Intrinsically Disordered Side of the Zika Virus Proteome

Rajanish Giri[1]*, Deepak Kumar[1], Nitin Sharma[1] and Vladimir N. Uversky[2,3]*

[1] School of Basic Sciences, Indian Institute of Technology Mandi, Mandi, India, [2] Department of Molecular Medicine and Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, FL, USA, [3] Laboratory of Structural Dynamics, Stability and Folding of Proteins, Institute of Cytology, Russian Academy of Sciences, Saint Petersburg, Russia

Over the last few decades, concepts of protein intrinsic disorder have been implicated in different biological processes. Recent studies have suggested that intrinsically disordered proteins (IDPs) provide structural plasticity and functional diversity to viral proteins that are involved in rapid replication and immune evasion in host cells. In case of Zika virus, the roles of protein intrinsic disorder in mechanisms of pathogenesis are not completely understood. In this study, we have analyzed the prevalence of intrinsic disorder in Zika virus proteome (strain MR 766). Our analyses revealed that Zika virus polyprotein is enriched with intrinsically disordered protein regions (IDPRs) and this finding is consistent with previous reports on the involvement of IDPs in shell formation and virulence of the *Flaviviridae* family. We found abundant IDPRs in Capsid, NS2B, NS3, NS4A, and NS5 proteins that are involved in mature particle formation and replication. In our view, the intrinsic disorder-focused analysis of ZIKV proteins could be important for the development of disorder-based drugs.

Keywords: Zika virus, intrinsically disordered proteins (IDPs), intrinsically disordered protein regions (IDPRs), molecular machinery, viral proteome, cellular proteome

INTRODUCTION

In 1947, Zika virus (ZIKV) was first identified in Uganda through a monitoring network of sylvatic yellow fever in rhesus monkeys (Dick et al., 1952). Outbreaks of ZIKV-related disease have been recorded throughout southern Africa with a high number of birth defects and abnormalities including microcephaly, intracranial calcification, and fetal death (Petersen et al., 2016). As many other members the *Flaviviridae* family, ZIKV is an arbovirus transmitted through the infected arthropods (by the bites of the infected mosquitoes from the *Aedes* genus, *Ae. aegypti* and *Ae. albopictus*). Therefore, distribution of Zika infection is mainly associated with the distribution of *Aedes* mosquito vectors that can be found in different parts of the world (Wikan and Smith, 2016). World Health Organization (WHO) has declared Zika related problems as public health emergency of international concern. Therefore, it is of utmost urgency to find out mechanism of pathogenesis of this virus and to develop therapeutics. A recent study on immunocompetent mouse model has strengthened the previous observations that ZIKV infection might cause neurological defects in fetuses (Lazear et al., 2016). This virus is transmitted through mosquitos, as well as via blood transfusion and also from mother to fetus during pregnancy (Wikan and Smith, 2016). Reports also suggest the possibility of sexual transmission (Grischott et al., 2016).

ZIKV belongs to the *Flaviviridae* family, genus *Flavivirus*, which includes several important human pathogens, such as West Nile virus (WNV), Dengue virus (DENV), Yellow fever virus

(YFV), and Japanese encephalitis virus (JEV). Recently, ZIKV structure has been solved by cryo-electron microscopy (Sirohi et al., 2016). Genome of this virus includes a single-stranded RNA consisting 10794 bases along with two non-coding regions known as the 5′ NCR and the 3′ NCR. The open reading frame (ORF) of the ZIKV, concerning the protein expression order is as follows: 5′-C-prM-E-NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5-3′. It codes for a single polyprotein that is posttranslationally cleaved into three structural proteins (Capsid (C), Precursor membrane (prM) protein, and Envelope (E) protein), and seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) (see **Figure 1**).
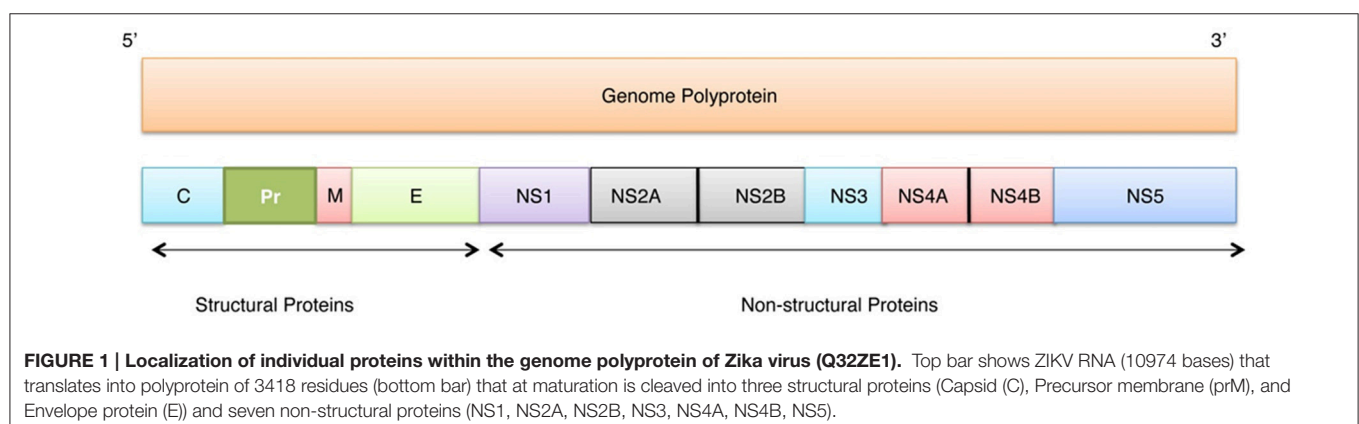
The major focus of this article is on intrinsic disorder-centered analysis of ZIKV proteome. In general, intrinsically disordered proteins (IDPs) are recently recognized class of proteins that lack stable three-dimensional structure in their native state but still functional. Structurally, these proteins are highly heterogeneous and include random coils, pre-molten globules, molten globules, proteins with large flexible linkers, and hybrid proteins containing ordered and disordered regions (Wright and Dyson, 2015). Lack of structure in IDPs and intrinsically disordered protein regions (IDPRs) allow the interaction of proteins with several partners, thereby regulating multiple signaling pathways. These multiple interactions are attributed to fast binding kinetics of IDPs/IDPRs that is regulated through coupled folding and binding mechanisms (Gianni et al., 2012). Disordered regions provide greater capture radii that increase the probability to interact with partners. This mechanism is known as the fly casting mechanism (Shoemaker et al., 2000). Commonly, viruses have highly compact genome and contain disordered protein regions. This could be one of the reasons for high mutagenic capacity of viruses (Xue et al., 2014). Recent studies have reported that core proteins of the viruses from the *Flaviviridae* family contain significant amount of IDPRs (Xue et al., 2014; Goh et al., 2016). A functional correlation between intrinsic disorder and protein function has been established in proteomes of other *Flaviviridae* family members, such as Dengue Virus (Meng et al., 2015) and Hepatitis C Virus (HCV) (Fan et al., 2014).

In our view, the intrinsic disorder-focused analysis of ZIKV proteins could be important for the development of new disorder-based drug strategies (Cheng et al., 2006; Uversky, 2010, 2012). Current drug development strategies have shown to utilize combination of conventional drug design and computational approaches to target dynamic ensembles of IDPs (Ambadipudi and Zweckstetter, 2016). In the current work, we have analyzed the penetrance of intrinsic disorder in the ZIKV proteome. Further, we have correlated the abundance of structural disorder with functionality of ZIKV proteins. This study provides a novel direction for elucidating the mechanism of virus-host interaction. Some proteins have been already used as drug targets in other flaviviruses, such as NS3 helicase, envelope glycoproteins, NS2B-NS3 (serine protease) and NS5 (RNA-directed RNA Polymerase) (Li et al., 2008; Poh et al., 2009; Mayhoub et al., 2011). Therefore, results of our current work should be considered before inhibitor designing.

MATERIALS AND METHODS

Reviewed and experimentally validated polyprotein sequence (UniProt ID: Q32ZE1) of the Zika virus strain Mr766 was used for the disorder analysis. There are several protein intrinsic disorder predictors developed, such as multiple members of the PONDR® family [e.g., PONDR® FIT (Xue et al., 2010), PONDR® VLXT (Romero et al., 2001), and PONDR® VSL2 (Obradovic et al., 2005), IUPred (Dosztanyi et al., 2005a), GlobPlot (Linding et al., 2003b), DisoPred (Ward et al., 2004), SPRITZ (Vullo et al., 2006), DisEMBL (Linding et al., 2003a)], etc. Many of these predictors have been assessed for accuracy within the frames of the Critical Assessment of Protein Structure Prediction (CASP) (Deng et al., 2012). Since many of these predictors are considering phenomenon of intrinsic from different angles, it is advisable to use several computational tools while looking for the abundance of intrinsic disorder in query proteins. Therefore, PONDR® FIT (Xue et al., 2010), PONDR® VLXT (Romero et al., 2001), PONDR® VSL2 (Obradovic et al., 2005), and IUPred (Dosztanyi et al., 2005a) were used in our study for disorder prediction in polyprotein of ZIKV. We also extended our analysis over each individual protein derived from the ZIKV polyprotein. Here, scores above 0.5 are considered to correspond to the disordered residues/regions. PONDR® VSL2B is one of the more accurate stand-alone disorder predictors

FIGURE 1 | Localization of individual proteins within the genome polyprotein of Zika virus (Q32ZE1). Top bar shows ZIKV RNA (10974 bases) that translates into polyprotein of 3418 residues (bottom bar) that at maturation is cleaved into three structural proteins (Capsid (C), Precursor membrane (prM), and Envelope protein (E)) and seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5).

(Peng et al., 2005; Peng and Kurgan, 2012; Fan and Kurgan, 2014), PONDR® VLXT is known to have high sensitivity to local sequence peculiarities and can be used for identifying disorder-based interaction sites (Dunker et al., 2001), whereas a metapredictor PONDR-FIT is moderately more accurate than each of the component predictors (Xue et al., 2010), PONDR® VLXT (Dunker et al., 2001), PONDR® VSL2 (Peng et al., 2005), PONDR® VL3 (Peng et al., 2006), FoldIndex (Prilusky et al., 2005), IUPred (Dosztanyi et al., 2005a), TopIDP (Campen et al., 2008). IUPred was designed to recognize IDPRs from the amino acid sequence alone based on the estimated pairwise energy content, where it was hypothesized that globular proteins are composed of amino acids which have the potential to form a large number of favorable interactions, whereas IDPs/IDPRs do not have unique 3D structure because their amino acid composition does not allow sufficient favorable interactions to form (Dosztanyi et al., 2005a,b).

Often, IDPs/IDPRs are involved in protein-protein interactions and molecular recognitions (Dunker et al., 2002a,b, 2008; Tompa, 2002; Oldfield et al., 2005; Dunker and Uversky, 2008; Uversky and Dunker, 2010; Uversky, 2013b). There are numerous reports emphasizing that IDPs/IDPRs are able to undergo at least partial disorder-to-order transitions upon binding, which is crucial for recognition, regulation, and signaling. Among these potential functional sites are short order-prone motifs within long disordered regions that are able to undergo disorder-to-order transition during the binding to a specific partner. These motifs are known as molecular recognition feature (MoRF), and they can be identified computationally (Oldfield et al., 2005; Cheng et al., 2007). We used ANCHOR algorithm to identify potential disorder-based binding sites (Dosztanyi et al., 2009; Meszaros et al., 2009). This approach relies on the pairwise energy estimation approach developed for the general disorder prediction method IUPred (Dosztanyi et al., 2005a,b), being based on the hypothesis that long regions of disorder contain localized potential binding sites that cannot form enough favorable intrachain interactions to fold on their own, but are likely to gain stabilizing energy by interacting with a globular protein partner (Dosztanyi et al., 2009; Meszaros et al., 2009).

RESULTS AND DISCUSSION

Intrinsic Disorder in ZIKV Polyprotein

In crystallography-based protein structure characterization it is assumed that the disordered regions cannot crystalize, being present in the form of regions with missing electron density, and only ordered regions have the propensity for crystal formation (Uversky, 2013a). In addition to this, crystallization conditions frequently contain various additives (presence of PEG, high salt concentrations, etc.), which make these conditions to be different from the natural environment. Therefore, a computational analysis of disorder based on the amino acid sequence alone using various programs may provide a great advantage to analyze the disorder in proteins (Uversky, 2013a).

Despite being a big threat, the holistic understanding of ZIKV proteins in both ordered and disordered perspective has not been established as of yet. Despite obvious interest to Zika virus, crystallographic data are currently available only for four ZIKV proteins, NS1, NS2B-NS3 protease (residues 49 to 95 of NS2B covalently linked via Gly$_4$-Ser-Gly$_4$ to the N-terminal protease domain (residues 1 to 170) of NS3), NS3 (only Helicase domain), M and E proteins (see below).

In this study we have computationally evaluated the predisposition of ZIKV polyprotein for intrinsic disorder (see **Figure 2**) and also studied intrinsic disorder propensity of all individual proteins derived from this polyprotein: Capsid protein C (residues 2-122, which includes protein C (residues 2-104) and the ER anchor for the protein C (residues 105-122), which is removed in mature form by serine protease NS3), precursor membrane protein prM (residues 123-290, which is further divided to peptide pr (residues 123-215) and small envelope protein M (residues 216-290)), an envelope protein E (residues 291-790), and seven non-structural proteins: NS1 (residues 791-1142), NS2A (residues 1143-1368), serine protease subunit NS2B (residues 1369-1498), serine protease NS3 (residues 1499-2115),
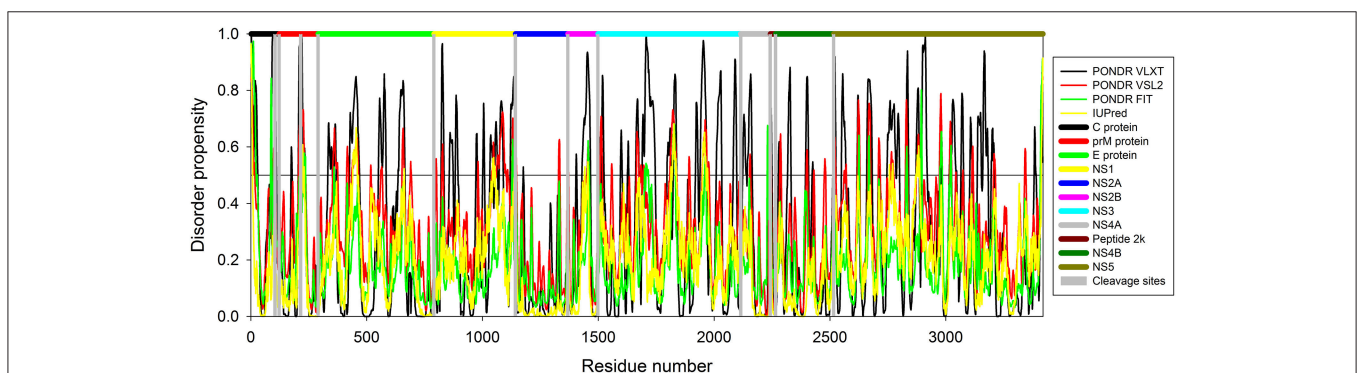
FIGURE 2 | Intrinsic disorder predisposition of Zika virus polyprotein (Q32ZE1). Disorder propensity is evaluated by PONDR® VLXT (black line), PONDR® VSL2 (red line), PONDR® FIT (green line) and IUPred (yellow line). Colored bars at the top of plot shows localization of the individual proteins. Corresponding cleavage sites leading for the generation of mature individual proteins are shown by gray bars. Disorder scores above the threshold 0.5 characterize residues/regions predicted to be disordered.

NS4A (residues 2116-2242), peptide 2k (residues 2243-2265), NS4B (2266-2516), and RNA-directed RNA polymerase NS5 (residues 2517-3419). It is recognized that disordered regions provide structural flexibility leading to binding promiscuity often involved in cellular regulations (Xue et al., 2014; Wright and Dyson, 2015). Often, disordered or flexible regions contain sites of proteolytic cleavage, since proteolytic digestion is known to occur much faster in unstructured than in structured protein regions (Fontana et al., 1986, 2004; Novotny and Bruccoleri, 1987; Iakoucheva et al., 2001). This suggests that the sites of preferential cleavage should be preferentially located within the regions that lack stable structure or possess high structural flexibility. Earlier, this specific use of intrinsic disorder/flexibility for generation of mature viral proteins from the polyprotein was reported for two other representatives of the *Flaviviridae* family, HCV (Fan et al., 2014) and Dengue Virus (Meng et al., 2015). In agreement with these considerations, **Figure 2** shows that, within the polyprotein, cleavage sites leading to the generation of mature ZIKV proteins are preferentially located within the disordered or flexible regions or at least in close proximity to the regions with increased flexibility. **Figure 3** further zooms into this phenomenon and shows that in the vicinity of the vertical gray lines that correspond to the cleavage sites, the per-residue disorder propensity scores evaluated by at least one of the disorder predictors used in this study typically spike to relatively high values.

Analysis of the IDPRs in ZIKV Structural Proteins

C Protein

Structural proteins in ZIKV consist of Capsid, prM and Envelope proteins. Capsid protein contains 102 residues and forms icosahedral capsid (30 nm in diameter) of the virus (Kuno and Chang, 2007), where the genomic RNA of ZIKV is encapsulated. Disordered regions of capsid protein include its N- and C-termini, with an overall predicted percent of intrinsic disorder (PPID) of 33.3% calculated from outputs of four predictors used in our study (**Table 1** and **Figure 4A**).

These observations are consistent with the results of previous studies on the disorder predisposition of proteins in flaviviruses (Goh et al., 2016). *In vitro* studies have implicated the role of intrinsic disorder for chaperone-like activities, such as viral genome packaging (Ivanyi-Nagy et al., 2008). In the case of Dengue virus, detailed functional study of capsid protein revealed that its N-terminal disordered region is responsible for carrying out multiple interactions necessary for mature virus particle formation (Martins et al., 2012). Furthermore, disordered N-terminal region interacts with phospholipids of lipid droplets (Martins et al., 2012). Similar studies have correlated the role of capsid disorder with the diverse functions in other flaviviruses, such as YFV and WNV.(Ivanyi-Nagy and Darlix, 2010) In other flaviviruses, it was shown that high virulence is correlated with the disorder levels of capsid protein (Goh et al., 2016). We assume that high abundance of disorder in capsid of ZIKV may provide an insight to uncover the mechanism of pathogenesis of this virus.

prM Protein

Next to capsid is the protein known as prM that acts as a chaperone for envelope protein. prM protein shows a central role in transition of immature virus particle to mature form, which is infectious, virulent, fusogenic, and can adhere to host cell membrane (Zhang et al., 2003). Immature viral particles of flaviviruses are noninfectious and are characterized by their "spiky" shape, possessing 60 trimeric E-prM heterodimer spikes (Zhang et al., 2003). The mature viral particle is smooth and contains 90 dimeric E:M heterodimers (Kuhn et al., 2002; Zhang et al., 2013). The low-pH environment of the trans-Golgi network is crucial for the maturation of viral particles, since it leads to the conformational changes of the surface glycoproteins needed for the cleavage of prM by the host protease furin to generate the pr peptide and mature protein M (Kuno and Chang, 2007). Subsequent removal of the pr peptide leads to the exposure of the ~12–amino acid-long fusion loop on the E protein, which, in the immature virus, is protected by the pr peptide (Yu et al., 2008). With exposed fusion loop, the virus is prepared for the low pH–mediated endosomal fusion (Yu et al., 2008). Therefore,

TABLE 1 | Some physicochemical and intrinsic disorder properties of Zika virus proteins.

Protein name	Length (M.W., kDa)	pI	PPID$_{VLXT}$	PPID$_{VSL2}$	PPID$_{FIT}$	PPID$_{IUPred}$	PPID$_{mean}$
C	103 (11.73)	12.0	40.8	34.0	36.9	12.6	33.3
prM	168 (19.01)	8.55	23.8	22.6	16.7	10.1	19.0
Pr	93 (10.51)	6.05	20.4	17.2	14.0	6.6	16.1
M	75 (8.51)	10.11	28.0	29.3	20.0	14.7	24.0
E	500 (54.09)	6.48	27.4	18.0	3.8	4.6	7.6
NS1	352 (40.08)	6.17	33.0	26.7	7.1	6.0	10.8
NS2A	226 (23.97)	10.34	7.5	7.5	9.7	3.1	5.7
NS2B	130 (13.77)	4.44	37.7	14.6	17.7	7.7	16.2
NS3	617 (68.41)	8.22	36.6	18.2	9.4	9.2	12.8
NS4A	127 (13.70)	5.82	40.2	12.6	23.6	6.3	16.5
NS4B	251 (26.94)	9.10	19.1	12.4	10.4	4.0	6.4
NS5	903 (103.02)	8.67	33.6	16.1	7.0	3.8	8.6

both proteins eventually generated from the prM are functionally important, where Pr peptide protects the fusion loop of the E protein, and M protein acts as a transmembrane protein in the mature viral particle (Yu et al., 2008).

ZIKV prM contains 168 amino acids. In addition to the disordered N-and C-termini there is a long disordered central region in prM located between residues 80 and 120. This region contains the furin cleavage site (see **Figure 4B** and **Figure 3**). A multitool computational disorder analysis shows that the overall disorder of prM is 19.0% (see **Table 1**). Since the proteolytic cleavage of prM generates two functionally important proteins, the pr peptide and M protein (Yu et al., 2008).The analysis of both proteins separately revealed the PPID of 16.1% and 24% for pr and M proteins respectively (see **Table 1**). Despite high level of predicted disorder, M protein is able to form stable structure, being complexed with the E protein (see **Figure 4D**). Structurally, ZIKV M protein is characterized by N-terminal soluble loop (M loop) which contain two short α-helices (residues 6–10 and 21–39) and two transmembrane α-helices (residues 40–52 and 56–71) connected by very short loop that forms the stem and the transmembrane part of this protein embedded in the lipid bilayer. (Liu et al., 2004; Sirohi et al., 2016). Curiously, although the soluble M loop is predicted to be mostly disordered and contains little regular secondary structure in the E-M complex, it is crucial for stabilization of the E-M dimer, being intercalated into the E protein structure. Therefore, it is likely that this region of the M protein undergoes functional disorder-to-order transition.

E Protein

At the next step, we analyzed the envelope protein E (500 residues) that participates in the membrane fusion between host late endosomes and virion. Protein E heterodimerizes with M, and this complex stabilizes E protein likely due to the chaperone-like activity of M (Hamel et al., 2015). The averaged PPID of 7.6% is observed in E protein (**Table 1**), which is rather low. However, several IDPRs of different length are predicted in the ZIKV E protein (**Figure 4C**). Viral capsid consists of 180 copies each of the E glycoprotein and the M protein anchored in a lipid membrane (Sirohi et al., 2016). Structurally, E protein consists of four domains; stem transmembrane domain and three ectodomains t external surface of the viral capsid and are mostly β-structural (see **Figure 4D**).

Previous studies on other flaviviruses (e.g., Dengue virus and West Nile virus) demonstrated that glycosylation of the E protein
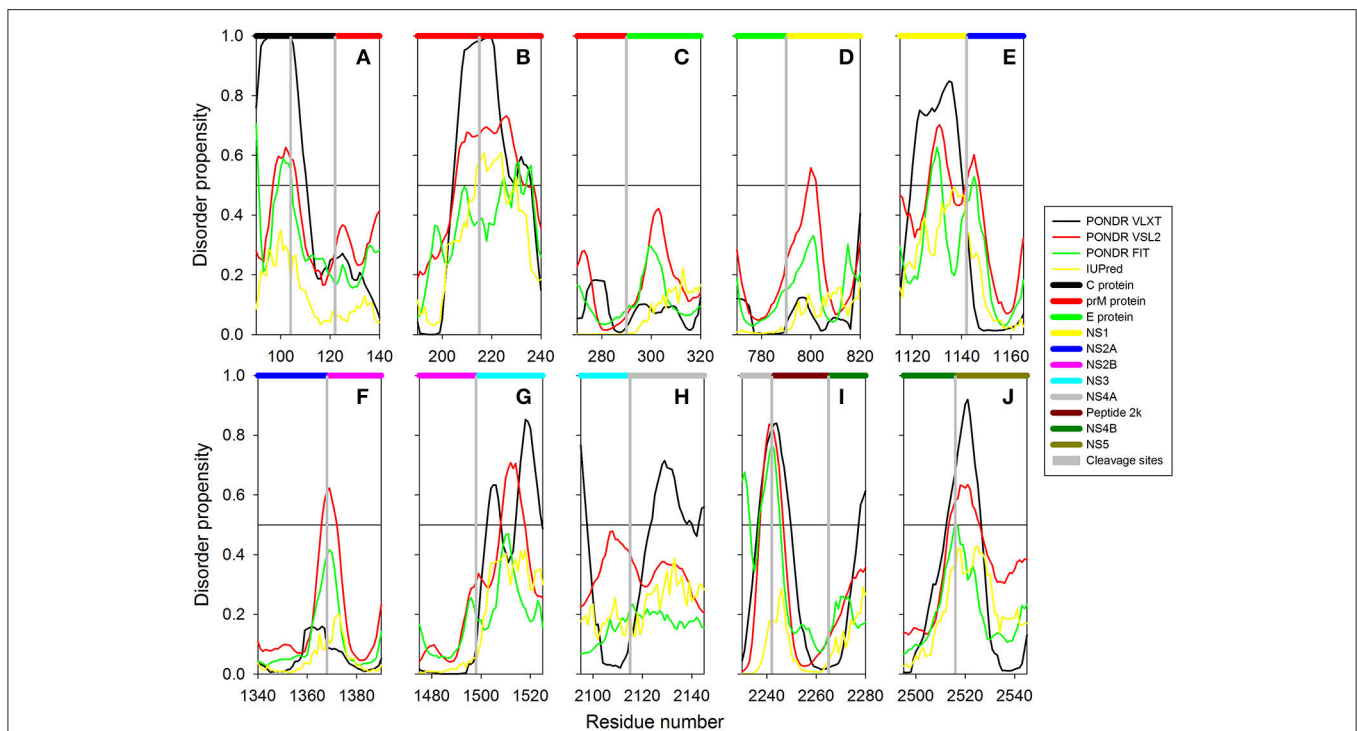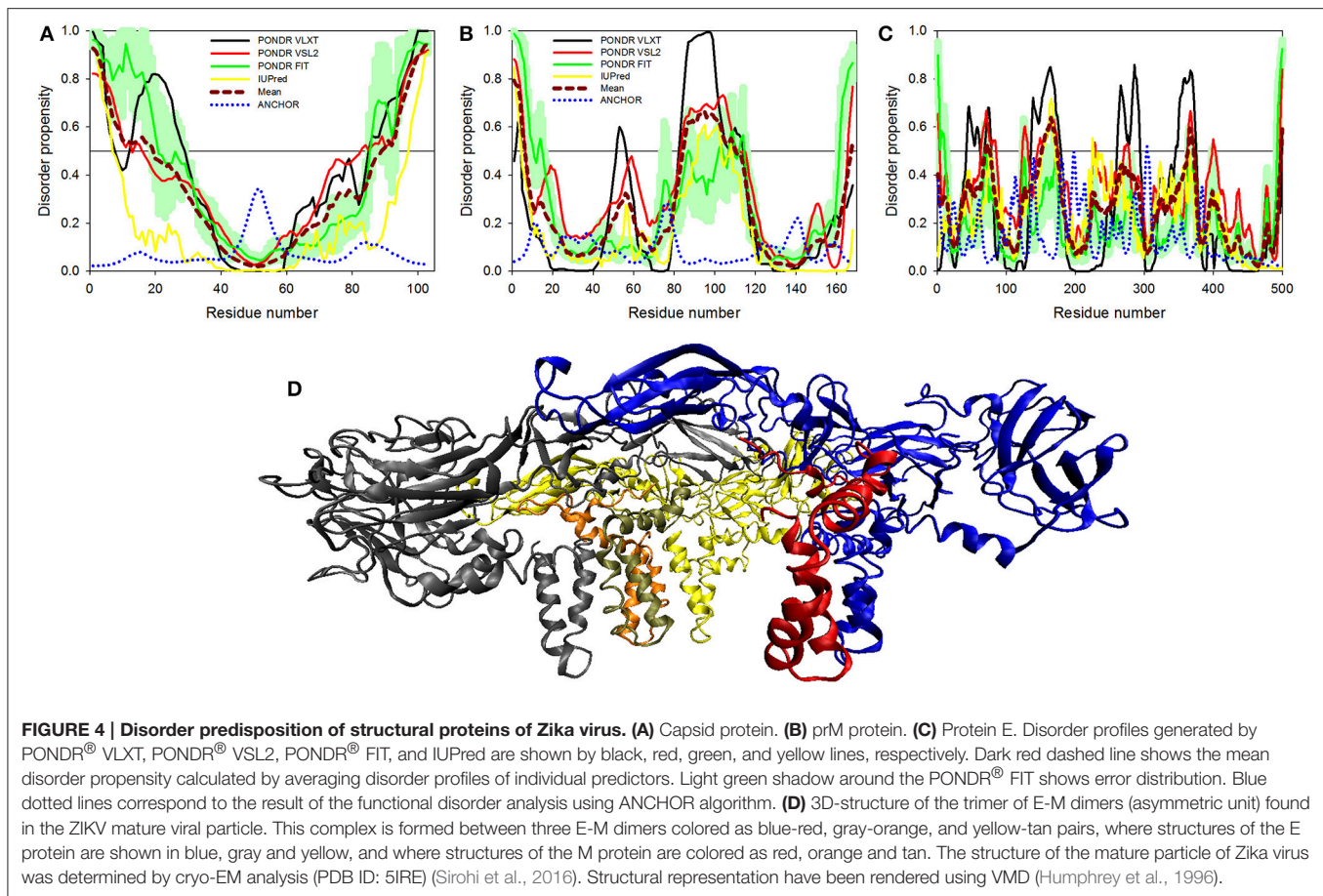

FIGURE 3 | The role of intrinsic disorder in maturation of individual proteins of Zika virus. Plots shows position of cleavage sites (gray vertical bars) in relation to disorder profiles at the junction between the individual proteins within the polyprotein. **(A)** A cleavage site between proteins C (black horizontal bar) and prM (red horizontal bar). A cleavage site at the position 104 within the pro-protein C leading to the removal of propeptide (residues 105–122) is also shown. **(B)** A cleavage site within the prM protein leading to generation of proteins Pr and M. **(C)** A cleavage site between the proteins prM (red horizontal bar) and E (green horizontal bar). **(D)** A cleavage site between the proteins E (green horizontal bar) and NS1 (yellow horizontal bar). **(E)** A cleavage site between the proteins NS1 (yellow horizontal bar) and NS2A (blue horizontal bar). **(F)** A cleavage site between the proteins NS2A (blue horizontal bar) and NS2B (pink horizontal bar). **(G)** A cleavage site between the proteins NS2B (pink horizontal bar) and NS3 (cyan horizontal bar). **(H)** A cleavage site between the proteins NS3 (pink horizontal bar) and NS4A (gray horizontal bar). **(I)** Cleavage sites between the protein NS4A (gray horizontal bar) and the peptide 2k (dark red horizontal bar) and the peptide 2k (dark red horizontal bar) and protein NS4B (dark green horizontal bar). **(J)** A cleavage site between the proteins NS4B (dark green horizontal bar) and NS5 (dark yellow horizontal bar).

FIGURE 4 | Disorder predisposition of structural proteins of Zika virus. (A) Capsid protein. **(B)** prM protein. **(C)** Protein E. Disorder profiles generated by PONDR® VLXT, PONDR® VSL2, PONDR® FIT, and IUPred are shown by black, red, green, and yellow lines, respectively. Dark red dashed line shows the mean disorder propensity calculated by averaging disorder profiles of individual predictors. Light green shadow around the PONDR® FIT shows error distribution. Blue dotted lines correspond to the result of the functional disorder analysis using ANCHOR algorithm. **(D)** 3D-structure of the trimer of E-M dimers (asymmetric unit) found in the ZIKV mature viral particle. This complex is formed between three E-M dimers colored as blue-red, gray-orange, and yellow-tan pairs, where structures of the E protein are shown in blue, gray and yellow, and where structures of the M protein are colored as red, orange and tan. The structure of the mature particle of Zika virus was determined by cryo-EM analysis (PDB ID: 5IRE) (Sirohi et al., 2016). Structural representation have been rendered using VMD (Humphrey et al., 1996).

at specific sites provides the ability to attach to different cell types (Beasley et al., 2005; Pokidysheva et al., 2006; Miller et al., 2008). Cryo-EM structure of mature ZIKV viral particle revealed that the glycosylation site is located at the loop region nearby the fusion peptide (Sirohi et al., 2016). Besides being needed for pH-mediated endosomal fusion, the ZIKV fusion loop (residues 98-110) can act as an epitope and react with different antibodies. **Figure 4C** shows that although this region is predicted to be mostly ordered, it has some weak tendency for the presence of disorder-based binding sites as evidenced by the output of the ANCHOR algorithm. In fact, the ANCHOR-based analysis of protein E revealed the absence of disordered binding regions that function via undergoing a disorder-to-order transition upon binding to a globular protein partner. However, the fusion loop is located in close proximity to the fifth highest spike found within the ANCHOR profile of protein E (see **Figure 4C**). Furthermore, our disorder analysis revealed that the longest disorder region found in ZIKV E protein (located within the 120-180 region) contains the glycosylation site (Asn154) (Sirohi et al., 2016). High disorder propensity of the loop encompassing the glycosylation site is in line with the reported sequence variability of this region among ZIKV strains (Faye et al., 2014) and in other flaviviruses (Sirohi et al., 2016), suggesting that local structural dynamics and sequence variability could be of functional importance for this protein. In fact, sites of various enzymatically-catalyzed

posttranslational modifications in proteins are commonly located within their IDPRs (Iakoucheva et al., 2004; Pejaver et al., 2014). Therefore, finding intrinsic disorder in functionally important regions of the ZIKV E protein signifies the potential therapeutic importance of its IDPRs.

Disorder Analysis of the ZIKV Non-structural Proteins

There are seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) in ZIKV. Six of the NS proteins (NS2A to NS5) in ZIKV are known to be engaged in the formation of a replication complex on the cytoplasmic side of the endoplasmic reticulum membrane (Song et al., 2016). In addition to the viral non-structural proteins this replication complex also contains several host cofactors (Salonen et al., 2005).

NS1 Protein

The glycoprotein NS1 is considered as a key molecule in replication, immune evasion and pathogenesis of flaviviruses. NS1 is also secreted out into extracellular space as hexameric lipoprotein particles that are involved in multiple interactions with various components of immune system and host cell molecules (Suthar et al., 2013). In fact, it is believed that NS1 can be responsible for the diverse clinical consequences of infection caused by flaviviruses (Kuno et al., 1998; Cheng et al.,

2009). Recently, the crystal structure of a C-terminal fragment of ZIKV NS1 (PDB ID: 5IY3) has been solved (Song et al., 2016). The sequence of this crystallized C-terminal domain (residues 176-351) along with the sequence of the full-length NS1 protein was used in our disorder analysis. This analysis revealed that the full-length NS1 is characterized by the PPID of 10.8% (see **Figure 5A** and **Table 1**), whereas the crystallizable C-terminal fragment (PDB ID: 5IY3) shows the PID of 14.1%. Structurally, NS1 is arranged in a rod-like head to head dimers. Each dimer has 20 β-strand showing ladder-like arrangement on one surface and a complex arrangement of several loops on an opposite surface (see **Figure 5B**) (Song et al., 2016). Typically, loops between the β-strands are short, with the exception to the spaghetti loop connecting strands β4 and β5 (residues 218–272), which lacks regular ordered structure and is predicted to contain significant amount of disorder (see **Figure 5A**). Furthermore, a potential glycosylation site is located within the short loop between the β3 and β4 strands (residues 207–209), which is also predicted to have high level of disorder. Therefore, the high abundance of disorder in NS1 is correlated with the complex arrangement of loops in 3D crystal structure (Song et al., 2016) (see **Figure 5B**) and has functional significance. A phylogenetic analysis based on the amino acid sequences of NS1 proteins from 10 flaviviruses revealed some unique sequence characteristics of ZIKV NS1 that positioned it in the individual phylogenetic group (Song et al., 2016). It was emphasized that there is a very large variability in positively and negatively charged surfaces in central loop regions of the NS1 proteins from DENV, WNV, and ZIKV (Song et al., 2016). This variability in loop regions of flaviviruses may be implicated in the diversity of their pathogenicity. In Dengue and WNV, these loops of NS1 have been implicated in immune pathogenesis, whereas in ZIKV, this mechanism needs to be explored (Suthar et al., 2013).
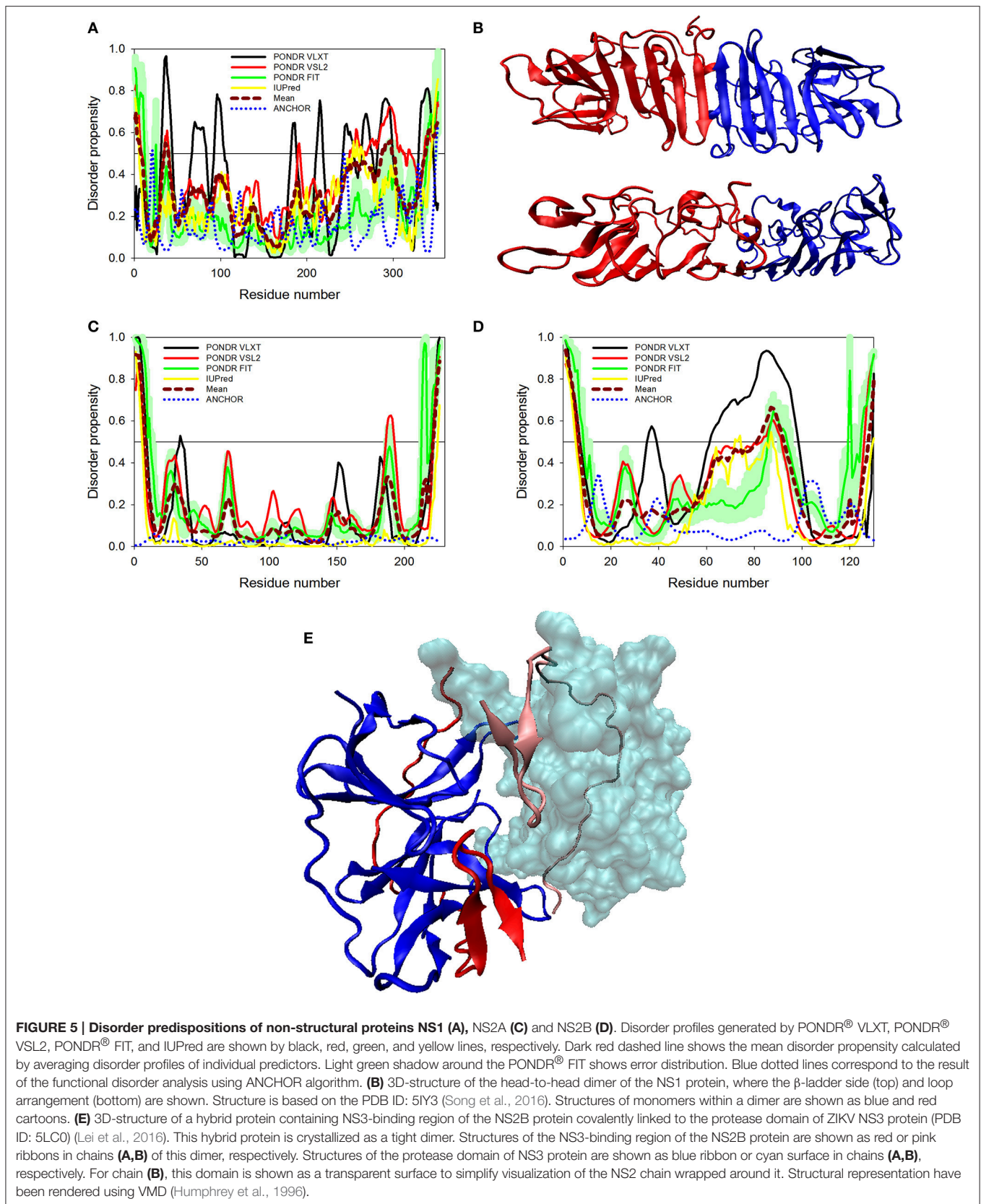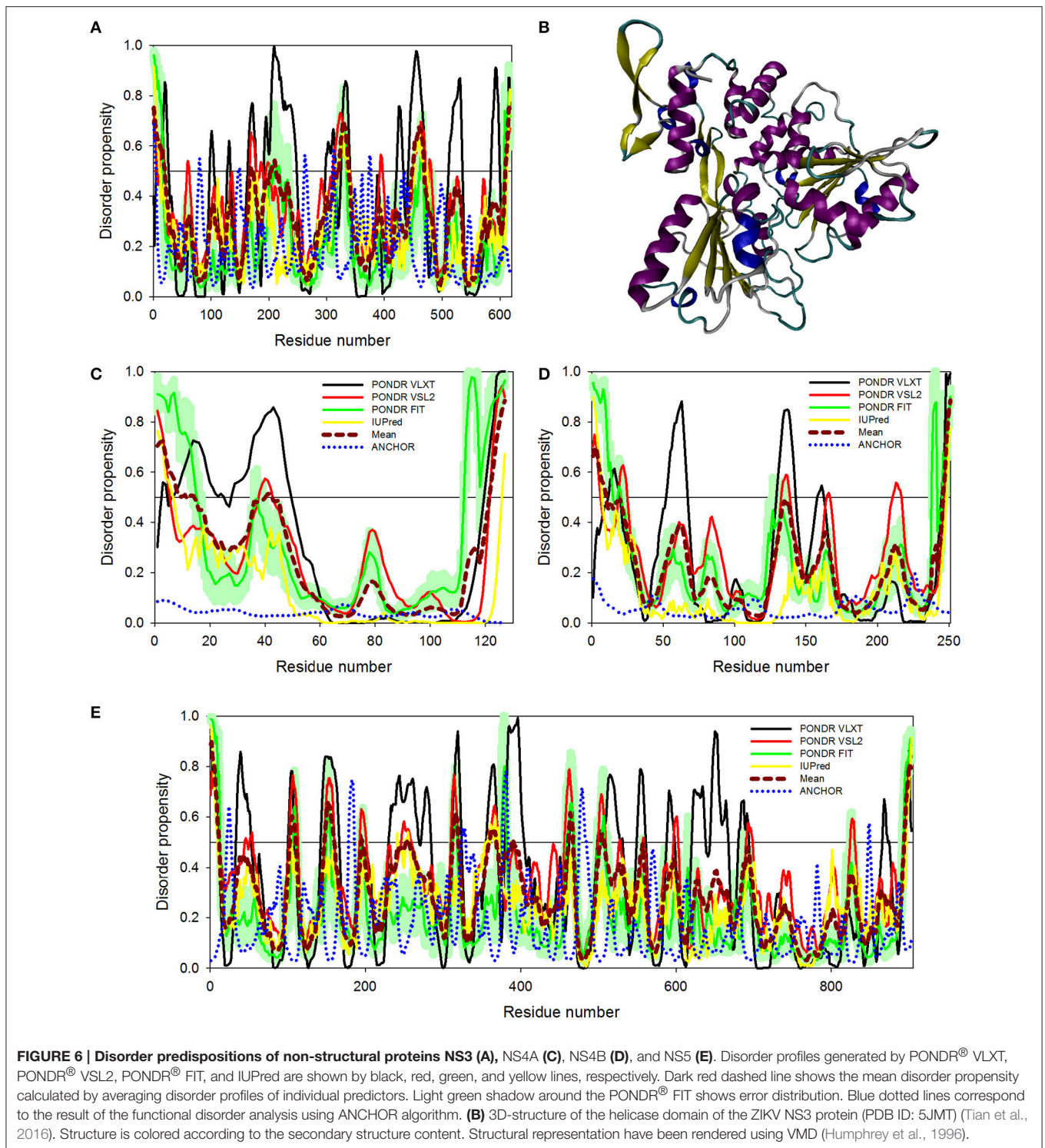
Proteins NS2A and NS2B

Non-structural proteins from NS2 to NS5 are involved in the formation of replication complex which is located on the endoplasmic reticulum membrane. NS2A possesses several important functions, such as involvement in viral RNA synthesis, virus-induced membrane formation, and inhibition of interferon α/β response (Xie et al., 2015). **Table 1** shows that, being the most ordered of the ZIKV proteins, NS2A has the PPID of 5.4%. However, NS2B is on the other side of spectrum, being characterized by a high abundance of disordered residues with PPID of 16.2 % (**Figures 5C,D**). **Figure 5D** shows that NS2B contains a central long disordered region of 37 residues (residues 62-98). This region is responsible for interaction with the NS3 protease (Murray et al., 2008). In the case of DENV it was shown that a central hydrophilic region of NS2B binds to NS3 and is needed for the formation of an active NS2B-NS3 protease complex. NS2B also stabilizes NS3 by acting as a chaperone. In previous studies on other members of the *Flaviviridae* family (Dengue virus and WNV), the structure of NS2B in a complex with NS3 was determined (Erbel et al., 2006). Curiously, recent crystallographic analysis of the hybrid protein constituting a central of the ZIKV NS2B (residues 49–95) covalently linked via the Gly$_4$-Ser-Gly$_4$ artificial linker to the N-terminal protease

domain of NS3 (residues 1–170) revealed that the NS3-binding region of NS2B wraps around the globular NS2B domain (see **Figure 5E**; PDB ID: 5LC0) (Lei et al., 2016). This structure clearly indicates that this NS2 region is disordered in its unbound form and folds upon binding to the globular protease domain of ZIKV NS3 protein. In the perspective of folding upon binding found for many IDPs, this seems to be a similar mechanism where protein acquires active conformation and function only by binding disordered partner as also evidenced in case of the model IDP systems such as KIX and cMyb (Gianni et al., 2012).

NS3 Protein

NS3 is a bifunctional enzyme that consists of two domains, such as the N-terminal protease domain (residues 1–167) and the C-terminal helicase domain (residues 168–617), which are essential for the polyprotein processing and the viral replication, respectively (Luo et al., 2015). NS3 consists of 617 residues, 12.6% of which are predicted to promote disorder. The longest disordered region found in this protein consists of 57 residues (residues 191–247) and is located within the helicase domain (**Figure 6A**). According to a recent report, Dengue virus NS3 contains the N-terminal proline-rich disordered region that plays a critical role in replication and virus particle formation (Gebhard et al., 2016). As it was already mentioned, the ZIKV NS3 protease domain (residues 1–170) was recently crystallized as a part of the hybrid protein containing the NS3-binding region of the ZIKV NS2B covalently linked by the Gly$_4$-Ser-Gly$_4$ peptide to N-terminus of the NS3 protein. No structural information was obtained for the 31 residue-long region connecting NS2B and NS3 and containing C-terminal tail of the NS2B, the Gly$_4$-Ser-Gly$_4$ linker, and the N-terminal 14 residues of the ZIKV NS3 containing the aforementioned proline-rich region (see **Figure 5E**; PDB ID: 5LC0) (Lei et al., 2016). Remaining part of the NS3 protease domain (except to its last 3 residues) was well-resolved and represents two β-barrels with strand orders AI-BI-CI-αI-DI-EIa-EIb-FI and AII-BIIa-BIIb-CII-DII-EIIa-EIIb-FII (Lei et al., 2016). Recently, a crystal structure of the NS3 helicase domain of ZIKV has been solved (PDB ID: 5JMT) (Tian et al., 2016) (see **Figure 6B**). In contrast to the DENV NS3 helicase, ZIKV has a monomeric helicase molecule. NS3 helicase is characterized by mostly ordered tertiary structure, with loops serving as connectors between the three domains of this protein (Tian et al., 2016). It was pointed out that the cleft between Domain I and II contains the NTPase active site and includes Walker A (or motif I or P-loop, residues 193–204) and B motifs (or motif II, residues 285–292) that play an important role in recognizing NTP and Mn^{2+} or Mg^{2+} cations (Caruthers and McKay, 2002; Tian et al., 2016). **Figure 6A** shows that both of these motifs are located within IDPRs. Although these loops are characterized by the considerable intrinsic flexibility, it was pointed out that they are highly conserved among flaviviruses and play an important role in binding and catalysis of NTP (Tian et al., 2016). Curiously, NS3 is one of the two ZIKV proteins that are predicted to have some disorder-based binding sites identified by ANCHOR. In fact, **Figure 6A** shows that there are 6 such sites in ZIKV NS3 (residues 1–4, 79–81, 147–150, 261–264, 311–314, and 374–376)

FIGURE 5 | Disorder predispositions of non-structural proteins NS1 (A), NS2A **(C)** and NS2B **(D)**. Disorder profiles generated by PONDR® VLXT, PONDR® VSL2, PONDR® FIT, and IUPred are shown by black, red, green, and yellow lines, respectively. Dark red dashed line shows the mean disorder propensity calculated by averaging disorder profiles of individual predictors. Light green shadow around the PONDR® FIT shows error distribution. Blue dotted lines correspond to the result of the functional disorder analysis using ANCHOR algorithm. **(B)** 3D-structure of the head-to-head dimer of the NS1 protein, where the β-ladder side (top) and loop arrangement (bottom) are shown. Structure is based on the PDB ID: 5IY3 (Song et al., 2016). Structures of monomers within a dimer are shown as blue and red cartoons. **(E)** 3D-structure of a hybrid protein containing NS3-binding region of the NS2B protein covalently linked to the protease domain of ZIKV NS3 protein (PDB ID: 5LC0) (Lei et al., 2016). This hybrid protein is crystallized as a tight dimer. Structures of the NS3-binding region of the NS2B protein are shown as red or pink ribbons in chains **(A,B)** of this dimer, respectively. Structures of the protease domain of NS3 protein are shown as blue ribbon or cyan surface in chains **(A,B)**, respectively. For chain **(B)**, this domain is shown as a transparent surface to simplify visualization of the NS2 chain wrapped around it. Structural representation have been rendered using VMD (Humphrey et al., 1996).

FIGURE 6 | Disorder predispositions of non-structural proteins NS3 (A), NS4A **(C)**, NS4B **(D)**, and NS5 **(E)**. Disorder profiles generated by PONDR® VLXT, PONDR® VSL2, PONDR® FIT, and IUPred are shown by black, red, green, and yellow lines, respectively. Dark red dashed line shows the mean disorder propensity calculated by averaging disorder profiles of individual predictors. Light green shadow around the PONDR® FIT shows error distribution. Blue dotted lines correspond to the result of the functional disorder analysis using ANCHOR algorithm. **(B)** 3D-structure of the helicase domain of the ZIKV NS3 protein (PDB ID: 5JMT) (Tian et al., 2016). Structure is colored according to the secondary structure content. Structural representation have been rendered using VMD (Humphrey et al., 1996).

which although were predicted by ANCHOR but then were filtered out by the algorithm because of their small size (they are ranging in length between 3 and 4 residues and ANCHOR uses the length threshold of 6 residues). Nevertheless, our analysis shows that ZIKV NS3 has several functionally important IDPRs.

NS4A and NS4B Proteins

Although non-structural proteins NS4A and NS4B from flaviviruses have not been crystallized as of yet, their biological roles in Dengue virus have been investigated experimentally.(Zou et al., 2015) NS4A functions by introducing rearrangements in membrane of host endoplasmic reticulum. These changes lead to

the formation of virus-induced membranous vesicles. NS4A also control the ATPase activity of the NS3 helicase, whereas NS4B inhibits the interferon-induced host STAT1 phosphorylation and nuclear translocation (Kuno and Chang, 2007). **Table 1** and **Figures 6C,D** show that although NS4A is predicted to be rather disordered, having the PPID of 16.5%, the level of intrinsic disorder in the NS4B protein is noticeably lower (its PPID is of 6.4%).

NS5 Protein

The NS5 protein is the largest ZIKV protein (903 residues) and has high degree of sequence similarity with the Dengue virus NS5 protein. In DENV, crystal structure of the full-length NS5 has been solved recently (PDB ID: 4V0Q) (Zhao et al., 2015), and structures of the methyltransferase domain (PDB ID: 1L9K) (Egloff et al., 2002), and the RNA-dependent RNA polymerase (RdRp) domain (PDB ID: 2J7W) of the NS5 (Yap et al., 2007) are also available. The NS5 is the most conserved protein amongst flaviviruses, exhibiting enzymatic activities that play vital roles in virus replication. In fact, it has the N-terminal domain (residues 1–262 in DENV3) that belongs to the S-adenosyl-L-methionine (SAM)-dependent methyltransferase (MTase) superfamily (Egloff et al., 2002) and the C-terminal domain (residues 273–900) which serves as the RNA-dependent RNA polymerase (RdRp) that synthesizes the anti-genome and progeny genome (Yap et al., 2007). Since NS5 contains two functional domains with several key enzymatic activities crucial for the viral RNA replication in the host cell, this protein represents one of the major targets for the design of antiviral inhibitors. Our analysis revealed that despite being the multifunctional enzyme, ZIKV NS5 contains numerous IDPRs (in a range of 20) and is characterized by the PPID of 8.6%. Furthermore, ANCHOR analysis revealed the presence of 4 disorder-based binding sites (residues 22–27, 179–186, 324–329, and 377–383), with two more disorder-based binding sites that were filtered out by the algorithm (residues 476–484 and 846–849). Our analysis also revealed that most of the NS5 disordered regions are located in the central part of the protein, with the longest disordered

region containing 40 amino acid residues (see **Figure 6E**). These findings of NS5 are also correlating well with the presence of long disordered region in DENV and JEV NS5 proteins.

CONCLUDING REMARKS

In conclusion, our analysis revealed that all ZIKV proteins contain disordered regions. These proteins are involved in diverse mechanisms of virus survival and immune evasion in other flaviviruses. Till now, the mechanisms of the ZIKV pathogenesis have not been deciphered in detail. Therefore, our study that considers ZIKV proteins from the protein intrinsic disorder perspective may provide novel insights that can help elucidating the molecular mechanisms of virus host interaction. In general, the highly dynamic and flexible nature of the disordered proteins or proteins containing IDPRs has shown to play major roles in disease development and progression (Uversky et al., 2008; Babu et al., 2011). As disordered regions are attractive and challenging drug targets, new drug development strategies should be developed to find small molecule inhibitors that could target IDPRs and could serve as effective antivirals. Therefore, detailed biophysical analysis of ZIKV disordered proteins is required to develop effective new antiviral therapeutics.

AUTHOR CONTRIBUTIONS

RG and VU: Conception and design, analysis and interpretation of data, writing and review of the manuscript and study supervision; DK and NS: acquisition of data, analysis and interpretation of data, writing of the manuscript.

FUNDING

REFERENCES

Ambadipudi, S., and Zweckstetter, M. (2016). Targeting intrinsically disordered proteins in rational drug discovery. *Expert Opin. Drug Discov.* 11, 65–77. doi: 10.1517/17460441.2016.1107041

Babu, M. M., van der Lee, R., de Groot, N. S., and Gsponer, J. (2011). Intrinsically disordered proteins: regulation and disease. *Curr. Opin. Struct. Biol.* 21, 432–440. doi: 10.1016/j.sbi.2011.03.011

Beasley, D. W., Whiteman, M. C., Zhang, S., Huang, C. Y., Schneider, B. S., Smith, D. R., et al. (2005). Envelope protein glycosylation status influences mouse neuroinvasion phenotype of genetic lineage 1 West Nile virus strains. *J. Virol.* 79, 8339–8347. doi: 10.1128/JVI.79.13.8339-8347.2005

Campen, A., Williams, R. M., Brown, C. J., Meng, J., Uversky, V. N., and Dunker, A. K. (2008). TOP-IDP-scale: a new amino acid scale measuring propensity for intrinsic disorder. *Protein Pept. Lett.* 15, 956–963. doi: 10.2174/092986608785849164

Caruthers, J. M., and McKay, D. B. (2002). Helicase structure and mechanism. *Curr. Opin. Struct. Biol.* 12, 123–133. doi: 10.1016/S0959-440X(02)00298-1

Cheng, H. J., Lin, C. F., Lei, H. Y., Liu, H. S., Yeh, T. M., Luo, Y. H., et al. (2009). Proteomic analysis of endothelial cell autoantigens recognized by anti-dengue virus nonstructural protein 1 antibodies. *Exp. Biol. Med. (Maywood).* 234, 63–73. doi: 10.3181/0805-RM-147

Cheng, Y., LeGall, T., Oldfield, C. J., Mueller, J. P., Van, Y. Y., Romero, P., et al. (2006). Rational drug design via intrinsically disordered protein. *Trends Biotechnol.* 24, 435–442. doi: 10.1016/j.tibtech.2006.07.005

Cheng, Y., Oldfield, C. J., Meng, J., Romero, P., Uversky, V. N., and Dunker, A. K. (2007). Mining alpha-helix-forming molecular recognition features with cross species sequence alignments. *Biochemistry* 46, 13468–13477. doi: 10.1021/bi7012273

Deng, X., Eickholt, J., and Cheng, J. (2012). A comprehensive overview of computational protein disorder prediction methods. *Mol. Biosyst.* 8, 114–121. doi: 10.1039/C1MB05207A

Dick, G. W., Kitchen, S. F., and Haddow, A. J. (1952). Zika virus. I. Isolations and serological specificity. *Trans. R. Soc. Trop. Med. Hyg.* 46, 509–520. doi: 10.1016/0035-9203(52)90042-4

Dosztányi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005a). IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 21, 3433–3434. doi: 10.1093/bioinformatics/bti541

Dosztányi, Z., Csizmók, V., Tompa, P., and Simon, I. (2005b). The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. *J. Mol. Biol.* 347, 827–839. doi: 10.1016/j.jmb.2005.01.071

Dosztányi, Z., Mészáros, B., and Simon, I. (2009). ANCHOR: web server for predicting protein binding regions in disordered proteins. *Bioinformatics* 25, 2745–2746. doi: 10.1093/bioinformatics/btp518

Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M., and Obradovic, Z. (2002a). Intrinsic disorder and protein function. *Biochemistry* 41, 6573–6582. doi: 10.1021/bi012159+

Dunker, A. K., Brown, C. J., and Obradovic, Z. (2002b). Identification and functions of usefully disordered proteins. *Adv. Protein Chem.* 62, 25–49. doi: 10.1016/S0065-3233(02)62004-2

Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., et al. (2001). Intrinsically disordered protein. *J. Mol. Graph. Model.* 19, 26–59. doi: 10.1016/S1093-3263(00)00138-8

Dunker, A. K., Silman, I., Uversky, V. N., and Sussman, J. L. (2008). Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.* 18, 756–764. doi: 10.1016/j.sbi.2008.10.002

Dunker, A. K., and Uversky, V. N. (2008). Signal transduction via unstructured protein conduits. *Nat. Chem. Biol.* 4, 229–230. doi: 10.1038/nchembio0408-229

Egloff, M. P., Benarroch, D., Selisko, B., Romette, J. L., and Canard, B. (2002). An RNA cap (nucleoside-2′-O-)-methyltransferase in the flavivirus RNA polymerase NS5: crystal structure and functional characterization. *EMBO J.* 21, 2757–2768. doi: 10.1093/emboj/21.11.2757

Erbel, P., Schiering, N., D'Arcy, A., Renatus, M., Kroemer, M., Lim, S. P., et al. (2006). Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. *Nat. Struct. Mol. Biol.* 13, 372–373. doi: 10.1038/nsmb1073

Fan, X., and Kurgan, L. (2014). Accurate prediction of disorder in protein chains with a comprehensive and empirically designed consensus. *J. Biomol. Struct. Dyn.* 32, 448–464. doi: 10.1080/07391102.2013.775969

Fan, X., Xue, B., Dolan, P. T., LaCount, D. J., Kurgan, L., and Uversky, V. N. (2014). The intrinsic disorder status of the human hepatitis C virus proteome. *Mol. Biosyst.* 10, 1345–1363. doi: 10.1039/c4mb00027g

Faye, O., Freire, C. C., Iamarino, A., Faye, O., de Oliveira, J. V., Diallo, M., et al. (2014). Molecular evolution of Zika virus during its emergence in the 20(th) century. *PLoS Negl. Trop. Dis.* 8:e2636. doi: 10.1371/journal.pntd.0002636

Fontana, A., de Laureto, P. P., Spolaore, B., Frare, E., Picotti, P., and Zambonin, M. (2004). Probing protein structure by limited proteolysis. *Acta Biochim. Pol.* 51, 299–321.

Fontana, A., Fassina, G., Vita, C., Dalzoppo, D., Zamai, M., and Zambonin, M. (1986). Correlation between sites of limited proteolysis and segmental mobility in thermolysin. *Biochemistry* 25, 1847–1851. doi: 10.1021/bi00356a001

Gebhard, L. G., Iglesias, N. G., Byk, L. A., Filomatori, C. V., De Maio, F. A., and Gamarnik, A. V. (2016). A proline-rich N-terminal region of the Dengue virus NS3 is crucial for infectious particle production. *J. Virol.* 90, 5451–5461. doi: 10.1128/JVI.00206-16

Gianni, S., Morrone, A., Giri, R., and Brunori, M. (2012). A folding-after-binding mechanism describes the recognition between the transactivation domain of c-Myb and the KIX domain of the CREB-binding protein. *Biochem. Biophys. Res. Commun.* 428, 205–209. doi: 10.1016/j.bbrc.2012.09.112

Goh, G. K., Dunker, A. K., and Uversky, V. N. (2016). Correlating Flavivirus virulence and levels of intrinsic disorder in shell proteins: protective roles vs. immune evasion. *Mol. Biosyst.* 12, 1881–1891. doi: 10.1039/C6MB00228E

Grischott, F., Puhan, M., Hatz, C., and Schlagenhauf, P. (2016). Non-vector-borne transmission of Zika virus: a systematic review. *Travel Med. Infect. Dis.* 14, 313–330. doi: 10.1016/j.tmaid.2016.07.002

Hamel, R., Dejarnac, O., Wichit, S., Ekchariyawat, P., Neyret, A., Luplertlop, N., et al. (2015). Biology of Zika virus infection in human skin cells. *J. Virol.* 89, 8880–8896. doi: 10.1128/JVI.00354-15

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. *J. Mol. Graph.* 14, 33–38, 27–38. doi: 10.1016/0263-7855(96)00018-5

Iakoucheva, L. M., Kimzey, A. L., Masselon, C. D., Bruce, J. E., Garner, E. C., Brown, C. J., et al. (2001). Identification of intrinsic order and disorder in the DNA repair protein XPA. *Protein Sci.* 10, 560–571. doi: 10.1110/ps.29401

Iakoucheva, L. M., Radivojac, P., Brown, C. J., O'Connor, T. R., Sikes, J. G., Obradovic, Z., et al. (2004). The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* 32, 1037–1049. doi: 10.1093/nar/gkh253

Ivanyi-Nagy, R., and Darlix, J. L. (2010). Intrinsic disorder in the core proteins of flaviviruses. *Protein Pept. Lett.* 17, 1019–1025. doi: 10.2174/092986610791498911

Ivanyi-Nagy, R., Lavergne, J. P., Gabus, C., Ficheux, D., and Darlix, J. L. (2008). RNA chaperoning and intrinsic disorder in the core proteins of *Flaviviridae*. *Nucleic Acids Res.* 36, 712–725. doi: 10.1093/nar/gkm1051

Kuhn, R. J., Zhang, W., Rossmann, M. G., Pletnev, S. V., Corver, J., Lenches, E., et al. (2002). Structure of dengue virus: implications for flavivirus organization, maturation, and fusion. *Cell* 108, 717–725. doi: 10.1016/S0092-8674(02)00660-8

Kuno, G., and Chang, G. J. (2007). Full-length sequencing and genomic characterization of Bagaza, Kedougou, and Zika viruses. *Arch. Virol.* 152, 687–696. doi: 10.1007/s00705-006-0903-z

Kuno, G., Chang, G. J., Tsuchiya, K. R., Karabatsos, N., and Cropp, C. B. (1998). Phylogeny of the genus Flavivirus. *J. Virol.* 72, 73–83.

Lazear, H. M., Govero, J., Smith, A. M., Platt, D. J., Fernandez, E., Miner, J. J., et al. (2016). A mouse model of Zika virus pathogenesis. *Cell Host Microbe* 19, 720–730. doi: 10.1016/j.chom.2016.03.010

Lei, J., Hansen, G., Nitsche, C., Klein, C. D., Zhang, L., and Hilgenfeld, R. (2016). Crystal structure of Zika virus NS2B-NS3 protease in complex with a boronate inhibitor. *Science* 353, 503–505. doi: 10.1126/science.aag2419

Li, Z., Khaliq, M., Zhou, Z., Post, C. B., Kuhn, R. J., and Cushman, M. (2008). Design, synthesis, and biological evaluation of antiviral agents targeting flavivirus envelope proteins. *J. Med. Chem.* 51, 4660–4671. doi: 10.1021/jm800412d

Linding, R., Jensen, L. J., Diella, F., Bork, P., Gibson, T. J., and Russell, R. B. (2003a). Protein disorder prediction: implications for structural proteomics. *Structure* 11, 1453–1459. doi: 10.1016/j.str.2003.10.002

Linding, R., Russell, R. B., Neduva, V., and Gibson, T. J. (2003b). GlobPlot: Exploring protein sequences for globularity and disorder. *Nucleic Acids Res.* 31, 3701–3708. doi: 10.1093/nar/gkg519

Liu, H., Chiou, S. S., and Chen, W. J. (2004). Differential binding efficiency between the envelope protein of Japanese encephalitis virus variants and heparan sulfate on the cell surface. *J. Med. Virol.* 72, 618–624. doi: 10.1002/jmv.20025

Luo, D., Vasudevan, S. G., and Lescar, J. (2015). The flavivirus NS2B-NS3 protease-helicase as a target for antiviral drug development. *Antiviral Res.* 118, 148–158. doi: 10.1016/j.antiviral.2015.03.014

Martins, I. C., Gomes-Neto, F., Faustino, A. F., Carvalho, F. A., Carneiro, F. A., Bozza, P. T., et al. (2012). The disordered N-terminal region of dengue virus capsid protein contains a lipid-droplet-binding motif. *Biochem. J.* 444, 405–415. doi: 10.1042/BJ20112219

Mayhoub, A. S., Khaliq, M., Kuhn, R. J., and Cushman, M. (2011). Design, synthesis, and biological evaluation of thiazoles targeting flavivirus envelope proteins. *J. Med. Chem.* 54, 1704–1714. doi: 10.1021/jm1013538

Meng, F., Badierah, R. A., Almehdar, H. A., Redwan, E. M., Kurgan, L., and Uversky, V. N. (2015). Unstructural biology of the Dengue virus proteins. *FEBS J.* 282, 3368–3394. doi: 10.1111/febs.13349

Mészáros, B., Simon, I., and Dosztányi, Z. (2009). Prediction of protein binding regions in disordered proteins. *PLoS Comput. Biol.* 5:e1000376. doi: 10.1371/journal.pcbi.1000376

Miller, J. L., de Wet, B. J., Martinez-Pomares, L., Radcliffe, C. M., Dwek, R. A., Rudd, P. M., et al. (2008). The mannose receptor mediates dengue virus infection of macrophages. *PLoS Pathog.* 4:e17. doi: 10.1371/journal.ppat.0040017

Murray, C. L., Jones, C. T., and Rice, C. M. (2008). Architects of assembly: roles of *Flaviviridae* non-structural proteins in virion morphogenesis. *Nat. Rev. Microbiol.* 6, 699–708. doi: 10.1038/nrmicro1928

Novotný, J., and Bruccoleri, R. E. (1987). Correlation among sites of limited proteolysis, enzyme accessibility and segmental mobility. *FEBS Lett.* 211, 185–189. doi: 10.1016/0014-5793(87)81433-3

Obradovic, Z., Peng, K., Vucetic, S., Radivojac, P., and Dunker, A. K. (2005). Exploiting heterogeneous sequence properties improves prediction of protein disorder. *Proteins* 61(Suppl. 7), 176–182. doi: 10.1002/prot. 20735

Oldfield, C. J., Cheng, Y., Cortese, M. S., Romero, P., Uversky, V. N., and Dunker, A. K. (2005). Coupled folding and binding with alpha-helix-forming molecular recognition elements. *Biochemistry* 44, 12454–12470. doi: 10.1021/bi05 0736e

Pejaver, V., Hsu, W. L., Xin, F., Dunker, A. K., Uversky, V. N., and Radivojac, P. (2014). The structural and functional signatures of proteins that undergo multiple events of post-translational modification. *Protein Sci.* 23, 1077–1093. doi: 10.1002/pro.2494

Peng, K., Radivojac, P., Vucetic, S., Dunker, A. K., and Obradovic, Z. (2006). Length-dependent prediction of protein intrinsic disorder. *BMC Bioinformatics* 7:208. doi: 10.1186/1471-2105-7-208

Peng, K., Vucetic, S., Radivojac, P., Brown, C. J., Dunker, A. K., and Obradovic, Z. (2005). Optimizing long intrinsic disorder predictors with protein evolutionary information. *J. Bioinform. Comput. Biol.* 3, 35–60. doi: 10.1142/S0219720005000886

Peng, Z. L., and Kurgan, L. (2012). Comprehensive comparative assessment of *in-silico* predictors of disordered regions. *Curr. Protein Pept. Sci.* 13, 6–18. doi: 10.2174/138920312799277938

Petersen, L. R., Jamieson, D. J., Powers, A. M., and Honein, M. A. (2016). Zika virus. *N. Engl. J. Med.* 374, 1552–1563. doi: 10.1056/NEJMra1602113

Poh, M. K., Yip, A., Zhang, S., Priestle, J. P., Ma, N. L., Smit, J. M., et al. (2009). A small molecule fusion inhibitor of dengue virus. *Antiviral Res.* 84, 260–266. doi: 10.1016/j.antiviral.2009.09.011

Pokidysheva, E., Zhang, Y., Battisti, A. J., Bator-Kelly, C. M., Chipman, P. R., Xiao, C., et al. (2006). Cryo-EM reconstruction of dengue virus in complex with the carbohydrate recognition domain of DC-SIGN. *Cell* 124, 485–493. doi: 10.1016/j.cell.2005.11.042

Prilusky, J., Felder, C. E., Zeev-Ben-Mordehai, T., Rydberg, E. H., Man, O., Beckmann, J. S., et al. (2005). FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics* 21, 3435–3438. doi: 10.1093/bioinformatics/bti537

Romero, P., Obradovic, Z., Li, X., Garner, E. C., Brown, C. J., and Dunker, A. K. (2001). Sequence complexity of disordered protein. *Proteins* 42, 38–48. doi: 10.1002/1097-0134(20010101)42:1<38::AID-PROT50>3.0.CO;2-3

Salonen, A., Ahola, T., and Kääriäinen, L. (2005). Viral RNA replication in association with cellular membranes. *Curr. Top. Microbiol. Immunol.* 285, 139–173. doi: 10.1007/3-540-26764-6_5

Shoemaker, B. A., Portman, J. J., and Wolynes, P. G. (2000). Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *Proc. Natl. Acad. Sci. U.S.A.* 97, 8868–8873. doi: 10.1073/pnas.160259697

Sirohi, D., Chen, Z., Sun, L., Klose, T., Pierson, T. C., Rossmann, M. G., et al. (2016). The 3.8 A resolution cryo-EM structure of Zika virus. *Science* 352, 467–470. doi: 10.1126/science.aaf5316

Song, H., Qi, J., Haywood, J., Shi, Y., and Gao, G. F. (2016). Zika virus NS1 structure reveals diversity of electrostatic surfaces among flaviviruses. *Nat. Struct. Mol. Biol.* 23, 456–458. doi: 10.1038/nsmb.3213

Suthar, M. S., Diamond, M. S., and Gale, M. Jr. (2013). West Nile virus infection and immunity. *Nat. Rev. Microbiol.* 11, 115–128. doi: 10.1038/nrmicro2950

Tian, H., Ji, X., Yang, X., Xie, W., Yang, K., Chen, C., et al. (2016). The crystal structure of Zika virus helicase: basis for antiviral drug design. *Protein Cell* 7, 450–454. doi: 10.1007/s13238-016-0275-4

Tompa, P. (2002). Intrinsically unstructured proteins. *Trends Biochem. Sci.* 27, 527–533. doi: 10.1016/S0968-0004(02)02169-2

Uversky, V. N. (2010). Targeting intrinsically disordered proteins in neurodegenerative and protein dysfunction diseases: another illustration of the D(2) concept. *Expert Rev. Proteomics* 7, 543–564. doi: 10.1586/epr.10.36

Uversky, V. N. (2012). Intrinsically disordered proteins and novel strategies for drug discovery. *Expert Opin. Drug Discov.* 7, 475–488. doi: 10.1517/17460441.2012.686489

Uversky, V. N. (2013a). A decade and a half of protein intrinsic disorder: biology still waits for physics. *Protein Sci.* 22, 693–724. doi: 10.1002/pro.2261

Uversky, V. N. (2013b). Intrinsic Disorder-based Protein Interactions and their Modulators. *Curr. Pharm. Des.* 19, 4191–4213. doi: 10.2174/1381612811319230005

Uversky, V. N., and Dunker, A. K. (2010). Understanding protein non-folding. *Biochim. Biophys. Acta* 1804, 1231–1264. doi: 10.1016/j.bbapap.2010.01.017

Uversky, V. N., Oldfield, C. J., and Dunker, A. K. (2008). Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annu. Rev. Biophys.* 37, 215–246. doi: 10.1146/annurev.biophys.37.032807.125924

Vullo, A., Bortolami, O., Pollastri, G., and Tosatto, S. C. (2006). Spritz: a server for the prediction of intrinsically disordered regions in protein sequences using kernel machines. *Nucleic Acids Res.* 34(Web Server issue), W164–168. doi: 10.1093/nar/gkl166

Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F., and Jones, D. T. (2004). Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* 337, 635–645. doi: 10.1016/j.jmb.2004.02.002

Wikan, N., and Smith, D. R. (2016). Zika virus: history of a newly emerging arbovirus. *Lancet Infect. Dis.* 16, e119–e126. doi: 10.1016/S1473-3099(16)30010-X

Wright, P. E., and Dyson, H. J. (2015). Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* 16, 18–29. doi: 10.1038/nrm3920

Xie, X., Zou, J., Puttikhunt, C., Yuan, Z., and Shi, P. Y. (2015). Two distinct sets of NS2A molecules are responsible for dengue virus RNA synthesis and virion assembly. *J. Virol.* 89, 1298–1313. doi: 10.1128/JVI.02882-14

Xue, B., Blocquel, D., Habchi, J., Uversky, A. V., Kurgan, L., Uversky, V. N., et al. (2014). Structural disorder in viral proteins. *Chem. Rev.* 114, 6880–6911. doi: 10.1021/cr4005692

Xue, B., Dunbrack, R. L., Williams, R. W., Dunker, A. K., and Uversky, V. N. (2010). PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. *Biochim. Biophys. Acta* 1804, 996–1010. doi: 10.1016/j.bbapap.2010.01.011

Yap, T. L., Xu, T., Chen, Y. L., Malet, H., Egloff, M. P., Canard, B., et al. (2007). Crystal structure of the dengue virus RNA-dependent RNA polymerase catalytic domain at 1.85-angstrom resolution. *J. Virol.* 81, 4753–4765. doi: 10.1128/JVI.02283-06

Yu, I. M., Zhang, W., Holdaway, H. A., Li, L., Kostyuchenko, V. A., Chipman, P. R., et al. (2008). Structure of the immature dengue virus at low pH primes proteolytic maturation. *Science* 319, 1834–1837. doi: 10.1126/science.1153264

Zhang, X., Ge, P., Yu, X., Brannan, J. M., Bi, G., Zhang, Q., et al. (2013). Cryo-EM structure of the mature dengue virus at 3.5-A resolution. *Nat. Struct. Mol. Biol.* 20, 105–110. doi: 10.1038/nsmb.2463

Zhang, Y., Corver, J., Chipman, P. R., Zhang, W., Pletnev, S. V., Sedlak, D., et al. (2003). Structures of immature flavivirus particles. *EMBO J.* 22, 2604–2613. doi: 10.1093/emboj/cdg270

Zhao, Y., Soh, T. S., Zheng, J., Chan, K. W., Phoo, W. W., Lee, C. C., et al. (2015). A crystal structure of the Dengue virus NS5 protein reveals a novel inter-domain interface essential for protein flexibility and virus replication. *PLoS Pathog.* 11:e1004682. doi: 10.1371/journal.ppat.1004682

Zou, J., Xie, X., Wang, Q. Y., Dong, H., Lee, M. Y., Kang, C., et al. (2015). Characterization of dengue virus NS4A and NS4B protein interaction. *J. Virol.* 89, 3455–3470. doi: 10.1128/JVI.03453-14