TEST Engineering & Management

January - February 2020
ISSN: 0193 - 4120 Page No. 4358 - 4362

Protein-Protein Interaction for Neuraminidase Enzyme using Graph Theory Software (Cytoscape)

A. Akil Nivetha [1], Dr. V. Jeyanthi [2]

[1,2,3]Department of Mathematics, Sri Krishna Arts and Science College, Coimbatore.

Article Info
Volume 82
Page Number: 4358 - 4362
Publication Issue:
January-February 2020

Abstract

There are seven biological classifications in influenza virus. Out of 7 only four types of influenza viruses cause the influenza fever in human, birds and mammals. Hemagglutinin [HA] and neuraminidase [NA] are the enzyme proteins available around the influenza virus A. Viral neuraminidase divided the sialic acid residues from the glycoprotein. Hemagglutinin replicates the influenza virus A. Structural information of 376 Neuraminidase enzyme proteins and 162 Hemagglutinin proteins are available in PDB database. Proteins are binds with one protein to many proteins. Interactions between proteins are very much useful for drug design research. Few of the software are available for protein-protein interaction research. In this paper we are using freely available IntAct database for bimolecular interaction like protein-protein interaction analysis. We are also using network analysis software Cytoscape for protein-protein interaction study. Using IntAct database we were obtained 575 interaction for neuraminidase protein database. We have submitted the viral neuraminidase data from IntAct to the network analysis software Cytoscape. From the 575 protein-protein interaction for neuraminidase enzyme, we have found most of the proteins have one to two interactions. But two of the neuraminidase enzymes proteins P03468 and P03452 have 20 and 35 protein-protein interactions respectively. Above two neuraminidase proteins have play an important role for the drug design in influenza virus A. Graph theory analysis also carried for 575 protein interaction.

Article History
Article Received: *18 May 2019*
Revised: *14 July 2019*
Accepted: *22 December 2019*
Publication: *21 January 2020*

Keywords: *Protein, Cytoscape, Neuraminidase and Hub protein*

I. INTRODUCTION

Influenza virus induces the contagious disease in human. Variety of two surface glycoprotein antigens three-dimensional structures of influenza virus are influenced the drug development researchers. Adsorption spectra of influenza haemagglutinins and enzyme activity of influenza viruses were analyzed on 1942 [1]. Hemagglutinin enzyme proteins [2] and the neuraminidase enzyme proteins [3] are coated around the glycoprotein surface. Comparative study of hemagglutinin and neuraminidase are analyzed by Smith and Palese [4]. Crystal structure of Neuraminidase enzyme were determined by X-ray diffraction methods [5] and this three-dimensional structures of viral neuraminidases enzyme protein are stored in PDB database. Influenza viruses are the RNA virus casus the infectious diseaselikeH1N1,

January - February 2020
ISSN: 0193 - 4120 Page No. 4358 - 4362

H5N1(bird flu).It affects birds and mammals. Influenza spreads around the world in some season, because of influenza virus 2 lakhs to 5 lakhs deaths every year around the world. Single-stranded Influenza A viruses (metameric RNA viruses) causes influenza flu in human. Protein function concept is arranged in order of rank. Biomolecular interactions helps to provide opportunity to understand the molecular systems biology. Study the molecular interactions are improve the quality of Protein function, cell machinery and new drug discovery. Up to 22000 protein-protein interaction map were published considerably last few years [6]. Based on the importance of protein interactions, large number of protein types and types of interacting proteins are classify into physical complexes and signaling pathways. The theory of protein -protein complex network analysis plays an important role in drug design, structure and function of protein, network analysis, graph theory and medical biology research. Applications of network analysis in Mathematical biology and medicine determining a protein function, diseases treatment and early diagnosis of disorders. In 2004 Pellegrini Matteo et.all. [7] discussed the Protein-protein interaction networks. The theory of protein -protein complex network analysis plays an important role in drug design, Computer science, graph theory and medical biology research. Applications of network analysis in Mathematical biology and medicine determining a protein function, diseases treatment and early diagnosis of disorders. Biological processes inside the cell of the protein are analyzed using the Protein-protein interaction. Network analysis software like Cytoscape deals the metabolic pathway and molecular modeling study. Correlation between various proteins, characteristics of proteins and their interaction are analyzed using this network analyzed softare. Sequence and structural information of proteins need to construct the networking map of bimolecular interaction. KEGG, EcoCyc, BioCyc [8] and metaTIGER [9] are the system biology software for biomolicular interaction. Metabolic network analysis lot of methods were developed.

IntAct [10] is a data model to collect the protein–protein interactions, DNA, RNA and small molecule interaction data from various research papers. IntAct database and uniprot are combined to collect the protein protein and gene gene interaction database from the various literature survey results using standard research journals. Protein sequence and its charactrists are information are collected from various lab and store in Uniprot KB database. Biomolicular interaction database also available this Uniprot KB server. [11]. Gene sequence and nucleic acids structure information are maintained by GenBank [12]. To import the IntAct data into graph theory analysis software, such as Cytoscape [13] to analysis the network interaction of molecules. Researchers interpret their large amount of protein protein interaction data to visualize the Network visualization tool like Cytoscape. In this paper we have understanding of protein networks to be need on a detailed knowledge of influenza disease from protein- protein interactions. It's specifying to identify the best target proteins for neuraminidase enzyme.

II. METHODS

In this paper we have collected 557 viral neuraminidase protein sequences from Uniprot. With the help of IntAct data we have to find the protein-protein interaction data for Neuraminidase sequences. Importing the protein – protein interactions through excel file into the Cytoscape. We are using the Cytoscape for further analysis. Cytoscape software involve to analysis the Direct and undirected network of protein interaction . Node and edge are important part in the graph theory. Parameters for node and edge are computed by Network analysis software. Filter this parameter and apply to the different visualization of protein-protein interaction. On this Cytoscape software we have change the parameters color for protein structure visualization. Bright color using for bimolecular

January - February 2020
ISSN: 0193 - 4120 Page No. 4358 - 4362

structural napping. Mostly green red and yellow color are the default color for bright, dark and middle color. Direct and undirected edges are analyzed by topological network anlyzer. Network interpretation was carried out using this analyzer. This interpretation is undirected in Cytoscape software output. Cytoscape software they are using many network parameters. Network connectivity properties and number of nodes connected each other are identified using this software. In the protein protein interaction software shortest path between nodes, diameter of the network and average shortest path length are analyzed. Based on the no of neighborhood node calculated the normalized parameters which implemented to network density values. Isolated nodes are identified and how it involve the density calculation also finalyzed. Hub nodes are evaluated from the network heterogeneity analysis. Clustering coefficient of nodes in the protein-protein interaction between 0 to 1. Average clustering coefficient between nodes involves to analysis the molecular interaction between edge and nodes, organization of biochemical path way networks. Distribution of protein protein sharing neighborhood also calculated using this software. Cytoscape software computes the more than one interaction between neighbor proteins using topological coefficient method. Algorithm for betweenness centrality[14] using random distribution are using this Cytoscape software. Other centrality like closeness centrality and stress centrality distribution are carried out this software. Few parameters are complex which can be fitted by fitting a line algorithm. First we install the Cytoscape software. Open the csv file from the output of IntAct database of neuraminidase enzyme proteins. Cytoscape software mapping the network analyzed result. Visualization of mapping parameter display on the screen. Cytoscape VizMapper visualize and fine tune the protein-protein interaction. Displayed color full network map we will be rotatable and designed using the software. Graph theory and topological parameter of protein-protein interaction data are plotted and analyzed by this software.

III. RESULT AND DISCUSSION

Cytoscape is a network analysis software for intermolecular interactions of edges, between nodes. We have exported the IntAct Protein interaction Database using Cytoscape's Core software and visualized the result. Fig. 1 illustrates key features through screenshots of Cytoscape uploading IntAct database.

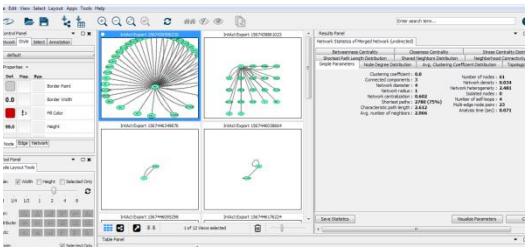

Fig. 1 Screen short of uploading IntAct file

Fig. 2 Neuraminidase protein have 20 interactions

575 viral neuraminidase protein sequence were collected from Uniprot KB. Out of 575 few interaction are available for neuraminidase enzyme, We have found few of the neuraminidase enzyme protein have more than one interaction. Only two neuraminidase enzymes proteins P03468 and P03452 have 20 and 35 interactions respectively, we have to display the network graph of P03468 and P03452 neuraminidase enzyme is shown in Fig.2 and Fig.3 respectively.

January - February 2020
ISSN: 0193 - 4120 Page No. 4358 - 4362

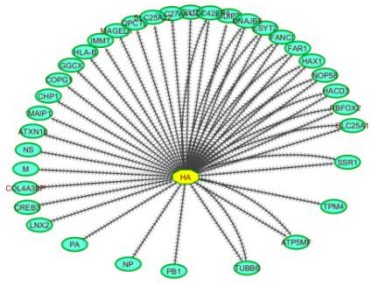

Fig. 3 Hemagglutinin protein has 35 interactions

P03468 is a Neuraminidase protein from influenza virus. This influenza virus A divided the glycoprotein using this ligand sialic acid. It will lead to find the new drug using molecular modeling methods. There is no 3D structure of this protein. If we find the structure and conformational angles of this protein will be suitable to find the new drug for influenza virus. This protein interact 20 different interactions with various proteins. This protein act like a hub protein.

P03452 is Hemagglutinin protein from influenza A virus. It binds to sialic acid-containing receptors on the cell surface, bringing about the attachment of the virus particle to the cell. This attachment induces virion internalization either through clathrin-dependent endocytosis or through clathrin- and caveolin-independent pathway. Plays a major role in the determination of host range restriction and virulence. This hub protein has 35 interaction with other proteins. After the finding of crystal structure of Hemagglutinin enzyme protein we know it has complex with receptor analogs. It will be spread through birds to the human. Some active site amino acids are easy to bind the human receptors.

Three dimensional structure of influenza hemagglutinin (1RVZ) shown in Fig. 5. It contains 327 amino acids. Ligand Beta-d-galactose, N-acetyl-d-glucosamine and O-sialic acid are linked with this enzyme. Most of the structure are Beta sheet shapes. Gamblin et. all. discovered this protein crystal structure on 2004.

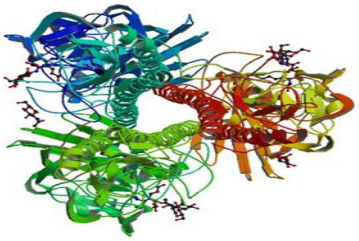

Fig. 5 3D structure of 1RVZ structure

The network visualization of merged neuraminidase protein-protein interactions are shown in Fig. 6. In this fig we know that two yellow edges indicate more number of protein interactions for viral

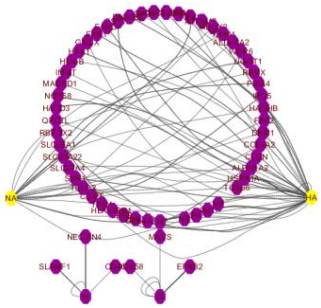

Fig. 6 Merged Network for Neuraminidase Protein – Protein interactions

neuraminidase protein. In the present study, network biology approach was done to identify the protein that are involved in Influenza virus. Topological parameters were evaluated.

The protein-protein interaction were clustered and validated based on ranking algorithms. These proteins can be used as significant therapeutic targets against this disease.

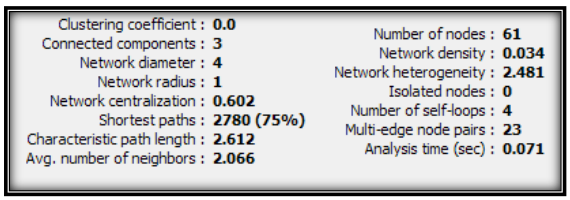

Fig. 7 Network analysis result from Cytoscape

Output of network analysis is shown in Fig. 7 Number of nodes we are using for our analysis is 61. We found 23 multi edge node pairs. From this Fig 7

January - February 2020
ISSN: 0193 - 4120 Page No. 4358 - 4362

5 % nodes are in shortest path. The value of average number of neighbours is 2.066.

IV. CONCLUSION

Protein-protein interaction of neuraminidase enzyme was analyzed by the software Cytoscape indicate to find the hub protein from the overall dataset. The above network analysis and graph theory analysis will be lead to find the new drug for various influenza viruses. From the big size protein database we have selected the particular protein for molecular modeling and drug design of influenza virus. In future first we will find the structure of the protein and will be doing structural and conformational analysis of the two particular proteins. This work defiantly leads the better research for influenza virus research.

REFERENCE

[1] Hirst GK. 1942. Adsorption of influenza haemagglutinins and virus by red blood cells. J Exp Med., Pp.195-209.

[2] Wharton SA, Weis W, Skehel JJ, Wiley DC. 1989. Structure, function, and antigenicity of the hemagglutinin of influenza virus. The influenza viruses, Pp. 153-173.

[3] Colman PM. 1989. Influenza virus neuraminidase: Enzyme and antigen. The influenza viruses., Pp. 175-218.

[4] Smith F1, Palese P. 1989. Variation in influenza virus genes. Epidemiological, pathogenic, and evolutionary consequences. The influenza viruses., Pp. 319-359.

[5] Varghese JN, Laver WG & Colman PM. 1983. Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 A resolution. Nature, Pp. 35-40.

[6] Giot,L, Bader,JS., Brouwer,C, Chaudhuri,A, Kuang,B, Li,Y, Hao,YL, Ooi CE, Godwin,B, Vitols E. 2003. A protein interaction map of Drosophila melanogaster. Science, Pp. 1727–1736.

[7] Pellegrini Matteo, Haynor David, Johnson JM. 2004. Protein interaction networks. Expert Review in Proteomics, Pp. 89-99.

[8] Karp PD, Ouzounis CA, Moore-Kochlacs C, Goldovsky L, Kaipa P, Ahren D, Tsoka S, Darzentas N, Kunin V, Lopez-Bigas N. 2005. Expansion of the BioCyc collection of pathway/genome databases to 160 genomes. Nucleic Acids Res., Pp. 6083-6089.

[9] Whitaker JW, Letunic I, McConkey GA, Westhead DR. 2009. metaTIGER: a metabolic evolution resource. Nucleic Acids Res., Pp. 531-538.

[10] Kerrien S, Alam-Faruque Y, Aranda B, Bancarz I, Bridge1 A, Derow C, Dimmer E, Feuermann1 M, Friedrichsen A, Huntley R, Kohler C, Khadake J, Leroy C, Liban A, Lieftink C, Montecchi-Palazzi L, Orchard S, Risse J, Robbe K, Roechert1 B, Thorneycroft D, Zhang Y, Apweiler R and Hermjakob H. 2007. IntAct—open source resource for molecular interaction data , Nucleic Acids Research, Pp. D561–D565.

[11] Apweiler,R, Bairoch,A, Wu,CH, Barker,WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M. 2004. UniProt: the Universal Protein knowledgebase. Nucleic Acids Res., Pp. D115–D119.

[12] Cochrane G, Aldebert P, Althorpe N, Andersson M, Baker W, Baldwin A, Bates K, Bhattacharyya S, Browne P, van den Broek A. 2006. EMBL Nucleotide Sequence Database: developments in 2005. Nucleic Acids Res., Pp D10–D15.

[13] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res., Pp. 2498–2504.

[14] Brandes U. 2001. A faster algorithm for betweenness centrality. J Math Sociol., Pp. 163-177.