GOR Method for Protein Structure Prediction using Cluster Analysis

Abstract

Protein structure prediction is one of the most important problems in modern computational biology. The emphasis here is on the use of computers because most of the tasks involved in genomic data analysis are highly repetitive or mathematically complex. The problem of this research focus on secondary structure prediction of amino acids. In the present research work, the GOR (Garnier, Osguthorpe, and Robson) Method is implemented so as to deal with amino acid residues to predict the 2D structure using different input formats of sequences. Combination of amino acids results in formation of protein through peptide bond. The practical implementation of protein structure prediction completely depends on the availability of experimental database. The analysis and interpretation of bioinformatics database which includes various types of data such as nucleotide and amino acid sequences, protein domains, and protein structures is an important step to determine and predict protein structure so as to understand the biological and chemical activities of organisms. GOR method uses the information theory to generate the code that relates amino acids sequence and secondary structure of proteins. Three scoring matrices are prepared in GOR method to calculate the probability of each amino acids present in every positions. Cluster analysis is used as data
mining model to retrieve the result

References


Index Terms

Computer Science
Artificial Intelligence

Keywords
Amino Acid  Protein  Polypeptide  DNA  RNA  DSSP  GOR