J Genet Mol Biol 2018, Volume 2

DECODING THE AMINO ACID SEQUENCE OF A PROTEIN TO EXTRACT ITS FOLDING INFORMATION

Takeshi Kikuchi
Ritsumeikan University, Japan

The problem how a unique and complex protein 3D structure forms is a long-standing problem in structural biology. It is very interesting how the folding information of a protein can be extracted from its sequence. When we make a plot of a pair of residues with the shorter average distance on a contact map so that the average distance of a residue pair <r> shows a scaling rule N1/3 in an implicit way, a region to be compact can be predicted along the amino acid sequence of a protein. Furthermore, an effective inter-residue potential can be defined based on the inter-residue average distance statistics. The contact frequency of a residue with other residues will provide information of a site of the initial folding events of a protein. We call the contact frequency F value. Thus, we consider that conserved hydrophobic residues during the evolution around a peak of the contact frequency plot of a protein have a significant role on the folding of a protein. In this study, the effectiveness of the present method is examined taking proteins with characteristic 3D structures such as lysozyme related proteins, β-trefoil proteins and so on. The results are compared with those of HD-exchange or Φ-value experiments. A, F, L, M, V, Y, W are taken as hydrophobic residues in this study. Furthermore, evolutionary analysis is also conducted with the phylogenetic tree of proteins in a superfamily. The common mechanisms of the proteins in a superfamily are discussed. We also discuss the general mechanism of protein folding.