ORIGINAL ARTICLE

ANALYSIS OF 3D STRUCTURE OF THE PROTEIN OF HAEMOPHILUS INFLUENZAE BY HOMOLOGY MODELLING HELPS IN PREDICTING BINDING SITES FOR SUBSTRATE, LEADS TO DESIGN ANTIBIOTIC

Rehana Rasool[1], Maria Shafiq[2], Samina Swati[3], Anila Farid[4], Sofia Shoukat[5], Madeeha Jadoon[6]

[1]Department of Community Medicine Abbottabad International Medical College Abbottabad, [2]Department of Physiology Ayub Medical College Abbottabad, [3]Department of Pathology Abbottabad International Medical College Abbottabad, [4]Department of Biochemistry Abbottabad International Medical College Abbottabad, [5]Department of Biochemistry Ayub Medical College Abbottabad, [6]Department of Biochemistry Women Medical and Dental College Abbottabad-Pakistan

Background: Haemophilus influenza persists as a well-known root of ill health in children throughout the entire world. Before the introduction of the vaccine, Haemophilus influenza remained a well-known and eminent source of septic arthritis, pneumonia, and epiglottitis. Haemophilus influenza, Neisseria meningitides, and staphylococcus pneumonia spreads through respiratory droplets and cause diseases such as meningitis, pneumonia, and other secondary infections related to respiratory diseases. Objective was to analyze the 3D structure of the protein of Haemophilus influenzae by homology modelling to design antibiotics. **Methods:** For the effective study of protein, computational tools were used to investigate protein structure and function, Comprehensive microbial resource (CMR) for comparative modelling, Interproscan, BLAST for sequence similarity searching, MODELLER 9.10 for homology modeling, Procheck and Protein Structure Analysis (ProSA) software for assessing model quality and structural validation. **Results:** The model showed that it consists of three alpha helices (red) and one beta-sheet. Ramachandran Plot statistics show that 97.4% of the debris is in the favoured region, 0% in the additional allowed region, 2.65% in the generally allowed part, and 0% in the disallowed part. Stability and energy were checked through ProSa. Z score was highly negative which showed that the model is highly stable. The greater the negative value, the more will be the stability of the model. **Conclusion:** Cell division protein H11025 was selected. The structure was modelled which has provided all the required information to design antibiotics to control the harmful effects regarding that protein.

Keywords: Homology modelling; Modeller; Haemophilus influenza; Prosa

Citation: Rasool R, Shafiq M, Samina, Farid A, Shoukat S, Jadoon M. Analysis of 3D structure of the protein of Haemophilus influenzae by homology modelling helps in predicting binding sites for substrate, leads to design antibiotic. J Ayub Med Coll Abbottabad 2024;36(2):323–5.

DOI: 10.55519/JAMC-02-12594

INTRODUCTION

Haemophilus influenza persists as a well-known root of ill-health in children throughout the entire world. It causes septic arthritis, pneumonia, and epiglottitis.[1] Nontypeable Haemophilus influenza (NTHi) had been 25-fold comparatively more active than staphylococcal thermonuclease in fluorescence resonance energy transfer (FRET) activity.[2] It spreads through respiratory droplets.[3] Antibiotics against this disease is only possible if we know the exact structure of disease-causing protein.

Computational access to create a three-dimensional structure of the protein is homology modelling. Computational methodology has contributed a lot in covering such gaps and limitations.[4]

3D structure of the protein is built by using a primary sequence of proteins and from knowledge and information gained from similarities related to structure with other proteins.[5]. Proteins have receptors on the surface at which drugs interact and inhibit harmful effects.[6] In one study, homology modelling helped in drug designing for GPCR. Based on the crystal structure of 16 different GPCRs, homology models of D2 Dopamine and Serotonin 5- HT2A receptors were constructed.[7] In another study, in some patients β-corona virus was found during the spread of respiratory diseases in Wuhan. It was found that PL-Pro, 3CL-Pro, and RdRp have been important targets for antiviral drug designing against 2019-nCOV.[8] Antibodies are key proteins produced by the immune system to target pathogen protein (antigen). It binds to a specific binding surface called an epitope.[9] Therefore, by analyzing the 3D structure of a protein of Haemophilus influenza by homology modelling, antibiotics can be designed to cure disease.

MATERIAL AND METHODS

It is a Quasi-experimental study carried out in controlled circumstances in the computer lab. Sample size is not required as no projection toward a population is required. Comprehensive microbial resource (CMR) is a powerful tool for downloading the complete sequence of Haemophilus influenza. Classification and functional annotation were done through Interproscan. Approximately 2000 proteins were found in H influenzae.456 hypothetical proteins were confirmed with unknown functions. All hypothetical proteins were separated and saved in Word document format. The three-dimensional structure of the protein was modelled through homology modelling. In model building, a template plays a very important role. The template was searched against Protein Data Bank (PDB) using the Basic Local Alignment Search Tool (BLAST).[10]

Modeller 9.10 was used to build models of the target protein based on the alignment with template structures obtained from the template-target alignment file. A total of 10 models were generated during the Modeller execution. To visualize the newly built model, we utilized DS Viewer software. Subsequently, the stability and reliability of the models were assessed using ProSA and Procheck, which provide insights into this overall quality and stereochemical correctness.[12,13]

RESULTS

Statics of plot:

97.4% residues in most favoured parts
0% in additional allowed parts
2.6% in generously allowed parts
0% in disallowed parts.

Ramachandran plot is a graphical representation of the dihedral angles of amino acid residues in protein structure. 97.4% of residues are present in the most favoured region indicating that the protein structure is predominantly adopting energetically favourable confirmation. 2.6% of residues fall within the generously allowed region, suggesting some deviations from the most favoured confirmations but still within acceptable ranges. There are no residues in the additional allowed parts, indicating that the structure is largely consistent with well-folded protein confirmations.

Description	Max score	Total score	Query cover	E value	Max ident	Accession
Chain A, The Crystal Structure Of The Bacterial Cell Division Protein Zapa >pdb|1W2E|B Chain B, The Crystal Structure Of The Bacterial Cell Division Protein Zapa	35.4	35.4	85%	0.001	27%	1W2E_A
Chain A, Unknown Conserved Bacterial Protein From Pseudomonas Aeruginosa Pao1 >pdb|1T3U|B Chain B, Unknown Conserved Bacterial Protein From Pseudc	35.4	35.4	85%	0.001	27%	1T3U_A
Chain A, Structure Of Uracil Phosphoribosyl Transferase >pdb|2EHJ|B Chain B, Structure Of Uracil Phosphoribosyl Transferase >pdb|2EHJ|C Chain C, Structure O	28.1	28.1	51%	0.92	31%	2EHJ_A
Chain A, Crystal Structure Of Clostridium Botulinum Neurotoxin Serotype F Catalytic Domain With An Inhibitor (Inh1) >pdb|3FIE|B Chain B, Crystal Structure Of Clost	27.3	27.3	25%	1.8	48%	3FIE_A
Chain A, Crystal Structure Of Clostridium Botulinum Neurotoxin Serotype F Light Chain	27.3	27.3	25%	1.8	48%	2A8A_A
Chain A, Crystal Structure Of Catalytic Domain Of Clostridium Botulinum Neurotoxin Serotype F >pdb|2A97|B Chain B, Crystal Structure Of Catalytic Domain Of Clos	27.3	27.3	25%	1.8	48%	2A97_A
Chain A, Crystal Structure Of A Putative Flavin Oxidoreductase With Flavin	26.2	26.2	56%	5.0	29%	1VHN_A

Figure-1: Blast result of *H. influenzae* hypothetical protein HI1025.

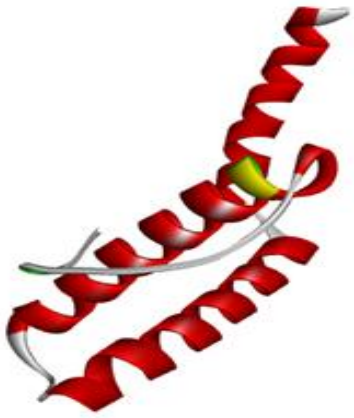

Figure-2: Modeller 9.10 was run. Homology model of *H. influenza* hypothetical protein HI1025. DS visualizer is used to see this model.

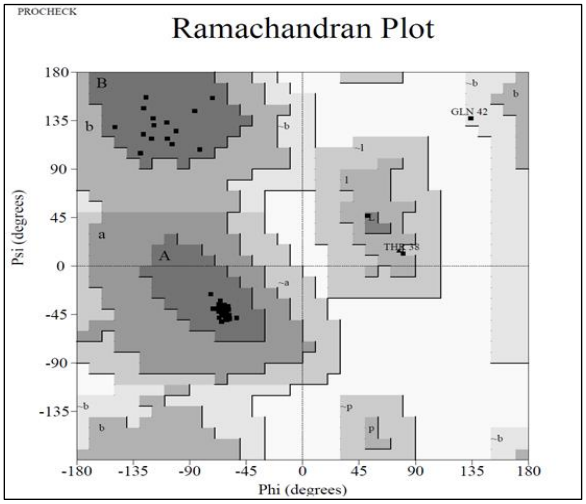

Figure-3: Ramachandran plot by procheck

DISCUSSION

Haemophilus influenzae causes pneumonia throughout the world. In this study, it was emphasized on the structural analysis of Haemophilus influenzae which plays a critical role in disease spreading. The methodology employed a combination of computational tools including CMR, BLAST, Modeller, and Procheck to predict and validate the 3D structure of the protein from Haemophilus influenzae. CMR and BLAST were utilized to provide a template for homology modeling. A template having accession number 1W2EA was selected as the template and the target has 85% identity. The structure of the cell division protein of Haemophilus influenzae H11025 was examined, delving into its intricate details. Modeller 9.10 was then employed to generate a 3D model based on the identified template. One best model was chosen. Model visualization was done through DS viewer. The model shows that it consists of three alpha helices (red) and one beta-sheet. Evaluation of the model was done through Procheck. Procheck facilitated the validation of the model's stereochemical quality and overall reliability. Procheck helped in deriving Ramachandran Plot statistics show that 97.4% of debris is in the favoured region,0% in the additional allowed region, 2.65% in the generally allowed part, and 0% in the disallowed part. Overall, these results suggest that the protein structure being analyzed is well-folded and energetically stable. The ProSA analysis of the predicted 3D structure of the protein from Haemophilus influenzae yielded crucial information regarding its overall quality and stability.

Z score was highly negative. Greater the negative value, the more will be the stability of the model. It lends confidence to the structural prediction generated by the homology modelling approach suggesting that the model accurately captures the essential features of the protein's 3D architecture. Overall, the Pro-SA analysis contributes to a more thorough understanding of the protein's 3D structure and its potential implications for substrate binding and antibiotic design.

CONCLUSION

Haemophilus influenzae causes long-term and severe diseases. Homology modelling of the protein structure of Haemophilus influenzae predicts binding sites for the substrate can provide valuable insights for designing antibiotics. By identifying these binding sites specific regions of protein can be targeted where substrate molecules interact, thus helping in the rational design of antibiotics that can inhibit the function of the protein, leading to the development of potential antimicrobial agents against Haemophilus influenzae.

AUTHORS' CONTRIBUTION

RR: Literature search, conceptualization of the study design. MS: Data collection. S: Data analysis. AF: Data interpretation. SS: Write-up. MJ: Proofreading.

REFERENCES

1. Hameduh T, Haddad Y, Adam V, Heger Z. Homology modeling in the time of collective and artificial intelligence. Comput Struct Biotechnol J 2020;18:3494–506.
2. Jaiteh M, Rodrignes, Espigaries I, Selent J, Carlsson J. Performance of virtual screening against GPCR homology models: Impact of template selection and treatment of binding site plasticity. PLoS Comput Biol 2020;16(3):e1007680.
3. Dong S, Sun J, Mao Z, Wang L, Lu YL, Li J. A guideline for homology modeling of the proteins from newly discovered beta coronavirus;2019 novel coronavirus(2019-nCOv). J Med Virol 2020;92(9):1542–8.
4. Bitencourt-Ferriera G, de Azevedo WF. Homology modeling of protein targets with modeler. Methods Mol Biol 2019;2053:231–49.
5. Muhammad MT, Aki-Yalcin E. Homology Modeling in drug discovery; Overview,current applications, and future perspectives. Chem Biol Drug Des 2019;93(1):12–20.
6. Franca TC. Homology modeling: an important tool for drug discovery. J Biomol Struct Dyn 2015;33(8):1780–93.
7. Wen S, Feng D, Chen D, Yang L, Xu Z. Molecular epidemiology and evolution of Haemophilus influenzae. Infect Genet Evol 2020;80:104205.
8. Jalalvand F, Su YC, Morgelin M, Brant M, Hallgren O, Westergren-Thorsson G, *et al*. Haemophilus influenza protein F mediates binding to laminin and human pulmonary epithelial cells. J Infect Dis 2013;207(5):803–13.
9. Desta IT, Kotelnikov S, Jones G, Ghani U, Abyzov M, Kholodov Y, *et al*. Mapping of antibody epitope based on docking and homology modeling. Proteins 2023;91(2):171–82.
10. Hung JH, Weng Z. Sequence alignment and Homology search. Cold Spring Harb Protoc 2016;2016(11):pdb-rot093088.
11. Meler A, Soding J. Automatic prediction of protein 3D. Structures by probabilistic multi-template homology modeling. PLoS Comput Biol 2015;11(10):e1004343.
12. Labib MM, Amin MK, Alzohairy AM, Elashtokhy MMA, Samir O, Saleh I, *et al*. In silico targeting,inhibition and analysis of polyketide synthase enzyme in Aspergillus ssp. Saudi J Biol Sci 2020;27(12):3187–98.
13. Fieg M. Local protein structure refinement via molecular dynamics simulations with locPREFMD. J Chem Inf Model 2016;56(7):1304–12.
14. Park H, Ovchinnikov S, Kim DE, DiMaio F, Baker D. Protein homology model refinement by large –scale energy optimization. Proc Natl Acad Sci 2018;115(12):3054–9.
15. Yu Z, Kang L, Zhao Wu S, Ding L, Zheng F, Liu J, *et al*. Identification of novel umami peptides from myosin via homology modeling and molecular docking. Food Chem 2021;344:128728.

Submitted: October 25, 2023	*Revised: April 19, 2024*	*Accepted: April 26, 2024*

Address for Correspondence:

Maria Shafiq, Department of Physiology, Abbottabad Ayub Medical College, Abbottabad-Pakistan
Cell: +92 332 992 8090
Email: mariasaboor1@gmail.com