Journal of Biochemistry and Cell Biology

Mini Review Open Access

A Short Note on Biochemical Functional and Protein Structures

Stark Johnson*

Department of Biochemistry and Molecular Genetics, Feinberg School of Medicine, North-western University, Chicago, IL 60611, USA

Abstract

The need for reliable computational methods to determine the biochemical function of these proteins is growing as the determination of protein sequences and structures through genome sequencing and structural genomics efforts grows exponentially. The efforts to address the problem of annotating the function of uncharacterized proteins at the molecular level are reviewed in this paper. The most recent trends in local structure-based methods have received less attention than the sequence- and three-dimensional structure-based methods for protein function prediction. This review focuses primarily on these local structure-based methods. The local spatial arrangements of these residues can be used to identify protein function, and computational methods have been developed to predict the residues that are essential for catalysis. In addition, for proteins with no known function, combining a variety of methods can aid in the acquisition of additional data and the improvement of function predictions. The various methods for predicting the function of proteins with no known function are being evaluated and tested by global initiatives like the Enzyme Function Initiative (EFI), COMputational Bridges to Experiments (COMBREX), and the Critical Assessment of Function Annotation (CAFA). By reducing the number of functional annotations that are either missing or incorrect, these initiatives and global collaborations will add significant value to the existing volume of structural genomics data and improve computational methods for predicting biochemical function [1, 2].

Keywords: Structural genomics; Protein function prediction; Local structure methods; Computational chemistry

Introduction

Since genome sequencing and high-throughput structure determination began, there has been a significant increase in the number of protein sequences and structures in databases like UniProt and the Protein Data Bank (PDB) [3]. The number of entries in the UniProt/TrEMBL database for protein sequences has increased by more than sixfold since January 2011 to over 89 million as of January 2015; Only a very small number of these proteins have a known purpose. Additionally, structural genomics projects like the Protein Structure Initiative (PSI) have resulted in the addition of over 13,000 structural genomics (SG) protein structures to the PDB. The purpose of the PSI, which was launched at the turn of the millennium by the National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health (NIH) in the United States, was to identify the three-dimensional structures of all families of proteins. The human genome project and the sequencing of many other organisms' genomes were finished at that time. The number of known protein structures has increased as a result of PSI and other SG programs' high throughput techniques [4, 5].

Discussion

Most of these protein structures lack reliable biochemical function-related information due to the PSI's primary focus on high volume structure determination and prompt public availability of protein structures; There are times when no functional annotation is provided [6, 7]. As a result, the majority of these proteins are given a putative function based on the closest match in sequence or structure; however, these assignments frequently contain incorrect functional labels, which can spread throughout databases. In 2010, the NIGMS began a brand-new PSI phase dubbed PSI: Biology. The purpose of this phase was to ascertain the SG proteins under structural investigation's biological functions. However, many functional annotations are still incomplete or incorrect. It is evident that improved computational techniques and biochemical experimentation-based verification are required. Computational methods that are dependable and accurate for predicting the function of proteins can improve the efficiency of experimental function verification and add significant value to genomics data. While several reviews of sequence-based and three-dimensional structure-based methods for function prediction have been published, this article focuses on more recent computational methods that use local structure to predict protein function at the molecular level; Prediction of the local spatial regions of the structure that are biochemically active serves as the foundation for each of these approaches. Last but not least, the research community as a whole's efforts to help test and verify functional predictions are investigated [8].

An examination of a protein's structure and shape is necessary for structure-based approaches to protein function prediction. By transferring the function of another similar protein with a known function, this analysis aids in determining where a ligand might bind. A protein's local site of biochemical activity can be used as a starting point for determining its function. Structure-based methods like geometric potential and geometric-based computational programs like Surfnet CASTp Ligsite PocketFinder look at the various properties of a protein surface or active site pocket to determine the identity and location of binding pockets. Based on PDB coordinate data, Surfnet generates a variety of protein surfaces, including van der Waals interactions, gaps between molecules, and pockets within a protein. Densities of the various surface output data are depicted as a grid. By applying a Gaussian function to the protein's atoms, these grids are created [9, 10]. The intensity of the densities is used to identify the specific important residues. Using the PDB, Swiss-Prot, and Online

***Corresponding author:** Stark Johnson, Department of Biochemistry and Molecular Genetics, Feinberg School of Medicine, North-western University, Chicago, IL 60611, USA, E-mail: Johnnson_stark@edu.in

Received: 02-Mar-2023, Manuscript No. jbcb-23-91626; **Editor assigned:** 06-Mar-2023, PreQC No. jbcb-23-91626 (PQ); **Reviewed:** 20-Mar-2023, QC No. jbcb-23-91626; **Revised:** 27-Mar-2023, Manuscript No. jbcb-23-91626 (R); **Published:** 30-Mar-2023, DOI: 10.4172/jbcb.1000179

Citation: Johnson S (2023) A Short Note on Biochemical functional and Protein structures. J Biochem Cell Biol, 6: 179.

Mendelian Inheritance in Man (OMIM) to identify active site residues, CASTp locates voids within a protein structure. Similarly, Ligsite can quickly analyze a large number of proteins by utilizing a collection of ligand–receptor complexes to locate pockets on a protein surface. The topological and geometric characteristics of the protein surface are also examined by PocketFinder and geometric potential. However, rather than searching the surface of a protein for pockets of various sizes, PocketFinder looks for ligand binding envelopes. In addition to global structural analysis, geometric potential employs local structural analysis to examine the pockets' residuals.

Conclusion

The process of annotating proteins with unknown or ambiguous functions remains difficult but essential for comprehending the enormous amount of data generated by genome sequencing and structural genomics initiatives. Methods for predicting functions that concentrate on the local spatial region of biochemical activity have the potential to enhance predictive ability. It is not always the case that proteins with high global sequence similarity also have high local sequence similarity at the active site. Conversely, proteins that share a low overall sequence similarity may share a high spatial similarity in the active site region. Too frequently, the function of a protein with a high global sequence similarity to a protein with no known function is transferred to the target protein without analysing the similarities in the local active site sequence.

Numerous research groups have developed methods for predicting protein function in an effort to provide useful information about enzymes whose functions are unknown. When making predictions, however, only one type of analysis—sequence- or structure-based analysis—is utilized, there is a greater chance of misannotation. The data gathered through genome projects can become more useful and complete as methods are improved and used simultaneously with one another. The difficulties associated with assigning functions to proteins can begin to be resolved with the assistance of these ground-breaking computational techniques mentioned above and others that will be developed in the future. It is evident that the field of protein function prediction will continue to expand, particularly as the quality and quantity of data continue to rise, despite the number of methods currently available to predict proteins' functions. Biochemical studies can be used to verify the predictions made while these computational methods are being improved. This kind of experimental verification is a big gap in the field right now. Biochemical studies may become more focused and require less time in the future as computational methods become more advanced and are validated experimentally. The fast, high-throughput functional annotation of these proteins will be made possible by future automation of computational methods, adding significant value to the growing repository of genomics data.

References

1. Tomlin JL, Sturgeon C, Pead MJ, Muir P (2000) Use of the bisphosphonate drug alendronate for palliative management of osteosarcoma in two dogs. Vet Rec 147: 129-32.
2. Psychas V, Loukopoulos P, Polizopoulou ZS, Sofianidis G (2009) Multilobular tumour of the caudal cranium causing severe cerebral and cerebellar compression in a dog. J Vet Sci 10: 81-3.
3. Loukopoulos P, Thornton JR, Robinson WF (2003) Clinical and pathologic relevance of p53 index in canine osseous tumors. Veterinary Pathology 40: 237-48.
4. Bech-Nielsen S, Haskins ME (1978) Frequency of osteosarcoma among first-degree relatives of St Bernard dogs. J Natl Cancer Inst 60: 349-53.
5. Wilkins RM, Cullen JW, Odom L, Jamroz BA, Cullen PM, et al. (2003) Superior survival in treatment of primary nonmetastatic pediatric osteosarcoma of the extremity. Ann Surg Oncol 10: 498-507.
6. Kundu ZS (2014) Classification, imaging, biopsy and staging of osteosarcoma. Indian J Orthop 48: 238-46.
7. Papalas JA, Balmer NN, Wallace C, Sangüeza OP (2009) Ossifying dermatofibroma with osteoclast-like giant cells: report of a case and literature review. Am J Dermatopathol 31: 379-83?
8. Gelberg KH, Fitzgerald EF, Hwang SA, Dubrow R (1995) Fluoride exposure and childhood osteosarcoma: a case-control study. Am J Public Health 85: 1678-83?
9. Luetke A, Meyers PA, Lewis A, Juergens H (2014) Osteosarcoma treatment where do we stand a state of the art review. Cancer Treat Rev 40: 523-532.
10. Dhaliwal J, Sumathi VP, Grimer RJ (2009) Radiation-induced periosteal osteosarcoma (PDF). Grand Rounds 10: 13-18.