IIOAB JOURNAL

COMMUNICATION

ANALYSIS OF RIBOVIRAL INTERACTIONS WITH PHOSPHATE AND SULPHATE IONS

Stecy Antony Selvakumar*, Arul Mugilan, Jeyasankari

Department of Physics, Kamarajar Government Arts College, Tamil Nadu, INDIA

ISSN 0976-3104

ABSTRACT

Computer modelling of ligand-protein interactions based on the structure is now an essential component of modern drug discovery. The design of small molecule drugs for the treatment of disease is made possible by the atomic resolution study of protein-ligand interaction. Here, we analyze several Riboviral proteins that briefly interacted with the ligands PO4 and SO4. The majority of the PDB data in the preceding study show that the amino acids Tyrosine and Tryptophan easily bind with PO4 and SO4 at about 2.85 and 2.56 Armstrong. Python is used to calculate bond lengths and RasMol is used to visualize the structure of various proteins. According to RasMol, Riboviral proteins favours Alpha helix and Beta Sheet structures. The analysis presented above is helpful for computational tasks as well as for experimental biologist working in molecular modelling and drug design.

KEY WORDS
Protein ligand interactions; Python; Rasmol

Received: 09 July 2023
Accepted: 30 July 2023
Published: 15 Oct 2023

INTRODUCTION

Proteins plays wide range of biological functions [1,2]. In addition to proteins and peptides, it interacts with nucleic acids, membranes, substrates, and small molecule ligands like oxygen and solvents[3,4]. One of the key characteristics of proteins is their ability to bind to other molecules with high specificity and affinity, which is determined by the shape and physicochemical properties of the residues [5-7]. In order for proteins to perform their functions, they must interact with other molecules, which is one of the most important types of protein function. Solving the complex structures using X-ray or nuclear magnetic resonance techniques is the ideal way to study the interaction. However, these experiments are usually time-consuming and expensive to carry out [8,9].

The structures of some large proteins and membrane proteins are extremely difficult or impossible to solve using traditional techniques [10-12]. Rather than solving the structure, an alternative method is to locate the binding sites by theoretically is the cheapest [13-17]. By solving the protein-ligand complex structure or determining the interactions by experiments is often time consuming and expensive, many computational efforts have been made to facilitate the study of the interactions [18-20]. Protein binding to small molecules (called ligands) is of particular interest and has a wide range of applications in structure-based drug design [21]. Many biological processes depend on interactions between proteins and ligands [22].

The functions of a protein can be regulated through ligand binding. Under physiological conditions, a protein can undergo conformational changes, which are responsible for the conformational transitions between low- and high-affinity states for the ligands [23-26]. This paper examines protein interactions with phosphate and sulphate. A phosphate group interacts with more than half of known proteins [27]. Various non-homologous protein families have evolved phosphate binding [28]. Molecular ligands with the phosphate group are involved in a large number of significant chemical reactions and molecular interactions that take place in the cell [27]. The binding of phosphoryl groups plays an essential role in a number of biological processes, including metabolism and biosynthesis, gene regulation, signal transduction, muscle contraction, and antibiotic resistance. [29]. Phosphate binding increases the stability of enzymes like aspartate aminotransferase and causes apoferritin to deposit iron [30].

The phosphate group is required for nucleotide recognition, and nucleotide-containing ligands were the first cofactors to bind to proteins [31]. It is therefore of utmost significance to characterize a protein for its propensity to interact with a phosphate or a phosphate-containing ligand [32]. For protein engineering studies, drug design, and structure prediction, a clear understanding of the factors governing the phosphate binding would be beneficial. Due to their negative charges, phosphate groups are predicted to bind to the protein's positively charged region [33]. The sulphate polysaccharide binds to a wide range of biologically active polypeptides, including enzymes, growth factors, cytokines, and viral proteins [34].

MATERIALS AND METHODS

Database

Our current study is based on the crystallographic data of Riboviral proteins from the Protein Data Bank PDB of Brookhaven National Laboratory [35-37]. The chosen proteins were nonhomologous, and their structures, as well as their interaction with ligands PO4 and SO4, were determined at great resolution. In

***Corresponding Author**
E-mail: astecy2705@gmail.com

THE IIOAB JOURNAL

this work, viral proteins were used with similarities between 30% (1508) and 90% (1050), their differences (458) being taken into account. Among these data sets, those with repeat protein sequences were removed, and the remaining IDs with ligand-interacting proteins were retained. Following protein selection, the next step is to determine bond length by building a Python program with a range of less than 13A˚[38,39]. It will show the outcomes of amino acids interacting with the ligands PO4 and SO4. The amino acids are separated in Excel, then the results are analyzed using a statistical approach [40]. Chi square test was performed using the formula [41],

$$X^2 = \frac{\{Observed - Expected\}^2}{Expected}$$

Graphical representations of chi square results show which amino acids are in the high preferential region and non-preferential region. Using RasMol, the shortest distances between amino acids that interact with the ligands PO4 and SO4 are visualized. Additionally, RasMol is used to analyse the structures of interacted amino acids [42]. The overall method is presented in the Figure. 1.

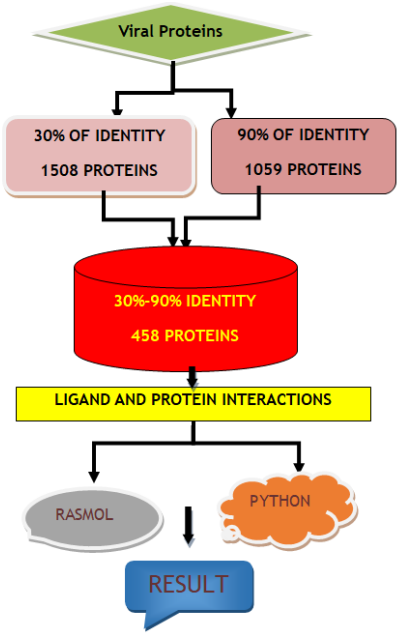

Fig. 1: Flowchart of the overall methodology.

RESULTS

The output has been extracted with the help of the program and is presented in graphical form.

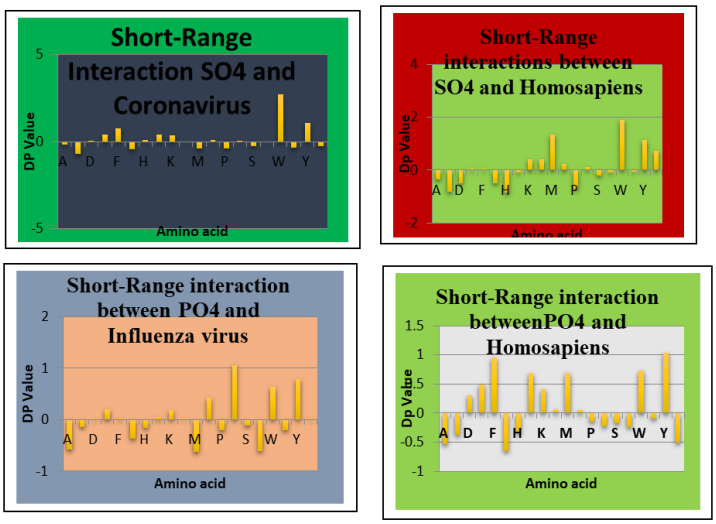

Fig. 2: Graphical representation of various protein IDs binding with SO4 and PO4 showing high preferences of amino acids binding with SO4 and PO4.

As shown in Figure-2, the graphical analysis is categorized into four categories: -1, which indicates that amino acids are not in range, and above 0.5, which indicates that amino acids are interacting with the ligand. Finally, -0.5 denotes the non-preferred region, while 1.5 and higher denote the highly preferred region of amino acids with ligand. Tyrosine is the most common residue found to interact with phosphate, followed by Phenylalanine, Tryptophan, and Methionine. Similarly, Tryptophan is the most abundant residue in sulphate, followed by Phenylalanine, Tyrosine, and Methionine. No conventional phosphate and sulphate binding site exists.

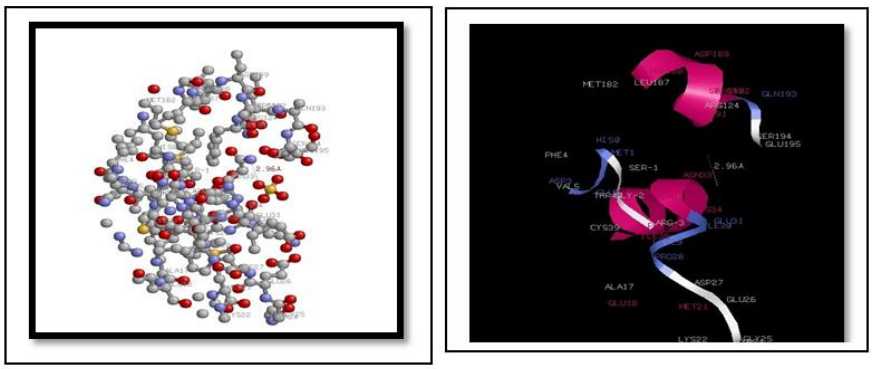

Fig. 3: Amino acid interactions and structure of 5WA7 using RasMol.

The Figure-3 represents the binding portion of amino acid residue with PO4 and structure. Similarly, the structure and binding sites of all these 1500 proteins were performed.

CONCLUSION

The traditional drug discovery process requires significant capital investments and takes a long time to develop a drug [43]. Because of the limitations and the more extended period of drug development, new techniques in drug discovery are required [44]. Python was chosen because of its widespread use and extensive utility in biosciences. To prepare data for further processing, it is collected, filtered, and altered. Python Programs are then used to analyse datasets to produce results. In this article Protein interactions with ligands, PO4 and SO4 are carried out. The main conclusion is that phosphate binding site residues are highly conserved relative to the alignment as a whole.The sulphate group is bound to the amino terminus of the alpha helix, whereas the phosphate group is bound to the Beta sheet. We anticipate that the precise prediction of ligand-binding sites and the improved ligand-protein complex structures will aid in other related studies, such as drug discovery.

CONFLICT OF INTEREST
Authors declare no conflict of interest.

ACKNOWLEDGEMENTS
We thank Prof. Mugilan for his support throughout the work.

FINANCIAL DISCLOSURE
None.

REFERENCES

[1] Steinbrecher T, Labahn A. [2010] Towards accurate free energy calculations in ligand protein-binding studies. Current Medicinal Chemistry. 17:767–785.
[2] Acuner Ozbabacan SE, Gursoy A, Keskin O, Nussinov R. [2010] Conformational ensembles, signal transduction and residue hot spots: Application to drug discovery. Current Opinion in Drug Discovery& Development. 13: 527–537.
[3] Perozzo R, Folkers G, Scapozza L. [2004] Thermodynamics of protein-ligand interactions: History, presence, and future aspects. Journal of Receptors and Signal Transduction Research. 24:1–52.
[4] Overington JP, Al-Lazikani B, Hopkins AL. [2006] How many drug targets are there?, Nature Reviews Drug Discovery. 5(12):993–996.
[5] Babu E, Bhuvaneswari J, Muthu Mareeswaran P, Thanasekaran P, Lee HM, Rajagopal S. [2019] Transition metal complexes-based aptamers as optical diagnostic tools for disease proteins and biomolecules. Coordination Chemistry Reviews.380:519–549.
[6] Lee H, Rho J, Messersmith PB. [2009] Facile Conjugation of Biomolecules onto Surfaces via Mussel Adhesive Protein Inspired Coatings. Advanced Materials. 21(4):431–434.

[7] Bock Joel R, Gough, David A. [2001] Predicting protein–protein interactions from primary structure. Bioinformatics.17:455–460.
[8] Opella SJ. [2013] Structure determination of membrane proteins by nuclear magnetic resonance spectroscopy. Annual Review of Analytical Chemistry.6:305–328.
[9] Cavalli A, Salvatella X, Dobson CM, Vendruscolo M. [2007] Protein structure determination from NMR chemical shifts. Proceedings of the National Academy of Sciences. 104(23):9615–9620.
[10] Arkin MR, Tang Y, Wells JA.[2014] Small-molecule inhibitors of protein-protein interactions: progressing toward the reality. Chemical & Biology. 21:1102-1114.
[11] Higueruelo AP, Jubb H, Blundell TL. [2013] Protein–protein interactions as druggable targets: recent technological advances. Current Opinion in Pharmacology. 13:791–796.
[12] Blundell TL, Patel S. [2004] High-throughput X ray crystallography for drug discovery. Current Opinion in Pharmacology. 4: 490–496.
[13] Brylinski M, Skolnick J. [2008] A threading-based method (FINDSITE) for ligand binding site prediction and functional annotation. Proceedings of the National Academy Sciences U.S.A. 105:129134.
[14] Capra JA, Laskowski RA, Thornton JM, Singh M, Funkhouser TA. [2009] Predicting protein ligand binding sites by combining evolutionary sequence conservation and 3D structure. PLoS Computational Biology. 5:e1000585.
[15] Roy A, Yang J, Zhang Y. [2012] COFACTOR: an accurate comparative algorithm for structure-based protein function annotation. Nucleic Acids Research.40:W471–W477.
[16] Yang J, Roy A, Zhang Y. [2013] Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. Bioinformatics.29:2588–2595.
[17] Heo L, Shin WH, Lee MS, Seok C. [2014] Galaxy Site: binding-site prediction by using molecular docking. Nucleic Acids Research.42:W210–W214.
[18] Roche DB, Buenavista MT, Mc Guffin LJ.[2013] The FunFOLD2 server for the prediction of protein-ligand interactions. Nucleic Acids Research.41:W303-W307.
[19] Pellecchia M, Bertini I, Cowburn D, Dalvit C, Giralt E, Jahnke W, Siegal G. [2008] Perspectives on NMR in drug discovery: a technique comes of age. Nature Reviews Drug Discovery. 7(9):738–745.
[20] Bertini I, McGreevy KS, Parigi G. [2012] NMR of Biomolecules: Towards Mechanistic Systems Biology. Wiley-VCH Verlag GmbH & Co. KGaA (1st ed.,pp.55-94).
[21] Malathi K, Ramaiah S. [2018] Bioinformatics approaches for new drug discovery: a review. Biotechnology and Genetic Engineering Reviews.1–18.
[22] Swain SS, Hussain T. [2021] Role of Bioinformatics in Early Drug Discovery: An overview and Perspective. Computation in Bioinformatics: Multidisciplinary Applications. doi.org/10.1002/9781119654803.ch3
[23] Dunn MF. [2010] Protein-Ligand Interactions: General Description. Encyclopedia of Life Sciences.
[24] Fukunishi Yoshifumi, Nakamura, Haruki. [2011] Prediction of ligand-binding sites of proteins by molecular docking calculation for a random ligand library. Protein Science.20:95–106.
[25] Zhao L, Zhu Y, Wang J, Wen N, Wang C, Cheng L. [2022] A brief review of protein–ligand interaction prediction. Computational-and-structural-biotechnology journal. 20:2831-2838.
[26] Zhao J, Cao Y, Zhang L. [2020] Exploring the computational methods for protein-ligand binding site prediction. Computational and Structural Biotechnology Journal. 18:417-426.
[27] Hirsch KH, Fischer FR, Diederich F. [2006] Phosphate recognition in structural biology. Angewandte Chemie International edition.46: 338–352.
[28] Saraste M, Sibbald PR, Wittinghofer A. [1990] The P-loop–a common motif in ATP- and GTP-binding proteins. Trends in Biochemical Sciences. 15:430–434.
[29] Martinez-Liarte JH, Iriarte A, Martinez-Carrion M. [1992] Inorganic phosphate binding and electrostatic effects in the active center of aspartate aminotransferase apoenzyme. Biochemistry. 31(10):2712–2719.
[30] Cheng YG, Chasteen ND.[1991] Role of phosphate in initial iron deposition in apoferritin. Biochemistry. 30(11): 2947–2953.
[31] Ji HF, Kong DX, Shen L, Chen LL, Ma BG, Zhang HY. [2007] Distribution patterns of small-molecule ligands in the protein universe and implications for origin of life and drug discovery. Genome Biology. 8:R176.
[32] Ghersi D, Sanchez R. [2009] EASYMIFS and SITEHOUND: a toolkit for the identification of ligand-binding sites in protein structures. Interdisciplinary Informatics Faculty Publications. 17. https://digitalcommons.unomaha.edu/interdiscipinformaticsfacpub/17.
[33] Hol WG J. [1985] The role of the α-helix dipole in protein function and structure. Progress in Biophysics and Molecular Biology.45: 149-195.
[34] Hol WGJ, Van Duijnen PT, Berendsen HJC. [1978] The α-helix dipole and the properties of proteins. Nature (London).273:443-446.
[35] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Bourne PE. [2000] The Protein Data Bank. Nucleic Acids Research.28:235–242.
[36] F C Bernstein, T F Koetzle, G J Williams, E F Meyer Jr, M D Brice, J R Rodgers, O Kennard, T Shimanouchi, M Tasumi. [1977] The Protein Data Bank: A Computer-based Archival File for Macromolecular Structures. Journal of Molecular Biology. 112(3):535–542.
[37] Burley SK, Berman HM, Kleywegt GJ, Markle JL, Nakamura H, Velankar S. [2017] Protein Data Bank (PDB): The Single Global Macromolecular Structure Archive. Protein Crystallography. 627–641.
[38] Ekmekci B, McAnany CE, Mura C. [2016] An Introduction to Programming for Bioscientists: A Python-Based Primer. PLoS Computational Biology. 12(6): 2e1004867.
[39] Bassi S. [2007] A Primer on Python for Life Science Researchers. PLoS Computational Biology. 3(11): e199.
[40] Patrick Aloy, Juan Cedano, Baldomero Oliva.[1997] 'TransMem': a neural network implemented in Excel spreadsheets for predicting transmembrane domains of proteins. Oxford University Press. Vol. 13 no. 3 Pages 231-234.
[41] Zhu P, Bowden P, Tucholska M, Marshall JG. [2011] Chi-square comparison of tryptic peptide-to-protein distributions of tandem mass spectrometry from blood with those of random expectation. Analytical Biochemistry. 409(2):189–194.
[42] Amber D Fraley, Katherine E Odegaard, Victor H Vilchiz, ChulHee Kang, Cathy Lee. [2022] Exploring Protein Structure with RasMol, a Free Online Tool. The American Biology Teacher. 84 (6): 369–373.
[43] Li J, Fu A, Zhang L. [2019] An Overview of Scoring Functions Used for Protein–Ligand Interactions in Molecular Docking. Interdisciplinary Sciences: Computational Life Sciences. 11(2): 320–328.
[44] Lin X, Li X, Lin X. [2020] A Review on Applications of Computational Methods in Drug Screening and Design. Molecules. 25(6): 1375.