INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

Elucidating Structural, Functional and Phylogenetic Relationship of Large Envelope Protein of Hepatitis B V irus

Ruchi Yadav[1]*

1 Amity Institute of Biotechnology, Amity University, Uttar Pradesh Lucknow Campus, 226028

OPEN ACCESS

Received: 22.03.2021

Accepted: 28.01.2022

Published: 11.03.2022

Citation: Yadav R (2022) Elucidating Structural, Functional and Phylogenetic Relationship of Large Envelope Protein of Hepatitis B V irus. Indian Journal of Science and Technology 15(10): 451-456. https://doi.org/10.17485/IJST/v15i10.491

* **Corresponding author**.

ryadav@lko.amity.edu

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objectives: To predict the structure of L-HBsAg envelope protein of Hepatitis B Virus and its evolutionary relationship. **Methods:** Bioinformatics Computational methods were used for the structure prediction and evolutionary conservation analysis of large envelope protein of Hepatitis B virus (L-HBsAg). Structure prediction L-HBsAg envelope protein was done using Multi-template homology modeling method using Schrodinger software. Phylogenetic tree was constructed by MEGA tool to identify protein similarity of L-HBsAg protein with other genotypes of HBV. **Findings:** Modelled structure of L-HBsAg protein shows that this protein has single alpha-helix and loop structure. Multiple sequence alignment result shows that L-HBsAg belongs to family of vMSA (pfam00695, major surface antigen from hepadnavirus) conserved domain. Phylogenetic tree reveals that L-HBsAg shows similarity with other proteins of HBV protein like preS protein and envelope protein. **Novelty:** Structural analysis identifies the binding properties of L-HBsAg protein. This study concluded that large envelope protein of HBV can be potent drug target for designing a novel drug.

Keywords: LHBsAg; HBV; evolutionary analysis; structure prediction; dane particles

1 Introduction

Hepatitis B is infection of liver caused by Hepatitis B Virus (HBV) and it infects millions of people worldwide.[(1)] Infections can be transferred from infected person to healthy person and can cause severe illness.[(2)] It can be caused by contaminated needles, body fluids, blood contact, mother to infants, semen's etc. Infected people show different symptoms like illness, vomiting, headache, fatigue, jaundice, fever, weakness etc.[(3)]

Hepatitis B Virus is a DNA virus and that encodes proteins responsible for its survival and development. HBV Genome carries genes that encodes outer envelope and inner core of the virus.[(4)] HBV outer core include surface proteins called as HBV surface antigen protein or HBsAg. It can be easily detected in blood of infected persons and positive test indicates infections with HBV virus.[(5)]

Inner core of HBV contains core protein called as HBcAg that encodes for the functional proteins that are responsible for HBV survival, replication and infection[(6)].

HBV genome and its lifecycle has been highly studied to understand the mechanism and mode of infection. HBV infects the liver and then its integrates into the nuclear genome of liver cell and uses its mechanism for the production of more HBVs and causes severe infection [7]. Even though the genome and HBV has been highly studied but its mode of infection is not completely understood [8]. Despite the availability of vaccine against HBV infection there are cases wherein people get infected hence more research is required to completely understand the genome complexity and mechanism of HBV.

It is important to study all the proteins that are expressed by HBV to unlock the potential of these proteins as drug targets and possibility that these target proteins can be used for drug development. HBV proteins can be categorized into core proteins, surface proteins, polymerase and Polymerase proteins [9]Figure 1 shows the HBV genome as retrieved from KEGG genome database it shows the different class of proteins that are expressed by HBV genome. The envelope is made up of three proteins the large HBsAg (L-HBsAg), middle (M-HBsAg), and major HBsAg (S-HBsAg) all encoded by S gene. Among it the L-HBsAg protein and the M-HBsAg are minor proteins and S-HBsAg is the major protein. [10] The three proteins are embedded in a lipid bilayer originating from the host cell. Each surface protein has a glycosylation site in the S domain. Additional modifications of the L and M proteins occur at the pre-S2 domain with an N-linked oligosaccharide and a myristic acid at the amino-terminal glycine residue of the pre-S1 domain. [11]

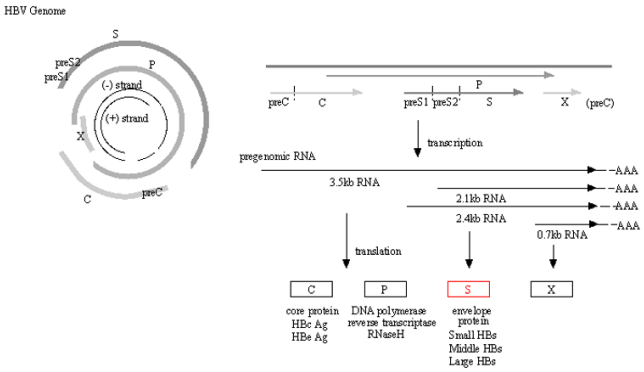

Fig 1. Genome of HBV retrieved from KEGG Database (PATHWAY: map05161, https://www.genome.jp/entry/pathway+map05161)

The large and the middle envelope protein are majorly involved in viral infection process. In HBV-infected patients, the content of circulating pre-Si antigen is paralleled by that of viral DNA suggesting that the LS protein has a role in the assembly of Dane particles. [12] The structures of the three envelope proteins suggest that they interact with each other to allow the assembly of particles, but the mode of this interaction and the role that each protein plays in the assembly process have not yet been elucidated. [13] Therefore, it is of interest to document the structural and evolutionary conservation analysis of large envelope protein molecular model of Hepatitis B virus (L-HBsAg) and Table 1 shows the list of proteins that are widely studied, and information has been submitted in protein database UniProt (https://www.uniprot.org/).

Table 1. List of HBV proteins and gene encoding, length of protein and UniProt Accession ID

S.No	Protein Name	Gene Name	Length (amino acid)	UniProt ID
1	HBc antigen	HBc C	183	A0A679DQ48
2	External core antigen (HBe antigen)	c	80	A0A0N9JUK8
3	Large Envelope Protein	S PreS1, preS1, preS1/S2/S, preS2	400	Q99HV3
4	Main protein	S OG17, OG18, OG22, OG25, OG27	226	Q9W966
5	Surface protein (S protein)	OG34 OG26, OG30, S, s	226	Q9DHB4
6	HBsAg	S s, HBVgp2	226	Q9PX80
7	Small surface antigen	s	226	Q04258
8	Polymerase protein	P	843	Q69028

Detailed study of HBV genomes and proteins in protein databases shows that despite of sequence deposition and detailed sequence information there is no structure information for proteins [14]. Protein structural details are very important to examine the functional property of proteins. Proteins structures gives detailed information about functional and binding property of protein. To elucidate the drug protein interaction prediction of protein structure becomes very important when

the experimental structure from X-ray or NMR is not available.

Sequence information decodes all the information required for the survival, production, and existence of any organism and so with viruses also. Sequence information can be used for prediction of sequence similarity and phylogenetic analysis. Current research had been done to understand the origin and diversity of HBV and to study genome conservation and variation.

Bioinformatics has enabled us to study the proteins' structure, function, evolutionary relationship etc. with significant analysis and careful investigation critical functions can be studied that can be helpful in wet lab experiments and can be used for drug design also. With this in this current research detailed analysis of HBV protein has been done to elucidate the protein structure function and evolutionary relationship of L-HBsAg.

2 Materials and Methods

The protein sequence of L-HBsAg envelope protein was retrieved from UniProt database (https://www.uniprot.org/) with UniProt id A8CEJ1 of L-HBsAg envelope protein. Structure prediction of L-HBsAg was done by multi-template homology modelling method using Schrodinger software. [15] Methodology along with the software's used for structure and phylogenetic tree prediction has been shown in Figure 2. Structural verification of modeled structure and statistical analysis was done by Ramachandran plot using SAVES server (https://saves.mbi.ucla.edu/) [16]

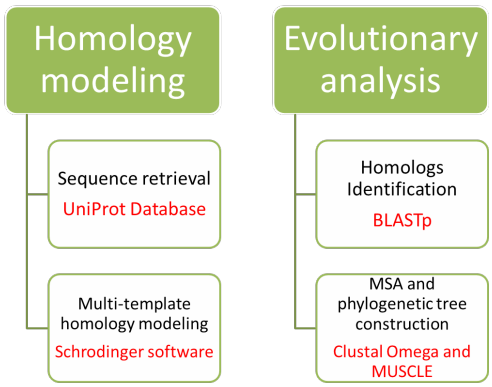

Fig 2. Methodology adopted for homology modeling and evolutionary analysis

BLASTP tool (https://blast.ncbi.nlm.nih.gov/) was used to identify homologs of L-HBsAg protein, and Multiple Sequence Alignment (MSA) was done between homologous sequences obtained from BLASTP using CLUSTAL omega tool (https://www.ebi.ac.uk/Tools/msa/clustalo/) [17]. MSA data was used to construct phylogenetic tree using clustal phylogeny tool (http://www.ebi.ac.uk/Tools).

3 Result and Discussion

3.1 Homology Modeling of L-HBsAg

Homology modeling of L-HBsAg protein was done using Multi Template Homology Modeling Method using Prime tool [18] of Schrodinger software. This method was used to build structure of query protein on the basis of three homologous proteins structures because different regions show similarity with different protein. So to obtain complete protein structure three templates were used. Three templates viz, 1WZ4_A, 1KCR_P and 4HIC_A were identified for structure prediction of L-HBsAg envelope protein. Details of these template proteins were shown in Table 2 and alignment between these templates are shown in Figure 3.

Figure 4 shows the modeled structure of L-HBsAg envelope protein. Structural analysis shows that this protein has single alpha-helix and loop structure. Structural verification of modeled L-HBsAg was done using Ramachandran plot analysis using SAVES server and shown in Figure 5. Ramachandran plot analysis of L-HBsAg envelope protein shows that only two amino acids viz Asp -133 and Gly-144 were in disallowed region and all the other residues are in allowed region. It shows that modeled structure is stable and can be used for further study like docking, ligand identification, ligand -protein interaction study etc.

Table 2. List of templates used for homology modeling of L-HBsAg envelope protein

S. no	Template id	Description	Identities	Positives
1	1WZ4_A	Solution Conformation of adr subtype HBV Pre-S2 Epitope	95%	100%
2	1KCR_P	Crystal Structure of antibody PC283 in complex with PS1 peptide	93%	93%
3	4HIC_A	Crystal structure of the potential transfer protein TraK from Gram-positive conjugative plasmid pIP501	92%	92%

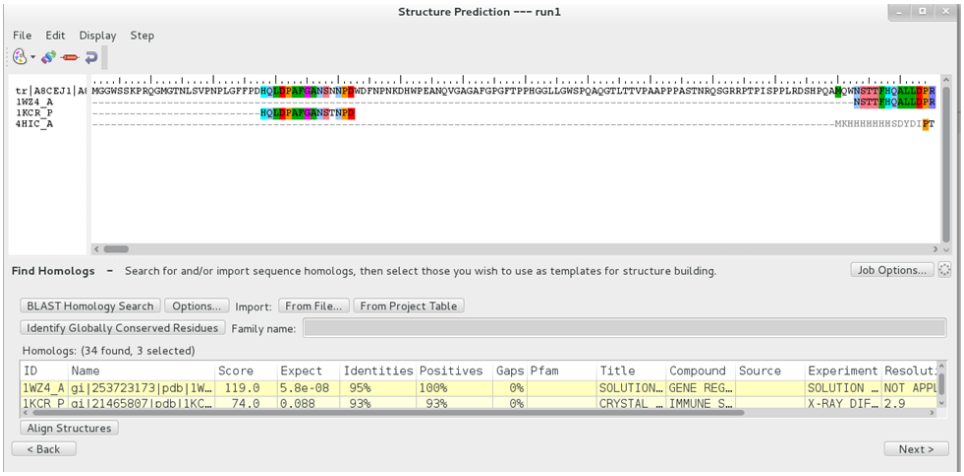

Fig 3. Multi templatehomology modeling of L-HBsAg envelopeprotein (UniProt Id A8CEJ1)

Fig 4. Modeled structure of L-HBsAg envelope protein

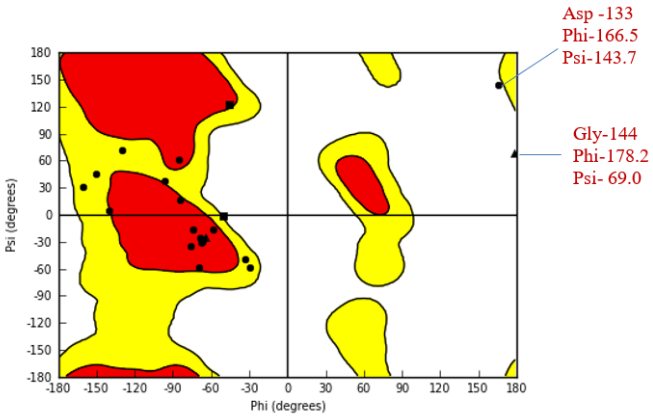

Fig 5. Ramachandran plotanalysis of L-HBsAg envelope protein for structural verification andidentification of outlier amino acids. As shown in Ramachandran plot only twoamino acid viz Asp -133 and Gly-144 were in disallowed region.

3.2 Multiple Sequence Alignment (MSA)

MSA of L-HBsAg protein was done using MEGA tool[19] and shown in Figure 6. Phylogenetic tree was constructed to identify protein similarity of L-HBsAg protein with other genotypes of HBV. BLASTP and Multiple Sequence Alignment (Figure 6) result shows that L-HBsAg belongs to family of vMSA (pfam00695, major surface antigen from hepadnavirus) conserved domain.

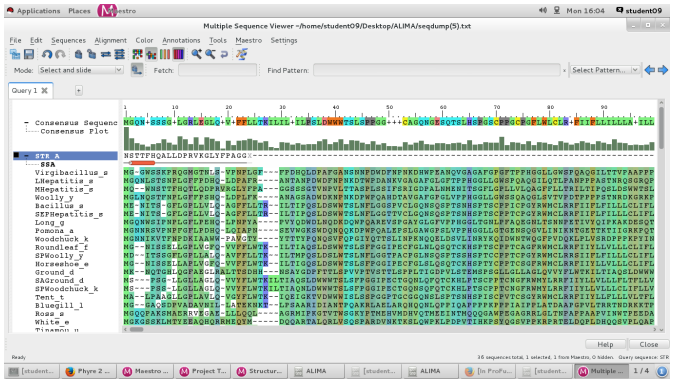

Fig 6. Multiple sequence alignment of L-HBsAg envelope protein that shows the conserved region (green) and gap region (white)

Phylogenetic tree was constructed using UPGMA method and detailed analysis of this evolutionary tree reveals that L-HBsAg shows similarity with other proteins of HBV protein like preS protein and envelope protein as shown in Figure 7

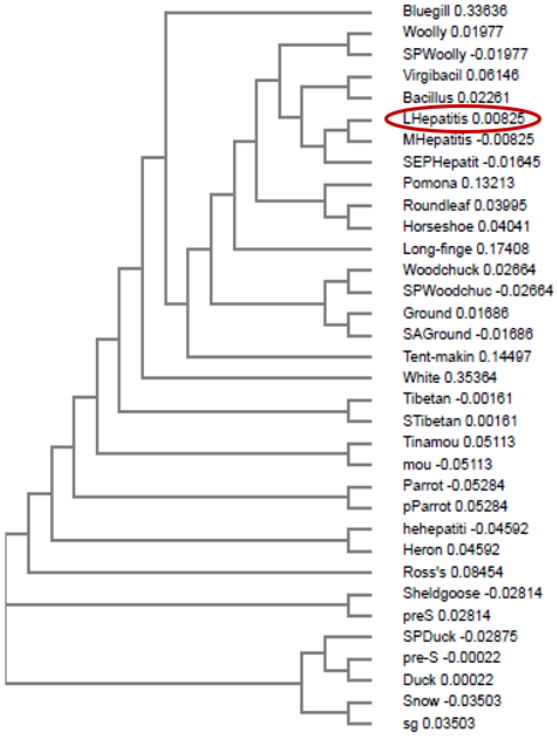

Fig 7. Phylogenetic tree of L-HBsAg envelope protein and its similarity with other organism

4 Conclusion

In this research, structural and evolutionary conservation analysis of large envelope protein molecular model of Hepatitis B virus (L-HBsAg) was done since is it is important protein that causes infection. Detailed analysis was done to gain insight in to

the evolutionary biology of L-HBsAg protein that can help to design potential inhibitors. Further Phylogenetic analysis shows that L-HBsAg protein is conserved in wide family of Hepatitis Viruses. Evolutionary analysis shows that L-HBsAg is evolved from Tibetan frog hepatitis B virus. This study concluded that large envelope protein of HBV can be a potent drug target for designing a novel drug for Hepatitis B.

5 Acknowledgement

The author acknowledges the sources of bioinformatics tools, software and databases that were used for this study.

References

1) Seto WK, Lo YR, Pawlotsky JM, Yuen MF. Chronic hepatitis B virus infection. *The Lancet*. 2018;392(10161):2313–2324. Available from: https://dx.doi.org/10.1016/s0140-6736(18)31865-8.
2) Shih C, Yang CC, Choijilsuren G, Chang CH, Liou AT. Hepatitis B Virus. *Trends in Microbiology*. 2018;26:386–387. Available from: https://dx.doi.org/10.1016/j.tim.2018.01.009.
3) Hu J, Protzer U, Siddiqui A. Revisiting Hepatitis B Virus: Challenges of Curative Therapies. *Journal of Virology*. 2019;93(20):1032–1051. Available from: https://dx.doi.org/10.1128/jvi.01032-19.
4) Revill PA, Tu T, Netter HJ, Yuen LKW, Locarnini SA, Littlejohn M. The evolution and clinical impact of hepatitis B virus genome diversity. *Nature Reviews Gastroenterology & Hepatology*. 2020;17(10):618–634. Available from: https://dx.doi.org/10.1038/s41575-020-0296-6.
5) Ko C, Chakraborty A, Chou WM, Hasreiter J, Wettengel JM, Stadler D, et al. Hepatitis B virus genome recycling and de novo secondary infection events maintain stable cccDNA levels. *Journal of Hepatology*. 2018;69(6):1231–1241. Available from: https://dx.doi.org/10.1016/j.jhep.2018.08.012.
6) Hu J, Protzer U, Siddiqui A. Revisiting Hepatitis B Virus: Challenges of Curative Therapies. *Journal of Virology*. 2019;93(20):1032–1051. Available from: https://dx.doi.org/10.1128/jvi.01032-19.
7) Budzinska MA, Shackel NA, Urban S, Tu T. Cellular Genomic Sites of Hepatitis B Virus DNA Integration. *Genes*. 2018;9(7):365–365. Available from: https://dx.doi.org/10.3390/genes9070365.
8) Martinez MG, Villeret F, Testoni B, Zoulim F. Can we cure hepatitis B virus with novel direct-acting antivirals? *Liver International*. 2020;40(S1):27–34. Available from: https://dx.doi.org/10.1111/liv.14364.
9) Zhao Q, Hu Z, Cheng J, Wu S, Luo Y, Chang J, et al. Hepatitis B Virus Core Protein Dephosphorylation Occurs during Pregenomic RNA Encapsidation. *Journal of Virology*. 2018;92(13):2139–2156. Available from: https://dx.doi.org/10.1128/jvi.02139-17.
10) Mak LY, Wong DKH, Cheung KS, Seto WK, Lai CL, Yuen MF. Review article: hepatitis B core-related antigen (HBcrAg): an emerging marker for chronic hepatitis B virus infection. *Alimentary Pharmacology & Therapeutics*. 2018;47:43–54. Available from: https://dx.doi.org/10.1111/apt.14376.
11) Mitra B, Wang J, Kim ES, Mao R, Dong M, Liu Y, et al. Hepatitis B Virus Precore Protein p22 Inhibits Alpha Interferon Signaling by Blocking STAT Nuclear Translocation. *Journal of Virology*. 2019;93(13):196–215. Available from: https://dx.doi.org/10.1128/jvi.00196-19.
12) Waheed Y, Siddiq M, Jamil Z, Najmi MH. Hepatitis elimination by 2030: Progress and challenges. *World Journal of Gastroenterology*. 2018;24:4959–4961. Available from: https://dx.doi.org/10.3748/wjg.v24.i44.4959.
13) Caballero A, Tabernero D, Buti M, Rodriguez-Frias F. Hepatitis B virus: The challenge of an ancient virus with multiple faces and a remarkable replication strategy. *Antiviral Research*. 2018;158:34–44. Available from: https://dx.doi.org/10.1016/j.antiviral.2018.07.019.
14) Gerlich WH, Glebe D, Kramvis A, Magnius LO. Peculiarities in the designations of hepatitis B virus genes, their products, and their antigenic specificities: a potential source of misunderstandings. *Virus Genes*. 2020;56(2):109–119. Available from: https://dx.doi.org/10.1007/s11262-020-01733-9.
15) Yadav R, Srivastava P. Establishment of resveratrol and its derivatives as neuroprotectant against monocrotophos-induced alteration in NIPBL and POU4F1 protein through molecular docking studies. *Environmental Science and Pollution Research*. 2020;27(1):291–304. Available from: https://dx.doi.org/10.1007/s11356-019-06806-3.
16) Rajbhar P, Singh D, Fatima R, Yadav R. Evolutionary relationship and repurposing of sars inhibitors against surface glycoprotein of SARS-COV-2. *International Journal of Applied Pharmaceutics*. 2021;2021:272–279. Available from: https://dx.doi.org/10.22159/ijap.2021v13i5.42652.
17) Yadav R, Srivastava P. Interaction Study of Antioxidants with Progressive Myoclonus Epilepsy by Molecular Docking Techniques. *Research Journal of Pharmacy and Technology*. 2019;12(2):584–584. Available from: https://dx.doi.org/10.5958/0974-360x.2019.00104.5.
18) Yadav R, Srivastava P. Interaction Study of Antioxidants with Progressive Myoclonus Epilepsy by Molecular Docking Techniques. *Research Journal of Pharmacy and Technology*. 2019;12(2):584–584. Available from: https://dx.doi.org/10.5958/0974-360x.2019.00104.5.
19) Nagy LG, Merényi Z, Hegedüs B, Bálint B. Novel phylogenetic methods are needed for understanding gene function in the era of mega-scale genome sequencing. *Nucleic Acids Research*. 2020;48(5):2209–2219. Available from: https://dx.doi.org/10.1093/nar/gkz1241.