Article

Open Reading Frame 4 protein as potential drug target for HEV: Structural evaluation through computational approaches

Zoya Shafat, **Shama Parveen**
Centre for Interdisciplinary Research in Basic Sciences, Jamia Millia Islamia, New Delhi, India
E-mail: zoya179695@st.jmi.ac.in, sparveen2@jmi.ac.in

Received 21 May 2023; *Accepted 28 June 2023*; *Published online 20 July 2023*; *Published 1 March 2024*

Abstract
Hepatitis E virus (HEV) is the main cause of acute hepatitis worldwide. The viral infection caused by G1 HEV in pregnant women has become a major health concern in the past few years. The mechanism underlying the pathogenesis of viral infection in HEV G1 isolates is attributed to four different open-reading frames (ORFs) i.e., ORF1, ORF2, ORF3 and ORF4. The present analysis has considered ORF4 protein as the molecular target due to its intrinsic disorder propensity. Intrinsically disordered regions (IDRs) are regions in proteins that do not possess stable secondary and tertiary structure and are prevalent in eukaryotes. IDRs are found to be closely associated with numerous human diseases, for instance, Parkinson and Alzheimer disease. The extreme flexibility and random coiled conformations of IDR allow it to undergo protein-protein interaction (PPI). The 3-dimensional (3D) structures of the target protein were designed using homology modelling algorithms. The generated models were assessed through structure verification tool PROCHECK. In this paper, we provide an overview of ORF4 protein structure–function relationship and its involvement in several biological processes through PPIs. Our results suggest that ORF4 protein has the potential to act as drug molecule, thus can accelerate the process of drug designing strategies against HEV.

Keywords Open Reading Frame 4; intrinsically disordered region; drug target; secondary structure; 3D model.

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 virus (HEV) is the major aetiological agent of Hepatitis E, also called enteric hepatitis (enteric means related to the intestines) infection. Worldwide, about 20 million HEV infections and 3.3 million symptomatic hepatitis E cases occur annually which results in 44,000 deaths (Hoofnagle et al., 2012; Kamar et al., 2017; Wedemeyer et al., 2012). HEV is a quasi-enveloped *Orthohepevirus*, with a single-strand, positive-sense RNA genome of around 7.2 kb in length and flanked with short 5′ and 3’ non-coding regions (NCR) (Meng, 2008; Mushahwar, 2008). Recently, a novel reading frame ORF4 has been identified in G1 of HEV entirely embedded within ORF1 in a different reading frame (Nair et al., 2016; Subramani et al., 2018).

It has been evidenced that mutation from leucine to proline in ORF4 protein occurs in fulminant hepatic failure (FHF) and acute hepatitis patients suggesting that HEV can produce proteasome-resistant ORF4 which may contribute to negative patient outcomes (Nair et al., 2016).

The major proportion of drug molecules targets the surface of natural substrate where most of its interactions and contacts can be imitated in order to control the protein function (Imming et al., 2006). For the past few years, much advancement has been observed in the field of protein-protein interactions (Wells and McClendon, 2007; Yin and Hamilton, 2005). Targeting protein-protein interactions is based on the fact that the energy is concentrated over the small contact area (area which is better contacted by drug molecule) rather than its distribution over the large surface region (Clackson and Wells, 1995). These interactions have emerged as viable systems for inhibition by small molecules and as potential drug targets. Intrinsically disordered regions (IDRs) in proteins lack well-defined stable structure in important regions or throughout the protein sequence. They are characterized by high flexibility and are rapidly interconverting structures which undergo coupled. Due to involvement of intrinsically disordered proteins (IDPs) in several biological processes through binding, it unlocks great potential in drug discovery (Metallo, 2010).

Till date, specific treatment against HEV strains has not been discovered. Only Hecolin, a prophylactic vaccine is licensed only in China (Haffar et al., 2015). Thus, further studies are required for the development of specific drug molecule to treat HEV infections against all strains. The HEV ORF4 protein has recently been linked to IDP (Shafat et al., 2021a; 2021b). Thus, targeting the ORF4 protein is ideal for devising treatment against the HEV. IDPs are often frequently associated with the progression of disease, and they constitute druggable-targets (Ruan et al., 2019; Nonell-Canals and Sanchez-Martinez, 2017; Neira et al., 2017; Santofimia-Castaño et al., 2020). In this article, we conducted computational analysis to provide insights into the structural characteristics of ORF4 region. The modelled structures were built through homology modelling approach and analyzed. The generated ORF4 models were validated using PROCHECK. Further, the models were scrutinized for various binding sites and intrinsic disorder content to recognize suitable drug target molecule against HEV.

2 Material and Methods

2.1 Retrieval of protein sequences

The ORF4 protein sequences with accession numbers, LC057248, KU168733, JN167538 and LC177791, were retrieved from NCBI (National Centre for Biotechnology Information) (GenBank Overview (nih.gov). These sequences were used as an input for further modelling analysis.

2.2 Structure prediction

To characterize a protein sequence in terms of its function remains one of the major problems encountered in biology. Due to the absence of an experimentally determined 3D structure of HEV ORF4 protein in protein data bank (PDB), we generated the ORF4 protein models through homology modelling approach, using the online available I-TASSER (Iterative Threading ASSEmbly Refinement) server (Zheng et al., 2021; Yang et al., 2015; Roy et al., 2010). The obtained models were further validated for drug target against HEV.

2.3 Structure validation

The generated 3D structure models of the HEV ORF4 were validated using PROCHECK (http://nihserver.mbi.ucla.edu/SAVES). PROCHECK is an efficient tool for assessing the quality of the predicted 3D model of the target protein molecule. The stereo-chemical property of the protein was evaluated by Ramachandran plot analysis. In addition, the amino acid residues in the allowed and disallowed regions including the overall G-factor were also analysed.

2.4 Intrinsic disorder evaluation

The DisProt (Database of Protein Disorder) webserver was employed for the prediction of intrinsic disorder in ORF4 sequences, using its default settings (Sickmeier et al., 2007). The protein residues which had predicted scores, between 0.2 and 0.5, were considered as flexible, while the residues with predicted scores, exceeding the 0.5 threshold value, were predicted as intrinsically disordered ones. The disorder variants were identified using the suggested criteria (Rajagopalan et al., 2011).

3 Results

Due to suggested indispensability of the ORF4 in replication and pathogenesis of HEV (Nair et al., 2016; Subramani et al., 2018; Shafat et al., 2021a), the ORF4 protein was evaluated as a drug target. In this context, the derived ORF4 protein sequences were used as query sequences to evaluate different structural and functional properties, using different *in-silico* approaches.

3.1 Homology modelled structure

The obtained tertiary structures for the target ORF4 protein were analysed by visualization through homology modelling algorithms (Fig. 1). The identified secondary structure components utilizing the generated modelled 3D ORF4 structures (through I-TASSER) using PDBsum analysis is illustrated in Fig. 2.

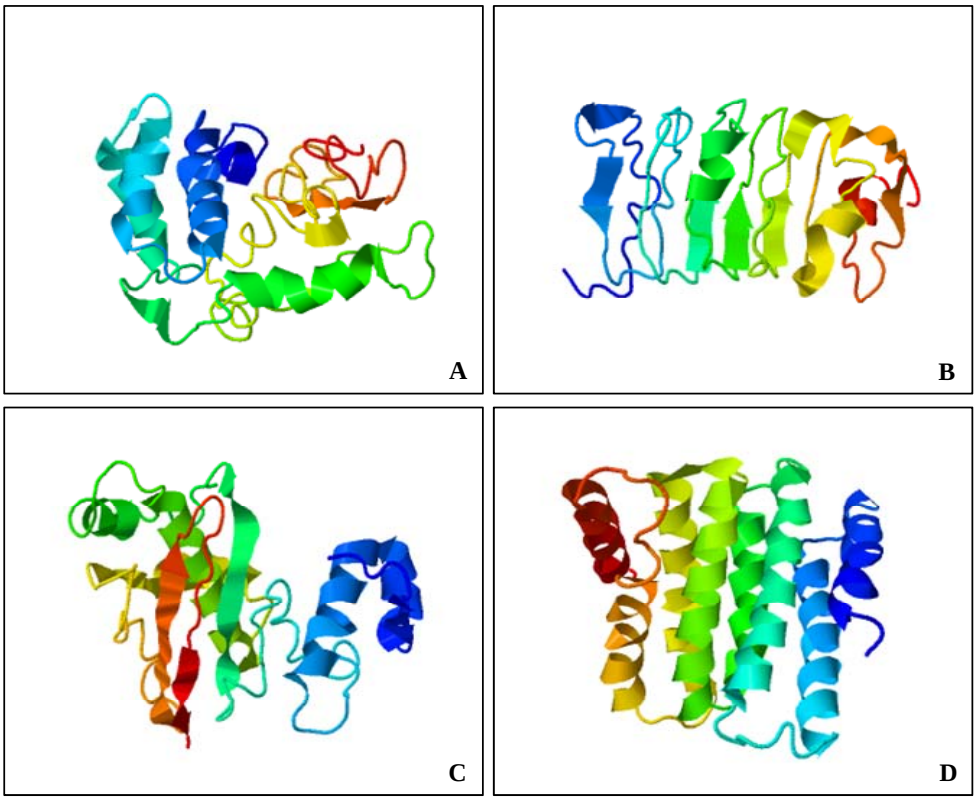

Fig. 1 Homology modelled three dimensional structure prediction of HEV ORF4. (A) LC057248, (B) KU168733, (C) JN167538 and (D) LC177791. The tertiary structure evaluation was conducted using I-TASSER webserver.

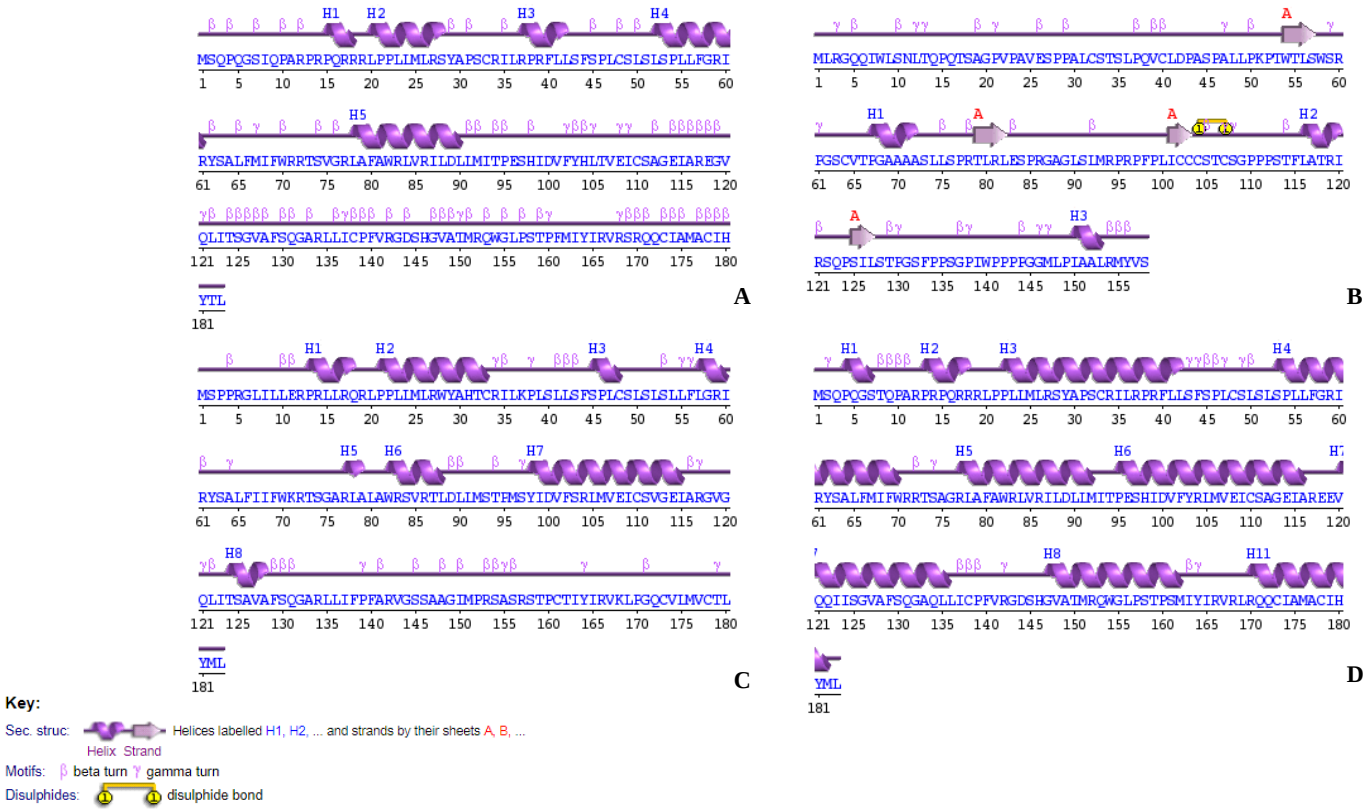

Fig. 2 Secondary structure elements identification in HEV ORF4 structured models. (A) LC057248, (B) KU168733, (C) JN167538 and (D) LC177791. The analysis was conducted using PDBsum webserver.

Comparative modelling among ORF4 protein sequences was undertaken to identify the best ORF4 model as a drug target. The details of the generated ORF4 models (by I-TASSER) are listed in the table (Table 1).

Table 1 Different statistics of the obtained ORF4 3D model by I-TASSER.

	LC057248	**KU168733**	**JN167538**	**LC177791**
Template	61btA	5wc0A	2ej0A	1eqfA
C-score	-4.36	-3.29	-3.41	-2.84
Estimated TM-score	0.26±0.08	0.35±0.12	0.36±0.12	0.39±0.13
Estimated RMSD (Å)	15.8±3.2	12.5±4.3	12.5±4.3	11.7±4.5

C-score represents a confidence score that estimates the quality of the predicted model by I-TASSER. The range is [0-1] and higher values indicate more confident predictions (Zheng et al., 2021; Yang et al., 2015; Roy et al., 2010). Its calculation is based on the significance of threading template alignments and the convergence parameters of the structure assembly simulations. C-score typically ranges between -5 and 2, wherein a C-score score having higher value signifies a higher confidence model and vice-versa (Zheng et al., 2021; Yang et al., 2015; Roy et al., 2010). TM-score and RMSD are known standards which measure structural similarity between two structures which are usually used to measure the accuracy of structure modelling when

the native structure is known (Zheng et al., 2021; Yang et al., 2015; Roy et al., 2010).

3.2 Validation of predicted 3D structure

The predicted 3D ORF4 models (generated through I-TASSER) were further assessed for Ramachandran Plot statistics through the PDBsum analysis. The overall protein's stereochemical quality, amino acid residues present in the favoured, additionally allowed, generously allowed and disallowed region including the G-factor were evaluated by PROCHECK statistics. The Ramachandran plots of ORF4 models showing the favorable, allowed and disallowed regions are illustrated in Fig. 3. The obtained information of the models is summarized in Table 2.

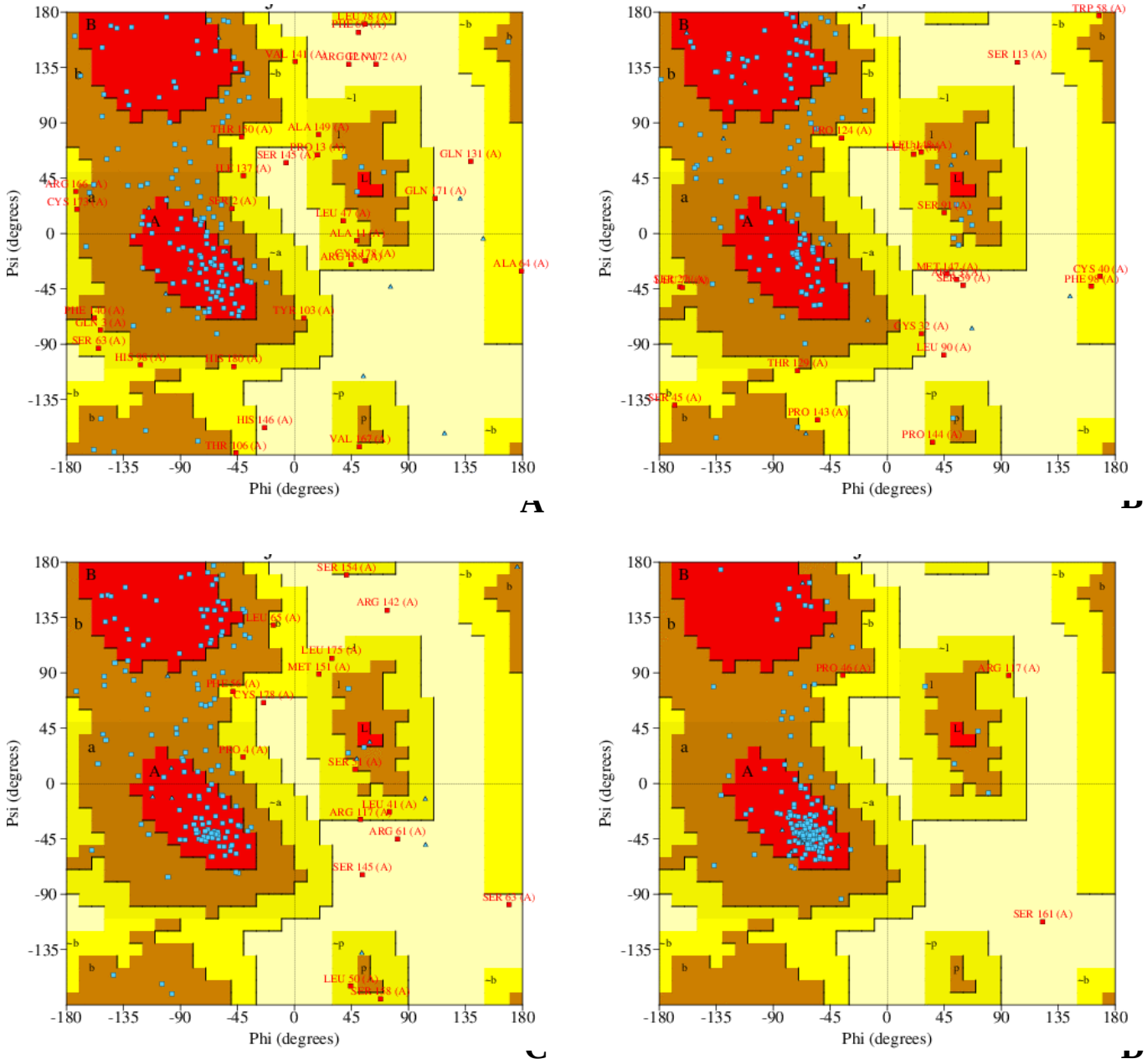

Fig. 3 Generated Ramachandran plots of HEV ORF4. (A) LC057248, (B) KU168733, (C) JN167538 and (D) LC177791. The plots were generated using PROCHECK statistics via PDBsum analysis.

Table 2 Custom-generated PDBsum page analysis of the ORF4 3D model.

	LC057248	**KU168733**	**JN167538**	**LC177791**
PDBsum analysis				
Clefts	10	10	10	10
Tunnels	1	4	3	4
Pores	-	2	-	1
ProMotif				
Sheet	--	1	--	--
Beta hairpins	--	--	--	--
Beta alpha beta units	--	2	--	--
Strands	--	4	--	--
Helices	5	3	8	11
Helix-helix interacs	6	1	7	23
Beta turns	63	22	25	12
Gamma turns	13	14	16	8
Disulphide	--	1	--	--
PROCHECK Statistics				
Ramachandran Plot statistics				
Most favoured regions (%)	44.6	51.3	55.3	82.9
Additional allowed regions (%)	37.6	34.8	35.2	15.8
Generously allowed regions (%)	13.4	9.6	5.7	0.6
Disallowed regions	4.5	4.3	3.8	0.6
G-Factors				
Overall average	-1.41****	-0.90*	-0.76*	-0.34

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.

All the ORF4 structured models were identified with several clefts with few tunnels and pores which suggested their interaction potential towards other molecules. Further, the models showed different type of promotifs, such as, helices, beta turn, gamma turn, etc.

3.3 ORF4 protein as a drug target

3.3.1 Quantifying disorder by calculating the predicted percentage of disordered residues

The HEV ORF4 protein sequences were further analyzed in terms of their intrinsic disorder content. The overall predicted percentage of intrinsic disorder (PPID) of the ORF4 proteins were calculated. On the basis of overall PPID score, the protein disorder variants are classified into three categories, i.e., highly ordered protein (PPID ≥ 30%), moderately disordered proteins (≤ 10 PPID < 30%) or highly disordered proteins (PPID ≥ 30%) (Rajagopalan et al., 2011).

3.3.2 Disorder analysis with DisProt

The predisposition for intrinsic disorder in HEV ORF4 proteins was evaluated using DisProt. Scores > 0.5 corresponded to disordered residues, wherein, green color was used to depict the disordered regions, i.e., areas in green are the predicted disordered protein regions in ORF4 proteins by DispProt (Sickmeier et al., 2007). The predicted intrinsic disorder profiles of HEV ORF4 polypeptides, generated from DisProt predictor, are represented in Figure 4. The DisProt analysis identified KU168733 with a PPID ≥30%, thus categorizing it as a highly disordered protein or intrinsically disordered protein (IDP) (Deiana et al., 2019; Van der Lee et al., 2014). Further, this is consistent with the different parameters of the obtained ORF4 3D models, that suggested the ORF4 protein (Accession number: KU168733), obtained from HEV-host Human, was identified as IDP as shown in the figure (Fig. 2). This model scored a C-value of -3.29 with RMSD value of 12.5±4.3 Å and had a percentage favorable region of 51.3% and G factor of -0.90 as shown in Tables 4 and 5. Importantly, this overall modelled protein structure was identified with maximum number of interacting sites in form of cleft, pore and tunnel, i.e., 10 clefts (Fig. 5), 2 pores (Fig. 6) and 4 tunnels (Fig. 7). Thus, the ORF4 protein (KU168733) was considered as a suitable drug target.

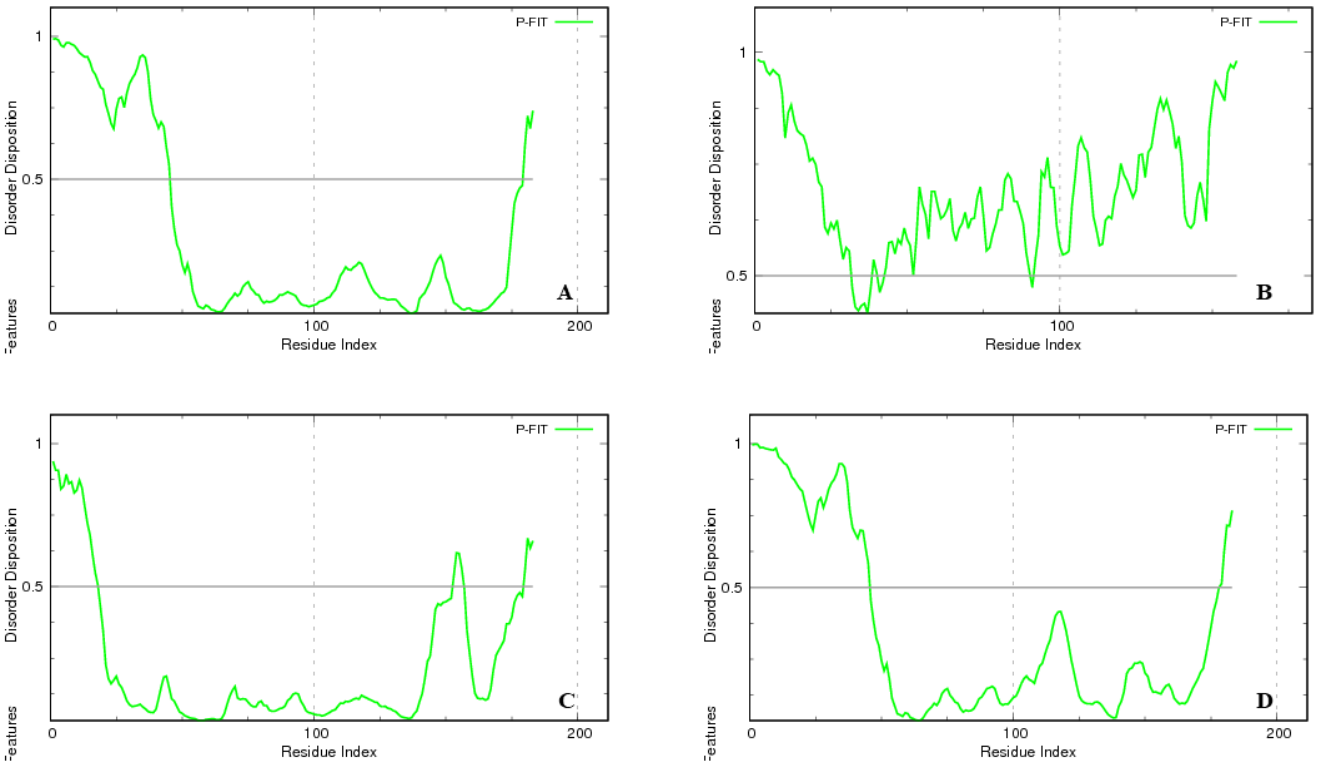

Fig. 4 Intrinsic disorder propensity analysis of HEV ORF4. (A) LC057248, (B) KU168733, (C) JN167538 and (D) LC177791. The resulting disorder profile evaluation was conducted using DisProt.

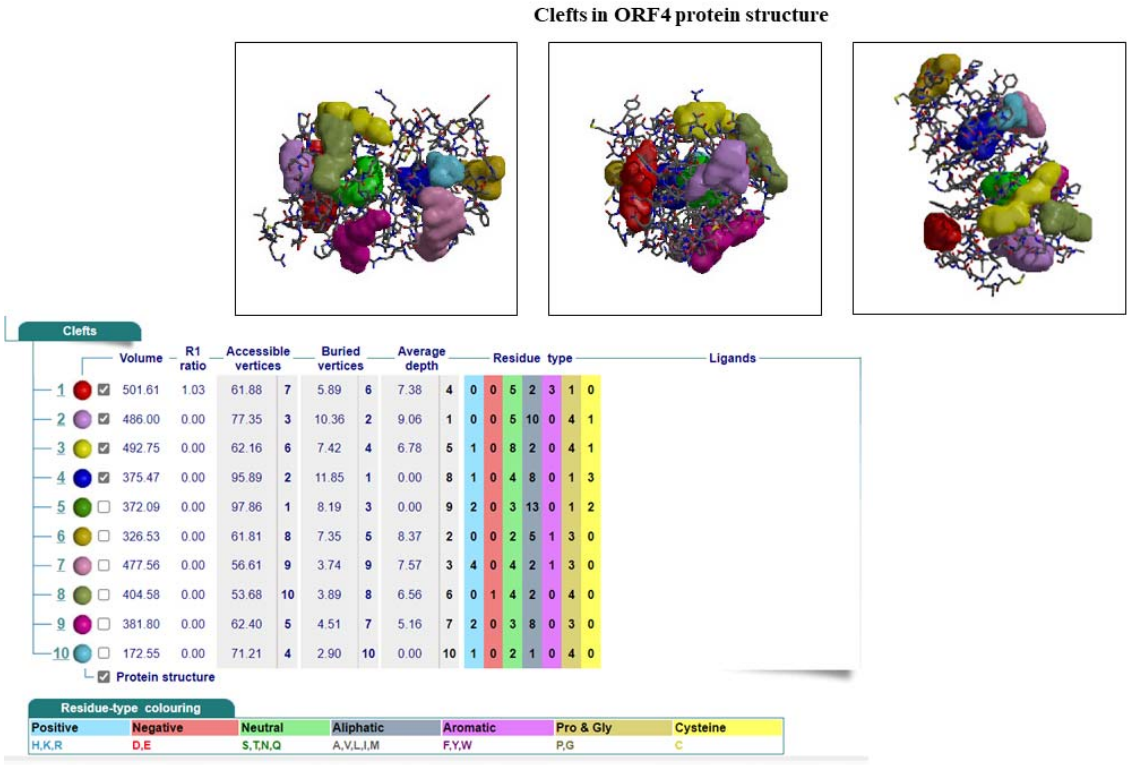

Fig. 5 Clefts calculated on whole ORF4 protein structure. (A) LC057248, (B) KU168733, (C) JN167538 and (D) LC177791. The analysis was processed via PDBsum analysis.

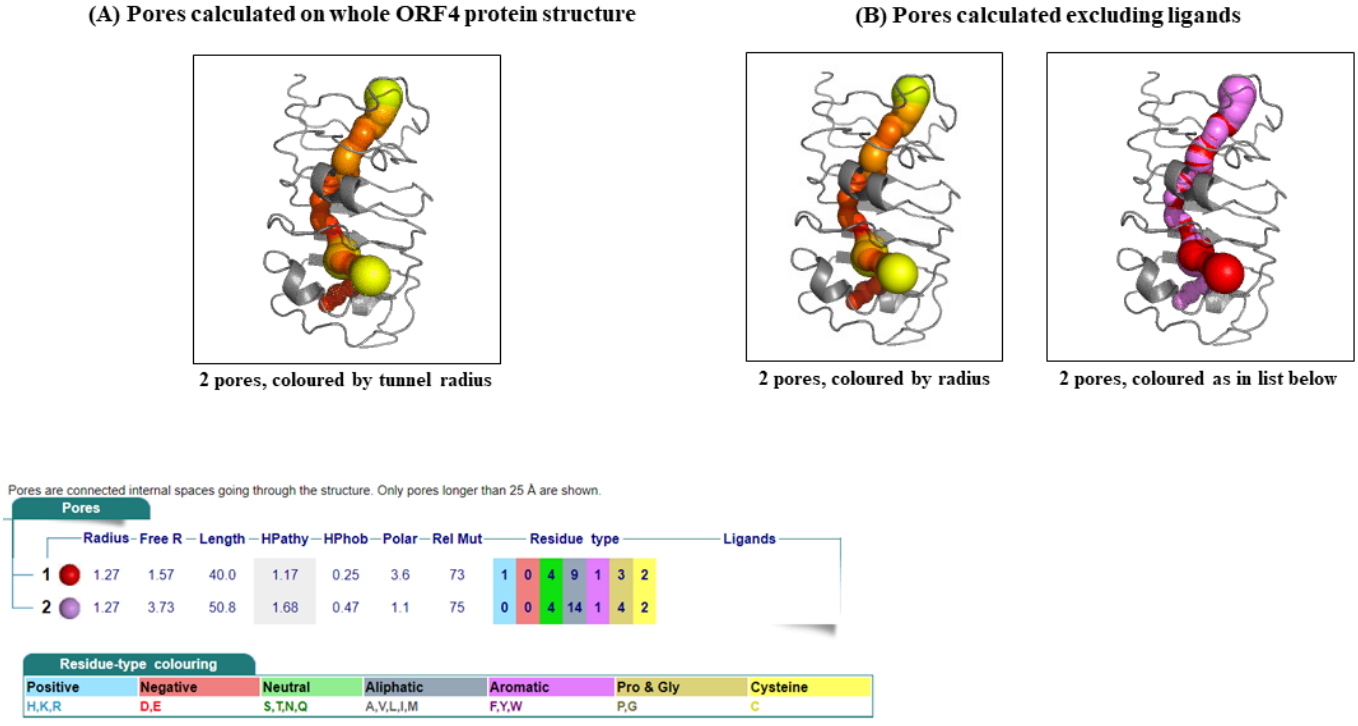

Fig. 6 Pores calculated on whole ORF4 protein structure. (A) LC057248, (B) KU168733, (C) JN167538 and (D) LC177791. The analysis was processed via PDBsum analysis.

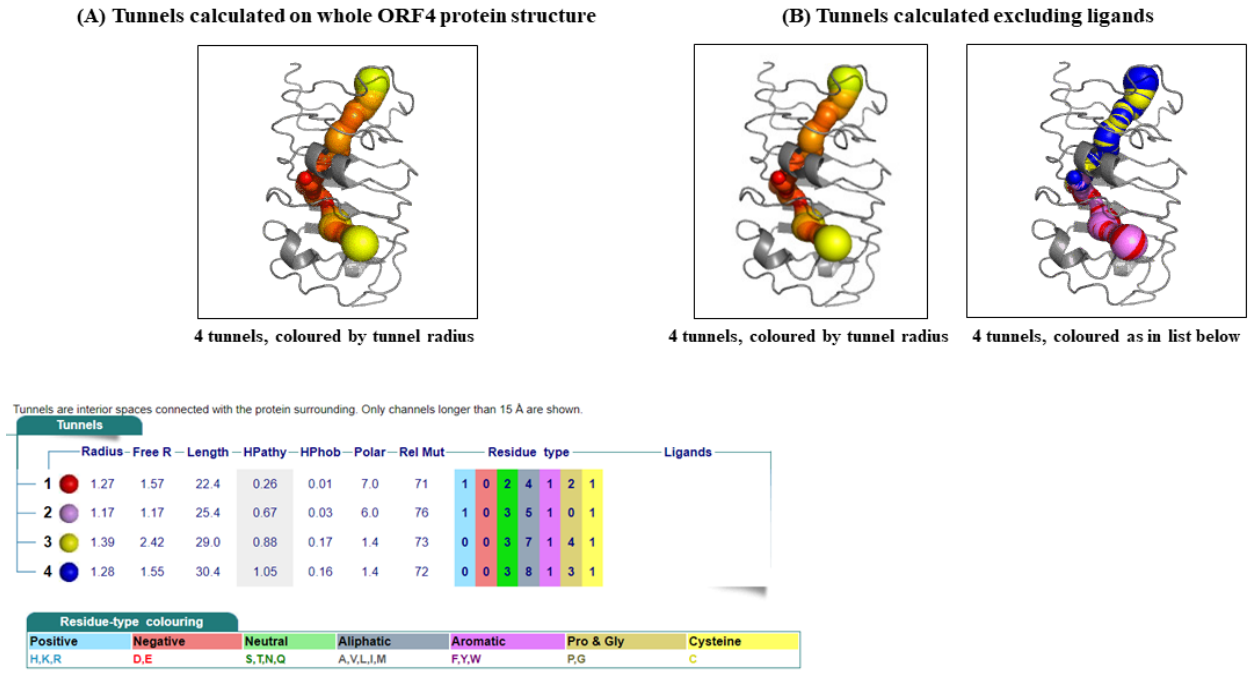

Fig. 7 Tunnels calculated on whole ORF4 protein structure. (A) LC057248, (B) KU168733, (C) JN167538 and (D) LC177791. The analysis was processed via PDBsum analysis.

4 Discussion

IDPs constitute the protein group that challenges the paradigm of a protein's folded structure. Most of the biologically active proteins possess definite 3D structures, in order to perform their specific function, however, IDPs does not follow this fundamental law of 'sequence-structure-function' (Uversky, 2013; Tompa, 2012). IDPs possess structural plasticity and thus perform various essential biological functions, such as, cell signalling and protein-protein interaction (Dyson, 2011), as they are characterized by lack of homogeneous 3D structures (Van Der Lee et al., 2014). IDR is defined by the term "disorder in disorders", to highlight the occurrence of disordered protein regions within the protein genes (Uversky et al., 2008). These disordered region in proteins are characterised by an important feature known as coupled binding and folding (Wright and Dyson, 2009). Thus, ID proteins are associated with large number of diseases, such as cancer, diabetes, neurodegenerative disease and cardiovascular disease. Many ID proteins become structured upon binding to a partner with the energy from specific interactions compensating for the entropic penalty from ordering (Dyson and Wright, 2005). As IDPs are frequently associated with the pathways or progression of human diseases they are considered as 'druggable-targets' (Ruan et al., 2019; Nonell-Canals and Sanchez-Martinez, 2017; Neira et al., 2017; Santofimia-Castaño et al., 2020; Fuertes et al., 2019).

The ORF4 component has recently been linked to regulatory roles in addition to its essentiality in replication of HEV (Nair et al., 2016; Subramani et al., 2018; Shafat et al., 2021c). Additionally, the ORF4 structure has recently been associated with intrinsically disordered protein region (IDPR) or IDP, therefore, it has application in drug designing strategies (Shafat et al., 2021a; 2021b; Metallo, 2019). Thus, targeting the ORF4 protein is ideal for devising treatment against the HEV. In this context, this chapter reports the structural analysis of the ORF4 proteins of HEV by employing bioinformatics approach. The ORF4-based data provided in the GenBank of NCBI has prompted us to predict its three-dimensional structure model from amino acid sequence. The present data will shed light on ORF4 implications in drug/vaccine development against HEV.

The secondary structure elements (helices and strands) are organized in different 3D spatial arrangement to form a tertiary structure of a protein. To perform structure-based drug-designing, it is quite essential to build a reliable model. The 3D structure model of the ORF4 protein was hypothesized using homology modelling approach. The homology modelling accurately predicts the structure of a given sequence theoretically which can be compared to the best results attained experimentally. The modelled structures of the HEV ORF4 proteins were obtained using an on-line platform, i.e., I-TASSER (Zheng et al., 2021; Yang et al., 2015). The I-TASSER server automatically generates a high-quality 3D structure model and biological function prediction of proteins using their amino acid sequences. This server implements I-TASSER based algorithms to predict protein structures and functions (Zheng et al., 2021; Yang et al., 2015). As ID proteins are considered as potential targets for drugs due to rapidly fluctuating structures (Metallo, 2010), therefore, the predisposition of intrinsic disorder of ORF4 proteins were analyzed. Further, as suggested, on the basis of PPID, the ORF4 protein sequences were grouped into three categories of protein disorder variants, i.e., structured protein, moderately disordered protein and highly disordered protein (Rajagopalan et al., 2011). In addition to this, the ORF4 proteins were categorically grouped into ordered protein (ORDP), IDPR and IDP (Deiana et al., 2019; Van der Lee et al., 2014). The identification of protein disorder variant in ORF4 for was carried out using DisProt (Sickmeier et al., 2007). DisProt, a database, gives information about those proteins which lack definite 3D structures in their native states (either entirety or in part) (Sickmeier et al., 2007). On implying PPID criterion, among the ORF4 proteins, the ORF4 protein (KU168733), i.e., (obtained from hot Human) was identified as a highly disordered protein as it consisted of highest fraction of intrinsic disorder in comparison to other ORF4 proteins (having accession numbers: LC057248 and JN167538 and LC177791). Our results show consistency with the previous report suggesting IDPs fail to arrange into a definite 3D structure under physiological conditions due to increased level of disordered-promoting residues (Shafat et al., 2021a, 2021b). Thus, out of several models, the obtained model from host Human can be considered as a reliable drug target due to its characteristic highly disordered (IDP) structure (Nonell-Canals and Sanchez-Martinez, 2017; Metallo, 2010). Identification of a cleft, pore or a tunnel accessible to ligand molecule is essential in the perspective of structure-based drug design process (Mbarek et al., 2019; Marques et al., 2017). Thus, the ORF4 modelled structures were scrutinized using PDBsum analysis to reveal the presence of varied binding sites. Interestingly, the ORF4 models were observed with several clefts along with few pores and tunnels. Enzyme active sites tend to be within sizeable depressions on the protein's surface, which are known as clefts or pockets. Clefts are present on protein's surface as gaps, which are essential in determining the protein interaction with other ligand molecules (Coleman and Sharp, 2006). The clefts also have tendency to form active enzyme sites through sizeable depressions (Coleman and Sharp, 2006). Our predicted 3D model revealed the presence of several clefts suggesting the commitment of ORF4 towards interaction with other target molecules. Tunnel connects interior of the protein molecule to its surrounding environment and influences the reactivity of the protein and determine the interaction nature and intensity (Brezovsky et al., 2018). Interestingly, it is important to mention that the maximum number of ligand accession sites, in terms of clefts, tunnels and pores, were identified in the ORF4 protein KU168733. Thus, the ORF4 protein (KU168733) had the maximum tendency to interact with ligands as compared to other ORF4 proteins. This again substantiated our intrinsic disorder findings suggesting KU168733 ORF4 protein as a IDP variant. Thus, the presence of clefts and tunnels further strengthen our present hypothesis. Altogether, from these observations it could be interpreted that ORF4 had propensity towards interaction with other target molecules, suggesting its role in protein–protein interactions. Thus, ORF4 could probably contribute to PPI due to possession of intrinsic disorder and, thus, may act as a target molecule in drug designing strategies against G1 HEV infections.

5 Conclusions

The significant parameters obtained from the predicted 3D models can be employed in future to understand the structural and functional properties of the ORF4 protein. We should enlarge our view of what constitutes human disease research, recognizing that the discoveries that have the most profound impact on disease treatments emanate from basic research on model organisms, rather than from studies of highly complex human diseases. The studies on IDPs belong to this basic research, which could provide us novel pathways for drug discovery. However, this theoretical knowledge based on IDPs novel features is not sufficient to design new drugs. Therefore, it is significant to collect more information on IDPs including its sequence, structure, dynamics, biophysical and interaction network, by both computational and experimental approaches. Drug discovery based on these novel features may reveal a bright new path.

Acknowledgment

The authors would like to acknowledge the Maulana Azad National Fellowship (MANF), University Grant Commission (UGC) and Council of Scientific and Industrial Research (CSIR) (37(1697)17/EMR-II) supported by the Government of India.

References

Hoofnagle JH, Nelson KE, Purcell RH. 2012. Hepatitis e. New England Journal of Medicine, 367(13): 1237-1244

Kamar N, Izopet J, Pavio N, Aggarwal R, Labrique A, Wedemeyer H, Dalton HR. 2017. Hepatitis E virus infection. Nature Reviews Disease Primers, 3(1): 1-6

Wedemeyer H, Pischke S, Manns MP. 2012. Pathogenesis and treatment of hepatitis e virus infection. Gastroenterology, 142(6): 1388-1397

Meng XJ. 2008. Hepatitis E Virus. Encyclopedia of Virology (Third Edition). 377-383, Academic Press, USA

Mushahwar IK. 2008. Hepatitis E virus: molecular virology, clinical features, diagnosis, transmission, epidemiology, and prevention. Journal of Medical Virology, 80(4): 646-658

Nair VP, Anang S, Subramani C, Madhvi A, Bakshi K, Srivastava A, Shalimar, Nayak B, CT RK, Surjit M. 2016. Endoplasmic reticulum stress induced synthesis of a novel viral factor mediates efficient replication of genotype-1 hepatitis E virus. PLoS Pathogens, 12(4): e1005521

Subramani C, Nair VP, Anang S, Mandal SD, Pareek M, Kaushik N, Srivastava A, Saha S, Shalimar, Nayak B, Ranjith-Kumar CT. 2018. Host-virus protein interaction network reveals the involvement of multiple host processes in the life cycle of hepatitis E virus. MSystems, 3(1): e00135-17

Imming P, Sinning C, Meyer A. 2006. Drugs, their targets and the nature and number of drug targets. Nature Reviews Drug Discovery, 5(10): 821-834

Wells JA, McClendon CL. 2007. Reaching for high-hanging fruit in drug discovery at protein–protein interfaces. Nature, 450(7172): 1001-1009

Yin H, Hamilton AD. 2005. Strategies for targeting protein–protein interactions with synthetic agents. Angewandte Chemie International Edition, 44(27): 4130-4163

Clackson T, Wells JA. 1995. A hot spot of binding energy in a hormone-receptor interface. Science, 267(5196): 383-386

Metallo SJ. 2010. Intrinsically disordered proteins are potential drug targets. Current Opinion in Chemical Biology, 14(4): 481-488

Haffar S, Bazerbachi F, Lake JR. 2015. Making the case for the development of a vaccination against hepatitis

E virus. Liver International, 35(2): 311-316

Shafat Z, Ahmed A, Parvez MK, Parveen S. 2021a. Role of ORF4 in Hepatitis E virus regulation: analysis of intrinsically disordered regions. Journal of Proteins and Proteomics, 12(4): 289-306

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

Ruan H, Sun Q, Zhang W, Liu Y, Lai L. 2019. Targeting intrinsically disordered proteins at the edge of chaos. Drug Discovery Today, 24(1): 217-27

Nonell-Canals A, Sanchez-Martinez M. 2017. Intrinsically disordered proteins as drug targets. MOJ Proteomics and Bioinformatics, 5(2): 00157

Neira JL, Bintz J, Arruebo M, Rizzuti B, Bonacci T, Vega S, Lanas A, Velázquez-Campoy A, Iovanna JL, Abián O. 2017. Identification of a drug targeting an intrinsically disordered protein involved in pancreatic adenocarcinoma. Scientific Reports, 7(1): 1-5

Santofimia-Castaño P, Rizzuti B, Xia Y, Abian O, Peng L, Velázquez-Campoy A, Neira JL, Iovanna J. 2020. Targeting intrinsically disordered proteins involved in cancer. Cellular and Molecular Life Sciences, 77(9): 1695-1707

Zheng W, Zhang C, Li Y, Pearce R, Bell EW, Zhang Y. 2021. Folding non-homologous proteins by coupling deep-learning contact maps with I-TASSER assembly simulations. Cell Reports Methods, 1(3): 100014

Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. 2015. The I-TASSER Suite: protein structure and function prediction. Nature Methods, 12(1): 7-8

Roy A, Kucukural A, Zhang Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. Nature Protocols, 5(4): 725-738

Sickmeier M, Hamilton JA, LeGall T, Vacic V, Cortese MS, Tantos A, Szabo B, Tompa P, Chen J, Uversky VN, Obradovic Z. 2007. DisProt: the database of disordered proteins. Nucleic Acids Research, 35(suppl_1): D786-793

Rajagopalan K, Mooney SM, Parekh N, Getzenberg RH, Kulkarni P. 2011. A majority of the cancer/testis antigens are intrinsically disordered proteins. Journal of cellular biochemistry, 112(11): 3256-3267

Deiana A, Forcelloni S, Porrello A, Giansanti A. 2019. Intrinsically disordered proteins and structured proteins with intrinsically disordered regions have diferent functional roles in the cell. PLoS One, 14(8): e0217889

van der Lee R, Lang B, Kruse K, Gsponer J, de Groot NS, Huynen MA, Matouschek A, Fuxreiter M, Babu MM. 2014. Intrinsically disordered segments affect protein half-life in the cell and during evolution. Cell Reports, 8(6): 1832-1844

Uversky VN. 2013. Unusual biophysics of intrinsically disordered proteins. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 1834(5): 932-951

Tompa P. 2012. Intrinsically disordered proteins: a 10-year recap. Trends in Biochemical Sciences, 37(12): 509-516

Dyson HJ. 2011. Expanding the proteome: disordered and alternatively folded proteins. Quarterly Reviews of Biophysics, 44(4): 467-518

Van Der Lee R, Buljan M, Lang B, Weatheritt RJ, Daughdrill GW, Dunker AK, Fuxreiter M, Gough J, Gsponer J, Jones DT, Kim PM. 2014. Classification of intrinsically disordered regions and proteins. Chemical Reviews, 114(13): 6589-6631

Uversky VN, Oldfield CJ, Dunker AK. 2008. Intrinsically disordered proteins in human diseases: introducing the D2 concept. Annual Review of Biophysics, 37: 215-246

Wright PE, Dyson HJ. 2009. Linking folding and binding. Current Opinion in Structural Biology, 19(1): 31-38

Dyson HJ, Wright PE. 2005. Intrinsically unstructured proteins and their functions. Nature Reviews Molecular Cell Biology, 6(3): 197-208

Fuertes G, Nevola L, Esteban-Martín S. 2019. Perspectives on Drug Discovery Strategies Based on IDPs. Intrinsically Disordered Proteins. 275-327, Academic Press, USA

Shafat Z, Tazeen A, Ahmed M, Parvez MK, Parveen S. 2021c. Understanding Hepatitis E viruses by exploring the structural and functional properties of ORF4. Network Biology, 11(4): 274-276

Metallo SJ. 2010. Intrinsically disordered proteins are potential drug targets. Current Opinion in Chemical Biology, 14(4): 481-488

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

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

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

Brezovsky J, Kozlikova B, Damborsky J. 2018. Computational analysis of protein tunnels and channels. Methods in Molecular Biology, 1685: 25-42