

Amino acid sequence fingerprints in divergent evolution of proteins

Monday, February 16th, at 11:00

UKB MU Kamenice 5/A3, seminar room

Dr. Štefan Janeček DrSc.

Laboratory of Protein Evolution, Department of Protein Structure and Function,
Institute of Molecular Biology, Slovak Academy of Sciences, SK-84551 Bratislava, Slovakia
e-mail: stefan.janecek@savba.sk

Proteins may basically be related by divergent or convergent evolution. In divergent evolution, i.e. the evolution from a common ancestor, there is only a small percentage of residues in the amino acid sequence of a protein, that have to be conserved in order the protein retains its function. The vast majority of the sequence can tolerate the amino acid substitution. This is in agreement with the general rule that the three-dimensional tertiary structure of a protein is more conserved than its primary structure. The tertiary structure is able to accommodate various sequence amendments without a substantial change of the spatial arrangement of the protein molecule so that if the essential residues of the active site are preserved, the resulting protein function and activity can remain without any observable modifications. The essential amino acid residues are mostly parts of the so-called conserved sequence regions that cover the isolated segments belonging to the active site of the protein. In the case the proteins and/or enzymes form families with members related by their function (e.g., the members can differ by their detailed enzyme specificity), then it should be possible to identify the amino acid residues responsible for the subtle differences among the family members. This means that the conserved sequence regions can be used as the sequence fingerprints of the individual members of the protein/enzyme family. The amino acid sequence fingerprints in divergent evolution of proteins will be illustrated using the alpha-amylase enzyme family as a model of divergently evolved protein family consisting of almost 30 different enzyme specificities from hydrolases, transferases and isomerases, and more than several thousand sequences available covering a wide spectrum of taxonomic sources from all the three domains of life (Bacteria, Archaea and Eucarya).