acid sequence. The feature extraction step is based on the concept of Chou's pseudo amino acid composition and an encoding technique that combines the physicochemical amino-acid properties with the residue couple model. A reduced set of classifiers, radial basis function support vector machines, are selected by running the sequential forward floating selection, where the objective function is the minimization of the error rate in the training set. The validity of the novel approach is proved by the comparison with other state-of-the-art methods in the tested problem.

The success rate of predicting protein structural class is a challenge project that is related to the accurate extraction of structural and functional information.

The research article entitled "Prediction of protein structure class by coupling improved genetic algorithm and support vector machine" by Drs. Z.-C. Li, X.-B. Zhou, Y.-R. Lin, and X.-Y. Zou (professor, School of Chemistry and Chemical Engineering, Sun Yat-Sen University, Guangzhou, China) presented a novel method for the prediction of protein structure class by coupling the improved genetic algorithm (GA) with the support vector machine (SVM). Improved GA is applied to the selection of optimised feature subset and the optimisation of SVM parameters. The predicted results are very encouraging, indicating that the approach might hold a high potential to become a useful tool in bioinformatics.

Because it is extremely labour-expensive and even impossible to determine in a timely manner the structures for hundreds of thousands of protein sequences by experiments alone, it is highly desirable to develop computational methods for classifying the quaternary structures of proteins from their primary sequences. In the research article entitled "Prediction of protein quaternary structure with a novel method of the segmented pseudo amino acid composition", Drs. Shao-Wu Zhang (Associate Professor), Wei Chen, Feng Yang and Quan Pan (Professor) of the College of Automation, Northwestern Ploytechnical University, Xi'an, China, developed a new computational method based on the concept of Chou's pseudo amino acid composition (i.e., sequence-segmented pseudo amino acid composition) to predict protein quaternary structures. The results show that the new approach is quite promising for predicting protein quaternary structure.

It has been known that a disordered protein is related to some diseases such as Alzheimer's and Parkinson's diseases. The paper entitled "DPROT: Prediction of disordered proteins using evolutionary information" by Drs. Deepti Sethi, Aarti Garg, and G.P.S. Raghava (Professor, Institute of Microbial Technology Chandigarh, India) describes a method developed for predicting disordered proteins with high accuracy. A number of SVM

models have been developed using various features of a protein, including evolutionary information in the form of a PSSM profile. In order to serve the scientific community, the web server DPROT has been developed for predicting disordered proteins from their amino acids sequence.

The paper by Dr. Yonge Feng and Professor Liaofu Luo (Laboratory of Theoretical Biophysics, Faculty of Science and Technology, Inner Mongolia University, Hohhot, China) develops a novel sequence-based method, the so-called "tetra-peptide-based increment of diversity with quadratic discrimination analysis (TPIDQD for short)" for protein secondary structure psrediction. The results show those tetra-peptide signals can indeed reflect some relationship between amino acid sequence and its secondary structure, indicating the importance of tetra-peptide signals as the protein folding code in the protein structure prediction.

The paper by Dr.Tze-Chuen Lee (Bioinformatics Institute, Singapore), Dr. Ann SG Lee (National Cancer Center, Singapore), and Dr. Kuo-Bin Li (Center for Systems and Synthetic Biology, National Yang-Ming University, Taipei, Taiwan) proposes a phylogenetic approach to predict the effects of missense mutations. The results obtained by their approach demonstrate that the use of amino acid properties can enhance the prediction of biochemical and structurally important residues and thus also predict the significance of missense mutations.

Finally, the review article by Dr. Zhi-Ping Liu, Dr. Ling-Yun Wu, Dr.Yong Wang, Dr. Xiang-Sun Zhang (Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing), and Professor Luonan Chen (Institute of Systems Biology, Shanghai University, Shanghai, China) is focussed on the topic of in silico annotation of protein function by largely accumulated sequence and structure data. In particular, the authors emphasized that the newly structure-based methods are able to identify the locally structural motifs and reveal their relationship with protein functions.

As one can see from the above contributions in this special issue, the inter-penetration between protein structure, bioinformatics, network biology, system biology, and biomedicine is an inexorable trend, particularly in the beginning of this era.