Article

Analysis of ORF5 protein signifies its importance in Norway rat Hepatitis E virus

Zoya Shafat[1], **Anwar Ahmed**[2], **Mohammad K. Parvez**[3], **Asimul Islam**[1], **Shama Parveen**[1]

[1]Centre for Interdisciplinary Research in Basic Sciences, Jamia Millia Islamia, New Delhi, India

[2]Centre of Excellence in Biotechnology Research, College of Science, King Saud University, Riyadh, Saudi Arabia

[3]Department of Pharmacognosy, College of Pharmacy, King Saud University, Riyadh, Saudi Arabia

E-mail: zoya179695@st.jmi.ac.in, anahmed@ksu.edu.sa, mohkhalid@ksu.edu.sa, sparveen2@jmi.ac.in

Received 20 November 2021; *Accepted 30 December 2021*; *Published 1 June 2022*

Abstract

Hepatitis E virus (HEV) is the chief cause of hepatitis E (inflammation of liver) across the globe. The Norway rat HEV genome consists of six open reading frames (ORFs), i.e., ORF1, ORF2, ORF3, ORF4, ORF5 and ORF5. The additional reading frame encoded protein ORF5 protein's structure and function remain to be explored. Therefore, the presented study was conceptualized to analyze the ORF5 protein for its physiochemical properties, primary structure, secondary structure, tertiary structure and functional characteristics using bioinformatics tools. The initial analysis revealed ORF5 protein as unstable, thermostable, hydrophilic and highly basic in nature. The primary structural analysis revealed higher percentages of amino acids Arg, Leu, Pro, Ser and Gly, which suggested that the ORF5 protein is richly endowed with some regulatory amino acids (Leu, Pro and Gly). The secondary structure of ORF5 protein showed all three major components (alpha-helix, beta-strand and random coil). The tertiary structure generated through homology modelling revealed mixed α/β structural fold with subsequently higher percentage of strands and abundance of coils. Moreover, the surface analysis revealed the several clefts and tunnels along with few pores, clearly suggested the ability of ORF5 protein towards interaction with other molecules. The ORF5 protein was also identified with several post-translationally modified sites including glycosylation, phosphorylation and myriystoylation. The presence of these modified sites indicated the role of ORF5 protein in regulation. Thus, our analyses taken together interpret the ORF5 protein's essentiality in HEV. This data will help in exploring the prospective role of this additional genomic component of rat HEV through the sequence, structure and functional annotation of ORF5 protein.

Keywords rat HEV; open reading frame 5 (ORF5); physicochemical parameters; structural analysis; homology modeling; functional analysis.

Network Biology
ISSN 2220-8879
URL: http://www.iaees.org/publications/journals/nb/online-version.asp
RSS: http://www.iaees.org/publications/journals/nb/rss.xml
E-mail: networkbiology@iaees.org
Editor-in-Chief: WenJun Zhang
Publisher: International Academy of Ecology and Environmental Sciences

1 Introduction

Hepatitis E is inflammation of the liver which is caused by the Hepatitis E virus (HEV) (Kumar et al., 2013). Worldwide, about 20 million HEV infections and 3.3 million symptomatic hepatitis E cases occur annually which results in 44,000 deaths (Khuroo and Khuroo, 2016). HEV is of the family *Hepeviridae* and belongs to the genus *Orthohepevirus* (Takahashi et al., 2010). The genome if HEV is a single stranded positive-sense RNA (7.2 kb in length), which is flanked with short 5′ and 3' non-coding regions (NCR) (Tam et al., 1991). The HEV genome comprises three open reading frames (ORFs): ORF1, ORF2 and ORF3. The ORF1, ORF2 and ORF3 encode the non-structural polyprotein (pORF1), capsid protein (pORF2) and the pleotropic protein (pORF3) respectively (Kenney and Meng, 2019).

The HEV ORF1 consists of seven domains: methyltransferase (MTase/MeT), Y (undefined), papain-like cysteine protease (PCP), hypervariable region/proline-rich hinge (HVR/PPR), X (macro), helicase (Hel/NTPase), and RNA-dependent RNA polymerase (RdRp) (Ansari et al., 2000; Ropp et al., 2000; Suppiah et al., 2011; Parvez, 2017). Rodent also plays a role in the epidemiology of hepatitis E. In the year 2009, two complete nucleotide sequences were analyzed from Norway rats in Germany which suggested a completely separate genotype for these HEV strains (Johne et al., 2010). These nucleotide sequences had high divergence to other HEV strains, i.e., HEV G1, HEV G2, HEV G3, HEV G4 and avian HEV (Johne et al., 2010). It was predicted through software that the genome in these rat HEV sequences was organized into a total of six reading frames (ORF1, ORF2, ORF3, ORF4, ORF5 and ORF5). i.e., rat HEV genome consisted of three additional ORFs (ORF4, ORF5 and ORF5). It was also identified that unlike typical HEV genomic organization, the ORFs ORF1 and ORF3 do not overlap in these two rat HEVs. Three additional putative ORFs of 280 - 600 nt that overlap with ORFs 1 or 2 were predicted for each rat HEV genome (Johne et al., 2010).

Though reports have demonstrated the structural and functional characteristics of some of the HEV ORF encoded proteins using computational approaches (Shafat et al., 2021a, 2021b, 2021c), however, the additional frame encoded protein ORF5 in Norway rat HEVs has not been explored. Therefore, the presented study describes the characteristics of the ORF5 protein of rat HEVs in terms of its structure and function. Using computational approach, we provide novel insights into the ORF5 protein's physiochemical properties, primary structure, secondary structure, tertiary structure and functional characteristics.

2 Material and Methods

2.1 Sequence retrieval

The rat HEV ORF5 amino acid sequence (Accession number: GU345043) was retrieved from the NCBI (National Center for Biotechnology Information) GenBank.

2.2 Physicochemical properties analysis

The amino acid sequences of the ORF5 protein in FASTA format was used as query in for the determination of physiochemical parameters. The various physical and chemical parameters of the retrieved sequences were computed using ProtParam (Expasy), a web-based server (Gasteiger et al., 2005). The ProtParam tool employed various parameters such as, instability index (II – protein stability) (Guruprasad et al., 1990), aliphatic index (AI – relative volume occupied by protein's aliphatic side chains) (Ikai, 1980), extinction coefficients (EC – protein-protein/protein-ligand interactions quantitative study) (Gill and Hippel, 1989), Grand Average of Hydropathicity (GRAVY - sum of all hydropathicity values divided by number of residues in a sequence) (Kyte and Doolittle, 1982), theoretical pI, half-life (Gonda et al., 1989), and number of positive and negative residues.

2.3 Structural analysis

The analysis of primary structure in terms of amino acid percentage composition for the ORF5 protein was computed using the ProtParam (Expasy) tool and PSIPRED (PSIPRED Workbench (ucl.ac.uk). The analysis of the secondary structure for the ORF5 protein was carried out using the online server Phyre2 Protein Homology/AnalogY Recognition Engine) (http://www.sbg.bio.ic.ac.uk/phyre2). The analysis of tertiary structure of the ORF5 protein was also conducted using Phyre2. The generated ORF5 protein 3D model was validated using Ramachandran plot analysis (PROCHECK) (http://nihserver.mbi.ucla.edu/SAVES) for stereo-chemical property.

2.4 Functional analysis

The N-linked sites for glycosylation were predicted using NetNGlyc 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc/) server**,** provided by Centre for Biological Sequence Analysis, Technical University of Denmark (CBS DTU). The O-linked sites for glycosylation were predicted using NetOGlyc 4.0 (http://www.cbs.dtu.dk/services/NetOGlyc/) server, provided by Centre for Biological Sequence Analysis, Technical University of Denmark (CBS DTU). The phosphorylation sites were predicted using NetPhos3.1 (NetPhos - 3.1 - Services - DTU Health Tech) server, provided by Centre for Biological Sequence Analysis, Technical University of Denmark (CBS DTU). For phosphorylation studies, we performed both generic and kinase specific predictions. ANTHEPROT v.6.9.3 predicted phosphorylation and other modified sites in the ORF5 protein. Location of signal peptide cleavage in the ORF5 protein was predicted using Signal P-4.1 (SignalP - 5.0 - Services - DTU Health Tech).

3 Results

3.1 Analysis of physicochemical properties characteristics

The various physiochemical parameters of the ORF5 protein are summarized in Table 1.

Table 1 Physicochemical parameters of the ORF5 protein of rat HEV.

Physicochemical properties	**ORF5**
Number of amino acids	186
Molecular weight	21279.08
Theoretical pI	11.92
Total number of negatively charged residues (Asp + Glu)	4
Total number of positively charged residues (Arg + Lys)	36
Formula	$C_{949}H_{1530}N_{304}O_{235}S_{10}$
Total number of atoms	3028
Extinction coefficient (assuming all Cys pairs residues form cystines)	62825
Extinction coefficient (assuming all Cys pairs residues are reduced)	62450
Estimated half-life	30 hours (mammalian reticulocytes, in vitro) > 20 hours (yeast, in vivo) > 10 hours (*Escherichia coli*, in vivo)
Instability index	65.75
Aliphatic index	77.15
Grand average of hydropathicity (GRAVY)	-0.490

Physicochemical analysis showed that the ORF5 polypeptide is of 186 amino acids. The isoelectric point (pI) of 11.92, showed it to be highly basic. The instability index of 65.75 indicated the unstable nature of the ORF5 protein (as >40 value implies unstable protein). Also, the positive aliphatic index value (77.15) predicted ORF5 protein as thermostable. Additionally, the negative GRAVY score (-0.490) indicated the hydrophilic nature of the ORF5 protein (as positive GRAVY value indicates hydrophobicity). Thus, taken together it can be interpreted that the ORF5 protein was found to be unstable, thermostable, hydrophilic and highly basic in nature.

3.2 Analysis of structural characteristics

Proteins differ from one another in their structure, primarily in their sequence of amino acids. The linear sequence of the amino acid polypeptide chain refers to its primary structure. The amino acid composition of ORF5 protein is summarized in Table 2 (Fig. 1).

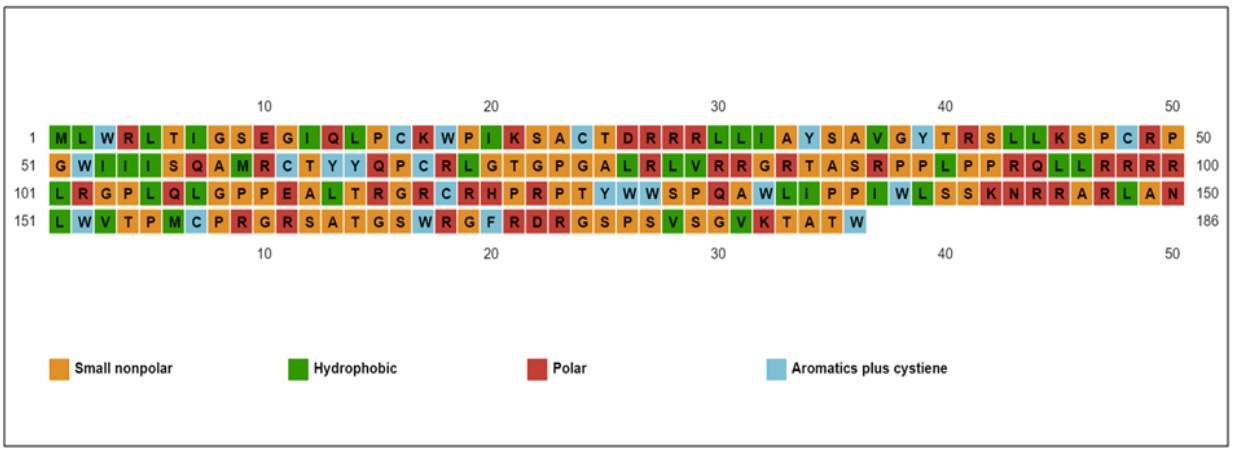

Fig. 1 Representation of amino acid composition in ORF5 protein. The analysis was conducted using PSIPRED.

Table 2 Amino acid composition of the ORF5 protein of rat HEV.

Amino acid	ORF5
Ala (A)	6.5
Arg (R)	16.7
Asn (N)	1.1
Asp (D)	1.1
Cys (C)	3.8
Gln (Q)	3.2
Glu (E)	1.1
Gly (G)	8.6
His (H)	0.5
Ile (I)	4.8
Leu (L)	11.3
Lys (K)	2.7

Met (M)	1.6
Phe (F)	0.5
Pro (P)	11.3
Ser (S)	8.1
Thr (T)	6.5
Trp (W)	5.4
Tyr (Y)	2.7
Val (V)	2.7
Pyl (O)	0
Sec (U)	0

*The values are represented as percentages.

Arg was observed as the top contributing amino acid to the ORF5's polypeptide chain. The top five amino acids that contributed to the polypeptide chain of ORF5 were included Arg, Leu, Pro, Ser and Gly (Fig. 1).

The predicted elements of secondary structure in the ORF5 protein are shown in Fig. 2. It was revealed that all three secondary structure components were observed in the ORF5 protein, i.e., alpha helix (α), beta strand (β) and random coil (Fig. 2). However, the percentage of β-strand was much greater than α-helix.

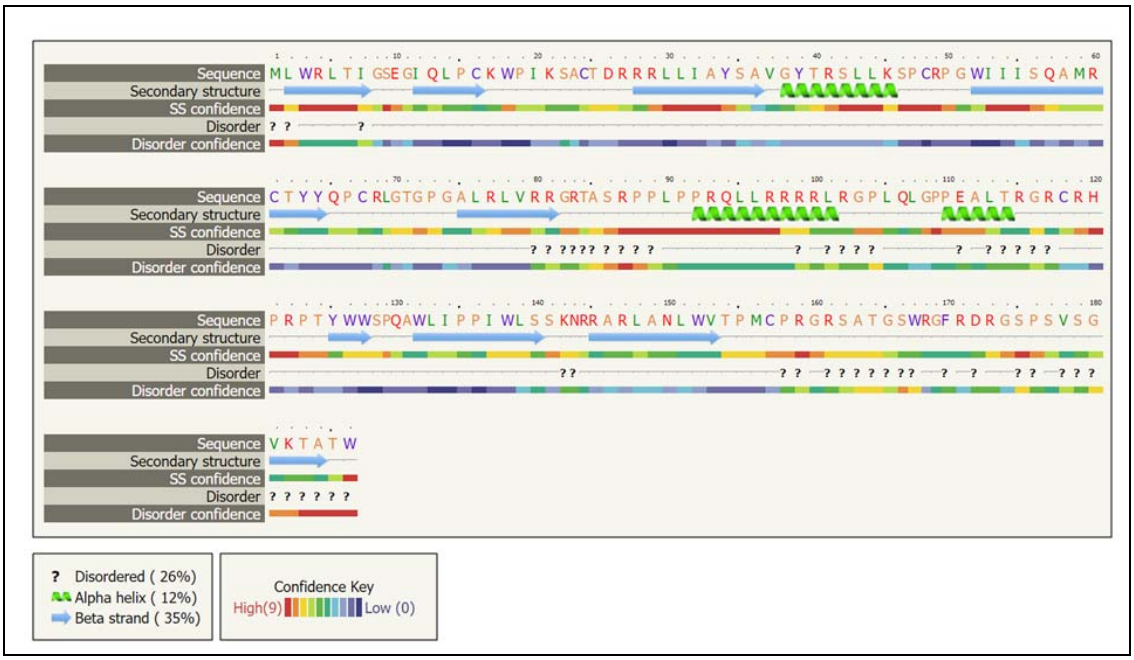

Fig. 2 Secondary structure elements of ORF5 protein of rat HEV. The analysis was conducted using Phyre2 server.

The generated 3D tertiary structure of the ORF5 protein (via Phyre2) was analyzed by visualization through homology modelling approach (Table 3) (Fig. 3). The elements of secondary structure form a three-dimensional structure of the protein by different combination patterns. The modeled 3D structure was in

agreement with the secondary structure prediction and revealed higher percentage of strands with predominance of coils.

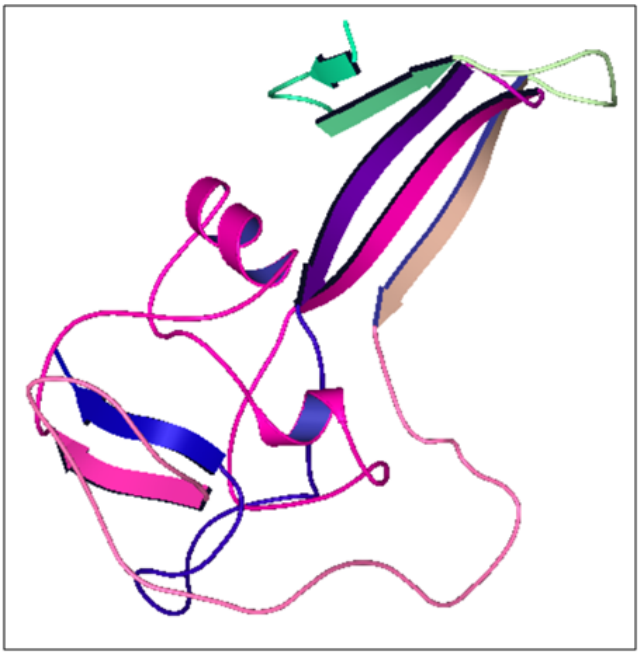

Fig. 3 Tertiary structure of the ORF5 protein of rat HEV. The analysis was conducted using the webserver Phyre2.

Table 3 Properties of the modelled 3D structure generated through Phyre2

Model dimensions (Å)	Template	Secondary structure and disorder prediction
X: 39.420	C2e6iA	Disordered (26%)
Y: 88.492		Alpha helix (12%)
Z: 66.812		Beta strand (25%)

Further, the obtained 3D ORF5 model (generated through Phyre2) was evaluated using PROCHECK (Fig. 4). The overall protein's stereochemical quality, amino acids present in the allowed, disallowed region and the G-factor was evaluated by Ramachandran map (Table 4).

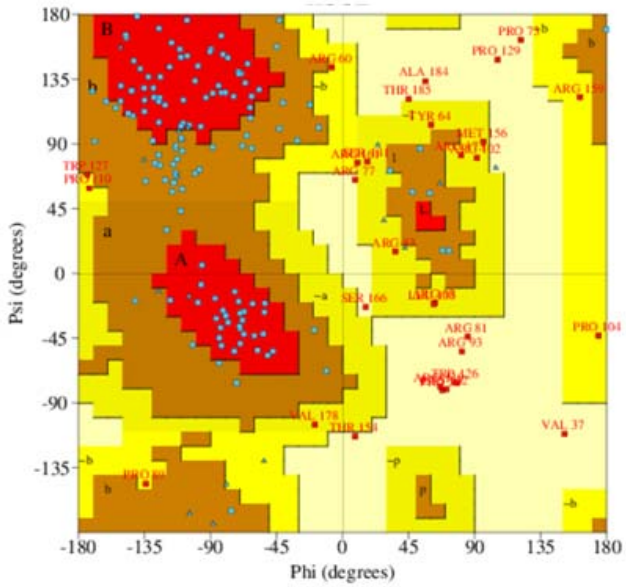

Fig. 4 Ramachandran plot of the ORF5 protein of rat HEV showing the favoured region**s.** The analysis was conducted using PROCHECK.

Table 4 Statistics for the obtained 3D ORF5 model.

PDBsum analysis	
Clefts	10
Pores	2
Tunnels	6
PROCHECK analysis	
Ramachandran Plot statistics	
Most Favoured Regions	48.3%
G-Factors	
Overall average	-2.63

Based on an analysis of **118** structures of resolution of at least **2.0** Angstroms and *R*-factor no greater than **20.0** a good quality model would be expected to
have over **90%** in the most favoured regions [A,B,L].
***G-factors** provide a measure of how **unusual**, or out-of-the-ordinary, a property is.
Values below -0.5* - unusual
Values below **-1.0**** - highly unusual

Further, the obtained 3D structural model was also identified with several clefts and tunnels in addition to few pores (Fig. 5).

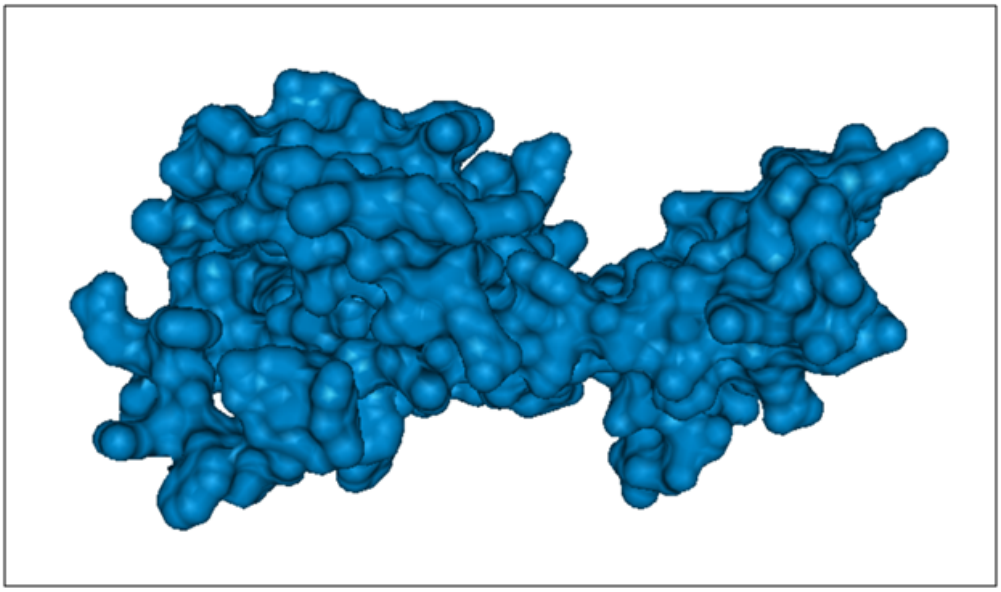

Fig. 5 Surface representation of the modeled 3D structure of the ORF5 protein of rat HEV.

3.3 Analysis of functional characteristics

None of the N-linked for glycosylation were identified in the ORF5 protein. However, 10 O-linked possible sites for glycosylation were predicted in protein sequence. Additionally, several phosphorylation sites including 14 Ser, 7 Thr and 2 Tyr residues were identified in the ORF5 protein using NetPhos3.1 server (Fig. 6).

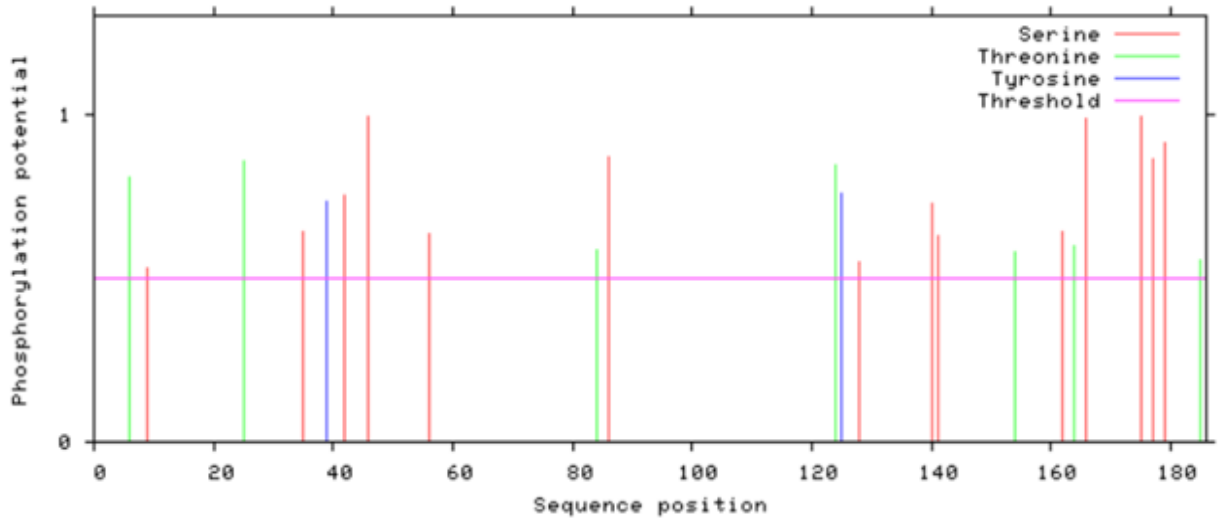

Fig. 6 Predicted phosphorylation sites in the rat HEV ORF5 protein sequence using NetPhos3.1.

The potential cleavage site for signal peptide was found to be absent in the ORF5 protein sequence (Fig. 7).

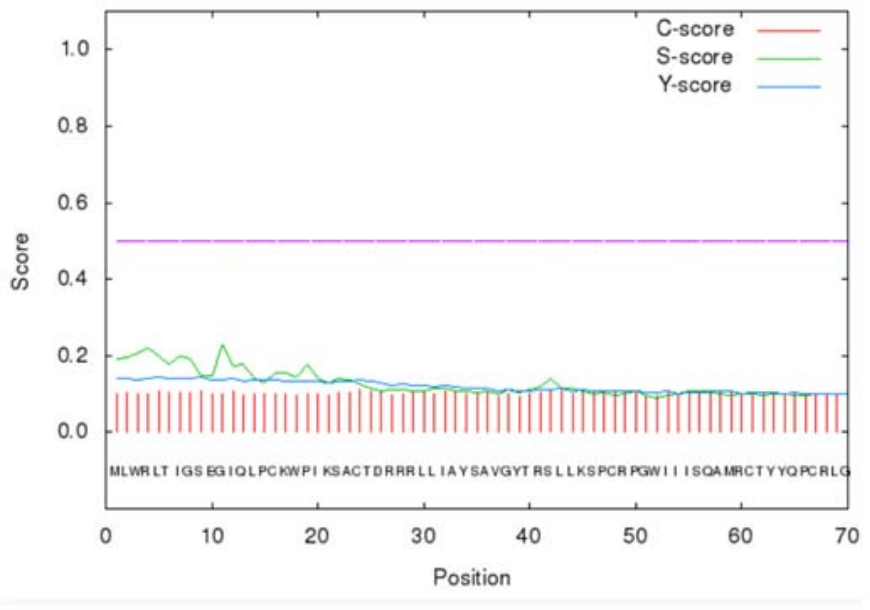

Fig. 7 SignalP-4.1 prediction. The signal peptide likelihood in the ORF5 protein sequence was found to be absent.

Additionally, it was revealed through ANTHEPROT that the ORF5 protein contained protein kinase C phosphorylation sites and N-myristoylation sites (Table 5).

Table 5 Motif regions present in the rat HEV ORF5 protein sequence.

Motifs	Number of sites	Amino acid residues
Protein kinase C phosphorylation site	3	24 - 26 139 - 141 165 - 167
N-myristoylation site	3	69 - 74 164 - 169 179 - 184

*The analysis was conducted using ANTHEPROT software.

4 Discussion

The Norway rat HEV genome consists of additional reading frames, i.e., a total of six ORFs (ORF1, ORF2, ORF3, ORF4, ORF5 and ORF5) (Johne et al., 2010). Regardless of the contribution of ORF5 protein's to rat HEV genomic component, its characteristics remain to be elucidated (Johne et al., 2010). In the present study, we determined the structural and functional characteristics of the ORF5 protein through assessing its physicochemical properties, primary structure, secondary structure, tertiary structure and functional analysis using a set of different computational methods.

The physiochemical parameters, such as molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of

hydropathicity (GRAVY) (Gasteiger et al., 2005), are imperative in decoding the uniqueness of a particular protein and therefore, these properties of ORF5 protein were computed using ProtParam (ExPasy server). The half-life is the amount of time which is taken by the protein to disappear its half of the amount after its synthesis in the cell (Bachmair et al., 1986; Gonda et al., 1989; Tobias et al., 1991; Ciechanover and Schwartz, 1989; Varshavsky, 1997). In this study, the half-life of all the proteins was 30 h. Instability index is another crucial factor which estimates the protein's stability in a test tube. A protein having instability index value is above the value of 40 is predicted as unstable, while a value smaller than 40 predicts it as stable. Our higher instability index (> 60) value indicated the unstable nature of the ORF5 protein (Guruprasad et al., 1990). The aliphatic index is another important factor that governs the protein nature (Ikai, 1980). It is defined as the relative volume occupied by aliphatic side chains (Ala, Val, Iso and Leu). The positive aliphatic index score is directly proportional to the increased thermostability of the globular proteins, i.e., proteins having higher aliphatic indices are comparatively more thermally stable in comparison to proteins having lesser aliphatic indices (Ikai, 1980). Thus, high aliphatic index value indicated high thermostability possessed by the ORF5 proteins due to the presence of some amino acids having aliphatic side chains (Ala, Val, Ile and Leu). Additionally, GRAVY is considered as an important factor for protein in determining its physiochemical properties (Kyte and Doolittle, 1982). The negative GRAVY score indicated the hydrophiliic nature of the ORF5 protein (as value between - 0.310 and - 0.514 and lower values have been shown to have good interaction with water indicating the hydrophilic nature of the given protein) (Kyte and Doolittle, 1982). The amino acid composition revealed the higher abundance of the amino acids Leu, Arg, Ile, Pro and Ser/Thr/Gly in the ORF5 proteins. Leu is categorized into the group "regulatory" as this group consists of eight most potent amino acids, such as Tyr, Phe, Gln, Pro, His, Trp, and Met (Mortimore and Pösö, 1987; Garlick, 2005). Charged amino acids (such as, Arg) contribute to the process of ligand binding (Claverie and Notredame, 2007). The amino acid Pro has also been documented in important functional and structural implications. The Pro performs significant roles like, molecular recognition, intracellular signaling and essential signaling cascades (Kay et al., 2000). Though amino acid Ser is recognized as nutritionally nonessential (dispensable), however, it plays crucial role in different cellular processes (Kalhan and Hanson, 2012). Thus, our amino acid compositional analysis suggested ORF5's role in various regulatory functions due to the significant percentages of these observed amino acid residues (Leu, Pro and Ser).

After the initial compositional analysis, structural analysis was carried out to predict the different structural levels of the ORF5 protein. The secondary structure showed all the three major components which included α-helix, β-strand and random coil. The amino acids structural diversity plays a vital role in the formation of protein self-assembly. The three-dimensional spatial arrangement of amino acid residues in a protein is known as the tertiary structure. To perform structure-based drug-designing, it is quite essential to build a reliable model. The modeled 3D structure showed the higher percentage of β-strand with predominance of coils. Taken together, our analysis revealed mixed α/β structural-fold of the ORF5 protein. It has been suggested that the presence of clefts, tunnels and pores in proteins, accessible to ligand molecules, is essential in the context of structure-based drug design process (Mbarek et al., 2019; Marques et al., 2017). Clefts present on protein's surface are sizeable depressions and are important in determining the protein interaction with other molecules as they possess tendency to be enzyme active sites (Coleman and Sharp, 2006). The size of clefts is considered as primary factors in governing the interaction between the receptor protein with the target molecules (Coleman and Sharp, 2006). Tunnels are defined as access paths which connects the interior of the protein molecule to the surrounding environment. Furthermore, tunnels influence the reactivity of the protein and determine the interaction nature and intensity (Brezovsky et al., 2018). Thus, we carried out the scrutinization for the presence of clefts, tunnels and pores in the ORF5 protein.

Interestingly, the modelled 3D structure was identified with several clefts and tunnels in addition to few pores, which clearly suggests the presence of binding sites on ORF5 protein. Thus, the presence of clefts and tunnel also strengthens our analysis, revealing the interactive ability of ORF5 protein with other ligand molecules.

Furthermore, the ORF5 protein model was predicted with some post-translationally modified sites. Post-translational modifications (PTMs) are various different type of modifications such as, phosphorylation, glycosylation, ubiquitnation, acetylation, etc, and known to contribute to cellular signal transduction regulation, transcription and translation (Keck et al., 2015; Duan and Walther, 2015). Presence of glycosylation has been shown to modulate the intracellular signaling machinery (Arey, 2012). Also, phosphorylation constitutes an essential mechanism for the proper establishment of an infection cycle in several intracellular pathogens (Marks, 1996; Zor et al., 2002). Phosphorylation is required for protein folding, signal transduction, intracellular localization PPIs, transcription regulation, cell cycle progression, survival and apoptosis (Vihinen et al., 2001; Li et al., 1990; Keck et al., 2015). Previous reports have suggested that the attachment of a myristoyl group regulates cellular signaling pathways in several biological processes (Udenwobele et al., 2017). Thus, these identified modified sites further substantiate our present hypothesis. Taken together, it can be interpreted from our findings that the ORF5 protein performs crucial functions by interacting with the other viral and host components, thus signifying its essentiality in rat HEV pathogenesis.

5 Conclusions

The Norway rat HEV ORF5 encoded protein is an essential component of its genome with unknown function. We document the physicochemical, structural and functional characteristics of the ORF5 encoded protein of Norway rat HEV using standard bioinformatics tools. Our analysis revealed that the ORF5 protein was unstable, thermostable, hydrophilic and basic in nature. The secondary structural analysis revealed the presence of all three major components (helices, strands and coils). The modeled 3D structure showed the presence of mixed α/β structural fold with the predominance of coils. The presence of several clefts and tunnels in addition to modified sites, such as glycosylation, phosphorylation and myristoylation, further signifies the essentiality of ORF5 protein in the pathogenesis of rat HEV. Knowledge on the structure of the ORF5 protein will provide insights into its functional role in the viral pathogenesis. This theoretical knowledge could further assist in exploring the characteristics of ORF5 protein of rat HEV. Furthermore, thorough experimental confirmations of these analyses are envisaged towards better understanding of the biology of HEV.

Acknowledgment

The authors would like acknowledge Maulana Azad National Fellowship (MANF) supported by the Government of India.

References

Ansari IH, Nanda SK, Durgapal H, Agrawal S, Mohanty SK, Gupta D, Jameel S, Panda SK. 2000. Cloning, sequencing, and expression of the hepatitis E virus (HEV) nonstructural open reading frame 1 (ORF1). Journal of Medical Virology, 60(3): 275-283

Arey BJ. 2012. The role of glycosylation in receptor signaling. Glycosylation, 26(10): 50262

Bachmair A, Finley D, Varshavsky A. 1986. In vivo half-life of a protein is a function of its amino-terminal

residue. Science, 234: 179-186

Brezovsky J, Kozlikova B, Damborsky J. 2018. Computational Analysis of Protein Tunnels and Channels. 25-42, Humana Press, New York, NY, USA

Ciechanover A and Schwartz AL. 1989. How are substrates recognized by the ubiquitin-mediated proteolytic system? Trends in Biochemical Sciences, 14: 483-488

Claverie JM, Notredame C. 2007. Bioinformatics for Dummies (2nd ed). Wiley Publishing, New York, USA

Coleman RG, Sharp KA. 2006. Travel depth, a new shape descriptor for macromolecules: application to ligand binding. Journal of Molecular Biology, l 362(3): 441-458

Duan G, Walther D. 2015. The roles of post-translational modifcations in the context of protein interaction networks. PLoS Computational Biology, 11(2): e1004049

Garlick PJ. 2005. The role of leucine in the regulation of protein metabolism. Journal of Nutrition, 135(6): 1553S-1556S

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. 2005. Protein identification and analysis tools on the ExPASy Server. 571-607, The Proteomics Protocols Handbook. Humana Press, USA

Gill SC, Hippel PHV. 1989. Calculation of protein extinction coefficient from amino acid sequence data. Analytical Biochemistry, 182: 319-326

Gonda DK, Bachmair A, Wunning I, Tobias JW, Lane WS, Varshavsky A. 1989. A Universality and structure of the N-end rule. Journal of Biological Chemistry, 264: 16700-16712

Guruprasad K, Reddy BVB, Pandit MW. 1990. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. Protein Engineering, Design and Selection, 4: 155-161

Ikai A. 1980. Thermostability and aliphatic index of globular proteins. The Journal of Biochemistry, 88: 1895-1898

Johne R, Heckel G, Plenge-Bönig A, Kindler E, Maresch C, Reetz J, Schielke A, Ulrich RG. 2010. Novel hepatitis E virus genotype in Norway rats, Germany. Emerging Infectious Diseases, 16(9): 1452

Kalhan SC, Hanson RW. 2012. Resurgence of serine: an often neglected but indispensable amino Acid. Journal of Biological Chemistry, 287(24): 19786-19791

Kay BK, Williamson MP, Sudol M. 2000. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. The FASEB Journal, 14(2): 231-241

Keck F, Ataey P, Amaya M, Bailey C, Narayanan A. 2015. Phosphorylation of single stranded RNA virus proteins and potential for novel therapeutic strategies. Viruses, 7(10): 5257-5273

Kenney SP, Meng XJ. 2019. Hepatitis E virus genome structure and replication strategy. Cold Spring Harbor Perspectives in Medicine, 9(1): a031724

Khuroo MS, Khuroo MS. 2016. Hepatitis E: an emerging global disease–from discovery towards control and cure. Journal of Viral Hepatitis, 23(2): 68-79

Kumar S, Subhadra S, Singh B, Panda BK. 2013. Hepatitis E virus: the current scenario. International Journal of Infectious Diseases, 17(4): e228-e233

Kyte J, Doolittle RF. 1982. A simple method for displaying the hydropathic character of a protein. Journal of Molecular Biology, 157: 105-132

Li G, La Starza MW, Hardy WR, Strauss JH, Rice CM. 1990. Phosphorylation of Sindbis virus nsP3 in vivo and in vitro. Virology, 179(1): 416-427

Marks F. 1996. Protein Phosphorylation. VCH, USA

Udenwobele DI, Su RC, Good SV, Ball TB, Varma Shrivastav S, Myristoylation SA. 2017. An important

protein modifcation in the immune response. Frontiers in Immunology, 8: 751

Marques SM, Daniel L, Buryska T, Prokop Z, Brezovsky J, Damborsky J. 2017. Enzyme tunnels and gates as relevant targets in drug design. Medicinal Research Reviews, 37(5): 1095-1139

Mbarek A, Moussa G, Leblond Chain J. 2019. Pharmaceutical applications of molecular tweezers, clefts and clips. Molecules, 24(9): 1803

Mortimore GE, Pösö AR. 1987. Intracellular protein catabolism and its control during nutrient deprivation and supply. Annual Review of Nutrition, 7(1): 539-568

Parvez MK. 2017. The hepatitis E virus nonstructural polyprotein. Future Microbiology, 12(10): 915-924

Ropp SL, Tam AW, Beames B, Purdy M, Frey TK. 2000. Expression of the hepatitis E virus ORF1. Archives of Virology, 145(7): 1321-1337

Shafat Z, Ahmed A, Parvez MK, Parveen S. 2021. Role of "dual-personality" fragments in HEV adaptation-analysis of Y-domain region. Journal of Genetic Engineering and Biotechnology, 19: 154

Shafat Z, Ahmed A, Parvez MK, Parveen Sb. 2021. Role of ORF4 in Hepatitis E virus regulation: analysis of intrinsically disordered regions. Journal of Proteins and Proteomics, 12: 289-306

Shafat Z, Ahmed A, Parvez MK, Parveen Sc. 2021. Sequence to structure analysis of the ORF4 protein from Hepatitis E virus. Bioinformation, 17(9): 818-828

Suppiah S, Zhou Y, Frey TK. 2011. Lack of processing of the expressed ORF1 gene product of hepatitis E virus. Virology Journal, 8(1): 1-5

Takahashi M, Tanaka T, Takahashi H, Hoshino Y, Nagashima S, Mizuo H, Yazaki Y, Takagi T, Azuma M, Kusano E, Isoda N. 2010. Hepatitis E Virus (HEV) strains in serum samples can replicate efficiently in cultured cells despite the coexistence of HEV antibodies: characterization of HEV virions in blood circulation. Journal of Clinical Microbiology, 48(4): 1112-1125

Tam AW, Smith MM, Guerra ME, Huang CC, Bradley DW, Fry KE, Reyes GR. 1991. Hepatitis E virus (HEV): molecular cloning and sequencing of the full-length viral genome. Virology, 185(1): 120-131

Tobias JW, Shrader TE, Rocap G, Varshavsky A. 1991. The N-end rule in bacteria. Science, 254, 1374-1377

Varshavsky A. 1997. The N-end rule pathway of protein degradation. Genes to Cells, 2(1): 13-28

Vihinen H, Ahola T, Tuittila M, Merits A, Kääriäinen L. 2001. Elimination of phosphorylation sites of Semliki Forest virus replicase protein nsP3. Journal of Biological Chemistry, 276(8): 5745-5752

Zor T, Mayr BM, Dyson HJ, Montminy MR, Wright PE. 2002. Roles of phosphorylation and helix propensity in the binding of the KIX domain of CREB-binding protein by constitutive (c-Myb) and inducible (CREB) activators. Journal of Biological Chemistry, 277(44): 42241-42248