New Algorithm for Protein Catalytic Site Recognition, Based on Graph Matching

Sina Sameti, Seyed Shahriar Arab*
Department of Biophysics, Faculty of Biological Sciences, Tarbiat Modares University, Tehran, Iran
*sh.arab@modares.ac.ir

Abstract: With the increasing amount of protein structural information in related databases such as PDB[1] development of methods which analyze and use this structural information in order to detect and predict protein catalytic sites has become very important. For this reason, researchers always need to improve existing algorithms, or designing new algorithms which should search for a huge amount of data over a short period of time, with great precision[2]. In general, methods which identify and predict protein catalytic sites are divided into two parts[3] Sequence-based methods and Structure-based methods. Although sequence-based methods are widely used as reliable tools for predicting protein function, but these methods are doomed to failure in cases where there is little sequence similarity between proteins. In fact, there are proteins, despite their negligible similarity in the sequence, have the same structure and function [4]. Because of that reason, the use of structure based methods, such as those are in this article, has become more and more important.Information of the protein catalytic sites is available in databases such as CSA[5]. Here we use this information, by converting spatial information into a graphic representation of the protein catalytic sites and matching them with the whole protein structural graph, to identify functional residues of the protein(s). Now we have a secondary database from sructural data of protein catalytic site(s), by converting the three-dimensional data to two-dimensional graphs[6]. Finally, the graphs will align the catalytic position with the general building block of the protein. This algorithm is used to generate a bioinformatics tool which detects different catalytic locations that are existed in the structure of protein.

Keywords: Protein Catalytic Sites; Graph Theory; Graph Matching; Local Alignment; Structure-based Algorithms

References

[1] Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanuchi, T. and Tasumi, M. (1977), The Protein Data Bank. European Journal of Biochemistry, 80: 319–324.

[2] Chatterjee, U. K. Roy, and D. Halder, “Protein Active Site Structure Prediction Strategy and Algorithm,” vol. 7, no. 3, pp. 1092–1096, 2017

[3] Watson JD, Laskowski RA, Thornton JM. Predicting protein function from sequence and structural data. *Curr Opin Struct Biol* 2005, 15(3):275–284.

[4] Wang K, Samudrala R. FSSA: a novel method for identifying functional signatures from structural alignments. *Bioinformatics* 2005, 21(13):2969– 2977

[5] Porter, C. T., Bartlett, G. J., & Thornton, J. M. (2004).The Catalytic Site Atlas: a resource of catalytic sites and residues identified in enzymes using structural data. *Nucleic Acids Research*, *32*(Database issue), D129–D133.

[6] T. Fober, M. Mernberger, G. Klebe, and E. Hüllermeier, “Graph-based methods for protein structure comparison,” *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* vol. 3, no. 5, pp. 307–320, 2013.