Amino Acids (2007) 33: I–LXX
DOI 10.1007/s00726-007-0578-0
Printed in The Netherlands

Amino Acids

10th International Congress on Amino Acids and Proteins (ICAAP)

Kallithea, Greece
August 20–25, 2007

Abstracts

Presidents:

M. Liakopoulou-Kyriakides, Aristotle University, Greece

R. Harris, University of Chichester, U.K.

K. C. Chou, Gordon Life Science Institute, U.S.A.

Honorary President:

G. Lubec, University of Vienna, Austria

predict subcellular localization of proteins is to identify sequence motifs such as signal peptide or nuclear localization signal. The main limitation of motif-based methods is that all proteins residing in a compartment do not have universal motifs. In order to overcome these limitations, researcher is using residue composition of proteins that includes amino acid, pseudo amino acid and dipeptides composition. Recently, evolutionary information in form of PSSM profiles has been used to predict subcellular localization of proteins. One of the major advantages of composition-based approach is that it provides fixed number of features irrespective of length of protein. This allows researchers to use machine-learning techniques like artificial neural network (ANN), support vector machines (SVM), nearest neighbors' method (KNN), etc.

Polymorphic proteins and natural selection in wood mice (*Apodemus*): hypotheses from bioinformatics and population models

P. R. Ramsey[1] and **J. J. Purger**[2]

[1] School of Biological Sciences, Louisiana Tech University, Louisiana, USA
[2] Department of Animal Ecology, University of Pécs, Pécs, Hungary

Gene and protein databases allow testing of hypotheses about the factors which maintain or eliminate protein polymorphisms. Enhanced data banks for structural and functional bioinformatics have stimulated reexamination of organism performance associated with protein variation. Persistence of polymorphisms, differential fitness among genotypes, and the selective advantage(s) of allozyme heterozygosity are being investigated from the expanded perspectives of single nucleotide polymorphisms, gene-expression breadths, and catalytic geometries. Species' histories and current environments make wood mice from the Balkan region ideal for investigations of natural selection and protein functional structure. Four of the five species of wood mice (genus *Apodemus*) in Europe are in the Serbian province of Vojvodina, east and north of the Sava and Danube rivers, where original biotopes of steppes and woodlands are now subdivided by canals and agriculture into a mosaic of disturbed areas and enlarged ecotones. Northern Serbia was along the permafrost limit during the Quaternary glaciations, and two closely related, resident species of *Apodemus* had different (Iberian vs. Balkan) refugia. To assess population subdivision and potential introgression, we collected 260 specimens from 5 localities across 3 years, surveyed variation in 3 structural and 15 enzymatic (both housekeeping and tissue-specific) protein loci. In spite of operative factors that should reduce variation, there is greater protein variation in Serbian *Apodemus* than in other examined populations and geographic patterns of variation in allelomorph frequencies and heterozygosity, yet lack of isolation by distance among populations. To test for selection on candidate genes, we used our data and previously published data on *Apodemus* allozymes, plus mitochondrial and nuclear gene sequences, to test hypotheses that the majority of the variation in heterozygosity and the uniform array of polymorphic systems are explained by nonsynonomous-to-synonomous (K_a/K_s) amino acid substitutions, protein quaternary structure, and expression breadth.

Multi-layered network structure of amino acid (AA) metabolism characterized by each essential AA-deficient condition

N. Shikata[1], **Y. Maki**[2], **Y. Noguchi**[3], **M. Mori**[3], **T. Hanai**[1], **M. Takahashi**[3], and **M. Okamoto**[1]

[1] Graduate School of Systems Life Sciences, Kyushu University, Japan
[2] Department of Digital Media, Fukuoka International University, Japan
[3] Research Institute for Health Fundamentals, Ajinomoto Co. Inc., Japan

The concentrations of free amino acids in plasma change coordinately and their profiles show distinctive features in various physiological condition, however, some of the changes observed in response to physiological status can not always be explained by the conventional flow-based metabolic pathway network. In this study, we have inferred the network structure of plasma amino acids with threshold-test analysis and multi-level-digraph analysis methods, without including the prior knowledge of metabolic pathway, using the plasma samples of rats which are fed diet deficient in single essential amino acid.

In the inferred network, we could draw some interesting interrelations between plasma amino acids as follows: 1) Lysine is located at the top control level and has effects on almost all of the other plasma amino acids. 2) Threonine plays a role in a hub in the network, which has direct links to the most number of other amino acids. 3) Threonine and methionine are interrelated to each other and form a loop structure.

Based on the inferred network structure, the dynamic analysis based on S-system is now being performed and recent advances will be discussed.

Structural bioinformatics and computer aided design of novel drugs and functional proteins

D.-Q. Wei, L. Li, and **K.-C. Chou**

Department of Bioinformatics, College of Life Sciences and Biotechnology, Shanghai Jiaotong University, Shanghai, China

Genome projects are yielding protein sequences for which there is no knowledge about function or conformation. The draft human genome has now been available and the exploitation of this unique source of knowledge is a major challenge for biology. In parallel with these sequencing projects, there are structural genomics initiatives involving the determination of the conformations of uncharacterized proteins in the genomes with the particular aim of determining function. In addition, gene expression arrays are providing a mass of data that require analysis to relate protein sequences to activity under different conditions and in different cellular locations. Knowledge of the structure and function of the relevant proteins, revealed by crystallography and NMR, is central to the interpretation and exploitation of this pool of biological information. Three-dimensional structure can guide further experiments to probe activity and direct the systematic design of therapeutic agents for diseases. The experimental determination of protein structure remains difficult. There are large disparities between the numbers of protein sequences. Protein modeling provides a valuable approach to maximize the biological knowledge that can be obtained given these disparities. Accordingly, the task of the Structural Bioinformatics is the development of protein modeling algorithms and the application of the technology to systems of interests. Methodologies based on the known protein structures from experiments have been developed to predict 3D structures from a 1D sequence information, which is known as the homology modeling. It has become a reliable tool as long as the homologue with experimental 3D structure of significant similarity (usually greater than 35%) could be found. Other methods, for example, ab initio molecular dynamics, is still too slow to predict structure of a real protein. Often the structures of many drug targets are not available. The structural bioinformatics tools are needed to generate them readily to initiate drug design processes. We have predicted many protein structures using homology modeling to facilitate the computer aided drug design, which include database screening, pharmacophore search, docking and molecular dynamics conformation search. To join the worldwide efforts against H5N1 viruses which experience rapid mutation and become increasingly drug-resistant. A homology model of the H5N1-NA from the highly pathogenic chicken H5N1 A viruses isolated during the 2003–2004 influenza outbreaks in Japan was built based on the crystal structure of N9-NA complexed with DANA (PDB code: 1F8B). It was found that the traditional constituent residues around the active site of NA family are highly conserved in the H5N1-NA. However, a partially lipophilic pocket composed by Ala248