Preference of dipeptides for adopting helix conformation

S. Fazli, Z. Rostami, N. Badri, K. Khalifeh & E. Heshmati*
Department of Biology, University of Zanjan, Zanjan, Postal code: 4537138791, IRAN
*Heshmati@znu.ac.ir

Abstract: Upon publication of levinthal paradox [1] and interesting works of Anfinsen [2], an important question has been raised about the correlation betweenc the sequence and ultimate unique 3D structure of protein and it became as one of the most active research fields in bioinformatics and structural biology. In order to address this problem, Chou and fasman were the first researchers who statistically analysed the frequency distribution of all 20 natural amino acids in the secondary structureal elements of proteins using structural information in structural databases. However, in spite of passing more than 40 years after pioneering work of chou-fasman for attributing the structural features to the sequence of protein [3], the lows governing the protein folding is still under serious questions [4]. Since the method proposed by chou-fasman was based on the preferences of single residues in any secondary structure, its prediction capability is limited [5]. Accordingly, it seems that updating this method to higher order building blocks rather than single amino acids may help computational biologists to obtain more accurate results in structural prediction. In line with this assumption, we statistically analyzed the relative frequencies of 20 naturally amino acides as well as all 400 possible dipeptides in helix conformation. Our dataset was obrtained from protein data bank consists of 500 proteins which were selected by applying appropriate filtering criteria in advanced search menu. After conversion of PDB file format to DSSP one, the results were used as input to PARS program which has been developed by our gruop [6]. The output of PARS, then were summarized and statistically analyzed. Comparison of our finding on the distribution of single amino acids shows high correlation with those obtained by chou-fasman indicating the similarity of two methods of analysis. For example, Ala, Glu, Lue and Gln can be considered as helix-former residues, while; Gly and pro are helix-breaker residues. Upon analysis of the distribution of dipeptides on the helix conformation including 3_{10}, π and α-helix, it was revealed that there are a kind of the neighbor effect [7] between different residues in the first and second positions of some representative dipeptides which divert them to adopt this conformation. These findings demonstrate that physico-chemical properties in dipeptides are not the same that is expected from its constituents residues. Resulting data of current work can be used for prediction the mutational effects of replacing a residue with other one on the local conformation around a given dipeptide. It may also be used for updating the chou-fasman method for predicting the secondary structure of proteins.

Key words: dipeptides; adopting helix

References

[1] C. Levinthal, Are there pathways for protein folding?, J. Chim. Phys. Physico-Chimie Biol. 65 (1968) 44–45.
[2] C.B. Anfinsen, Principles that Govern the Folding of Protein Chains, Science (80-.). 181 (1973) 223–230.
[3] P.Y. Chou, G.D. Fasman, Prediction of protein conformation., Biochemistry. 13 (1974) 222–245.
[4] K.A. Dill, S.B. Ozkan, T.R. Weikl, J.D. Chodera, V.A. Voelz, The protein folding problem: when will it be solved?, Curr. Opin. Struct. Biol. 17 (2007) 342–346.
[5] J. Kyngäs, J. Valjakka, Unreliability of the Chou-Fasman parameters in predicting protein secondary structure., Protein Eng. 11 (1998) 345–348.
[6] A. Fathinavid, K. Khalifeh, E. Heshmati, PARS a software for Protein Assignment Regarding Secondary structure, Unpubl. Data. (n.d.).
[7] M. Ghadimi, E. Heshmati, K. Khalifeh, Distribution of dipeptides in different protein structural classes: An effort to find new similarities, Eur. Biophys. J. (2017).