Calculation of amino acids propensity for secondary structures

Seyede Niloufar Nouri, Seyed Shahriar Arab*, Javad Zahiri
Department of Biophysics, Faculty of Biological Sciences, Tarbiat Modares University,Tehran, Iran*sh.arab@modares.ac.ir

Abstract: The concept of "propensity" in fact indicates the frequency of each amino acid presence in each of three structures: Helix, Extended and Coil. Calculation of the amino acid propensities for secondary structures was first used in 1974 by Chou and Fasman[1]. The original dataset used by them was very small and included 15 proteins. The most recent calculation of the propensity for amino acids was performed on 2000 proteins in 2006 by Susan Costantini and Giovanni Colonna[2]. Nowadays, in structure prediction methods, the calculation of the propensity for different purposes has a special place and is used in high-confidence reliable bioinformatics algorithms [3][4]. By updating and increasing data, it is necessary to recalculate the propensity of each residue to be used in the new developed algorithms and databases. First of all, the dataset was created. From the entire PDB database, which included 134251 structures, duplicate proteins were deleted by using the CD-HIT software and the dataset was generated containing 34825 proteins. By running a program that reads DSSP files, the information of the secondary structure of the dataset was determined. Statistical analyzes on the distribution of amino acids and their secondary structures were obtained on the created dataset. Finally, the propensity table of each amino acid for placement in different secondary structures was updated.

Keywords*:* Propensity; Secondary structure prediction; Amino acids distribution

References

[1] P.Y. Chou, G.D. Fasman, Prediction of protein conformation, Biochemistry 13 (1974) 222–245.
[2] S.Costantini,G.Colonna, Amino acid propensities for secondary structures are influenced by the protein structural class, Biochemical and Biophysical Research Communications 342 (2006) 441–451
[3] W. Kabsch, C. Sander, How good are predictions of protein secondary structure? FEBS Lett. 155 (1983) 179–182.
[4] J. Kyngas, J. Valjakka, Unreliability of the Chou–Fasman parameters in predicting protein secondary structure, Protein Eng. 11 (1998) 345– 348.