How to Cite:
Farid, A., Jadoon, M., Shoukat, S., Faryal, U., Karim, B., & Shahzad, A. (2023). Homology modeling – a step towards vaccine development by analyzing structure of haemophilus influenza protein, transcriptional regulator H10994. *International Journal of Health Sciences*, *7*(S1), 2344–2349. https://doi.org/10.53730/ijhs.v7nS1.14495

Homology modeling – a step towards vaccine development by analyzing structure of haemophilus influenza protein, transcriptional regulator H10994

Anila Farid
Lecturer, Department of Biochemistry, Abbottabad International Medical College, Abbottabad

Dr Madeeha Jadoon
Assistant Professor, Department of Biochemistry, Women Medical and Dental College, Abbottabad
Corresponding author email: Dr.jadoon73@gmail.com

Dr Sofia Shoukat
Assistant Professor, Department of Biochemistry, Ayub Medical College, Abbottabad

Prof Dr Uzma Faryal
Professor and HOD, Department of Biochemistry, Women Medical and Dental College, Abbottabad

Dr Bilal Karim
Assistant Professor Biochemistry, Islamabad Medical and Dental College, Islamabad

Dr Anwar Shahzad
Associate Professor, Community Medicine and Public Health AIMI, Abbottabad

***Abstract*---**Introduction: Haemophilus influenza is a type of bacterium that is non motile, gram negative and causes poisoning and infection including pneumonia,bronchitis etc. In order to study the resistivity of H.influenza protein, transcription regulator: HI0433, homology modeling is an important step to predict structure._Material and methods: Bioinformatics such as CMR, BLAST, modeller Prcheck and Prosa was carried out to find 3D structure of protein. Results and Discussion: H.influenza has 1792 proteins. Out of these, 456 hypothetical proteins were found. Homology modeling of transcriptional regulator H10994 was done it consists of 8 helices and

7 beta sheets. Ramachandran plot has shown that it consists of 95.2% particles in maximum allowed regions, 2.9 % particles in fewer allowed region, 1.4% particles in inadequate allowed region, 5% particles in disallowed region. Conclusion: By homology modeling of *H. influenza, transcriptional regulator protein (*HI0433), structure was designed which has provided enough information for vaccine development to control its transcription for causing disease.

***Keywords*---**Haemophilus influenza, Homology modeling, vaccine development.

Introduction

Haemophilus influenza is a type of bacterium that is non motile, gram negative and causes poisoning and infection including pneumonia, bronchitis etc[1]. It is transmitted through respiratory droplets or from close contact. Naturewise, It is generally aerobic .[2] Different types of H.influenza are known. Mostly six or seven types are talked about.[3] Numerous antibiotics like ceftriaxone and ceftotoxamine are introduced into the blood when infection is too serious.[4] It exhibits much resistance to Macrolide.[5] H.influenzae consists of many strains which are highly resistant to antibiotics. Thus making treatment highly challenging. Bioinformatics aids in analyzing large scale genomic data and understanding different purposes of proteins.[6]

In order to study the resistivity of H.influenza protein, **transcription regulator: HI0433**, homology modeling is an important step to predict structure. It is also known as comparative modeling which is preferred to predict 3D structure of protein.[7] It is based on its amino acid sequence and known structure of protein (template) .It gives important evidence that how resistivity can be reduced and vaccine is developed to cure infection effectively. Homology modeling is very fast technique to make or construct protein models. Template is used to build models of target proteins.Protein Data Bank (PDB) helps in searching template.[8,9]

The process involves aligning target with template sequence and then using that alignment for unknown protein model building. The level of resemblance between template and target sequence determines the accuracy of homology modeling. The predicted structure is more reliable if similarity is high.[10] A protein transcriptional regulator is responsible for infection. So it is selected for homology modeling in order to develop vaccine against it.Homology modeling helped a lot in discovery of ligands to modulate their activity.[11] Recognition of template from PDB data bank of known protein,template and querry sequence alignment, building of model on alignment basis and last refignment of built model are four steps for homology modeling.[12]

Material and Methods

Complete protein sequence of Haemophilus influenzae was downloaded from database of CMR (comprehensive microbial resources (http; cmr, jcvi. org). It is free website which put on view details entire prokaryotic genome. H.I contains

many proteins which do not show clear function (hypothetical protein) .We distinct hypothetical protein from total genome and put aside in word document format.

Interproscan is tool helps in visualizing function of hypothetical protein sequence. Different structure of proteins were modeled after function annotation of protein. In model building, template plays essential part. Blast against PDB (Protein date bank) helps in finding template. The template which has maximum identity, Low E -value, high query coverage is preferred to be selected with equivalent target proteins.

Target template alignment file helps in building target protein model. Blast helps in target and template alignment. Modeller 9.10 program is used to construct 3D homology model of protein. DS viewer helps in visualizing newly build model. 10 models were created. One best model was selected. Procheck helped in checking reliability of model and the energy is checked through Prosa A.

Results and Discussion

BLAST worked for alignment *E.Coli* Yebc protein (PDB i.d. 1KON), crystal structure *E.Coli* Yebc protein was selected as template for homology modeling. E value, querry coverage and sequence identity were the criteria for the selection of template. Similarity was 41% between target and template. 75% sequence identity and 100% query coverage was shown by both sequences. Modeller 9.10 was run to obtain 10 models. Out of 10 modelss, the best model was selected. The model was visualized by DS viewer. It consists of 8 alpha helices (red) and 7 beta sheets (blue).

The reliability of models was examined through Procheck.10 files were set up by Procheck. A valuable accent among them was Ramachandran Plot .It has 4 regions which are distributed into maximum, fewer, inadequately and disallowed regions.95.2% debris present in maximum allowed region, 2.9 % in fewer allowed region, 1.4% in inadequate allowed region and in disallowed region, 5% particle were found. The energy was accessed through Prosa which was highly negative (3e-140). It showed the stability of model.

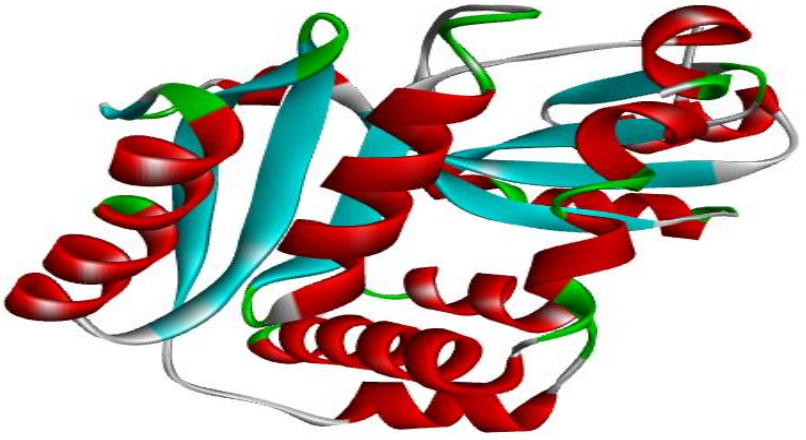
Sequences producing significant alignments:

Select: All None Selected:0

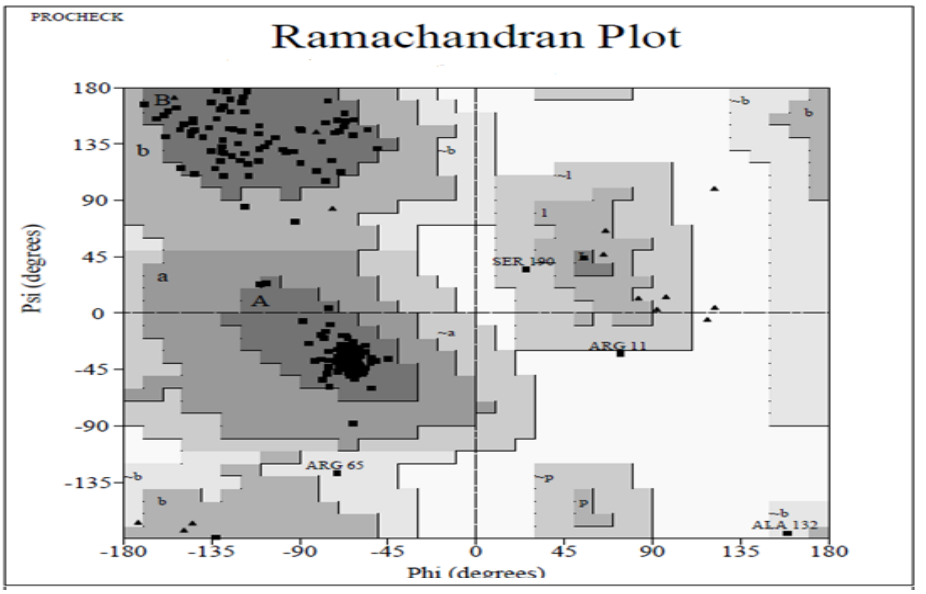
Alignments Download GenPept Graphics Distance tree of results Multiple alignment

Description	Max score	Total score	Query cover	E value	Max ident	Accession
Chain A, Crystal Structure Of E.Coli Yebc	397	397	100%	3e-140	75%	1KON_A
Chain A, Structure Of A Yebc Family Protein (Cbu_1566) From Coxiella Burnetii	274	274	97%	1e-91	55%	4F3Q_A
Chain A, Crystal Structure Of A Conserved Hypothetical Protein Aq1575 From Aquifex Aeolicus	231	231	99%	5e-75	48%	1LFP_A
Chain A, X-Ray Structure Of Y162_helpy Northeast Structural Genomics Consortium Target Pr6	100	100	92%	7e-25	28%	1MW7_A
Chain A, Chromo Shadow Domain From Fission Yeast Swi6 Protein >pdb|1E0B|B Chain B, Chromo Shadow Domain From Fission Yeast Swi6 Protein.	27.7	27.7	7%	2.1	47%	1E0B_A
Chain V, Model Of The Yeast F1fo-Atp Synthase Dimer Based On Subtomogram Average >pdb|4B2Q|VV Chain v, Model Of The Yeast F1fo-Atp Synthase Dimer Ba	27.3	27.3	14%	2.4	37%	4B2Q_V
Chain C, Subcomplex Of The Stator Of Bovine Mitochondrial Atp Synthase >pdb|2CLY|F Chain F, Subcomplex Of The Stator Of Bovine Mitochondrial Atp Synthase	27.3	27.3	14%	2.7	37%	2CLY_C
Chain A, Solution Structure Of Subunit F6 From The Peripheral Stalk Region Of Atp Synthase From Bovine Heart Mitochondria	27.3	27.3	14%	2.9	37%	1VZS_A
Chain V, The Structure Of The Membrane Extrinsic Region Of Bovine Atp Synthase >pdb|2WSS|Z Chain Z, The Structure Of The Membrane Extrinsic Region Of Bov	27.3	27.3	14%	3.2	37%	2WSS_V
Chain D, Methanococcus Maripaludis Phosphoseryl-Trna Synthetase	28.9	28.9	35%	4.1	28%	2ODR_D
Chain B, Methanococcus Maripaludis Phosphoseryl-Trna Synthetase	28.9	28.9	35%	4.2	28%	2ODR_B
Chain A, Methanococcus Maripaludis Phosphoseryl-Trna Synthetase	28.9	28.9	35%	4.2	28%	2ODR_A
Chain C, Methanococcus Maripaludis Phosphoseryl-Trna Synthetase	28.9	28.9	35%	4.2	28%	2ODR_C
Chain A, Human Retinoic Acid Receptor Rxr-Gamma Ligand-Binding Domain >pdb|2GL8|B Chain B, Human Retinoic Acid Receptor Rxr-Gamma Ligand-Binding	28.5	28.5	23%	4.7	34%	2GL8_A

Blast result of *H. influenzae*strain 86 028NP hypothetical protein HI0433



Three Dimensional structure of *H. influenza* hypothetical protein HI0433. (8 helices 7 beta sheets)



Ramachandran plot of H. influenza hypothetical protein HI0433 using PROCHECK

Plot statistics

Maximum allowed regions	**95.2% particles**
Fewer allowed region	**2.9 % particles**
Inadequate allowed region	**1.4% particles**
Disallowed region	**5% particles**

Conclusion

By homology modeling of *H. influenza, transcriptional regulator protein (*HI0433), structure was designed which has provided enough information for vaccine development to control its transcription for causing disease.

References

1. Arora, D.R. (2007). Textbook of Microbiology, 2nd Edition
2. Kuhnert, P., Christensen, H. (2008).Pasteurellaceae: Biology, Genomics and Molecular Aspects. *Caister Academic Press*. ISBN 978-1-904455-34-9.
3. Ryan, K.J., Ray, C.G., (2004). *Sherris Medical Microbiology* (4th ed.). McGraw Hill. pp. 396–401. ISBN 0-8385-8529-9.
4. Chang, C.M., Lauderdale, T.L., Lee, H.C., Lee, N.Y., Wu, C.J., Chen, P.L., Lee C.C., Chen, P.C., Ko, W.C., (2010). Colonisation of fluoroquinolone-resistant Haemophilus influenzae among nursing home residents in southern Taiwan". *J. Hosp. Infect.* 75 (4): 304–8.
5. Roberts, M. C. and Soge, O.O., (2011). Characterization of macrolide resistance genes in Haemophilus influenzae isolated from children with cystic fibrosis. *J. Antimicrob. Chemother.* 66 (1): 100–4.
6. Bourne, P. E. and Weissig, H. (2003). Structural Bioinformatics. Wiley-Liss

7. Mosimann, S., Meleshko, R. and James, M.N.G. (1995). A critical assessment of comparative molecular modeling of tertiary structures of proteins. *Proteins,* 23, 301-317.
8. Deng W,Liou SR, Plunkett G of human organic cation, Mayhew GF, Rose DJ, Burland V, et al. (2003). Comparative genomics of Salmonella enteric serovar Typhi strains Ty2 and CT18.Journal of bacteriology. 1;185(7):2330-7.
9. Janson G, Zhang C, Prado MG, Paiardini A. (2017). PyMod 2.0:improvement in protein sequence-structure analysis and homology modeling within PyMOL. Bioinformatics. Feb 1;33(3):444-6.
10. Dakal TC, Kumar R, Romator D. (2017). Structural modeling of human organic cation transporters.Computational Biology and Chemistry. 1;68:153-6
11. Costanzi S. (2011). Homology modeling of class ag proteincoupled receptors. In Homology Modeling (pp.259-279).Humana Press. 16. Qiu J, Zang S,Ma Y,Owusu L,Zhou L ,Jiang
12. Meier A, Soding J. (2015). Automatic prediction of protein 3D. structures by probabilistic multi-template homology modeling. PLoS Comput Biol. 23;(10):e1004343.