Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

International Journal of Biological Macromolecules 144 (2020) 892–908

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

journal homepage: www.elsevier.com/locate/ijbiomac

Understanding the penetrance of intrinsic protein disorder in rotavirus proteome

Deepak Kumar [a,1], Ankur Singh [a,1], Prateek Kumar [a], Vladimir N. Uversky [b], C. Durga Rao [c,*], Rajanish Giri [a,d,*]

[a] Indian Institute of Technology Mandi, VPO Kamand, Himachal Pradesh 175005, India
[b] Department of Molecular Medicine and Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, FL, United States
[c] SRM University, AP – Amaravati, Neerukonda, Mangalagiri Mandal Guntur District, Mangalagiri, Andhra Pradesh 522502, India
[d] BioX Center, Indian Institute of Technology Mandi, Himachal Pradesh, India

ARTICLE INFO

Article history:
Received 2 July 2019
Received in revised form 9 September 2019
Accepted 20 September 2019
Available online 15 November 2019

Keywords:
Rotavirus
Intrinsically Disordered Proteins (IDPs)
Reoviridae
Diarrhea
MoRF

ABSTRACT

Rotavirus is a major cause of severe acute gastroenteritis in the infants and young children. The past decade has evidenced the role of intrinsically disordered proteins/regions (IDPs)/(IDPRs) in viral and other diseases. In general, (IDPs)/(IDPRs) are considered as dynamic conformational ensembles that devoid of a specific 3D structure, being associated with various important biological phenomena. Viruses utilize IDPs/IDPRs to survive in harsh environments, to evade the host immune system, and to highjack and manipulate host cellular proteins. The role of IDPs/IDPRs in Rotavirus biology and pathogenicity are not assessed so far, therefore, we have designed this study to deeply look at the penetrance of intrinsic disorder in rotavirus proteome consisting 12 proteins encoded by 11 segments of viral genome. Also, for all human rotaviral proteins, we have deciphered molecular recognition features (MoRFs), which are disorder based binding sites in proteins. Our study shows the wide spread of intrinsic disorder in several rotavirus proteins, primarily the nonstructural proteins NSP3, NSP4, and NSP5 that are involved in viral replication, translation, viroplasm formation and/or maturation. This study may serve as a primer for understanding the role of IDPs/MoRFs in rotavirus biology, design of alternative therapeutic strategies, and development of disorder-based drugs.

1. Introduction

Diarrheal diseases are a major cause of morbidity and mortality in children aged below 5 years, predominantly in low-income nations [1]. According to WHO reports, diarrheal diseases cause more than 500,000 deaths per year worldwide. Rotaviruses belong to *Reoviridae* family and are documented as the primary causative agents of acute viral diarrhea in humans, animals, and avian species [2]. Among numerous serotypes of rotavirus, only five, such as G1, G2, G3, G4, and G9, are highly infectious. G3 serotype of rotavirus was highly infectious during 1994–1995 [3]. In 1995–1996, an outbreak of G9 serotype of rotavirus occurred, and this serotype was involved in more than 50% of cases of rotavirus infections during that period. From 1996 to 1999, most of the rotaviral infections were caused by G1 and G2 serotypes, with both serotypes being accountable for nearly 50% infections [3,4]. Rotavirus is a non-enveloped RNA virus with a complex concentric triple-layered capsid that encloses the genome of 11 segments of double-stranded RNA (dsRNA) [5]. Each RNA segment codes for a single protein but the 11th segment encodes two non-structural proteins (NSP5 and NSP6) from overlapping open reading frames. As a result, the 11 segments of the viral genome encode 12 proteins, which are six non-structural proteins (NSP1, NSP2, NSP3, NSP4, NSP5, and NSP6) and six structural viral proteins (VP1, VP2, VP3, VP4, VP6, and VP7) [5] (Fig. 1).

The structural proteins of the virion govern cell entry, host specificity, antigenic specificities and enzymatic functions associated with viral genome replication and transcription. The non-structural proteins play vital roles in genome replication and evasion of innate immune response of the host [5]. Cryo-electron microscopy structure (PDB ID: 4V7Q) of rotavirus reveals the architectural organization of the triple-layered concentric capsid that surrounds the segmented RNA genome [6]. The diameter of mature rotavirion, which is characterized by the T = 13 icosahedral symmetry, is ~1,000 Å. The outer shell contains multiple copies of two proteins, VP4 & VP7, whereas the intermediate and inner

* Corresponding authors at: Indian Institute of Technology Mandi, VPO Kamand, Himachal Pradesh 175005, India.
E-mail addresses: cdrmcb@gmail.com (C.D. Rao), rajanishgiri@iitmandi.ac.in (R. Giri).
[1] DK and AS contributed equally.

https://doi.org/10.1016/j.ijbiomac.2019.09.166
0141-8130/

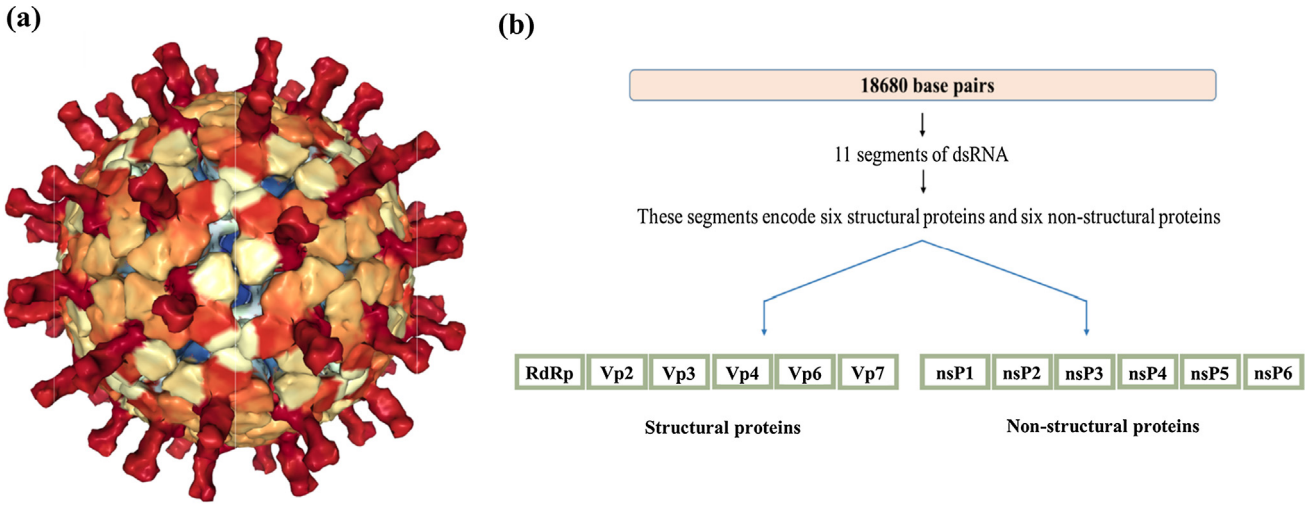

Fig. 1. **(a)** Atomic model of Rotavirus taken from PDB ID: 4V7Q [6] **(b)** Schematic representation of all structural and non-structural proteins encoded by rotavirus genome segments. Rotavirus genomic dsRNA 18,680 base pairs (top bar, light pink color), encodes six structural (green color bar) and six non-structural proteins (light orange color bar).

capsids consist of VP6 and VP2, respectively. The outer surface of rotavirus exhibits 120 Å-long 60 spikes composed of VP4. The capsid features 132 aqueous channels with ~140 Å in length, spanning two outer capsid layers [7]. During virus entry, a sequence of molecular transformations in the outer layer involving the spike protein VP4 facilitates the internalization of the virus and release of the 'double-layered particle' (DLP) in the cytoplasm with the removal of the outer capsid [6]. Removal of the outer layer results in activation of the transcriptase and initiates synthesis and extrusion of the capped mRNAs into the cytosol [5,6].

Although crystallographic structures and functions of most of the rotavirus proteins are known rather well, the intrinsically disordered aspects of these proteins have not been evaluated as of yet. Also, the crystallography provides only the static view of a protein inside crystal packing where disordered regions may undergo disorder to order transitions that ultimately hide the native flexibility and disordered propensity [8]. Therefore, our prime focus is to fill this gap by evaluating the natural content of intrinsic protein disorder inside rotavirus proteome. In general, disordered proteins (which are proteins structural features cannot be determined by the traditional structure determination methods, such as electron cryo-microscopy and X-ray crystallography) are represented by 'Intrinsically Disordered Proteins' (IDPs) and/or proteins containing 'Intrinsically Disordered Protein Regions' (IDPRs) that devoid of specific 3D structure in their native, biologically active forms [9–11]. IDPs/IDPRs have numerous crucial functions, being involved in the interaction with multiple partners and regulating many cellular processes. Recent studies implicated the direct role of IDPs/IDPRs in various human diseases, such as cancer, amyloidosis, diabetes, cardiovascular diseases, viral diseases, and many more [11]. In general, viruses, due to their limited protein-coding capacity, could gain several functional advantages from IDPs/IDPRs in terms of assembly of replication complexes involving a large number of viral and host proteins and survival in extreme host environments [12]. In recent studies, we have shown the prevalence of IDPs/IDPRs and their functional significance in Zika virus (ZIKV) and Chikungunya virus (CHIKV) [13,14], as well as in Hepatitis C virus (HCV) [15], human immunodeficiency virus-1 (HIV-1) [16], various human papillomaviruses (HPVs) [17], respiratory syncytial virus (RSV) [18], Dengue virus [19], MERS-CoV [20], and Alkhurma virus (ALKV) [21].

IDPs/IDPRs are promiscuous binders in their nature, as they can have numerous interactions with distinct partners at a time, and can form important protein-protein interaction network hubs to control several signaling pathways at the same time [22]. In recent studies, it is reported that several IDPs/IDPRs display disorder-to-order transitions after interaction with their partners [23]. For example, the intrinsically disordered transactivation domain of c-Myb attains an α-helical conformation after association with KIX [24–27], whereas C-terminal regulatory domain of p53 oncoprotein undergoes transitions from its disordered form into α-helical, β-structural, or two irregular forms with different configurations at binding to Cyclin A, Sirtuin, CBP, and the S100ββ dimer [23,28]. To enable such folding at binding interactions with their partners, many IDPs/IDPRs are equipped with molecular recognition features (MoRFs) [29,30], which are short disorder-based binding regions that regulate the functional mechanism under physiological conditions. Generally, viral proteins such as those of Zika, HPV, and Chikungunya virus have a higher frequency of MoRF regions, where MoRFs are identified as important factors in regulating the functional protein-protein interactions, having implications in drug discovery [31–34].

In this work, we have examined the prevalence of intrinsic disorder in SA11 rotavirus proteome via multiple predictors and also performed PONDR® VSL2-based comparative intrinsic disorder analysis of different rotavirus serotypes (G1, G2, G3, G4, & G9). We have also predicted several MoRF regions in non-structural proteins of SA11 strain. We found a higher percentage of disordered regions in non-structural proteins, suggesting that their IDPRs might be involved in the regulation of assembly of replication complexes and processes and maturation of virus particles. This study could provide the important regions of rotavirus proteome that could be used as improtant targets for designing disordered-based small molecule inhibitors to block virus-host protein interactions for alternate therapy against the viral disease.

2. Materials and methods

We obtained all the reviewed protein sequences of simian rotavirus strain SA11 from UniProt which is most studied reference strain of rotavirus [35]. It has been widely used for experimental characterization of rotaviral pathogenesis globally [36,37]. Therefore this strain seems to be most appropriate for the intrinsic disorder MoRF analysis. Furthermore, we have also obtained the reviewed protein sequences of most common human pathogenic rotavirus serotypes from UniProt to examine the comparative

penetrance of intrinsic disorder. The list below provides the UniProt IDs for different proteins of these rotaviral serotypes: for G1 serotype, the UniProt IDs for NSP1-VP7 are (B3SRS0, B3SRS2, B3SRS1, Q9YJN7, B3SRS5, B3SRS6, B1NKR3, B1NKR4, B1NKR5, B3SRR9, B1NKR6, and P11853 respectively); for G2 serotype, the UniProt IDs for NSP1-VP7 are (P35423, Q03240, A1YTU9, B3SRT2, P23048, B3SRT4, A7J3A6, B1NKR8, Q6WNW5, P11196, A7J3A8, and P11850 respectively); for G3 serotype, the UniProt IDs for NSP1-VP7 are (B3SRV2, B3SRV4, B3SRV3, B3SRV6, B3SRV7, B3SRV8, B1NKS9, B1NKT0, B1NKT1, P11195, B1NKT2, and P11854 respectively); for G4 serotype, the UniProt IDs for NSP1-VP7 are (Q82045, B3SRX0, Q82053, Q82035, B3SRX3, B3SRX4, B1NKT7, B1NKT8, B1NKT9, P11200, B1NKU0, and P10501 respectively); and for G9 serotype, the UniProt IDs for NSP1-VP7 are (B3SRX6, B3SRX8, Q82049, B3SRY0, B3SRY1, B3SRY2, B1NKU1, B1NKU2, B1NKU3, B3SRX5, B1NKU4, and B3SRX9 respectively).

For intrinsically disordered regions (IDPRs) and MoRF analysis, numerous technologically advanced predictors are available, such as IUPred [38], ANCHOR [39], MoRFPred [40], MoRFchibi_web [41], DisEMBL [42], GlobPlot [43], DisoPred [44], PONDR® pool [including predictors, such as PONDR® VLXT [45], PONDR® VLS2 [46], and PONDR® VL3 [47], as well as a meta-predictor PONDR® FIT [48], and several others [49]. Critical Assessment of Protein Structure Prediction (CASP) was used to examine the accuracy of several disorder predictors, and several of the software mentioned above were compared using CASP [50]. All the available predictors are developed to identify the presence of intrinsically disordered residues and regions in query proteins, yet they all have diverse perspectives to examine the occurrence of intrinsic disorder. Therefore, it is sensible to use several computational tools to comprehend the abundance of intrinsic disorder in proteins [11]. All the predictors have limited accuracies of the prediction of intrinsic disorder in proteins, and the predictors whose accuracies are above 75% are considered as precise predictors. Characteristically, if the mean disordered score of any protein/region is above the threshold of 0.5, such proteins/regions are considered as IDPs/IDPRs. In addition, the iso-electric point of proteins has also been predicted by putting sequence in ExPasy (ProtParam tool) server [51].

Based on that standard, we selected four predictors from the PONDR family [PONDR® VLXT, [45] PONDR® VLS2, [46] PONDR® VL3, [47], PONDR® FIT [48]] and one more widely used server IUPred to analyze the prevalence of intrinsic disorder in rotavirus proteome. A meta-predictor PONDR® VL-XT integrates three predictors, the VL1 predictor trained to recognize long IDPRs, the amino-terminus-based predictor (XN), and the carboxyl-terminus-based predictor (XC) both trained to identify short disordered regions based on the X-ray crystallographic studies [45]. PONDR® VSL2 predictor uses a linear support vector machine (SVM) to predict both long and short IDPRs with 81% accuracy [46,52]. PONDR® VL3 is trained for sequence-based predictions of long IDPRs (>30 amino acid residues) by calculating the means of 10 artificial neural networks and choose the final prediction [47]. A meta-predictor PONDR® FIT represents a combination of PONDR® VLXT, PONDR® VSL2, PONDR® VL3, FoldIndex, IUPred, and TopIDP that gives considerably enhanced precision in the collective as related to its different constituent predictors [48]. For each protein, we used the outputs of these five disorder predictors to calculate the corresponding PPID (predicted percent of intrinsic disorder) values, which correspond to the percent of residues of a query protein that were predicted to be disordered by a given predictor. We also calculated the mean PPID values for query proteins by averaging the outputs of all five predictors.

Apart from IDP analysis, we have also performed the MoRF predictions for each protein of SA11 rotavirus strain. In general, MoRFs are the short (5–25 residues) protein binding sites within disordered region sometimes longer also (more than 50 residues) and may have the potential to undergo disorder to order transitions upon interaction with other partners [53]. In our study, four standard MoRF predictors were used as: ANCHOR [39], MoRFPred [40], MoRFchibi_web [41] and Disopred [54]. All these predictors utilize different types of algorithms to predict MoRFs, so we have used four predictors to completely characterized these regions in rotavirus proteome. ANCHOR based predictions utilize estimated energy calculations relative to biophysical properties whereas MoRFpred identifies on the basis of sequence alignment and support vector machine (SVM) [39,40]. Disopred provide disorder region based upon sequence property training sets trained through neural networks [54]. MoRFchibi_web is the most advanced and accurate predictor available; it gives the combined output of MoRFchibi and MoRFDC by following the Bayes rules [55]. Therefore, we have used only MoRFchibi_web-based predictions to map the regions on available crystal structures of rotavirus proteome.

We have also investigated the multiple sequence alignment of G1, G2, G3, G4 and G9 serotypes with the SA11 strain of rotavirus by Clustal Omega, web-based server [56]. All sequences were analyzed on the basis of three annotations which includes, conserved positions, strong polymorphisms and weak polymorphisms. The conserved regions have similar amino acid sequence throughout all strains. The strong polymorphisms are represented by space in the output file, which means substitutions among strongly dissimilar amino acid residues according to PAM 250 matrix. The weak polymorphisms are regions which have conserved residues among similar sequences based on their properties. These are denoted by colon and period symbols based on their scores having a threshold of 0.5 in PAM 250 substitution matrix.

The predicted MoRF regions were also checked for functional annotations by using eukaryotic linear motif (ELM) server [57] which gives the information of short linear motifs (SLiMs) (3–10 residues) within IDPRs having conserved functional sites. These motif sequences provide an interface for functional interaction with other binding partners with low binding affinity. In this study we have predicted functional sites in rotavirus proteins from strain SA11 using ELM (Eukaryotic Linear Motif) server, which constitutes the manually annotated instances (matched sequences) curated from experimental literature [58].

3. Results and discussion

3.1. Examination of intrinsic disorder status of rotavirus SA11 proteins

Proteins and protein regions that do not possess stable and precise three-dimensional (3D) structures under physiological conditions are known as IDPs/IDPRs. As explained before, during co-crystallization of proteins with their natural partners, interaction with said partners can stimulate binding-induced folding of IDPRs resulting in the appearance of specifically structured conformations, suggesting that the actual content of intrinsic disorder in proteins could be noticeably higher than the quantity projected from the missing electron density. Although the crystal structures of some of the rotavirus proteins are deposited in the Protein Data Bank (PDB), in the absence of complete structural knowledge pertaining to these proteins in solution, the use of computational means represents the best way to detect IDPRs in a given protein, thereby providing an important way of generating crucial information regrading protein structure and fucntion.

3.2. Examination of intrinsic disorder in rotavirus structural proteins

3.2.1. RNA-dependent RNA polymerase (RdRp/VP1)

RdRp, also known as VP1, is encoded by the gene segment 1 in all rotavirus strains. It is a compact globular protein with a

Table 1
Analysis of intrinsically disordered properties of SA11 rotavirus proteins.

Protein	$PPID_{VLXT}$	$PPID_{VSL2}$	$PPID_{VL3}$	$PPID_{FIT}$	$PPID_{IUPRED}$	$PPID_{MEAN}$
RdRp/VP1	11	13.4	6.6	2.2	0.4	3.1
VP2	17.4	18.5	14.9	11.9	7.7	14.2
VP3	4.6	4.2	0	2.0	0	0
VP4	4	11.2	0	15.7	6.4	12.8
VP6	10.3	10.3	0	4.8	0.5	0.5
VP7	15.3	9.2	0	7.1	0	0
NSP1	10.7	11.1	8.3	6.9	0.4	6.7
NSP2	9.2	15.8	7.9	9.1	2.2	6.6
NSP3	30.2	36.8	25.7	7.9	0.6	17.5
NSP4	39.4	51.4	45.7	22.3	9.1	32
NSP5	54	96.5	100	86.9	56.4	84.3
NSP6	6.5	10.9	0	15.2	0	4.3

diameter of 70 Å. Its 3D crystallographic structure is known (PDB ID: 2R7Q), containing three domains, such as N-terminal domain (residues 1–332), polymerase domain comprising finger subdomains, palm subdomain, and thumb subdomain (residues 333–778) and a C-terminal domain (residues 779–1089) [59]. A priming loop (residues 557–567), also known as a "primer grip," situated at the joining site of thumb and palm subdomains, has a flexible structure that interacts with triphosphates during initiation. The catalytic activity of VP1 is enhanced by the interaction between the residue Leu138 of NSP1 protein with NSP2/NSP5 nonstructural proteins under physiological condition [60]. Our analysis revealed that the mean PPID of the RdRp protein is rather low (3.1%) **{**Table 1**}**, this protein is having compact globular structure with few predicted MoRFs in its polymerase domain (residues 393–407 and 463–471) that may be responsible for the polymerase activity {Fig. 2.1(a) and Fig. 2.1(a1)} {Table 2 and Fig. 3(a)}. Since no MoRFs have been mapped in predicted disordered region (residues 390–404) therefore we have not given the functional annotation of MoRFs through ELM server for VP1 (Table S1).

3.2.2. Structural protein VP2

The structural protein VP2, encoded by gene segment 2, forms the innermost capsid of the triple-layered viral shell. This layer encapsidates the VP1 and VP3 structural proteins in the vicinity of the 5-fold axes, covering the viral genome, and RNA processing enzymes [61]. It is not only anchored with VP1, but the 5-fold axes of VP2 are also an indispensable cofactor for the initiation of genome replication [62]. T = 1 symmetry of VP2 is needed for the initiation of the process of dsRNA synthesis [63]. Although the function of the N-terminal region (residues 1–132) is not understood well, this region seems to be important for the activation/localization of the RNA-dependent RNA polymerase (VP1) and guanylyl transferase-methylase (VP3) into the virion center [64]. Our intrinsic disorder predisposition analysis revealed that the mean PPID of VP2 is 14.2% {Table 1}, indicating that it can be classified as moderately disordered protein. This analysis also showed that the N-terminal region of the VP2 protein is predicted to be completely disordered (residues 1–120) {Fig. 2.1((a)} and also possess MoRF regions predicted by several algorithms {Table 2 and Fig. 3(b)}. **For example, MoRFchibi_web has predicted 5 MORF regions.** Experimental studies showed that the deletion of first 25 N-terminal residues of VP2 (a region that is predicted by ANCHOR and DISOPRED to contain MoRFs) entirely prevents the

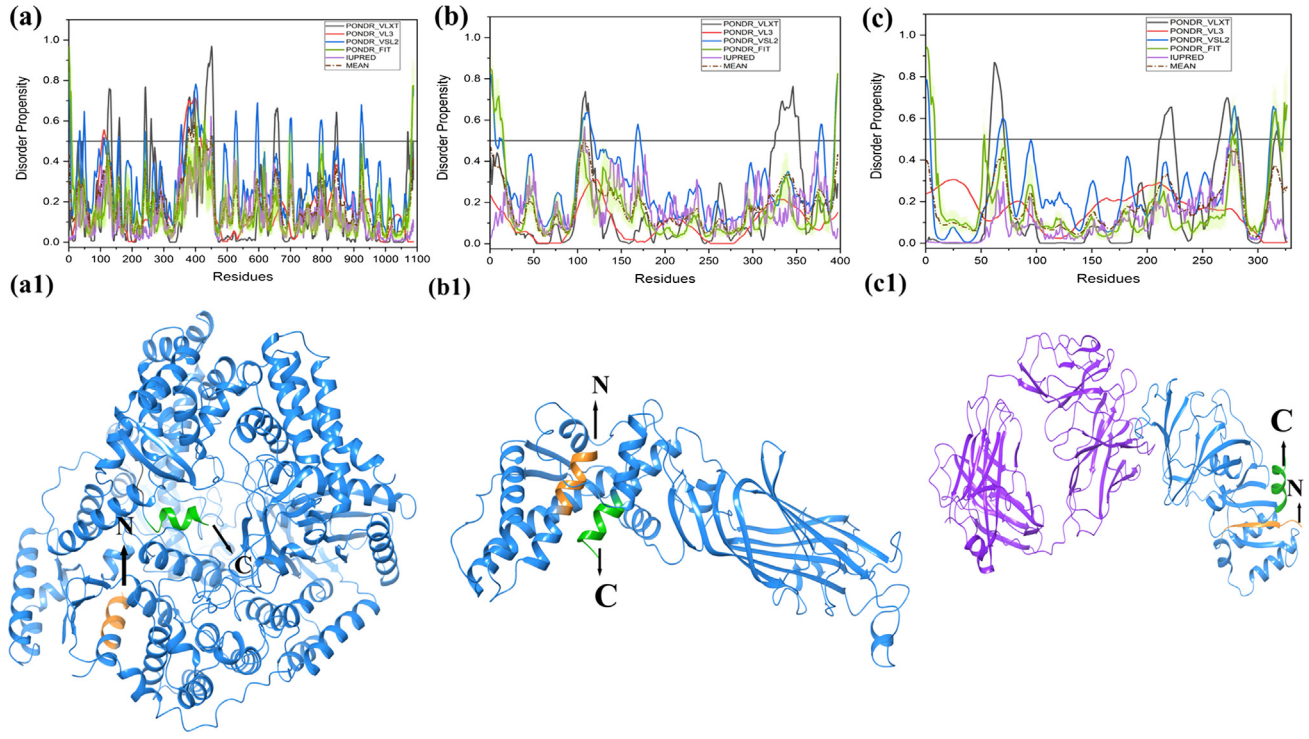

Fig. 2.1. Schematic illustration of the structural and intrinsic disorder characteristics of rotaviral structural proteins. The outpus of PONDR® VL-XT, PONDR® VL3, PONDR® VSL2, PONDR® FIT, and IUPred predictors are shown by black, red, blue, olive, and violet colors, respectively, whereas short dashed lines of wine color represent the mean disorder. Light olive shadow around PONDR® FIT curves represents the error distribution. The Morf regions is shown in grey color in the RdRp protein structure (a1). Plots signify disorder status of (**a**) 1,088 residue-long RdRp/VP1 that contains several IDPRs at its N-terminus; (**b**) 397 residue-long VP6 that possesses several IDPRs at its N- and C-terminal regions; **(c)** 326 residue-long VP7 is characterized by the presence of an IDPR at its N-tail and several IDPRs within the C-terminal domain. X-ray crystal structres are shown for RdRp protein (PDB ID: 2R7Q) **(a1)**, VP6 protein (PDB ID: 1QHD) **(b1)**, and VP7 protein (PDB ID: 3FMG) is structurally determined in complex with light and heavy chains of Fab of neutralizing antibody 4F8 **(c1)**. The N and C-terminals are denoted in orange and green colors with arrows respectively. Sequence of experimentally determined structure is shown (disordered residues in red color). Some residues are missing in the structures that are shown in plum color in the sequence.

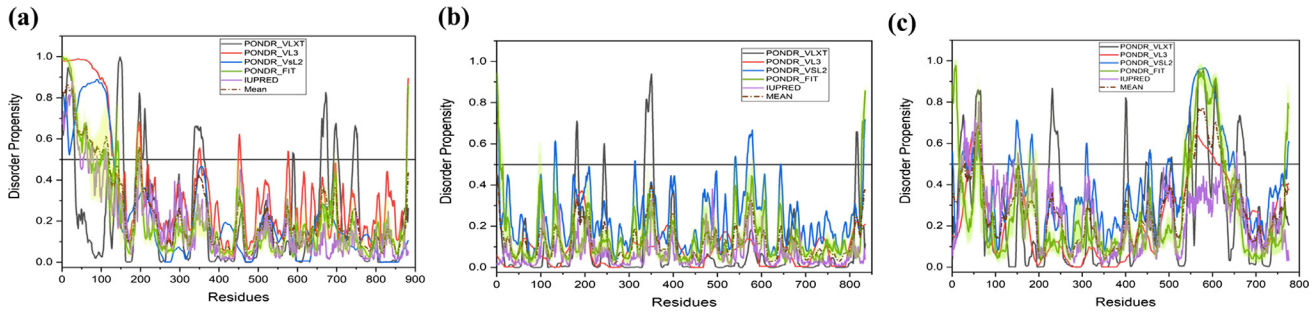

Fig. 2.2. Schematic illustration of the structural and intrinsic disorder characteristics of rotaviral structural proteins. (**a**) 822 residue-long VP2 that includes long IDPR at its N-terminal region; (**b**) 835 residue-long VP3 that shows least level of disorder with no IDPRs; (**c**) 755 residue-long VP4 consisting of VP8* (residues 1–231) and VP5* (residues 248–776) that displays higher disorder levels at its C-terminus. The color schemes are same as Fig. 2.1.

Table 2
MoRF analysis of all proteins of SA11 rotavirus strain by four different tools. Red numeric values show the MoRF regions that are common between two or more predictors.

Name	Protein length (AA)	Predicted disordered regions corresponding to MoRFs*	ANCHOR		MoRFchibi_web		MoRFpred		DISOPRED	
			Number of MoRFs ≥ 5 residues	MoRF-forming residues	Number of MoRFs ≥ 5 residues	MoRF-forming residues	Number of MoRFs ≥5 residues	MoRF-forming residues	Number of MoRFs ≥5 residues	MoRF-forming residues
RdRp/VP1	1088	390-404	1	463-471	2	23-29 1064-1077	2	9-13 410-414	3	393-407 788-794 1075-1088
VP2	882	1-119	2	1-8 39-45	5	74-80 106-114 122-138 144-155 161-173	1	208-212	1	1-70
VP3	835	-	None	None	1	820-825	None	None	1	1-11
VP4	775	-	4	1-16 77-88 103-108 201-209	1	762-767	1	8-13	None	None
VP6	397	-	None	None	1	381-396	1	121-125	1	1-5
VP7	326	-	None	None	2	292-301 319-326	1	319-325	None	None
NSP1	496	466-496	None	None	1	481-492	None	None	1	483-496
NSP2	317	297-317	None	none	2	38-61 282-294	1	308-317	None	None
NSP3	315	1-9 95-125	1	119-126	1	66-76	None	None	1	1-12
NSP4	175		None	None	2	32-37 124-131	None	None	None	None
NSP5	198	1-105 119-175 194-198	5	1-13 40-54 72-81 104-126 174-198	4	3-12 34-62 141-164 174-197	5	10-15 45-51 73-78 106-114 146-162	2	8-33 131-142

*We have shown only the predicted disordered regions of each protein where MoRF residues are lying within that region.

RNA binding activity and incorporation of VP1 and VP3 into virus-like particles (VLPs) [64]. The principal domain of VP2 (residues ~ 100–880) associated with the RdRp polymerase activity also exhibits several IDPRs (residues 100–120, 140–144, and 192–200) {Fig. 2.2((a)}, which have clear functional implications in virus maturation [61]. Also, the MoRFs present in these disordered regions indicates the essential disordered based binding sites are regulating the overall function and also predicted their functional annotations with ELM server (Table S1).

3.2.3. Structural protein VP3

Structural protein VP3 is encoded by genome segment 3. It is a multi-functional capping enzyme of 835 amino acids [65]. While the N-terminal region of ~690 residues contains domains needed for mRNA capping functions (a variable N-terminal domain, a guanine-N7-methyltransferase (N7-MTase) domain, a ribose-2′-O -methyltransferase domain, a guanylyltransferase and RNA 5′-triphosphatase (GTPase/RTPase) domain), the C-terminal region of ~150 amino acids possesses a 2′,5′-phosphodiesterase (PDE) activity that cleaves 2′-5′-oligoadenylate [66,67], thereby preventing RNase L activation [66,67]. Apart from capping-associated activities, VP3 is also associated with virulence. In its N7-MTase and GTase/RTPase domains, VP3 has some sequence motifs that are conserved among the rotavirus and orbivirus species. The conserved residues in the N7-MTase domain are located in region 134–139, whereas in GTase/RTPase domain, residues from 552 to 555 and Ser527, Arg531, Trp571, His610 are highly conserved [66,67]. Even though this multifunctional protein has a mean PPID score of 0% {Table 1}, individual predictors, such as PONDR® VLXT, PONDR® VSL2, and PONDR® FIT have predicted short stretches of

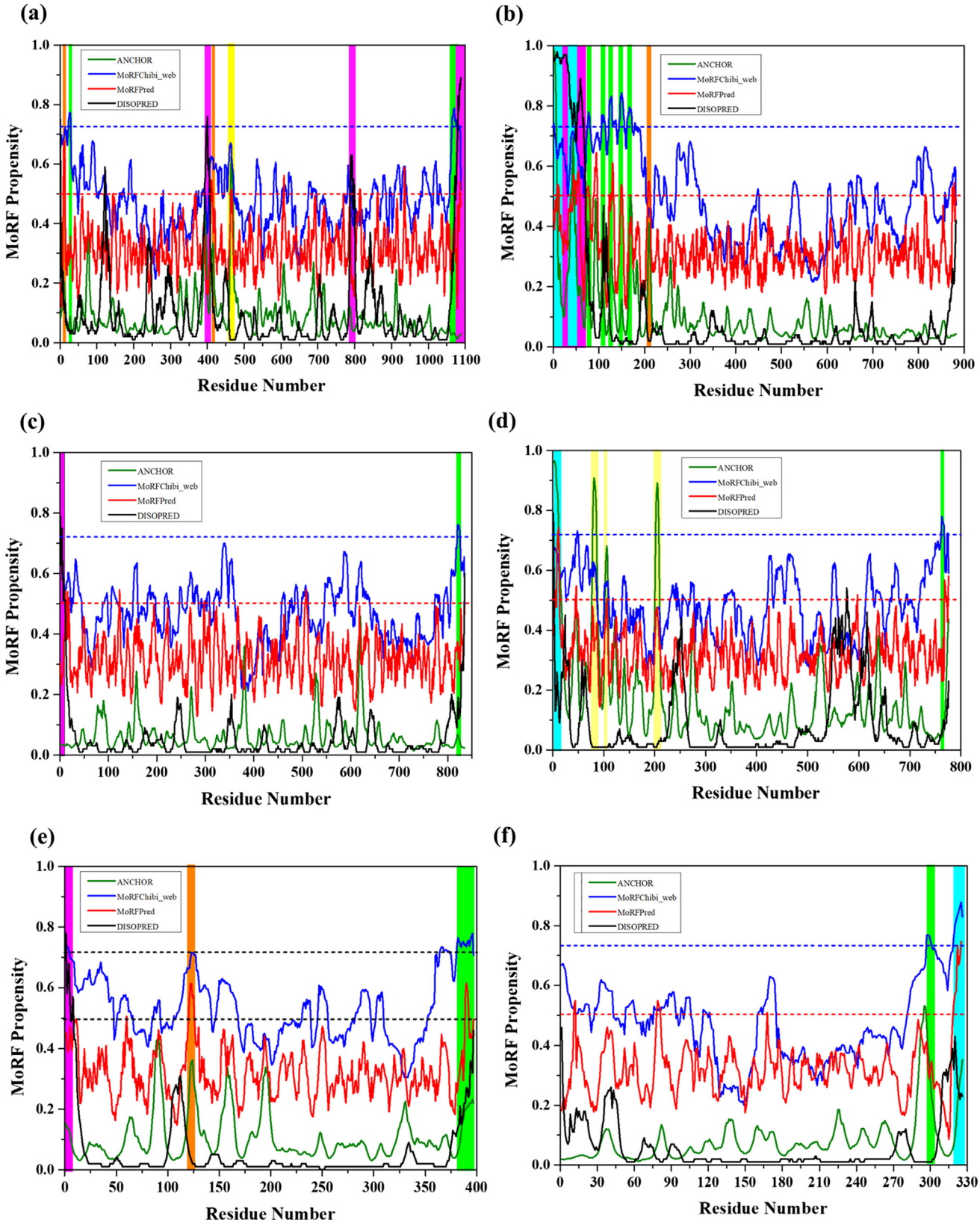

Fig. 3. Evaluation of disorder-based interactivity of the rotavirus structural proteins. The presence of MoRF regions was evaluated by ANCHOR (olive lines), DISOPRED (black line), MoRFpred (red line) and MoRFchibi_web blue lines). The threshold for MoRF predictions by ANCHOR, DISOPRED, MoRFPred are 0.5 and of MoRFchibi_web is 0.725. These thresholds are shown as dashed red lines and blue line, respectively. Positions of MoRFs predicted by ANCHOR is shown by a yellow bar, orange bar by MoRFPred, pink bar by DISOPRED, overlapped common MoRF positions are shown by cyan bar, and MoRFchibi_web are shown by olive bars, respectively. Plots signify MoRF analysis of **(a)** VP1 **(b)** VP2 **(c)** VP3 **(d)** VP4 **(e)** VP6 **(f)** VP7.

disorder at residues 179–184, 242–246, 335–355, 814–819 (aa residues), 1–5, 131–135, 313–314, 540–542, 568–582, 644, and 832–835 (aa residues) and 1–8 and 827–835 (aa residues), respectively while IUPred has not predicted any region as disordered. However, two MoRF regions {Table 2 and Fig. 3(c)} were observed with functional annotations given by ELM server (Table S1) seems important for the functionality of this protein {Fig. 2.2((b)}.

3.2.4. Structural protein VP4

VP4, a non-glycosylated outer capsid spike protein, is encoded by the 4th genome segment. The trypsin-mediated proteolytic cleavage of VP4 into the N-terminal VP8 (molecular weight ~28,000 Da) and C-terminal VP5 (molecular weight ~60,000 Da) fragments enhance the virus infectivity [68]. The negatively charged VP4 protein comprises several random coils and turns in its N-terminal region [69]. This N-terminal region possesses MoRF forming residue {Table 2 and Fig. 3(d)}. Four conserved cysteine residues; i.e., Cys203, Cys216, Cys318, and Cys380, are present in all the rotavirus strains [70]. Trypsin cleavage at Arg241 and Arg247 yields the receptor binding VP8* fragment and VP5* fragment, that enhances the viral infectivity [70–72]. The 265–474 region of VP5* possesses membrane permeabilization property and is highly disordered as predicted by all individual predictors {Fig. 2.2((c)} [73,74] VP4 is known to undergo structural rearrangements during membrane penetration [73]. Overall, VP4 is associated with multiple functions, such as cell surface interaction, neutralization, penetration, virulence, host range, and hemagglutination [73–77]. VP4 also binds microtubules and actin [78,79]. The mean PPID score for VP4 is only 12.8% {Table 1} but the individual predictors PONDR® VLXT, PONDR® VL3, PONDR® VSL2, PONDR® FIT, and IUPred have predicted higher percentage of disordered regions in total 157, 97, 209, 123, and 50 residues respectively {Fig. 2.2(c)}. The N-terminal and C-terminal regions of this protein exhibit intrinsic disorder and MoRF sites that help in spike formation at the outer layer of the virus particle [6].

3.2.5. Structural protein VP6

The intermediate capsid protein VP6 is encoded by the sixth gene segment and serves as the major structural protein of the virion [5]. It consists of 397 residues that are highly conserved among group A rotavirus strains. VP6 determines the group and subgroup antigenic specificities of the virus [80]. The protein has a well-defined 3D structure (PDB ID: 1QHD) {Fig. 2.1(b1)} that self assembles into a trimer with two different domains: a β-barrel domain, situated away to the center and a proximal α-helix domain, positioned closer to the center. VP6 plays a critical role in the organization/maturation of the virion by interacting with an assembly of the inner and outer capsid layers, and packaging of the genomic RNA in the inner capsid [7]. The region between amino acids 105 and 328 near the center of the protein mediates Trimerization, and the C-terminal region between residues 251 and 397 mediates association of the protein with VP2 of the single-shelled particle [81]. Antibodies against VP6 inhibit viral mRNA synthesis, suggesting a role for the protein in mRNA synthesis within the DLPs [82]. Though this protein has a crystal structure and a mean PPID value of 0.5% {Table 1}, the N-terminal region is predicted to be intrinsically disordered by the individual predictors, PONDR® VLXT (residues 103–115), PONDR® VSL2 (residues 1–6, 11–14, and 104–119 residues), and PONDR® FIT (residues 1–15) while IUPred has predicted only two residues 107 and 108 to be disordered. {Fig. 2.1 and 2.1(b)}. Three MoRF predictors have shown only single MoRF region in VP6 protein (Table2, Fig. 3e). MoRFchibi_web predicted MoRF residues at C-terminal (381–396) have shown to play a crucial role in the maturation and organization of the virus particle [81].

3.2.6. Structural protein VP7

VP7 is the second most abundant capsid protein of the virion and is encoded by 9th gene segment of the rotavirus strain SA11. Both VP7 and VP4 are highly immunogenic and induce neutralizing antibodies, and VP7 exhibits several antigenic epitopes [83,84]. The mean PPID value of this protein is (0%) {Table 1}, and VP7 has a well-defined crystal structure (PDB ID: 3FMG) {Fig. 2.1(c1)} that is enriched in α-helices and β-sheets. VP7 contributes the stability of the VP4 spikes at the outer surface of the virion and in the process of infection [83,84]. Although it does not show any predicted disordered regions, the crystallographic conformation of VP7 demonstrates that the N-terminal arm (residues 58–78) defines a grip that forms a connection between VP7 and VP6, and this region is disordered in nature [83,85]. Two of the five predictors {PONDR® VLXT (residues 58–71) and PONDR® VSL2 (residues 66–73)} suggest that these regions are disordered in nature. The C-terminal region (residues 313–325) of the VP7 protein is also disordered in nature and also has MoRF regions that may associate with the local dyad within its trimer {Table 2 and Fig. 3(f)}. The ELM server has predicted many SLiMs for VP7 and the MoRF residues have been functionally annotated in these SliMs (Table S1). The same residues are predicted to be disordered by the individual predictors, PONDR® VLXT (residues 317–319), PONDR® VSL2 (residues 312–319 and 323–326), and PONDR® FIT (residues 313–319 and 323–326). In crystal structure of VP7 (residues 50–330), predicted disorder residues (310–328) are missing and hence are in agreement with our disorder predictions. Experimental studies have suggested a role for the disordered N- and C-terminal domains in the organization and maturation of the virions by undergoing significant conformational changes, where the aqueous interface of inner surface of VP7 allows protein interaction with the outer surface of VP6 by one pair of van der Waals interaction (Pro279 and Thr281 of VP7 interacting with Pro313 of VP6) and one side chain H-bond (Gln305-Asn310) [83].

3.3. Exploration of the intrinsic disorder status of the rotaviral non-structural proteins

3.3.1. Non-structural protein NSP1

NSP1 is a non-structural RNA binding protein, encoded by the 5th genome segment. It is a basic protein and the most divergent among the rotaviral proteins. The N-terminal region of 150 residues is relatively more conserved than the rest of the protein and contains three basic regions at residues 10–39, 79–91 and 111–126 and a conserved cysteine-rich motif between residues 42 and 75 that forms a zinc-finger. A second non-conserved Cys-rich motif is also present between residues 314 and 327 [86–88]. Mutant viruses containing the NSP1 gene coding for only the amino-terminal 50 amino acids are able to replicate, suggesting that the region downstream of residue 50 not be essential for virus replication [88]. The partial crystal structure of NSP1 has been determined (PDB ID: 5JER) that has C-terminal stretch (475–500) bound to IRF3 protein {Fig. 4.1(a1)}. Analysis of the complete sequence of NSP1 protein (1–495 residues long) using the five different predictors showed the mean PPID score of 6.7% {Table 1}. PONDR® VSL2 predicted disordered regions (311–319 residues) that comes under zinc-binding site (314–327 residues) {Fig. 4.1 (a)}. Overall the C-terminal region is predicted to be disordered by all the predictors {Fig. 4.1(a1)} containing MoRFs precited by MoRFchibi_web and DISOPRED (residues 481–492 and, 483–497) {Table 2 and Fig. 5(a)}. The ELM server predictions have shown many binding sites for MoRF regions of NSP1 as shown in supplementary file (Table S1).

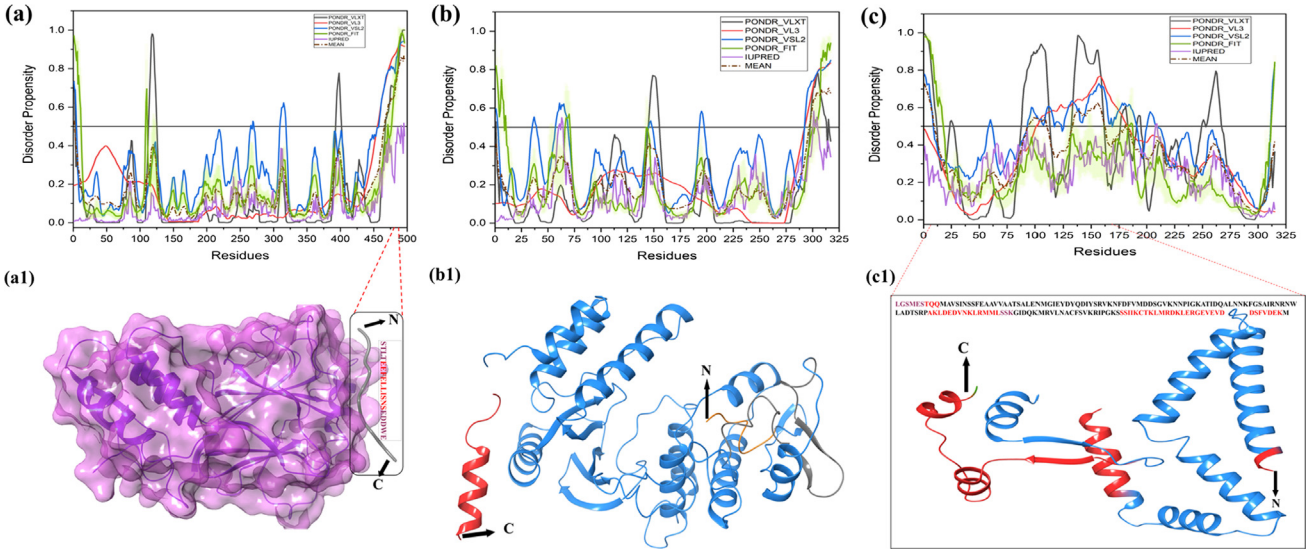

Fig. 4.1. Graphical illustration of the structural and intrinsic disorder characteristics of rotaviral structural proteins. The outputs of PONDR® VL-XT, PONDR® VL3, PONDR® VSL2, PONDR® FIT and IUPred predictors are shown by black, red, blue, olive, and violet colors, respectively, whereas short dashed lines of wine color represent the mean disorder. Light olive shadow around PONDR® FIT curves represents the error distribution. The red color indicates the mean disorder region present in the protein structures and the grey color represents the MoRF regions present in NSP1 (a1), NSP2 (b1), and NSP3 (c1). The N and C-terminals are denoted in orange and green colors with arrows respectively. Plots signify the disorder status of NSP1 (**a**); NSP2 (**b**); and NSP3 **(c).** X-ray crystal structures are shown for NSP1 protein (PDB ID: 5JER) in complex with Interferon regulatory factor 3 (IRF3, shown in violet color) **(a1)**; NSP2 (PDB ID: 6CY9) **(b1)**; and NSP3 protein (PDB ID: 1KNZ) **(c1)**.

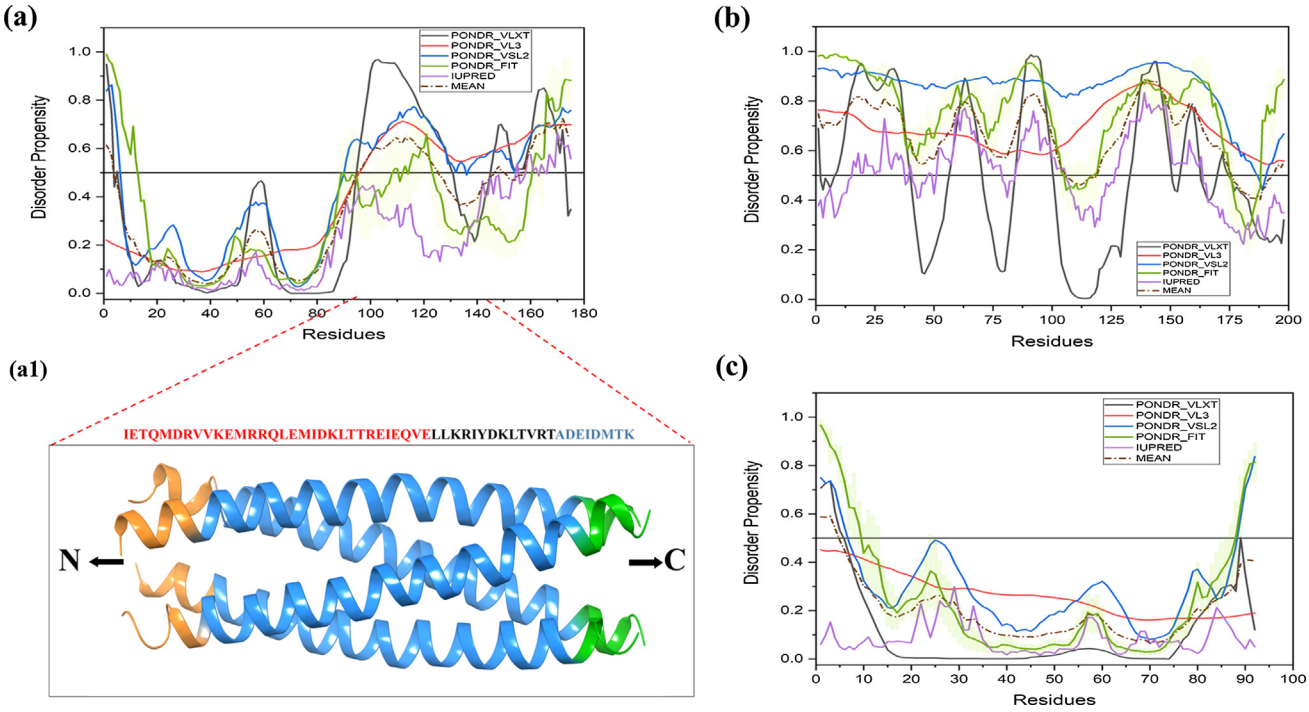

Fig. 4.2. Graphical illustration of the structural and intrinsic disorder characteristics of rotaviral structural proteins. Plots indicate the disorder status of NSP4 (**a**); NSP5 (**b**); and NSP6 **(c)**. Crystallized structure of diarrhea-inducing domain of NSP4 (PDB ID: 2O1J, residues 95–140) is shown in **(a1)**.

3.3.2. Non-structural protein NSP2

NSP2 is encoded by the 8th gene segment in SA11 strain of rotavirus. It is an essential nucleating protein of viroplasms, which function as viral factories, where viral genome replication and assembly of double-layered particles (DLPs) take place [89]. Its interaction with NSP5 leads to viroplasm formation [90]. NSP2 interacts with several host proteins and sequesters them into the viroplasm [91]. It also interacts with tubulin and is associated with the perinuclear localization of the viroplasms [92,93]. Crystallographic structure (PDB ID: 6CY9) of NSP2 has revealed that it forms a large doughnut-shaped octamer through tail-to-tail tetramer assemblies {Fig. 4.1(b1)} [94]. It possess non-specific sequence-independent RNA-binding activity and helps in efficient viral mRNA replication in combination with the viral polymerase complex [89,94–96]. NSP2 also possesses Mg^{2+}-dependent NTPase, helix-unwinding, and NDP kinase activities [96–98]. Residues

Table 3
The infection rate (in percentage) of different serotypes of rotavirus based on the geography and temporal regions.

Regions	Serotypes				
	G1	G2	G3	G4	G9
South America	34	23	2	9	15
Africa	23	2	21	4	7
Asia	34	13	1	20	12
North America	73	11	6	4	3
Europe	71.6	9	2	11	4
Australia/Oceania	82.4	14	1	2	0.5

responsible for NDP kinase activity, such as Thr 56, His 59, Phe 64 are located within the MoRF region (38–61), which also includes some of the residues 53 to 76 that form part of RNA-binding grooves {Table 2 and Fig. 5(b)} This multifunctional protein showed a mean PPID value of about 6.6% {Table 1}. {Fig. 4.1(b1)} shows that the C-terminal RNA-binding domain (residues 295–317) is predicted to be mostly disordered, suggesting that it undergoes the conformational changes after interacting with the viral polymerase complex. Since this protein interacts with host proteins, our ELM predictions annotated several binding sites for NSP2 MoRFs as shown in supplementary file (Table S1).

3.3.3. Non-structural protein NSP3

The 7th genome segment encodes NSP3 protein [99]. The protein is slightly acidic and binds specifically the conserved tetranucleotide sequence present at the 3′ end of the rotaviral mRNAs [100]. There is a crystal structure available for NSP3 (PDB ID:1KNZ) {Fig. 4.1(c1)} with a dimerization region (residues 150–206), an eIF4G-binding region (residues 274–313) and an RNA binding domain (residues 83–149) [101–103]. The RNA-binding domain helps the protein to act as dimer *in vivo* and is implicated in facilitating viral mRNA translation by competing with and evicting the poly(A)-binding protein (PABP) from eIF4GI and shut-off of host cellular protein synthesis [104–107]. The overall mean PPID score of NSP3 protein is 17.5% {Table 1} with higher disorder content being found within its dimerization region (residues 150–206), and in the RNA binding domain (residues 83–149) {Fig. 4.1(c)} that are important for the protein function. Also, the residues 150–160 are missing in PDB structure and are predicted to be disordered through mean PPID region {Fig. 4.1(c1)}. The RNA binding domain also comes under MoRF region predicted by ANCHOR {Table 2 and Fig. 5(c)}. Experimental studies have shown that mutations in the flexible region (four amino acid residues, W170A, K171E, R173E, and R187E: K191E) reduced the stable dimer formation and altered the protein stability [107]. In agreement with these observations, disorder prediction analysis has also shown that this is characterized by the enhanced disorder predisposition.

3.3.4. Non-structural protein NSP4

NSP4 (20.26 kDa) is a 175 residue-long glycosylated protein encoded by the 10th genome segment [108]. It is localized to the endoplasmic reticulum (ER) through its N-terminal hydrophobic regions and functions as the ER-anchored receptor for DLPs for their maturation into TLPs in the ER and is essential for the morphogenesis of the virion [108]. It is also localized in other organelles as well as is secreted from the apical surface of the polarized epithelial cells [109]. It is the first identified viral enterotoxin [110,111]. The flexible C-terminal cytoplasmic region (residues 161–175) of NSP4 binds to VP6 of the DLPs and mediates their entry into the ER lumen where the virus becomes temporarily enveloped [112,113] and matures into TLP. NSP4 interacts with microtubules and immobilizes the early secretory pathway [114,115]. The NSP4 protein shows overall 32% mean PPID score {Table 1} with higher levels of disorder in cytoplasmic domain (residues 95–146), enterotoxic peptide region (residues 114–135), and the flexible C-terminal region (residues 139–175) {Fig. 4.2(a)}. Although, the region from residue 114 to 135 forms α-helical bundle in crystal structure {Fig. 4.2(a1)}, however it has been predicted to have high disorder propensity based on amino acid composition. This is in line with the known fact that under specific conditions used to promote protein crystallization process, disordered regions tend to form some secondary structure elements [116]. Also, in a study, NSP4 peptides (residue 120–147) has shown disordered type of CD spectra in buffer and undergo significant conformational transitions in membranous environment [117]. The aforementioned disordered regions of NSP4 play a key role in the initiation of diarrhea and binding to DLP [112,113]. The MoRF region (residues 124–131) resides in the region that is responsible for many functions such as residue 112–148 utilized as the VP4 binding domain residues 112–140 for Caveolin binding, and residues 112–135 used for Ca^{+2} binding and enterotoxin domain {Table 2 and Fig. 5(d)}. Functional annotations with predicted SliMs have also been shown in supplementary file (Table S1). Amino acid substitution mutations P168A, Y166S, and M175I in the unstructured C-terminus of NSP4 lead to the complete loss of interaction of the protein with DLPs [113]. This multifunctional protein is pleiotropic in nature because of the oligomeric state combined with high conformational dynamics of the disordered cytoplasmic domain of the protein as evident from the calcium-lacking pentameric (PDB ID: 1G1J & 2O1K) [118,119] and the calcium-containing tetrameric (PDB ID: 2O1J) [120–122] coiled-coil structures of the diarrhea-inducing region from ST3, and SA11 and I321 strains of rotavirus, respectively. NSP4 has a cholesterol recognition amino acid consensus sequence (CRAC), which is predicted to be a completely disordered region and it interacts with cholesterol and caveolin, leading to its localization to the plasma membrane [123].

3.3.5. Non-structural protein NSP5

The 11th gene segment encodes NSP5 protein (21.70 kDa) [124]. It undergoes O-linked glycosylation [125] and phosphorylation [126,127], and interaction with NSP2 leads to the hyperphosphorylation of NSP5 [90,128]. NSP5 is crucial for viroplasm formation and recruitment of other viral proteins [129]. Both the N-terminal (residues 1–33) and C-terminal (residues 131–198) regions of the NSP5 are important for interaction with NSP2, viroplasm formation, and virus replication [130]. NSP5 protein of rotavirus has the highest intrinsic disorder level among all the rotaviral proteins with mean PPID of 84.3% {Table 1}. {Fig. 4.2(b)} shows that NSP5 is predicted to have long IDPRs in its N-terminal region, and within the C-terminal domain that contains the C_XC motif (C171 and C174) responsible for the iron-sulfur cluster organization. There are various MoRF regions present in NSP5 protein of rotavirus SA11 and other strains that are involved in the functionality of potential protein kinase C phosphorylation sites (S-X-R/K-R/K-X-X-S and K/R-X-S) at serine residues 22, 30, 75, 100, and 136 [131] {Table 2 and Fig. 5(e)}. There are also serine residues at positions 56, 154, and 165 that are potential substrates of casein kinase II (S-X-X-D/E) {Table 2 and Fig. 5(e)}. Phosphorylation of Ser-67 within the SDSAS motif (amino acids 63–67) was required to trigger hyperphosphorylation by promoting the activation function, and mutation of all three Sers of motif a (Ser-63, Ser-65, and Ser-67) completely abolished activation function [132]{Table 2 and Fig. 5(e)}. Recent studies demonstrated cytoplasmic relocalization of several nuclear hnRNPs (heterogeneous nuclear ribonucleoproteins), and ARE-BPs (AU-rich element-binding proteins) during rotavirus infection, influencing viral gene expression/replication and progeny virus production [91]. Rotavirus was also demonstrated to induce atypical stress granules and P-bodies during

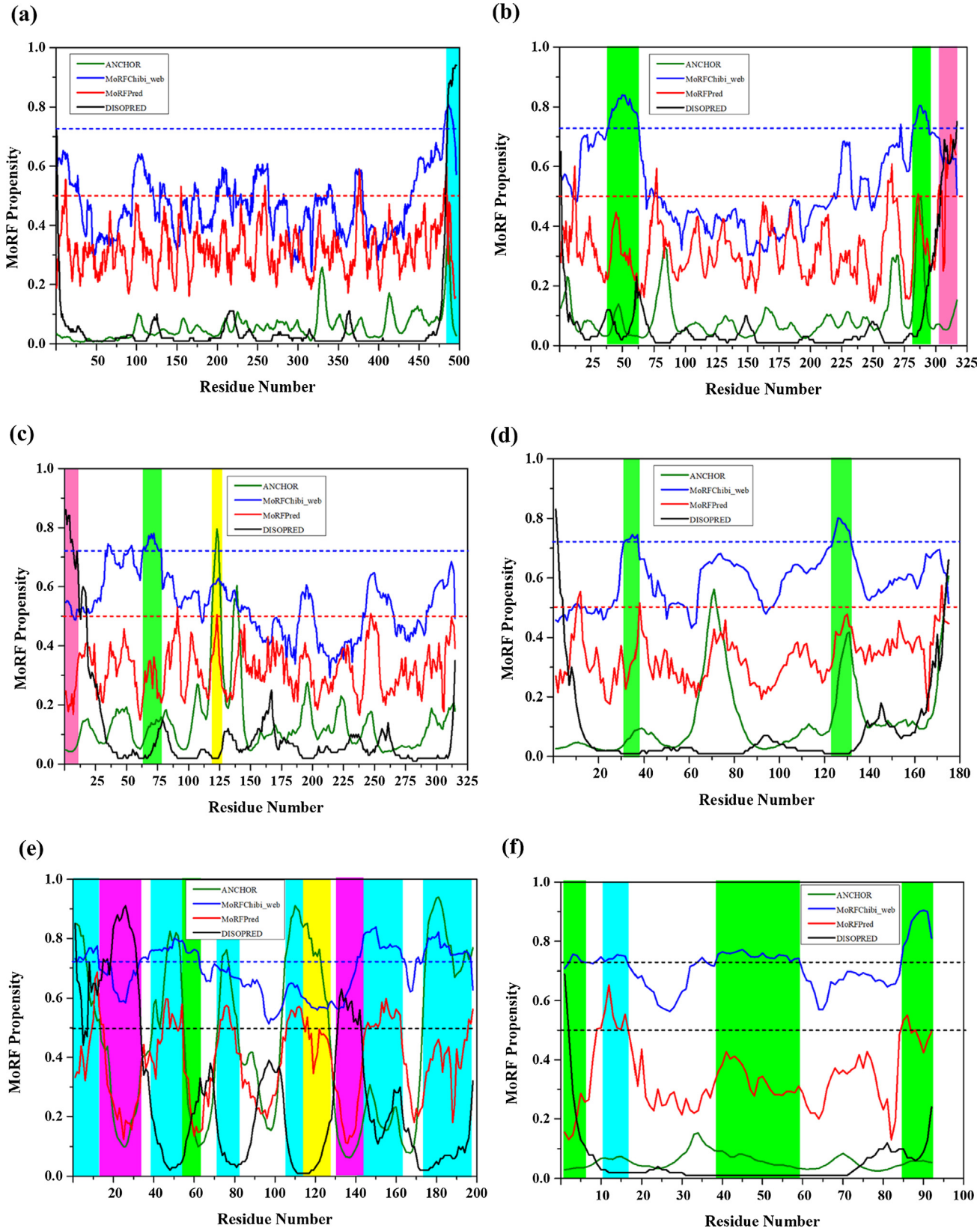

Fig. 5. Evaluation of disorder binding regions of the rotavirus non-structural proteins. The presence of MoRF regions was evaluated by ANCHOR (olive lines), DISOPRED (black line), MoRFpred (red line) and MoRFchibi_web blue lines). The threshold for MoRF predictions by ANCHOR, DISOPRED, MoRFPred are 0.5 and of MoRFchibi_web is 0.725. These thresholds are shown as dashed red lines and blue line, respectively. Positions of MoRFs predicted by ANCHOR is shown by a yellow bar, orange bar by MoRFPred, pink bar by DISOPRED, overlapped common MoRF positions are shown by cyan bar and MoRFchibi_web are shown by olive bars, respectively. Plots signify MoRF analysis of **(a)** NSP1 **(b)** NSP2 **(c)** NSP3 **(d)** NSP4 **(e)** NSP5 **(f)** NSP6.

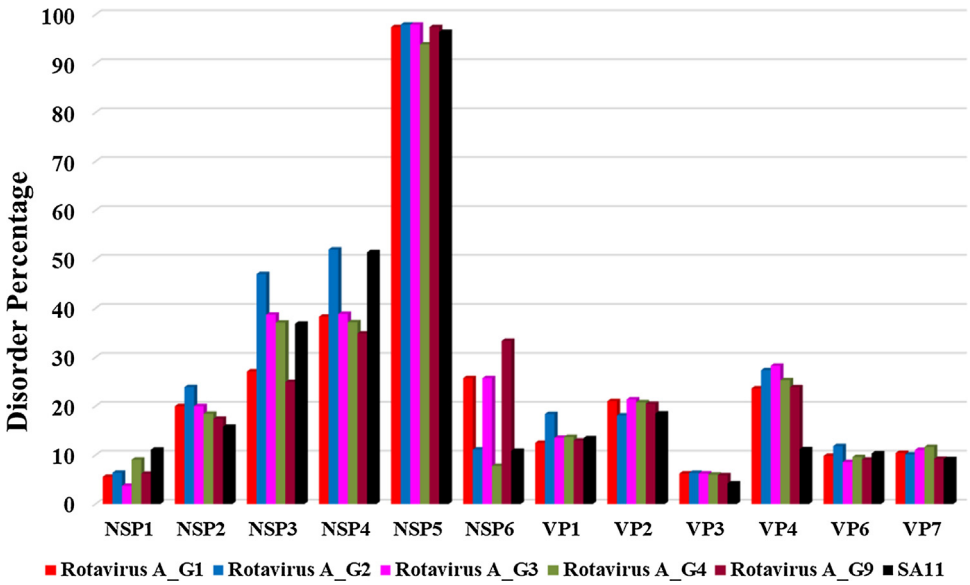

Fig. 6. Outcomes of the evaluation of the intrinsic disorder predisposition of rotavirus proteins from different serotypes in terms of the PONDR® VSL2-based PPID values. PPID scores for all proteins of rotavirus are specified on the y-axis. Results for proteins from G1, G2, G3, G4, and G5 serotypes are shown by red, blue, magenta, olive, and wine color bars, respectively. Corresponding results for the SA11 reference strain are shown by black bars.

infection [133]. These studies further demonstrated that majority of the relocalized nuclear proteins, some cytoplasmic proteins, as well as the atypical stress granules and P-bodies formed during rotavirus infection are sequestered in the viroplasms largely through the interaction of the host proteins with NSP2 and/or NSP5 in the viroplasm [91,133]. Intrinsic disorder appears to be critical for the multifunctionality of NSP5. The presence of high intrinsic disorder level in NSP5 protein is expected based on the peculiarities of its genetic makeup. In fact, viruses have evolved a complex genetic organization for optimal use of their limited genomes and production of all necessary structural and regulatory proteins, including rather frequent utilization of overlapping open reading frames [12,16,134,135]. It has been shown that proteins or protein fragments corresponding to overlapping genes are either disordered or possess complementary disorder distribution, where if a protein encoded by one ORF is ordered, then the protein encoded by the overlapped ORF will have high levels of intrinsic disorder [12,16,134,135].

3.3.6. Non-structural protein NSP6

NSP6 (11.01 kDa) is another viral protein produced from an overlapping open reading frame (ORF) formed by the gene duplication of segment 11 [136]. Interestingly, the NSP6 proteins has varying length of amino acid residues in different rotavirus strains such as SA11 has NSP6 with 92 amino acid residues where in some strains NSP6 is truncated with less than 40 residues (G1, G3 and G9) [137]. In comparison to other non-structural proteins, NSP6 is expressed at a very low level and is observed to be unstable [138]. It binds DNA and RNA in a sequence-independent manner with equal affinity. The first 20 residues at the N-terminal region are essential for mitochondrial targeting and altering the mitochondrial function during rotavirus infection [139]. NSP6 is localized to viroplasm after phosphorylation and regulates self-association of NSP5 [138]. The protein is only slightly disordered in nature and has the mean PPID score of 4.3% {Table 1}. {Fig. 4.2 (c)} demonstrates that NSP6 has two short, terminally-located IDPRs. The N-terminal disordered region (residues 1–6) also has MoRF forming residues (1–20) are essential for mitochondrial targeting and alter the mitochondrial function during rotavirus infection {Table 2 and Fig. 5(f)}. Recent reverse genetics experiments demonstrated that NSP6 is not essential for rotavirus replication in the cell culture [140].

3.4. Comparative analysis of intrinsic disorder status in proteins from different rotavirus serotypes

There are multiple different serotypes of rotavirus that cause the rotaviral infections worldwide, with the most causative serotypes being G1, G2, G3, G4, and G9, which are collectively responsible for more than 95% of rotavirus infections. The infection rate of these serotypes mainly dependent on the geography and temporal regions of the rotavirus (as shown in Table 3) [4].

To evaluate if there are noticeable differences in the intrinsic disorder predisposition of rotaviral proteins from different serotypes of this virus, we conducted a computational analysis of structural and non-structural proteins from the retrovirus serotypes G1, G2, G3, G4, and G9. This comparative analysis supported the occurrence of a higher percentage of disordered regions in non-structural proteins and the presence of short IDPRs in structural proteins {Figs. 6–8}. Although viral proteins from these serotypes have shown similar intrinsic disordered patterns as those examined in the SA11 reference strain, there also were some important differences. For example, the G2 serotype was characterized by slightly higher protein disorder percentage among all five serotypes, with the largest difference being observed for NSP2, NSP3, NSP4, NSP5, VP1, and VP6 of SA11 strain (see {Figs. 2.1, 2.2 and 4.1, 4.2}).

4. Conclusion

The functional mechanisms of some of the rotaviral proteins are known, but not completely deciphered because of the non-availability of complete structural information. The main objective of this analysis is to fill the gap to some extent by providing the data on the prevalence of intrinsic disorder in the form of intrinsically disordered proteins/regions (IDPs/IDPRs) in the rotavirus proteome. Based on several computational and experimental studies, it is evident that IDPRs or IDPs and MoRFs are crucial for viral repli-

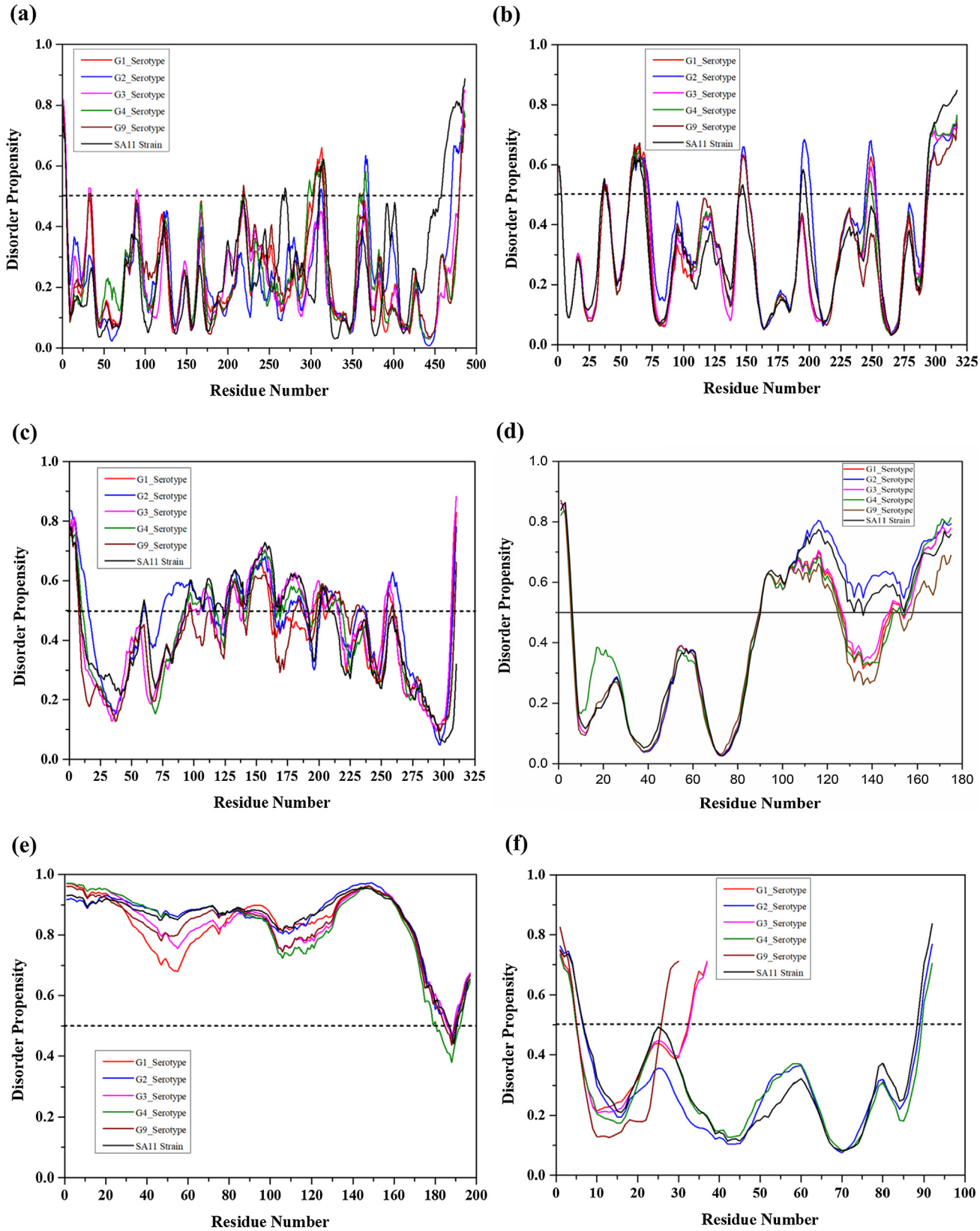

Fig. 7. Comparative analysis of the intrinsically disorder predisposition of the non-structural rotaviral proteins from 5 different serotypes of human rotavirus (G1, G2, G3, G4, and G9) with reference strain SA11. PONDR® VSL2 predictor was utilized in this study. The PONDR® VSL2-based disorder profiles of non-structural proteins from each serotype (G1, G2, G3, G4, and G9) are represented by red, blue, magenta, olive, and wine colors, respectively. Black curves show corresponding results for the proteins from the SA11 reference strain. Plots represent corresponding per-residue disorder propensities for **(a)** NSP1, **(b)** NSP2, **(c)** NSP3 **(d)** NSP4, **(e)** NSP5, and **(f)** NSP6. The NSP6 protein found to be truncated in G1, G3 and G9 serotypes having residue length less than 40 amino acids.

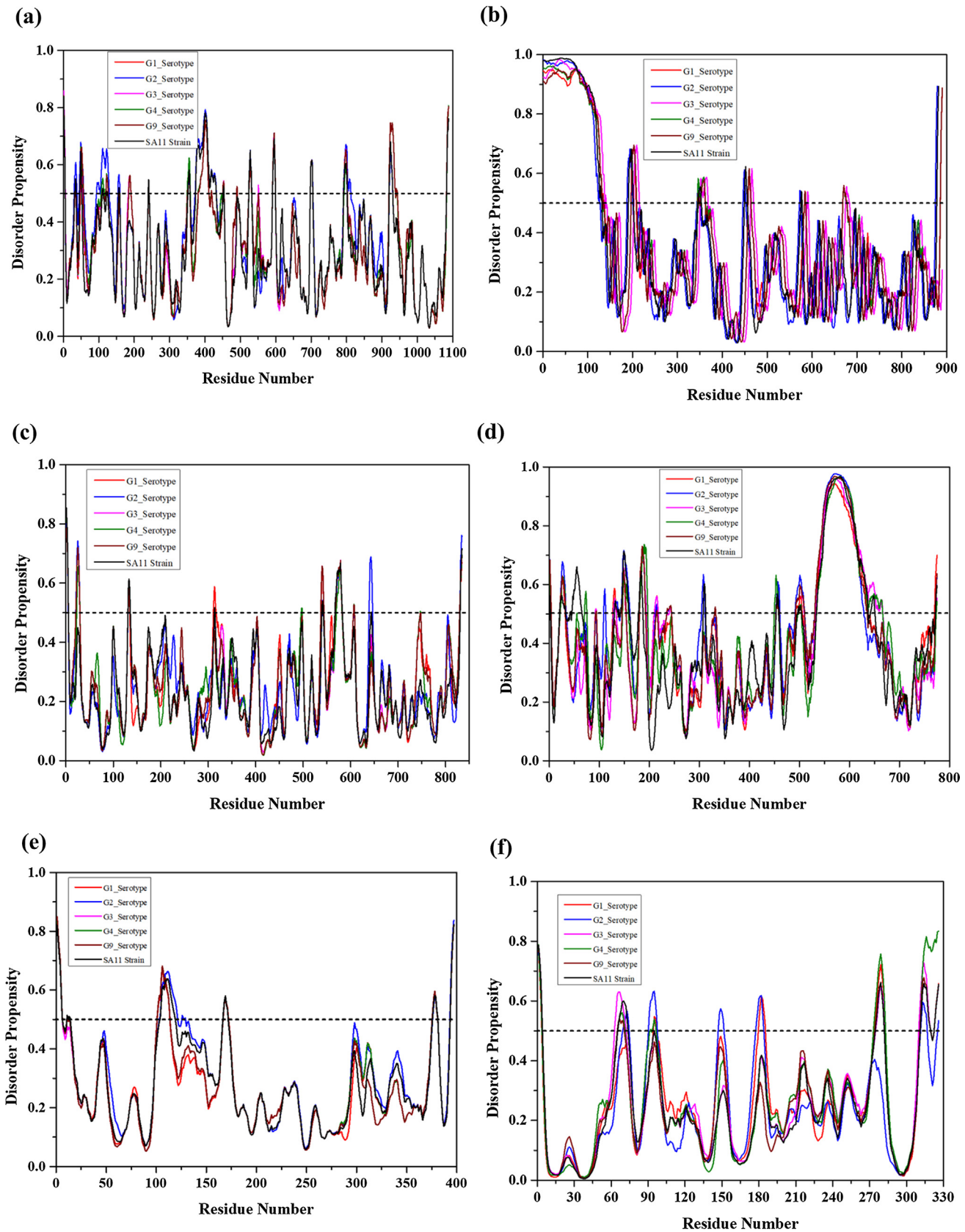

Fig. 8. Comparative analysis of the intrinsically disorder predisposition of the structural rotaviral proteins from 5 different serotypes of human rotavirus (G1, G2, G3, G4, and G9) with reference strain SA11. PONDR® VSL2 predictor was utilized in this study. The PONDR® VSL2-based disorder profiles of non-structural proteins from each serotype (G1, G2, G3, G4, and G9) are represented by red, blue, magenta, olive, and wine colors, respectively. Black curves show corresponding results for the proteins from the SA11 reference strain. Plots represent the corresponding per-residue disorder propensities for **(a)** VP1/RdRp, **(b)** VP2, **(c)** VP3, **(d)** VP4, **(e)** VP6, and **(f)** VP7.

cation and pathogenesis. Intrinsic disorder is also known to confer flexible conformations to viral proteins, as well as provide important mechanistic means for their binding promiscuity; i.e., their ability to be engaged in interactions with many viral and host proteins, as demonstrated for rotaviral NSP2 and NSP5 within the viroplasm [91,133] and for proteins of many other viruses [141], leading to the subversion/hijacking of host cellular processes, including the protein synthesis machinery and nuclear transport to promote virus propagation [142,143]. To compensate for the lack of the detailed information on the dark side of the rotaviral proteome, by employing several computational tools, here we show that rotavirus (the reference strain (SA11) and all five serotypes (G1, G2, G3, G4, and G9) analyzed in this study) contains several short and a few long IDPR in many structural and non-structural proteins, with NSP5 being expected to be disordered as a whole. Comparatively to other rotaviral proteins, NSP4 and NSP5 showed highest levels of intrinsic disorder that appear to be crucial for their multi-faceted functions. Also, the presence of multiple MoRFs predominantly in NSP5 indicates the functional importance of this protein for the rotavirus life cycle. Therefore, this study suggests that by virtue of its presence, intrinsic disorder (in the form of IDPRs with various length, or in the form of total disorder of NSP5) provides significant conformational plasticity and functional diversity to the rotaviral proteins and contributes to the assorted biological functions involving communication with various host proteins and protein-protein interactions within the virus. A well-known functional mechanism adopted by disordered proteins is disorder-to-order transitions upon binding to a partner that could be disrupted by small molecules. A new approach for drug design based on the concept of ligand-mediated disorder-to-order transition is slowly growing traction over traditional structure-based drug screening studies. Recent studies on the identification of small molecules that inhibit viroplasm formation by targeting protein-protein interactions [144] in this supramolecular rotavirus replication organelle could form the basis for the drug discovery based on the disordered regions of NSP2 and NSP5 that are essential for nucleation of viroplasm and recruitment of a large number of viral and host proteins into the viroplasm [91,133]. Since high mutation rate among RNA viruses is the leading cause of vaccine failure, targeting disordered regions with small molecules for inhibiting either binding to their partners or preventing their conformational transition may provide a potential alternate therapeutic approach. We believe that understanding the penetrance of the IDPs/IDPRs in in rotavirus proteome would provide a new direction to decipher the mechanisms of virus pathogenesis, morphogenesis, virus survival, and drug discovery.

Acknowledgment

This work was partially supported by DST grant, India (YSS/2015/000613) to RG and IIT-Mandi, India to RG. DK is grateful to ICMR fellowship.

Author's contributions

RG, conception, design and study supervision; RG, VU, CDR, AS, PK and DK.: acquisition, analysis, and interpretation of data, writing, and review of the manuscript. # DK and AS contributed equally.

Declartation of Competing Interest

The authors declare no competing financial interest.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijbiomac.2019.09.166.

References

[1] C. Troeger, M. Forouzanfar, P.C. Rao, I. Khalil, A. Brown, R.C. Reiner, N. Fullman, R.L. Thompson, A. Abajobir, M. Ahmed, M.A. Alemayohu, N. Alvis-Guzman, A. T. Amare, C.A. Antonio, H. Asayesh, E. Avokpaho, A. Awasthi, U. Bacha, A. Barac, B.D. Betsue, A.S. Beyene, D.J. Boneya, D.C. Malta, L. Dandona, R. Dandona, M. Dubey, B. Eshrati, J.R.A. Fitchett, T.T. Gebrehiwot, G.B. Hailu, M. Horino, P.J. Hotez, T. Jibat, J.B. Jonas, A. Kasaeian, N. Kissoon, K. Kotloff, A. Koyanagi, G.A. Kumar, R.K. Rai, A. Lal, H.M.A. El Razek, M.A. Mengistie, C. Moe, G. Patton, J.A. Platts-Mills, M. Qorbani, U. Ram, H.S. Roba, J. Sanabria, B. Sartorius, M. Sawhney, M. Shigematsu, C. Sreeramareddy, S. Swaminathan, B. A. Tedla, R.T.-M. Jagiellonian, K. Ukwaja, A. Werdecker, M.-A. Widdowson, N. Yonemoto, M. El Sayed Zaki, S.S. Lim, M. Naghavi, T. Vos, S.I. Hay, C.J.L. Murray, A.H. Mokdad, Estimates of global, regional, and national morbidity, mortality, and aetiologies of diarrhoeal diseases: a systematic analysis for the global burden of disease study 2015, Lancet Infect. Dis. 17 (2017) 909–948, https://doi.org/10.1016/S1473-3099(17)30276-1.
[2] WHO | Rotavirus, WHO, 2018.
[3] H.F. Clark, D.A. Lawley, A. Schaffer, J.M. Patacsil, A.E. Marcello, R.I. Glass, V. Jain, J. Gentsch, Assessment of the epidemic potential of a new strain of rotavirus associated with the novel G9 serotype which caused an outbreak in the United States for the first time in the 1995–1996 season, J. Clin. Microbiol. 42 (2004) 1434–1438.
[4] M. O'Ryan, The ever-changing landscape of rotavirus serotypes pediatr, Infect. Dis. J. 28 (2009) S60–S62, https://doi.org/10.1097/INF.0b013e3181967c29.
[5] M.K. Estes, J. Cohen, Rotavirus gene structure and function, Microbiol. Mol. Biol. Rev. 53 (1989) 410–449.
[6] E.C. Settembre, J.Z. Chen, P.R. Dormitzer, N. Grigorieff, S.C. Harrison, Atomic model of an infectious rotavirus particle, EMBO J. 30 (2010) 408–416, 10.1038/.
[7] B.V. Prasad, G.J. Wang, J.P. Clerx, W. Chiu, Three-dimensional structure of rotavirus, J. Mol. Biol. 199 (1988) 269–275.
[8] V.N. Uversky, A.K. Dunker, Understanding protein non-folding, Biochim. Biophys. Acta. 1804 (2010) 1231–1264, https://doi.org/10.1016/j.bbapap.2010.01.017.
[9] A. Bhowmick, D.H. Brookes, S.R. Yost, H.J. Dyson, J.D. Forman-Kay, D. Gunter, M. Head-Gordon, G.L. Hura, V.S. Pande, D.E. Wemmer, P.E. Wright, T. Head-Gordon, Finding our way in the dark proteome, J. Am. Chem. Soc. 138 (2016) 9730–9742, https://doi.org/10.1021/jacs.6b06543.
[10] N. Perdigão, J. Heinrich, C. Stolte, K.S. Sabir, M.J. Buckley, B. Tabor, B. Signal, B. S. Gloss, C.J. Hammang, B. Rost, A. Schafferhans, S.I. O'Donoghue, Unexpected features of the dark proteome, Proc. Natl. Acad. Sci. 112 (2015) 15898–15903, https://doi.org/10.1073/pnas.1508380112.
[11] V.N. Uversky, Intrinsically disordered proteins from A to Z, Int. J. Biochem. Cell Biol. 43 (2011) 1090–1103, https://doi.org/10.1016/j.biocel.2011.04.001.
[12] B. Xue, D. Blocquel, J. Habchi, A.V. Uversky, L. Kurgan, V.N. Uversky, S. Longhi, Structural disorder in viral proteins, Chem. Rev. 114 (2014) 6880–6911, https://doi.org/10.1021/cr4005692.
[13] R. Giri, D. Kumar, N. Sharma, V.N. Uversky, Intrinsically disordered side of the zika virus proteome, Front. Cell. Infect. Microbiol. 6 (2016) 144, https://doi.org/10.3389/fcimb.2016.00144.
[14] A. Singh, A. Kumar, R. Yadav, V.N. Uversky, R. Giri, Deciphering the dark proteome of Chikungunya virus, Sci. Rep. 8 (2018) 5822, https://doi.org/10.1038/s41598-018-23969-0.
[15] X. Fan, B. Xue, P.T. Dolan, D.J. LaCount, L. Kurgan, V.N. Uversky, 23713779, Mol. BioSyst. 10 (2014) 1345–1363, https://doi.org/10.1039/C4MB00027G.
[16] B. Xue, M.J. Mizianty, L. Kurgan, V.N. Uversky, Protein intrinsic disorder as a flexible armor and a weapon of HIV-1, Cell. Mol. Life Sci. 69 (2012) 1211–1259, https://doi.org/10.1007/s00018-011-0859-3.
[17] B. Xue, K. Ganti, A. Rabionet, L. Banks, V.N. Uversky, Disordered interactome of human papillomavirus, Curr. Pharm. Des. 20 (2014) 1274–1292.
[18] J.N. Whelan, K.D. Reddy, V.N. Uversky, M.N. Teng, Functional correlations of respiratory syncytial virus proteins to intrinsic disorder, Mol. Biosyst. 12 (2016) 1507–1526, https://doi.org/10.1039/C6MB00122J.
[19] F. Meng, R.A. Badierah, H.A. Almehdar, E.M. Redwan, L. Kurgan, V.N. Uversky, Unstructural biology of the Dengue virus proteins, FEBS J. 282 (2015) 3368–3394, https://doi.org/10.1111/febs.13349.
[20] G.K.-M. Goh, A.K. Dunker, V. Uversky, Prediction of intrinsic disorder in MERS-CoV/HCoV-EMC supports a high oral-fecal transmission, PLoS Curr. (2013), https://doi.org/10.1371/currents.outbreaks.22254b58675cdebc256dbe3c5aa6498b.
[21] E.M. Redwan, A.A. AlJaddawi, V.N. Uversky, Structural disorder in the proteome and interactome of Alkhurma virus (ALKV), Cell. Mol. Life Sci. (2018), https://doi.org/10.1007/s00018-018-2968-8.
[22] M.M. Babu, The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease, Biochem. Soc. Trans. 44 (2016) 1185–1200, https://doi.org/10.1042/BST20160172.
[23] C.J. Oldfield, J. Meng, J.Y. Yang, M.Q. Yang, V.N. Uversky, A.K. Dunker, Flexible nets: disorder and induced fit in the associations of p53 and 14–3-3 with

their partners, BMC Genomics 9 (2008) S1, https://doi.org/10.1186/1471-2164-9-S1-S1.
[24] A. Toto, C. Camilloni, R. Giri, M. Brunori, M. Vendruscolo, S. Gianni, Molecular recognition by templated folding of an intrinsically disordered protein, Sci. Rep. 6 (2016) 21994, https://doi.org/10.1038/srep21994.
[25] D. Kumar, N. Sharma, R. Giri, Therapeutic interventions of cancers using intrinsically disordered proteins as drug targets: C-myc as model system 1176935117699408 Cancer Inform. 16 (2017), https://doi.org/10.1177/1176935117699408.
[26] M. Brunori, S. Gianni, R. Giri, A. Morrone, C. Travaglini-Allocatelli, Morphogenesis of a protein: folding pathways and the energy landscape, Biochem. Soc. Trans. 40 (2012) 429–432, https://doi.org/10.1042/BST20110683.
[27] R. Giri, A. Morrone, A. Toto, M. Brunori, S. Gianni, Structure of the transition state for the binding of c-Myb and KIX highlights an unexpected order for a disordered system, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 14942–14947, https://doi.org/10.1073/pnas.1307337110.
[28] C.J. Oldfield, Y. Cheng, M.S. Cortese, P. Romero, V.N. Uversky, A.K. Dunker, Coupled folding and binding with α-helix-forming molecular recognition elements †, Biochemistry 44 (2005) 12454–12470, https://doi.org/10.1021/bi050736e.
[29] A. Mohan, C.J. Oldfield, P. Radivojac, V. Vacic, M.S. Cortese, A.K. Dunker, V.N. Uversky, Analysis of molecular recognition features (MoRFs), J. Mol. Biol. 362 (2006) 1043–1059, https://doi.org/10.1016/J.JMB.2006.07.087.
[30] Y. Cheng, C.J. Oldfield, J. Meng, P. Romero, V.N. Uversky, A.K. Dunker, Mining?-helix-forming molecular recognition features with cross species sequence alignments?, Biochemistry 46 (2007) 13468–13477, https://doi.org/10.1021/bi7012273.
[31] P.M. Mishra, V.N. Uversky, R. Giri, Molecular recognition features in zika virus proteome, J. Mol. Biol. (2017), https://doi.org/10.1016/j.jmb.2017.10.018.
[32] A. Singh, A. Kumar, V.N. Uversky, R. Giri, Understanding the interactability of chikungunya virus proteins *via* molecular recognition feature analysis, RSC Adv. 8 (2018) 27293–27303, https://doi.org/10.1039/C8RA04760J.
[33] M. Aarthy, D. Kumar, R. Giri, S.K. Singh, E7 oncoprotein of human papillomavirus: Structural dynamics and inhibitor screening study, Gene 658 (2018) 159–177 (accessed March 23, 2018) https://www.sciencedirect.com/science/article/pii/S0378111918302646#f0030.
[34] N. Sharma, A. Murali, S.K. Singh, R. Giri, Epigallocatechin gallate, an active green tea compound inhibits the Zika virus entry into host cells via binding the envelope protein, Int. J. Biol. Macromol. 104 (2017) 1046–1054, https://doi.org/10.1016/J.IJBIOMAC.2017.06.105.
[35] S. López, C.F. Arias, Simian rotavirus SA11 strains, J. Virol. 66 (1992) 1832.
[36] M. Hewish, Y. Takada, B. Coulson, Integrins a2b1 and a4b1 can mediate SA11 rotavirus attachment and entry into cells, J. Virol. (2000), https://doi.org/10.1128/JVI.74.1.228-236.2000.Updated.
[37] P. Isa, C.F. Arias, S. López, Role of sialic acids in rotavirus infection, Glycoconj. J. (2006), https://doi.org/10.1007/s10719-006-5435-y.
[38] Z. Dosztanyi, V. Csizmok, P. Tompa, I. Simon, IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content, Bioinformatics 21 (2005) 3433–3434, https://doi.org/10.1093/bioinformatics/bti541.
[39] Z. Dosztányi, B. Mészáros, I. Simon, ANCHOR: web server for predicting protein binding regions in disordered proteins, Bioinformatics 25 (2009) 2745–2746, https://doi.org/10.1093/bioinformatics/btp518.
[40] F.M. Disfani, W.-L. Hsu, M.J. Mizianty, C.J. Oldfield, B. Xue, A.K. Dunker, V.N. Uversky, L. Kurgan, MoRFpred, a computational tool for sequence-based prediction and characterization of short disorder-to-order transitioning binding regions in proteins, Bioinformatics 28 (2012) i75–i83, https://doi.org/10.1093/bioinformatics/bts209.
[41] N. Malhis, M. Jacobson, J. Gsponer, MoRFchibi SYSTEM: software tools for the identification of MoRFs in protein sequences, Nucleic Acids Res. 44 (2016) W488–W493, https://doi.org/10.1093/nar/gkw409.
[42] R. Linding, L.J. Jensen, F. Diella, P. Bork, T.J. Gibson, R.B. Russell, Protein disorder prediction: implications for structural proteomics, Structure 11 (2003) 1453–1459 (accessed January 17, 2017) http://www.ncbi.nlm.nih.gov/pubmed/14604535.
[43] R. Linding, R.B. Russell, V. Neduva, T.J. Gibson, GlobPlot: Exploring protein sequences for globularity and disorder, Nucleic Acids Res. 31 (2003) 3701–3708, https://doi.org/10.1093/nar/gkg519.
[44] J.J. Ward, J.S. Sodhi, L.J. McGuffin, B.F. Buxton, D.T. Jones, Prediction and functional analysis of native disorder in proteins from the three kingdoms of life, J. Mol. Biol. 337 (2004) 635–645, https://doi.org/10.1016/j.jmb.2004.02.002.
[45] P. Romero, Z. Obradovic, X. Li, E.C. Garner, C.J. Brown, A.K. Dunker, Sequence complexity of disordered protein, Proteins 42 (2001) 38–48 (accessed January 17, 2017) http://www.ncbi.nlm.nih.gov/pubmed/11093259.
[46] Z. Obradovic, K. Peng, S. Vucetic, P. Radivojac, A.K. Dunker, Exploiting heterogeneous sequence properties improves prediction of protein disorder, Proteins Struct. Funct. Bioinforma. 61 (2005) 176–182, https://doi.org/10.1002/prot.20735.
[47] K. Peng, S. Vucetic, P. Radivojac, C.J. Brown, A.K. Dunker, Z. Obradovic, Optimizing long intrinsic disorder predictors with protein evolutionary information, J. Bioinform. Comput. Biol. 3 (2005) 35–60.
[48] B. Xue, R.L. Dunbrack, R.W. Williams, A.K. Dunker, V.N. Uversky, PONDR-FIT: a meta-predictor of intrinsically disordered amino acids, Biochim. Biophys. Acta 1804 (2010) 996–1010, https://doi.org/10.1016/j.bbapap.2010.01.011.
[49] F. Ferron, S. Longhi, B. Canard, D. Karlin, S. Kim, M. Kanehisa, M. Cortese, J. Lawson, C. Brown, J.G. Sikes, C. Newton, A. Dunker, A practical overview of protein disorder prediction methods, Proteins Struct. Funct. Bioinforma. 65 (2006) 1–14, https://doi.org/10.1002/prot.21075.
[50] X. Deng, J. Eickholt, J. Cheng, A comprehensive overview of computational protein disorder prediction methods, Mol. Biosyst. 8 (2012) 114–121, https://doi.org/10.1039/c1mb05207a.
[51] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, A. Bairoch, Protein Identification and Analysis Tools on the ExPASy Server, in: Proteomics Protoc. Handb., 2009. doi:10.1385/1-59259-890-0:571.
[52] K. Peng, P. Radivojac, S. Vucetic, A.K. Dunker, Z. Obradovic, Length-dependent prediction of protein intrinsic disorder, BMC Bioinf. 7 (2006) 208, https://doi.org/10.1186/1471-2105-7-208.
[53] V. Vacic, C.J. Oldfield, A. Mohan, P. Radivojac, M.S. Cortese, V.N. Uversky, A.K. Dunker, Characterization of molecular recognition features, MoRFs, and their binding partners, J. Proteome Res. 6 (2007) 2351–2366, https://doi.org/10.1021/pr0701411.
[54] J.J. Ward, L.J. McGuffin, K. Bryson, B.F. Buxton, D.T. Jones, The DISOPRED server for the prediction of protein disorder, Bioinformatics (2004), https://doi.org/10.1093/bioinformatics/bth195.
[55] N. Malhis, E.T.C. Wong, R. Nassar, J. Gsponer, Computational identification of MoRFs in protein sequences using Hierarchical application of bayes rule, PLoS ONE (2015), https://doi.org/10.1371/journal.pone.0141603.
[56] F. Sievers, A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J.D. Thompson, D.G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, Mol. Syst. Biol. (2011), https://doi.org/10.1038/msb.2011.75.
[57] H. Dinkel, K. Van Roey, S. Michael, N.E. Davey, R.J. Weatheritt, D. Born, T. Speck, D. Krüger, G. Grebnev, M. Kubań, M. Strumillo, B. Uyar, A. Budd, B. Altenberg, M. Seiler, L.B. Chemes, J. Glavina, I.E. Sánchez, F. Diella, T.J. Gibson, The eukaryotic linear motif resource ELM: 10 years and counting, Nucleic Acids Res. (2014), https://doi.org/10.1093/nar/gkt1047.
[58] H. Dinkel, S. Michael, R.J. Weatheritt, N.E. Davey, K. Van Roey, B. Altenberg, G. Toedt, B. Uyar, M. Seiler, A. Budd, L. Jödicke, M.A. Dammert, C. Schroeter, M. Hammer, T. Schmidt, P. Jehl, C. McGuigan, M. Dymecka, C. Chica, K. Luck, A. Via, A. Chatr-Aryamontri, N. Haslam, G. Grebnev, R.J. Edwards, M.O. Steinmetz, H. Meiselbach, F. Diella, T.J. Gibson, ELM–the database of eukaryotic linear motifs, Nucleic Acids Res. 40 (2012) D242–D251, https://doi.org/10.1093/nar/gkr1064.
[59] X. Lu, S.M. McDonald, M.A. Tortorici, Y.J. Tao, R. Vasquez-Del Carpio, M.L. Nibert, J.T. Patton, S.C. Harrison, Mechanism for coordinated RNA packaging and genome replication by rotavirus polymerase VP1, Structure (2008) 1678–1688, https://doi.org/10.1016/j.str.2008.09.006.
[60] A.O. McKell, L.E.W. LaConte, S.M. McDonald, A temperature-sensitive lesion in the N-terminal domain of the rotavirus polymerase affects its intracellular localization and enzymatic activity, J. Virol. 91 (2017), https://doi.org/10.1128/JVI.00062-17.
[61] H. Jayaram, M. Estes, B.V.V. Prasad, Emerging themes in rotavirus cell entry, genome organization, transcription and replication, Virus Res. 101 (2004) 67–81, https://doi.org/10.1016/j.virusres.2003.12.007.
[62] J.T. Patton, M.T. Jones, A.N. Kalbach, Y.W. He, J. Xiaobo, Rotavirus RNA polymerase requires the core shell protein to synthesize the double-stranded RNA genome, J. Virol. 71 (1997) 9618–9626.
[63] S.M. McDonald, J.T. Patton, Rotavirus VP2 core shell regions critical for viral polymerase activation, J. Virol. 85 (2011) 3095–3105, https://doi.org/10.1128/JVI.02360-10.
[64] J.A. Lawton, C.Q.-y. Zeng, S.K. Mukherjee, J. Cohen, M.K. Estes, B.V. Venkataram Prasad, Three-dimensional structural analysis of recombinant rotavirus-like particles with intact and amino-terminal-deleted VP2: Implications for the architecture of the VP2 capsid layer downloaded from, 1997.
[65] M. Liu, P.A. Offit, M.K. Estes, Identification of the simian rotavirus SA11 genome segment 3 product, Virology 163 (1988) 26–32, https://doi.org/10.1016/0042-6822(88)90230-9.
[66] K.M. Ogden, M.J. Snyder, A.F. Dennis, J.T. Patton, Predicted structure and domain organization of rotavirus capping enzyme and innate immune antagonist VP3, J. Virol. 88 (2014) 9072–9085, https://doi.org/10.1128/JVI.00923-14.
[67] T. Brandmann, M. Jinek, Crystal structure of the C-terminal 2′,5′-phosphodiesterase domain of group a rotavirus protein VP3, Proteins Struct. Funct. Bioinforma. 83 (2015) 997–1002, https://doi.org/10.1002/prot.24794.
[68] W. Wang, B. Donnelly, A. Bondoc, S.K. Mohanty, M. McNeal, R. Ward, K. Sestak, S. Zheng, G. Tiao, The rhesus rotavirus gene encoding VP4 Is a major determinant in the pathogenesis of biliary atresia in newborn mice, J. Virol. 85 (2011) 9069–9077, https://doi.org/10.1128/JVI.02436-10.
[69] E. Kindler, E. Trojnar, G. Heckel, P.H. Otto, R. Johne, Analysis of rotavirus species diversity and evolution including the newly determined full-length genome sequences of rotavirus F and G, Infect. Genet. Evol. 14 (2013) 58–67, https://doi.org/10.1016/j.meegid.2012.11.015.
[70] S.E. Crawford, S.K. Mukherjee, M.K. Estes, J.A. Lawton, A.L. Shaw, R.F. Ramig, B. V. Prasad, Trypsin cleavage stabilizes the rotavirus VP4 spike, J. Virol. 75 (2001) 6052–6061, https://doi.org/10.1128/JVI.75.13.6052-6061.2001.
[71] S. Lopez, C.F. Arias, J.R. Bell, J.H. Strauss, R.T. Espejo, Primary structure of the cleavage site associated with trypsin enhancement of rotavirus SA11

infectivity, Virology 144 (1985) 11–19, https://doi.org/10.1016/0042-6822(85)90300-9.
[72] M.K. Estes, D.Y. Graham, B.B. Mason, Proteolytic enhancement of rotavirus infectivity: molecular mechanisms, J. Virol. 39 (1981) 879–888.
[73] P.R. Dormitzer, E.B. Nason, B.V. Venkataram Prasad, S.C. Harrison, Structural rearrangements in the membrane penetration protein of a non-enveloped virus, Nature 430 (2004) 1053–1058, https://doi.org/10.1038/nature02836.
[74] W. Dowling, E. Denisova, R. LaMonica, E.R. Mackow, Selective membrane permeabilization by the rotavirus VP5* protein is abrogated by mutations in an internal hydrophobic domain, J. Virol. 74 (2000) 6368–6376.
[75] Y. Hoshino, L.J. Saif, S.-Y. Kang, M.M. Sereno, W.-K. Chen, A.Z. Kapikian, Identification of Group A rotavirus genes associated with virulence of a porcine rotavirus and host range restriction of a human rotavirus in the gnotobiotic piglet model, Virology 209 (1995) 274–280, https://doi.org/10.1006/viro.1995.1255.
[76] M. Morelli, K.M. Ogden, J.T. Patton, Silencing the alarms: Innate immune antagonism by rotavirus NSP1 and VP3, Virology 479–480 (2015) 75–84, https://doi.org/10.1016/j.virol.2015.01.006.
[77] L. Fiore, H.B. Greenberg, E.R. Mackow, The VP8 fragment of VP4 is the rhesus rotavirus hemagglutinin, Virology 181 (1991) 553–563.
[78] M. Nejmeddine, G. Trugnan, C. Sapin, E. Kohli, L. Svensson, S. Lopez, J. Cohen, Rotavirus spike protein VP4 is present at the plasma membrane and is associated with microtubules in infected cells, J. Virol. 74 (2000) 3313–3320, https://doi.org/10.1128/JVI.74.7.3313-3320.2000.
[79] W. Condemine, T. Eguether, N. Couroussé, C. Etchebest, A. Gardet, G. Trugnan, S. Chwetzoff, The C-terminus of rotavirus VP4 protein contains an actin binding domain which requires co-operation with the coiled-coil domain for actin remodeling, J. Virol. (2018), https://doi.org/10.1128/JVI.01598-18.
[80] H. Greenberg, V. McAuliffe, J. Valdesuso, R. Wyatt, J. Flores, A. Kalica, Y. Hoshino, N. Singh, Serological analysis of the subgroup protein of rotavirus, using monoclonal antibodies, Infect. Immun. 39 (1983) 91–99.
[81] L.L. Clapp, J.T. Patton, Rotavirus morphogenesis: domains in the major inner capsid protein essential for binding to single-shelled particles and for trimerization, Virology 180 (1991) 697–708.
[82] D.I. Ginn, R.L. Ward, V.V. Hamparian, J.H. Hughes, Inhibition of rotavirus in vitro transcription by optimal concentrations of monoclonal antibodies specific for rotavirus VP6, J. Gen. Virol. 73 (1992) 3017–3022, https://doi.org/10.1099/0022-1317-73-11-3017.
[83] J.Z. Chen, E.C. Settembre, S.T. Aoki, X. Zhang, A.R. Bellamy, P.R. Dormitzer, S.C. Harrison, N. Grigorieff, Molecular interactions in rotavirus assembly and uncoating seen by high-resolution cryo-EM, (n.d.).
[84] S.D. Trask, P.R. Dormitzer, Assembly of highly infectious rotavirus particles recoated with recombinant outer capsid proteins, J. Virol. 80 (2006) 11293–11304, https://doi.org/10.1128/JVI.01346-06.
[85] M. Mathieu, I. Petitpas, J. Navaza, J. Lepault, E. Kohli, P. Pothier, B.V.V. Prasad, J. Cohen, F.A. Rey, Atomic structure of the major capsid protein of rotavirus: implications for the architecture of the viron, EMBO J. 20 (2001) 1485–1497, https://doi.org/10.1093/emboj/20.7.1485.
[86] J. Hua, E.A. Mansell, J.T. Patton, Comparative analysis of the rotavirus NS53 gene: Conservation of basic and cysteine-rich regions in the protein and possible stem-loop structures in the RNA, Virology 196 (1993) 372–378, https://doi.org/10.1006/VIRO.1993.1492.
[87] J. Hua, X. Chen, J.T. Patton, Deletion mapping of the rotavirus metalloprotein NS53 (NSP1): the conserved cysteine-rich region is essential for virus-specific RNA binding, J. Virol. 68 (1994) 3990–4000.
[88] K. Taniguchi, K. Kojima, N. Kobayashi, T. Urasawa, S. Urasawa, in: Viral Gastroenteritis, Springer Vienna, Vienna, 1996, pp. 53–58, https://doi.org/10.1007/978-3-7091-6553-9_6.
[89] C. Aponte, D. Poncet, J. Cohen, Recovery and characterization of a replicase complex in rotavirus-infected cells by using a monoclonal antibody against NSP2 downloaded from, J. Virol. 70 (1996) 985–991.
[90] C. Eichwald, J.F. Rodriguez, O.R. Burrone, Characterization of rotavirus NSP2/NSP5 interactions and the dynamics of viroplasm formation, J. Gen. Virol. 85 (2004) 625–634, https://doi.org/10.1099/vir.0.19611-0.
[91] P. Dhillon, V.N. Tandra, S.G. Chorghade, N.D. Namsa, L. Sahoo, C.D. Rao, Cytoplasmic relocalization and colocalization with viroplasms of host cell proteins, and their role in rotavirus infection, J. Virol. 92 (2018), https://doi.org/10.1128/JVI.00612-18.
[92] C. Eichwald, F. Arnoldi, A.S. Laimbacher, E.M. Schraner, C. Fraefel, P. Wild, O.R. Burrone, M. Ackermann, Rotavirus viroplasm fusion and perinuclear localization are dynamic processes requiring stabilized microtubules e47947 PLoS ONE 7 (2012), https://doi.org/10.1371/journal.pone.0047947.
[93] D. Martin, M. Duarte, J. Lepault, D. Poncet, Sequestration of free tubulin molecules by the viral protein NSP2 induces microtubule depolymerization during rotavirus infection, J. Virol. 84 (2010) 2522–2532, https://doi.org/10.1128/JVI.01883-09.
[94] H. Jayaram, Z. Taraporewala, J.T. Patton, B.V.V. Prasad, Rotavirus protein involved in genome replication and packaging exhibits a HIT-like fold, Nature 417 (2002) 311–315, https://doi.org/10.1038/417311a.
[95] F. Arnoldi, M. Campagna, C. Eichwald, U. Desselberger, O.R. Burrone, Interaction of rotavirus polymerase VP1 with nonstructural protein NSP5 is stronger than that with NSP2, J. Virol. 81 (2007) 2128–2137, https://doi.org/10.1128/JVI.01494-06.
[96] Z. Taraporewala, D. Chen, J.T. Patton, Multimers formed by the rotavirus nonstructural protein NSP2 bind to RNA and have nucleoside triphosphatase activity, J. Virol. 73 (1999) 9934–9943.
[97] Z.F. Taraporewala, X. Jiang, R. Vasquez-Del Carpio, H. Jayaram, B.V.V. Prasad, J. T. Patton, Structure-function analysis of rotavirus NSP2 octamer by using a novel complementation system, J. Virol. 80 (2006) 7984–7994, https://doi.org/10.1128/JVI.00172-06.
[98] Z.F. Taraporewala, J.T. Patton, Identification and characterization of the helix-destabilizing activity of rotavirus nonstructural protein NSP2, J. Virol. 75 (2001) 4519–4527, https://doi.org/10.1128/JVI.75.10.4519-4527.2001.
[99] C. Rao, M. Das, P. Ilango, R. Lalwani, B. Rao, K. Gowda, Comparative nucleotide and amino acid sequence analysis of the sequence-specific RNA-binding rotavirus nonstructural protein NSP3, Virology 207 (1995) 327–333, https://doi.org/10.1006/VIRO.1995.1087.
[100] D. Poncet, C. Aponte, J. Cohen, Rotavirus protein NSP3 (NS34) is bound to the 3′ end consensus sequence of viral mRNAs in infected cells, J. Virol. 67 (1993) 3159–3165.
[101] N.M. Mattion, J. Cohen, C. Aponte, M.K. Estes, Characterization of an oligomerization domain and RNA-binding properties on rotavirus nonstructural protein NS34, Virology 190 (1992) 68–83, https://doi.org/10.1016/0042-6822(92)91193-X.
[102] M. Piron, T. Delaunay, J. Grosclaude, D. Poncet, Identification of the RNA-binding, dimerization, and eIF4GI-binding domains of rotavirus nonstructural protein NSP3, J. Virol. 73 (1999) 5411–5421.
[103] R.C. Deo, C.M. Groft, K.R. Rajashankar, S.K. Burley, Recognition of the rotavirus mRNA 3′ consensus by an asymmetric NSP3 homodimer, Cell 108 (2002) 71–81, https://doi.org/10.1016/S0092-8674(01)00632-8.
[104] C.M. Groft, S.K. Burley, Recognition of eIF4G by rotavirus NSP3 reveals a basis for mRNA circularization, Mol. Cell. 9 (2002) 1273–1283, https://doi.org/10.1016/S1097-2765(02)00555-5.
[105] L. Padilla-Noriega, O. Paniagua, S. Guzmán-León, Rotavirus protein NSP3 shuts off host cell protein synthesis, Virology 298 (2002) 1–7, https://doi.org/10.1006/VIRO.2002.1477.
[106] M. Piron, P. Vende, J. Cohen, D. Poncet, Rotavirus RNA-binding protein NSP3 interacts with eIF4GI and evicts the poly(A) binding protein from eIF4F, EMBO J. 17 (1998) 5811–5821, https://doi.org/10.1093/emboj/17.19.5811.
[107] H.I. Contreras-Treviño, E. Reyna-Rosas, R. León-Rodríguez, B.H. Ruiz-Ordaz, T. D. Dinkova, A.M. Cevallos, L. Padilla-Noriega, Species A rotavirus NSP3 acquires its translation inhibitory function prior to stable dimer formation e0181871 PLoS ONE 12 (2017), https://doi.org/10.1371/journal.pone.0181871.
[108] K.S. Au, W.K. Chan, J.W. Burns, M.K. Estes, Receptor activity of rotavirus nonstructural glycoprotein NS28, J. Virol. 63 (1989) 4553–4562.
[109] A. Bugarcic, J.A. Taylor, Rotavirus nonstructural glycoprotein NSP4 is secreted from the apical surfaces of polarized epithelial cells, J. Virol. 80 (2006) 12343–12349, https://doi.org/10.1128/JVI.01378-06.
[110] J.M. Ball, P. Tian, C.Q. Zeng, A.P. Morris, M.K. Estes, Age-dependent diarrhea induced by a rotaviral nonstructural glycoprotein, Science 272 (1996) 101–104.
[111] M.R. Jagannath, M.M. Kesavulu, R. Deepa, P.N. Sastri, S. Senthil Kumar, K. Suguna, C. Durga Rao, N-and C-terminal cooperation in rotavirus enterotoxin: novel mechanism of modulation of the properties of a multifunctional protein by a structurally and functionally overlapping conformational domain, J. Virol. 80 (2006) 412–425, https://doi.org/10.1128/JVI.80.1.412-425.2006.
[112] C. Durga Rao, K. Pamidimukkala, J.R. Marathahalli, S. Kaza, C.D. Rao, Conformational differences unfold a wide range of enterotoxigenic abilities exhibited by rNSP4 peptides from different rotavirus strains, Open Virol. J. 5 (2011) 124–135, https://doi.org/10.2174/1874357901105010124.
[113] D. Rajasekaran, N.P. Sastri, J.R. Marathahalli, S.S. Indi, K. Pamidimukkala, K. Suguna, C.D. Rao, The flexible C terminus of the rotavirus non-structural protein NSP4 is an important determinant of its biological properties, J. Gen. Virol. 89 (2008) 1485–1496, https://doi.org/10.1099/vir.0.83617-0.
[114] W. Yang, M.A. McCrae, The rotavirus enterotoxin (NSP4) promotes re-modeling of the intracellular microtubule network, Virus Res. 163 (2012) 269–274, https://doi.org/10.1016/j.virusres.2011.10.011.
[115] A. Xu, A.R. Bellamy, J.A. Taylor, Immobilization of the early secretory pathway by a virus glycoprotein that binds to microtubules, EMBO J. 19 (2000) 6465–6474, https://doi.org/10.1093/emboj/19.23.6465.
[116] V.N. Uversky, A decade and a half of protein intrinsic disorder: biology still waits for physics, Protein Sci. 22 (2013) 693–724, https://doi.org/10.1002/pro.2261.
[117] H. Huang, F. Schroeder, M.K. Estes, T. Mcpherson, J.M. Ball, Interaction(s) of rotavirus non-structural protein 4 (NSP4) C-terminal peptides with model membranes, Biochem. J. (2004), https://doi.org/10.1042/bj20031789.
[118] A.R. Chacko, M. Arifullah, N.P. Sastri, J. Jeyakanthan, G. Ueno, K. Sekar, R.J. Read, E.J. Dodson, D.C. Rao, K. Suguna, Novel pentameric structure of the diarrhea-inducing region of the rotavirus enterotoxigenic protein NSP4, J. Virol. 85 (2011) 12721–12732, https://doi.org/10.1128/JVI.00349-11.
[119] A.R. Chacko, J. Jeyakanthan, G. Ueno, K. Sekar, C.D. Rao, E.J. Dodson, K. Suguna, R.J. Read, A new pentameric structure of rotavirus NSP4 revealed by molecular replacement, Acta Crystallogr. Sect. D Biol. Crystallogr. 68 (2012) 57–61, https://doi.org/10.1107/S0907444911049705.
[120] R. Deepa, C. Durga Rao, K. Suguna, Structure of the extended diarrhea-inducing domain of rotavirus enterotoxigenic protein NSP4, Arch. Virol. 152 (2007) 847–859, https://doi.org/10.1007/s00705-006-0921-x.
[121] G.D. Bowman, I.M. Nodelman, O. Levy, S.L. Lin, P. Tian, T.J. Zamb, S.A. Udem, B. Venkataraghavan, C.E. Schutt, Crystal structure of the oligomerization

domain of NSP4 from rotavirus reveals a core metal-binding site, J. Mol. Biol. 304 (2000) 861–871, https://doi.org/10.1006/JMBI.2000.4250.
[122] S. Kumar, R. Ramappa, K. Pamidimukkala, C.D. Rao, K. Suguna, New tetrameric forms of the rotavirus NSP4 with antiparallel helices, Arch. Virol. 163 (2018) 1531–1547, https://doi.org/10.1007/s00705-018-3753-6.
[123] R.D. Parr, S.M. Storey, D.M. Mitchell, A.L. McIntosh, M. Zhou, K.D. Mir, J.M. Ball, The rotavirus enterotoxin NSP4 directly interacts with the caveolar structural protein caveolin-1, J. Virol. 80 (2006) 2842–2854, https://doi.org/10.1128/JVI.80.6.2842-2854.2006.
[124] S.K. Welch, S.E. Crawford, M.K. Estes, Rotavirus SA11 genome segment 11 protein is a nonstructural phosphoprotein, J. Virol. 63 (1989) 3974–3982.
[125] S.A. González, O.R. Burrone, Rotavirus NS26 is modified by addition of single O-linked residues of N-acetylglucosamine, Virology 182 (1991) 8–16.
[126] I. Afrikanova, M.C. Miozzo, S. Giambiagi, O. Burrone, Phosphorylation generates different forms of rotavirus NSP5, J. Gen. Virol. 77 (1996) 2059–2065, https://doi.org/10.1099/0022-1317-77-9-2059.
[127] C. Eichwald, F. Vascotto, E. Fabbretti, O.R. Burrone, Rotavirus NSP5: mapping phosphorylation sites and kinase activation and viroplasm localization domains, J. Virol. 76 (2002) 3461–3470.
[128] I. Afrikanova, E. Fabbretti, M.C. Miozzo, O.R. Burrone, Printed in great britain rotavirus NSP5 phosphorylation is up-regulated by interaction with NSP2, J. Gen. Virol. 79 (2018) 2679–2686.
[129] R. Contin, F. Arnoldi, M. Campagna, O.R. Burrone, Rotavirus NSP5 orchestrates recruitment of viroplasmic proteins, J. Gen. Virol. 91 (2010) 1782–1793, https://doi.org/10.1099/vir.0.019133-0.
[130] K.V.K. Mohan, J. Muller, I. Som, C.D. Atreya, The N- and C-terminal regions of rotavirus NSP5 are the critical determinants for the formation of viroplasm-like structures independent of NSP2, J. Virol. 77 (2003) 12184–12192, https://doi.org/10.1128/JVI.77.22.12184-12192.2003.
[131] J. Blackhall, A. Fuentes, K. Hansen, G. Magnusson, Serine protein kinase activity associated with rotavirus, Phosphoprotein NSP5 (1997).
[132] T.R. Fuerst, E.G. Niles, F.W. Studier, B. Moss, O.R. Burrone, Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase, Proc. Natl. Acad. Sci. U. S. A. 83 (1986) 8122–8126, https://doi.org/10.1073/pnas.83.21.8122.
[133] P. Dhillon, C. Durga Rao, Rotavirus induces formation of remodeled stress granules and P-bodies and their sequestration in viroplasms to promote progeny virus production, J. Virol. (2018), https://doi.org/10.1128/JVI.01363-18.
[134] C. Rancurel, M. Khosravi, A.K. Dunker, P.R. Romero, D. Karlin, Overlapping genes produce proteins with unusual sequence properties and offer insight into de novo protein creation, J. Virol. 83 (2009) 10719–10736, https://doi.org/10.1128/JVI.00595-09.
[135] B. Xue, R.W. Williams, C.J. Oldfield, G.K.-M. Goh, A.K. Dunker, V.N. Uversky, Viral disorder or disordered viruses: do viral proteins possess unique features?, Protein Pept. Lett. 17 (2010) 932–951.
[136] D.B. Mitchell, G.W. Both, Simian rotavirus SA11 segment 11 contains overlapping reading frames, Nucleic Acids Res. 16 (1988) 6244.
[137] E.M. Heiman, S.M. McDonald, M. Barro, Z.F. Taraporewala, T. Bar-Magen, J.T. Patton, Group A human rotavirus genomics: evidence that gene constellations are influenced by viral protein interactions, J. Virol. (2008), https://doi.org/10.1128/jvi.01402-08.
[138] E.W. Rainsford, M.A. McCrae, Characterization of the NSP6 protein product of rotavirus gene 11, Virus Res. 130 (2007) 193–201, https://doi.org/10.1016/j.virusres.2007.06.011.
[139] G. Holloway, R.I. Johnson, Y. Kang, V.T. Dang, D. Stojanovski, B.S. Coulson, Rotavirus NSP6 localizes to mitochondria via a predicted N-terminal a-helix, 2018, doi:10.1099/jgv.0.000294.
[140] S. Komoto, Y. Kanai, S. Fukuda, M. Kugita, T. Kawagishi, N. Ito, M. Sugiyama, Y. Matsuura, T. Kobayashi, K. Taniguchi, Reverse genetics system demonstrates that rotavirus nonstructural protein NSP6 is not essential for viral replication in cell culture, J. Virol. 91 (2017), https://doi.org/10.1128/JVI.00695-17.
[141] R.E. Lloyd, Nuclear proteins hijacked by mammalian cytoplasmic plus strand RNA viruses, Virology 479–480 (2015) 457–474, https://doi.org/10.1016/j.virol.2015.03.001.
[142] D. Walsh, I. Mohr, Viral subversion of the host protein synthesis machinery, Nat. Rev. Microbiol. 9 (2011) 860–875, https://doi.org/10.1038/nrmicro2655.
[143] M.L. Yarbrough, M.A. Mata, R. Sakthivel, B.M.A. Fontoura, Viral subversion of nucleocytoplasmic trafficking, Traffic 15 (2014) 127–140, https://doi.org/10.1111/tra.12137.
[144] S. La Frazia, A. Ciucci, F. Arnoldi, M. Coira, P. Gianferretti, M. Angelini, G. Belardo, O.R. Burrone, J.-F. Rossignol, M.G. Santoro, Thiazolides, a new class of antiviral agents effective against rotavirus infection, target viral morphogenesis, inhibiting viroplasm formation, J. Virol. 87 (2013) 11096–11106, https://doi.org/10.1128/JVI.01213-13.