 doi:10.20944/preprints202208.0531.v1

Article

Looking At the Pathogenesis of the Rabies Lyssavirus Strain Pasteur Vaccins through A Prism of the Disorder-Based Bioinformatics

Surya Dhulipala [1] and Vladimir Uversky [1,2,3,*]

[1] Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL 33612, USA; sdhulipala@usf.edu (S.A.); vuversky@usf.edu (V.N.U.)
[2] USF Health Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, FL 33612, USA
[3] Protein Research Group, Institute for Biological Instrumentation of the Russian Academy of Sciences, Federal Research Center "Pushchino Scientific Center for Biological Research of the Russian Academy of Sciences", Pushchino, Moscow region, 142290, Russia
* Correspondence: vuversky@usf.edu

Abstract: Rabies is a neurological disease that causes between 40,000 and 70,000 deaths every year. Once a rabies patient has become symptomatic, there is no effective treatment for the illness, and in unvaccinated individuals, the case-fatality rate of rabies is close to 100%. French scientists Louis Pasteur and Émile Roux developed the first vaccine for rabies in 1885. If administered before the virus reaches the brain, the modern rabies vaccine imparts long-lasting immunity to the virus and saves more than 250,000 people every year. However, the rabies virus can suppress the host's immune response once it has entered the cells of the brain, making death likely. This study aims to make use of disorder-based proteomics and bioinformatics to determine the impact that intrinsically disordered protein regions (IDPRs) in the proteome of the rabies virus have on the infectivity and lethality of the disease. This study uses the proteome of Rabies Lyssavirus (RABV) strain Pasteur Vaccins (PV), one of the best understood strains due to its use in the first rabies vaccine, as a model. The study suggests that the high levels of intrinsic disorder in the phosphoprotein (P-protein) and nucleoprotein (N-protein) allow them to participate in creation of the Negri bodies and help this virus suppress the antiviral immune response in the host cells. Additionally, the study suggests that there is a link between disorder in the matrix (M) protein and the modulation of viral transcription. The disordered regions in the M protein have a possible role in initiating viral budding within the cell. Furthermore, we checked the prevalence of functional disorder in a set of 37 host proteins directly involved in the interaction with the RABV proteins. The hope is that these new insights will aid in the development of treatments for rabies that are effective after infection.

Keywords: Rabies; intrinsic disorder; intrinsically disordered protein; intrinsically disordered protein region; protein-protein interaction

1. Introduction

Rabies Lyssavirus is a bullet-shaped, negative-sense, unsegmented, single stranded RNA virus of the *Rhabdoviridae* family. There are 10 viruses in the Rabies serogroup, but most are not pathogenic to humans. Rabies Lyssavirus and Australian Bat Lyssavirus are the only two rhabdoviruses that have been known to cause disease in humans [1]. Rabies virus (RABV) is a zoonotic neurotropic virus that causes fatal neurological symptoms in almost all mammals and is spread through the bite of an infected mammal. The Rabies disease causes between 40,000 and 70,000 deaths every year worldwide. Once a Rabies patient has become symptomatic, there is no effective treatment for the illness. In fact, in unvaccinated individuals, the case-fatality rate of Rabies is close to 100% [2]. French scientists Louis Pasteur and Émile Roux developed the first vaccine for Rabies in 1885. If

 doi:10.20944/preprints202208.0531.v1

administered before the virus reaches the brain, the modern Rabies vaccine imparts long-lasting immunity to the virus. It saves more than 250,000 people every year. RABV is able to suppress the host's immune response once it has entered the cells of the brain, making death likely, if the vaccine is not administered [3]. The virus works by infecting muscle cells through saliva in bites before travelling to the brain. The virus travels through neuromuscular junctions and neural pathways using axioplasmic transport. Once the virus reaches the brain, it triggers endocytosis into the neurons. It then undergoes transcription and replication by polymerase shuttering [4]. The virus suppresses the antiviral response and infects more cells before eventually getting released and transported to the salivary glands. The high neuroinvasiveness of RABV has been attributed to the ability of the virus to evade immune responses through various means and to conserve the structures of neurons [3].

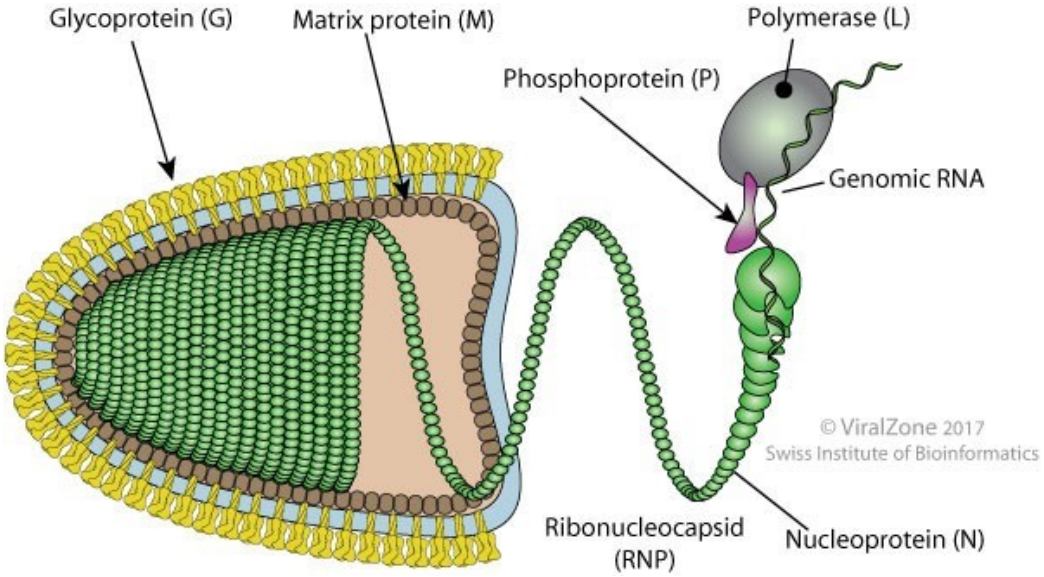

Figure 1. Structural features of the RABV. In the mature RABV, the Nucleoprotein, Phosphoprotein and viral Polymerase envelop the Genomic RNA in a structure known as the RIbonucleocapsid (RNP). The Matrix protein surrounds the RNP and determines the shape of the virus. The Matrix protein also anchors the Glycoprotein to the envelope [5] (original source of the image; Philippe Le Mercier, SIB Swiss Institute of Bioinformatics).

All RABVs encode for 5 major proteins, as well as 4 more isoforms of the phosphoprotein that are generated through alternative initiation. The major proteins are the phosphoprotein (P-protein), matrix protein (M-protein), glycoprotein (G-protein), nucleoprotein (N-protein) and the polymerase, or large protein (L-protein). A total of 5 phosphoproteins can be made through alternative initiation [6]. The RNA genome of the virus is roughly 11 kb in size. A graphical representation of the virus can be seen in **Figure 1**. The P-protein is bound tightly around the RNA of the virus to create the RNP (ribonucleoprotein) core, and the RNP condenses with the N- and L-proteins to create the helical nucleocapsid of the virus. The N-protein homomultimerizes to form the nucleocapsid and binds with the P-protein [7]. The P-protein forms a homotrimer when phosphorylated and bound to the L-protein, which positions the L-protein over the template RNA strands [8]. The M-protein forms a homomultimer between the nucleocapsid core and the glycoprotein spikes that surround the virus to connect the two. The G-protein spikes, which are homotrimers of the G-protein, form the sole ligand for the cellular receptor. Anti-rabies antibodies target the G protein homotrimers [9,10].

Once the virus reaches the cell, the G-protein initiates endocytosis. The low pH inside the endosome initiates the fusion of the membrane with the endosome, resulting in the release of the nucleocapsid into the cell. The P-protein acts as a cofactor to the L-protein,

 doi:10.20944/preprints202208.0531.v1

creating the functional viral polymerase [11]. The M-protein acts as a bridge between the plasma membrane and the generated nucleocapsids, providing a crucial role in virus budding [3,12,13]. The condensing of the nucleocapsid and the final shape is mediated by the M-protein [3].

The P-protein has an important role in the infection and propagation of the RABV. It binds to the dynein light chain proteins DYNLL1 and DYNLL2 in humans, allowing the virus to be transported along the axon to the nucleus. During transcription, alternative initiation using a ribosomal leaky scanning mechanism can create five variations of the P-protein, termed P1-5 [14]. This enables higher flexibility in the action of the P-protein. It interacts with the ribosomal protein L9 during the initial stages of RABV infection to inhibit RABV transcription during this period, controlling viral replication [15]. The P-protein also acts as an antagonist to STAT1 and STAT2 proteins within the cell. STAT proteins, which are activated by type-I interferon receptors in the type-I interferon-mediated innate immune response, are cytoplasmic signal transducers and transcription activators associated with the immune response [16,17]. The P-protein exploits microtubule processes within the cell to transport the virus and force STAT1 to use microtubule-inhibited mechanisms, suppressing nuclear transport of STAT and signaling [18]. STAT proteins are important for the establishment of the antiviral state in the cell. The STAT-targeting feature of interferon antagonists is considered a determining factor in pathogenicity of the virus. Inhibition of the interaction between the P-protein and STAT has been shown to severely reduce lethality and increase the immune response to RABV. The N-protein forms a complex with the P-protein to encapsulate the RNA genome of the virus and protect it from nuclease activity and phosphorylation [7]. The N-protein has been shown to help suppress the type-3 interferon-mediated immune system by evading the RIG-I-mediated antiviral response [19,20]. The L-protein, due to its activity as a viral polymerase, plays multiple enzymatic roles in the synthesis and processing of viral RNA [21]. It acts as a low fidelity polymerase, resulting in a high rate of mutations for the RABV.

The M-protein has been shown to be significant in regulating viral budding. It has been shown to bind to the plasma membrane even in the absence of other viral proteins, which suggests that the protein acts to connect the plasma membrane with the nucleocapsid. The M-protein has been shown to be able to induce vesicle budding without necessitating an interaction with other proteins [3,12]. The infection of a cell with M-protein-deficient RABV mutant was observed to result in smaller numbers of rod shaped or round viral particles rather than the bullet-shaped viruses that are standard with mature RABV [13]. The M-protein has been shown to inhibit viral transcription in many viruses including RABV, suggesting that it has a highly regulatory role in transcription and replication [12,22]. The M-protein has been shown to interact with RelAp43, a protein in the NF-κB family of proteins, which results in the further suppression of the antiviral immune response [23]. The G-protein is associated with accelerating the budding efficiency of the virus [12,13]. The G-protein also increases pathogenicity drastically because, as the only surface protein in the virus, it induces innate immune responses by binding to immune receptors [24]. This binding is shown to promote effective virus uptake, which drastically increases virulence.

As typical for other viruses, all RABV proteins are multifunctional. Such multifunctionality is crucial, as it allows a very small set of viral proteins to manage and control each and any aspect of the viral “life” from entry to replication to formation and exit of new infectious particles and regulate each and any aspect of virus interaction with the host. In fact, a typical protein repertoire of a typical virus includes a minimal set of specific structural proteins crucial for the viral particle assembly and a set of non-structural proteins needed for the hijacking of many functional pathways of the host cell. This is why many viral proteins are multifunctional. They are typically engaged in numerous interactions with various host cell components. In the protein universe, multifunctionality and binding promiscuity are typically associated with protein intrinsic disorder. This directly follows from the recognition that many protein functions do not need a unique 3-D struc-

 doi:10.20944/preprints202208.0531.v1

ture [25-33] and that such structure-less intrinsically disordered proteins (IDPs) and intrinsically disordered protein regions (IDPRs) are commonly found in various proteomes [34-38], where they are involved in regulation, signaling, and control pathways [27,30,32,39-41], thereby possessing functions that complement functional repertoire of the traditional ordered proteins [42-47]. IDPs/IDPRs and are often involved in various human diseases [48,49]. They have complex mosaic structures and show remarkable multilevel spatiotemporal heterogeneity, existing as dynamic conformational ensembles [25,27,31,40,42,50,51] where different parts of a protein can be (dis)ordered to different degrees [40,52,53]. Importantly, these differently (dis)ordered pieces of the protein structural mosaic might have well-defined and specific functions [53]. Therefore, IDPs/IDPRs are structurally and functionally heterogeneous complex systems whose functionality is described in terms of the protein structure-function continuum [53,54], where structural and functional diversification of a protein is defined by several factors determining the capability of a single gene to encode a set of distinct protein molecules, proteoforms [55]. This is achieved at several levels by altering the chemical structure of proteinaceous product(s) of a given gene via allelic variations at the DNA level (utilizing several specific means, such as single or multiple point mutations, indels, SNPs),by alternative splicing and other pre-translational mechanisms affecting mRNA, via a broad arsenal of countless post-translational modifications (PTMs) of a polypeptide chain, by the presence of intrinsic disorder, or by structural alterations induced by functioning [54].

Importantly, based on the computational analyses of the abundance of intrinsic disorder in various organisms it has been concluded that the proteomes of viruses have the largest variability in the content of disordered residues in comparison with all other kingdoms of life [37,38]. Abundance and functional importance of IDPs/IDPRs were systematically investigated for human papillomaviruses (HPVs) [56,57], human immunodeficiency virus 1 (HIV-1) [58], human hepatitis C virus (HCV) [59,60], Dengue virus [61], rotavirus [62], human respiratory syncytial virus [63], Zika virus [64,65], Chikungunya virus [66], Alkhurma virus (ALKV) [67], Japanese encephalitis virus [68], and SARS-CoV-2, Human SARS, and Bat SARS-Like coronaviruses [69]. These studies suggested that the presence of IDPRs in viral proteins is crucial for their functionality and represents an important means of the overall enhancement of viral propagation during the virus life cycle. To the best of our knowledge, there is no similar analysis of the disorder status of RABV proteome. The aim of this study is to fill this gap by conducting comprehensive computational analysis of the penetrance of intrinsic disorder in the proteins of Rabies Lyssavirus strain Pasteur Vaccins proteome.

2. Materials and Methods

The Universal Protein Resource (UniProt) is an annotated database of protein sequences [70]. Entries may be manually annotated with information extracted from literature and evaluated computational analysis (Swiss-Prot), or computationally analyzed (TrEMBL). The Rabies Lyssavirus strain Pasteur Vaccins proteome is a Swiss-Prot annotated entry. Each entry contains information including but not limited to the taxonomy, molecular function and included biological processes, subcellular location, potential modifications, pathology, interactions, and structures. The amino acid sequences and the structures are the most valuable resources out of these for the purposes of studying intrinsic disorder in the virus. A set of 37 human proteins involved in interaction with the RABV proteins was assembled through the literature search, with the majority of data being retrived from [71]. Amino acid sequences and basic disorder-related features of these proteins are provided in **Supplementary Figures S1**, **S2**, **S3**, **S4**, and **S5**.

The amino acid sequences can be analyzed through various methods to identify regions in the protein that have a predisposition towards intrinsic disorder. Through a comparison of these regions of intrinsic disorder and the function of the protein, an analysis can be made of the disorder-based functionality of the protein. Determining the structure

 doi:10.20944/preprints202208.0531.v1

of the protein is important for mapping intrinsic regions of disorder to regions on the protein. Intrinsically disordered regions are flexible. As a result, these regions are not recorded in a crystal structure of the protein. Regions of the protein that do not show up in the crystal structure are indicative of intrinsic disorder [72].

For this study, amino acid FASTA sequences for the Rabies Lyssavirus PV Proteome were gathered through UniProt. The sequences were run through a series of disorder prediction tools to generate an estimate of the intrinsic disorder of each residue. These predictions of intrinsic disorder were averaged to form an overall prediction of intrinsic disorder for that residue. **Table 1** shows a summary of these predictions, including the UniProt entry ID of each protein analyzed. The table also displays the protein length and the length of the longest region of disorder within the protein. Overall disorder was calculated based on the incidence of regions of high disorder. The FASTA sequences used in this study are reproduced below, in Supplementary Materials SM1. The underlined sections of each amino acid sequence correspond to regions of high disorder within the protein. The amino acid sequences presented here were taken from UniProt [70]. Regions of high intrinsic disorder (intrinsic disorder of more than 30%) are shown using bold underlined text. The per-residue intrinsic disorder propensity for each protein is calculated by taking the averages of the PONDR® VL3, PONDR® VLS2, PONDR® VLXT, PONDR® FIT, IUPred_Short, and IUPred_Long predictors.

Table 1. Proteins in the Rabies PV proteome and their evaluated percentage of intrinsically disordered residues as determined by the averaged predicted disorder.

Protein	UniProt entry ID [70]	Protein length (residues)	Longest disordered region (residues)	Percent of disordered residues
P (Phosphoprotein)	P06747	297	87	67.3%
M (Matrix protein)	P08671	202	54	43%
N (Nucleoprotein)	P06025	450	93	30.6%
G (Glycoprotein)	P08667	524	49	27%
L (Large protein)	P11213	2142	105	23%

Predictor of Natural Disordered Regions (PONDR®) is a meta-prediction software that can analyze an amino acid FASTA sequence on a per-residue basis to predict regions of intrinsic disorder. This study made use of the PONDR®VLXT, PONDR®VL3, PONDR®VLS2, and PONDR® FIT meta-predictors from the PONDR family, as well as IUPred_Short and IUPred_Long [73-75].

PONDR®VLXT works by applying three different neural networks: one for each terminal end of an intrinsically disordered sequence and one for the internal region of the sequence. Each network uses a specific dataset that contains only the amino acid residues that are present in that region. The result of the predictor averages the results of the three networks. Transitions between the prediction networks work by averaging the predictors in a short region of overlap at the boundary between the two. PONDR® VLXT is useful for predicting short regions of disorder but underestimates the occurrence of long disordered regions. PONDR® VL3 works by running the residue through ten neural networks and selecting the final prediction by taking the simple majority vote of the predictions. This meta-predictor is known to be useful for predicting longer regions of intrinsic disorder. PONDR® VSL2 combines neural network predictors for short and long disordered regions. The networks are trained using sequences of specific lengths, and the final prediction is a weighted average of the predictions for each length. Due to combining both short and long disordered regions, it is considered the most accurate predictor of the three [73-75].

 doi:10.20944/preprints202208.0531.v1

IUPred works off the assumption that globular and structured proteins have higher numbers of effective inter-residue interactions than disordered proteins do, which means that they have negative free-energy. Structured proteins have lower free-energy estimates compared to disordered proteins. The IUPred meta-predictor is able to use this biophysics based approach to estimate disorder by calculating the pairwise free-energy of the sequence [74,75]. IUPred Long predicts global structural disorder, or disorder in regions with more than 30 consecutive residues. IUPred Short is useful for predicting short, disordered regions, such as region corresponding to the missing residues in the X-ray structure of a largely globular protein. ANCHOR2 is used to predict context-dependent intrinsic disorder. Context-dependent intrinsic disorder may occur when the binding region of an IDPR is able to interact specifically with a globular protein. When bound, these regions adopt an ordered structure. Context-dependent intrinsic disorder may also occur when the change in disorder is due to a change in the redox state. These regions may change their disorder depending on their localization relative to the cell. For all query proteins, the presence of such context-dependent disordered regions, disorder-based binding regions, and molecular recognition features (MoRFs); i.e., disordered regions that fold at the interaction with partners, was analyzed by the ANCHOR algorithm [76,77].

A recently designed computational platform RIDAO (Rapid Intrinsic Disorder Analysis Online) was used to obtain intrinsic disorder-related characteristics of the query proteins [78]. This tool aggregates the results from a number of well-known disorder predictors: PONDR® VLXT [79], PONDR® VL3 [80], PONDR® VLS2B [81], PONDR® FIT [73], IUPred2 (Short) and IUPred2 (Long) [74,82]. Furthermore, RIDAO provides mean disorder profile (MDP) along with the standard errors. It also performs CH-CDF (charge-hydropathy – cumulative distribution function) analysis of query proteins [83-85] and yields data for the ΔCH-ΔCDF plots, which enables rapid discrimination between flavors of disorder [86].

The outputs of the per-residue predictors were averaged, and proteins was grouped based on their percent of predicted intrinsically disordered residues (PPIDR) using an accepted classification criteria [87]. Proteins with an average of intrinsically disordered residues below 10% are considered ordered or mostly ordered. Proteins containing between 10% and 30% of predicted disordered residues are considered moderately disordered. Proteins containing more than 30% of predicted disordered residues are considered highly disordered.

Complementary disorder evaluations together with important disorder-related functional information were retrieved from the D²P² database (http://d2p2.pro/) [88], which is a database of predicted disorder for a large library of proteins from completely sequenced genomes [88]. D²P² database uses outputs of IUPred [74,82], PONDR® VLXT [79], PrDOS [89], PONDR® VSL2B [81], PV2 [88], and ESpritz [90]. The visual console of D²P² displays 9 colored bars representing the location of disordered regions as predicted by these different disorder predictors. In the middle of the D²P² plots, the blue-green-white bar shows the predicted disorder agreement between nine disorder predictors (IUPred, PONDR® VLXT, PONDR® VSL2, PrDOS, PV2, and ESpritz), with blue and green parts corresponding to disordered regions by consensus. Above the disorder consensus bar are two lines with colored and numbered bars that show the positions of the predicted (mostly structured) SCOP domains [91,92] using the SUPERFAMILY predictor [93]. Yellow zigzagged bar shows the location of the predicted disorder-based binding sites (MoRF regions) identified by the ANCHOR algorithm [76], whereas differently colored circles at the bottom of the plot show location of various PTMs assigned using the outputs of the PhosphoSitePlus platform [94], which is a comprehensive resource of the experimentally determined post-translational modifications.

Information on the interactability of human proteins interacting with RABV proteins was retrieved using Search Tool for the Retrieval of Interacting Genes; STRING, http://string-db.org/. STRING generates a network of protein-protein interactions based on

 doi:10.20944/preprints202208.0531.v1

predicted and experimentally-validated information on the interaction partners of a protein of interest [95]. In the corresponding network, the nodes correspond to proteins, whereas the edges show predicted or known functional associations. Seven types of evidence are used to build the corresponding network, where they are indicated by the differently colored lines: a green line represents neighborhood evidence; a red line – the presence of fusion evidence; a purple line – experimental evidence; a blue line – co-occurrence evidence; a light blue line – database evidence; a yellow line – text mining evidence; and a black line – co-expression evidence [95]. In this study, STRING was utilized in three different modes: to create PPI networks centered at individual human proteins, to generate the internal network of protein-protein interactions (PPIs) among the human proteins involved in interactions with the RABV proteins, and to build a PPI network centered at the entire set.

Propensity of the RABV proteins to undergo liquid-liquid phase separation was evaluated by FuzDrop (Fuzzy Droplet Predictor, https://fuzdrop.bio.unipd.it/predictor) [96].

All computer-generated structures of the RABV proteins analyzed in this study were generated using SWISS-MODEL [97] and ExPasy. 3D structural models of human proteins interacting with the RABV proteins were generated by AlphaFold [98].

3. Results and Discussion

3.1. Functional disorder of the P-protein

The P-protein of the RABV, which is a 297 residue-long catalytic polymerase cofactor and regulatory protein that plays an important role in viral transcription and replication, was shown to display a significant amount of intrinsic disorder (see **Figure 2**). In fact, roughly 67% of the protein was found to consist of highly disordered regions. The protein consists of two intrinsically disordered regions, IDD1 and IDD2 that flank an oligomerization domain. Negri bodies, which are inclusions in the host cytoplasm used for viral replication, are formed from the interaction of the two intrinsically disordered regions and the C-terminal domain of P with the intrinsically disordered regions of the N-protein [99]. Additionally, this high disorder content suggests that the mechanism of action for the P-protein to suppress the Type-I interferon-mediated immune response may be to use its disordered status for interaction with the STAT proteins of the cells.

ANCHOR analysis showed that the P-protein contains four disorder-based binding sites, molecular recognition features (MoRFs), which are disordered regions that folds at interaction with specific partners. These are residues 1-21, 89-105, 111-134, and 222-227. The IDPRs that were found in this protein corresponded to ligand binding sites, which supports the idea that the IDPRs interact with the STAT1 and STAT2 proteins [16,17]. It was pointed out that because of these C-terminal domain-driven interactions of the RABV P-protein (CTD, residues 186-297 [100-103]) with host STAT proteins, the P-protein represents the major interferon antagonist of the lyssavirus, thereby affecting the type-I interferon (IFNα/β)-mediated innate immune response [16]. It was also pointed out that interaction of the RABV P-protein with STATs is crucial for the development of the lethal rabies disease [16].

Furthermore, another level of structural and functional complexity of this protein is given by the fact at it has multiple isoforms generated by the alternative initiation of the P-protein during viral transcription. In fact, alternative initiation generates isoforms P2, P3, P4, and P5, which differ from the canonical isoform P1 by missing N-terminal residues 1-19, 1-52, 1-68, and 1-82. An obvious consequence of this truncation is the elimination of the first MoRF region of P1, suggesting that these isoforms might be characterized by different interactability. Curiously, although P3, P4, and P5 all have lost an N-terminal MoRF as expected, P2 was predicted to behave differently. In fact, despite missing the N-terminal residues predicted as the MoRF in P1, this isoform gained three new N-terminal MoRFs (residues 1-14, 19-27, and 30-38).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2022 doi:10.20944/preprints202208.0531.v1

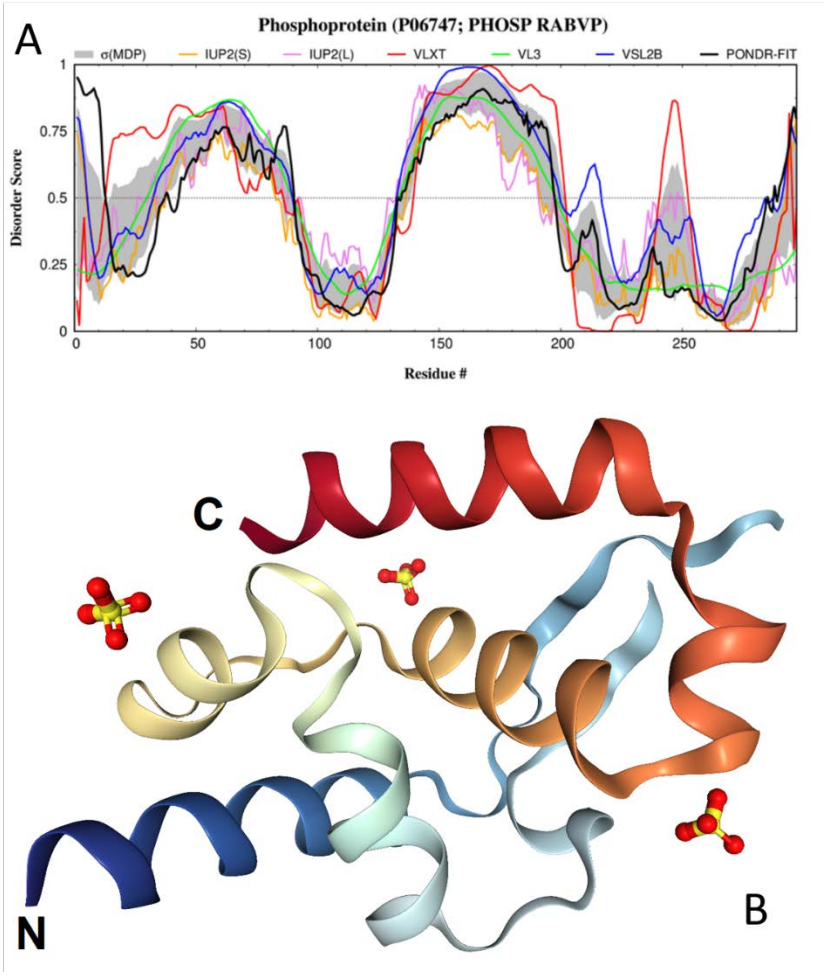

Figure 2. Structure and disorder in the RABV P-protein (UniProt ID: P06747). A. Intrinsic disorder profile generated using data aggregated by RIDAO platform. **B**. Crystal structure of the C-terminal region of the RABV P-protein (residues 192-295) (PDB ID: 3OA1).

Therefore, due to the presence of regions with high intrinsic disorder content, several MoRFs and several phosphorylation sites, and the usage of alternative initiation, it is possible for the P-protein to serve many roles within the virus [5]. Furthermore, the P-protein isoforms were shown to differ in nucleocytoplasmic localization and microtubule (MT) association, mediated by several functional motifs, including nuclear localization sequence (NLS, residues 211-214) and N- and C-terminally located nuclear export sequences (N-NES and C-NES, residues 49-58 and 223-232, respectively) [104-106]. For example, the shorter isoforms (P3 to P5) lacking the N-terminally located NES are more nuclear and are capable of binding and bundling MTs [104]. As per the outputs of PONDR® VSL2, N-NES and NLS are located within IDPRs (residues 33-89 and 208-216, respectively, see **Figure 2A**).

The functional diversity of the RABV P-protein is further increased by the phosphorylation of its serine residues S63 and S64 by an unknown kinase (denoted rabies virus protein kinase [RVPK], and residues S162, S210, and S271 by protein kinase C (PKC) [107], all located within the IDPRs. A recent study revealed that phosphorylation of the P3 isoform of the RABC P-protein at the S210 position resulted in the significant reduction of the nuclear localization modulated to MT binding/bundling of the P3 [108]

Figure 2A shows that intrinsic disorder is unevenly distributed within the P-protein sequence. It is preferentially concentrated at its N-terminal and central regions (residues 1-200), with the C-terminal domain being predicted to possess more ordered structure. In agreement with the results of the computational evaluation of intrinsic disorder predisposition of this protein, a crystal structure was solved for residues 192-295 (see **Figure 2B**)

 doi:10.20944/preprints202208.0531.v1

(PDB ID: 3OA1). Curiously, although a much longer fragment of the P-protein (residues 69-297) was used in the crystallization experiments, a very significant part of this polypeptide (residues 69-191, 220-221, and 296-297) was not observed in the resulting structure, representing regions of missing electron density; i.e., protein regions with high conformational flexibility that preclude them from being crystallized.

Since the P-protein participates in the formation of Negri bodies [99] and can bind PML-bodies [109] we also checked the liquid-liquid phase separation (LLPS) potential of this protein by FuzDrop. This analysis revealed that the longest isoform (P1) of the P-protein has a strong LLPS potential, being characterized by the p_{LLPS} = 0.5276 and containing a droplet-promoting region (DPR, residues 134-184) located within the long central IDPR of this protein (residues 133-199). Shorter isoforms of the P-protein P2, P3, P4, and P5 showed the p_{LLPS} values of 0.6561 (DPR, residues 115-166), 0.6832 (DPRs, residues 1-34 and 82-133), 0.4938 (DPRs residues 1-19 and 66-115), 0.3526 (DPR, residues 52-103), indicating that alternative initiation dramatically affect the LLPS potential of the P-protein. FuzDrop also indicated that P1 contains 9 regions with context-dependent interactions (residues 20-47, 52-64, 67-85, 94-101, 116-125, 129-138, 188-210, 252-257, and 273-282). These regions potentially can undergo disorder-to-order or disorder-to-disorder transitions but maintain the conformational heterogeneity in the bound state are show sensitivity to the cellular context or post-translational modifications, potentially serving as regulatory engines of cellular pathways [110].

3.2. Functional disorder in the M-protein

The M-protein is a 202-residue-long protein that plays a crucial role in the assembly and budding of the virion and engages in complete coverage of the ribonucleoprotein coil to keep it in a condensed bullet-shaped form (see **Figure 1**). It was found to be highly disordered as well, with 43% of this protein being composed of IDPRs (see **Figure 3A**). This suggests that the interactions of the M-protein with RelAp43 and other proteins in the NF-κB pathway [23] may induce suppression of the antiviral response via utilization of some advantages of intrinsic disorder. This is in line with the results of several studies on the potential roles of intrinsic disorder in proteins that form shells of several viruses (such as SARS-CoV-2, MERS-CoV, SARS-CoV, other CoVs, Nipah, Zika, HIV, and retroviruses) for viral transmission, immune evasion and virulence [111-119]. This intrinsic disorder in the M-protein may also allow for increased flexibility of the protein to aid in the regulation of virus budding. The M-protein has been shown to have the ability to create vesicles without any interaction with other viral proteins, suggesting that the flexibility of the protein allows it to induce budding by itself [3,13]. **Figure 3A** represents the disorder profile of the MATRX_RABVP protein and shows that a section in the middle of the protein, spanning roughly from residue 50 to residue 175, displays low disorder content and is likely to represent the structural domain of the protein.

Curiously, region 115-151, which is essential for the glycoprotein binding includes an IDPR (residues 129-141), indicating that intrinsic disorder of this region can contribute to its interactability. Furthermore, the M-protein contains a PPxY motif (residues 35-38), which is commonly found in viral proteins capable of manipulation of the autophagic machinery to prevent autophagic degradation of viruses [121]. This PPxY motif is included into the long N-terminal IDPR (residues 1-48) (see **Figure 3A**).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2022 doi:10.20944/preprints202208.0531.v1

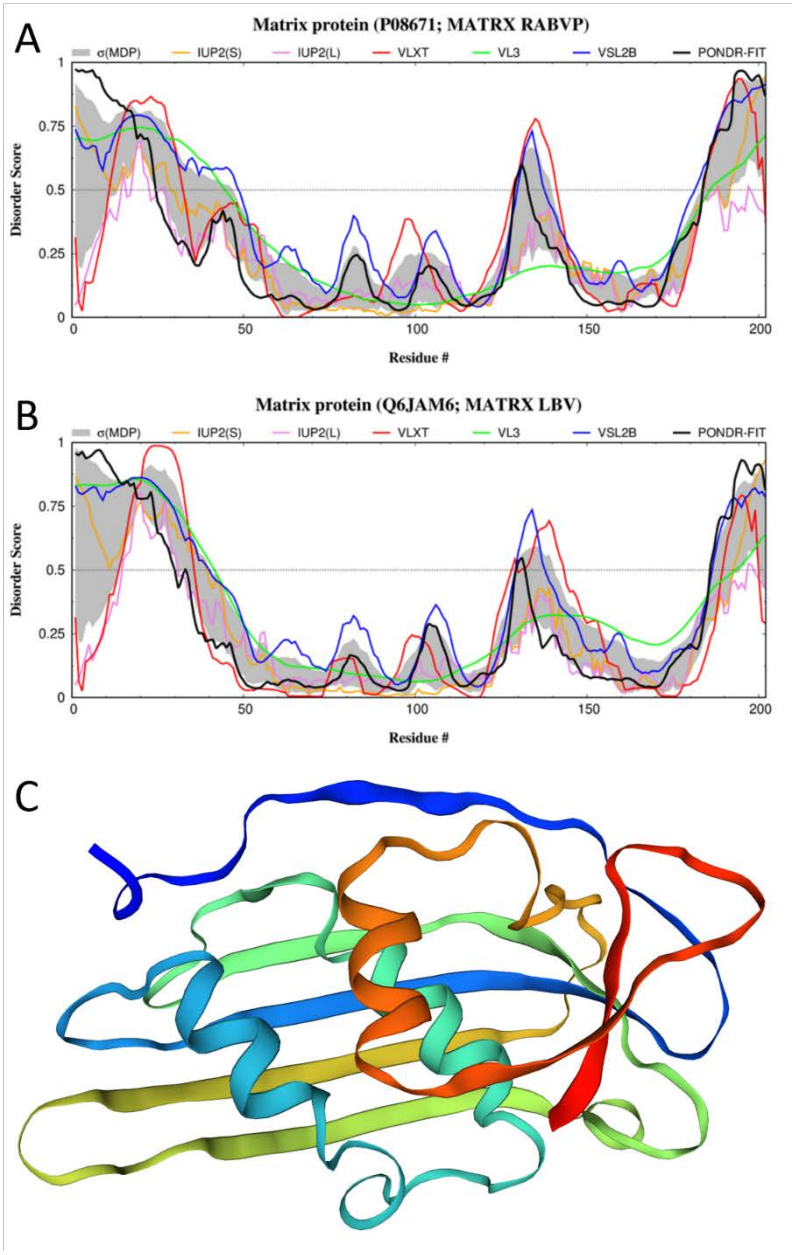

Figure 3. Structure and disorder in the M-protein from RABV (UniProt ID: P08671). **A**. Intrinsic disorder profile generated the M-protein from RABV strain VP by RIDAO. **B**. Intrinsic disorder profile generated for the Lagos bat virus matrix protein (UniProt ID: Q6JAM6) by RIDAO. **C**. A structural model for the 30-202 fragment of the RABV M-protein that was created by SWISS-MODEL [97] using as a template the structure of the Lagos bat virus matrix protein (PDB ID: 2W2S; [120]) that shows the sequence identity to the query M-protein of 76.73%.

Although no structural information is currently available for the M-protein of RABV, we used SWISS-MODEL (https://swissmodel.expasy.org/) to create homology models for this protein. **Figure 3C** shows a model for the 30-202 fragment of the RABV M-protein that was created using the structure of the Lagos bat virus matrix protein (PDB ID: 2W2S; [120]; UniProt ID: Q6JAM6) as a template with sequence identity to the query M-protein of 76.73%. Comparison of the **Figures 3B** and **3B** illustrates a remarkable similarity of the per-residue intrinsic disorder predispositions of the Lagos bat virus matrix protein and the RABV M-protein, thereby providing the intrinsic disorder-based validation of the selection of the Lagos bat virus matrix protein as a template.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2022 doi:10.20944/preprints202208.0531.v1

Finally, FuzDrop analysis indicated that although the M-protein shows low probability of the spontaneous LLPS (p_{LLPS} = 0.2220), this protein contains one C-terminally located DPR (residues 199-202) which is included into the IDPR (residues 182-202). FuzDrop also identified regions 16-29, 121-131, and 172-190 as regions with content-dependent interactions [110].

3.3. Functional disorder in the N-protein

The 450-residue-long N-protein is a viral RNA-binding protein that encapsulates the genome in a ratio of one N-protein per nine ribonucleotides. The long C-terminal IDPR and several shorter IDPRs make up roughly 31% of the protein (see **Figure 4A**).

The N-protein has been shown to form a complex with the P-protein during replication and encapsulate the viral genomic RNA to protect it against nucleases. It binds through the N-terminus of the protein, which is highly intrinsically disordered.

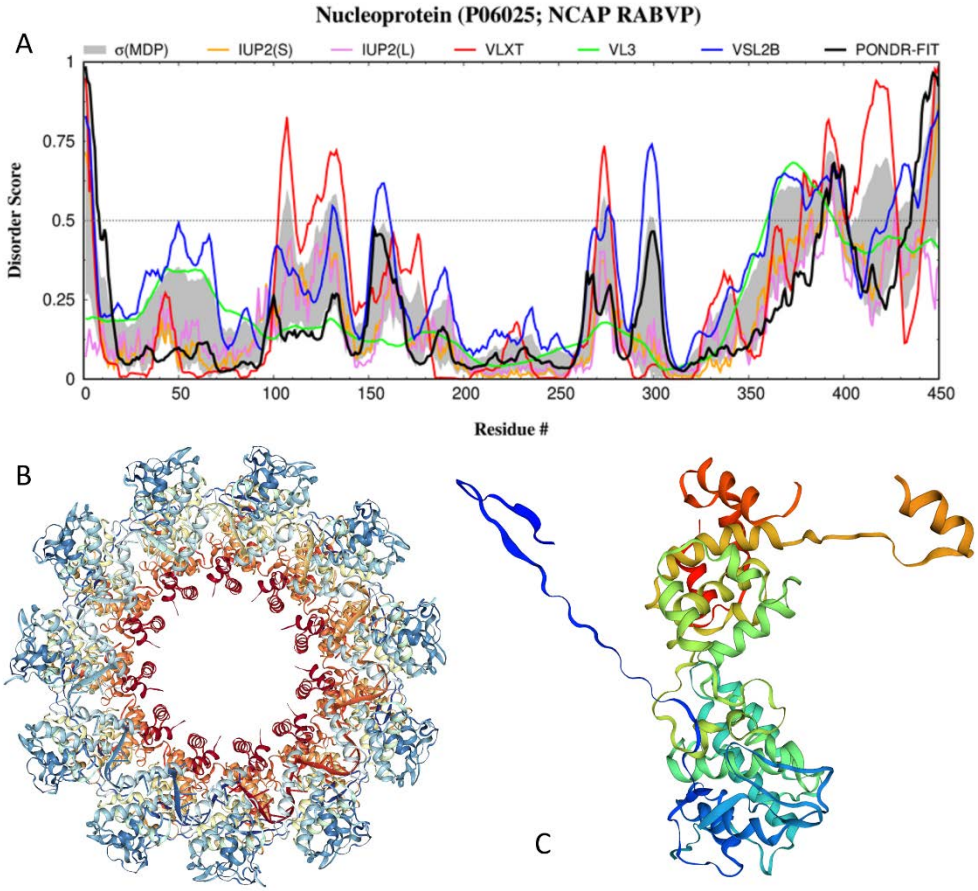

Figure 4. Structure and disorder in the N-protein from the RABV strain PV (UniProt ID: P06025). A. Intrinsic disorder profile generated the M-protein from RABV strain VP by RIDAO. B. Crystal structure of the N-protein from the RABV strain ERA (which is 99.11% identical to the N-protein from the RABV strain PV) in complex with RNA (PDB ID: 2GTT; [122]), where the protomers of the N-protein are organized in an undecameric ring. C. Crystal structure of the N-protein protomer taken computationally out the undecameric homo-oligomer and demonstrating the presence of two "arms" in the structure (residues 6-28 and 349-414).

The complex formed during replication, called a Negri body, is an inclusion in the host's cytoplasm that is formed from interactions of the highly disordered C terminal of the N-protein with the intrinsically disordered regions of the P-protein [123]. The complex prevents non-specific RNA binding and phosphorylation of the RNA [7]. The N-protein is predicted to have one MoRF (residues 406-411) located within the

 doi:10.20944/preprints202208.0531.v1

long C-terminal IDPR, which also includes a phosphoserine at the position 389. A large region in the middle of the protein with little disorder suggests that the protein serves a largely structural purpose. The N-protein also functions to prevent activation of the RIG-I mediated antiviral response [19,20]. The IDPR found in the N-protein may serve to aid in suppressing the antiviral response. **Figure 4A** shows the results of the disorder prediction for the NCAP_RABVP protein. Crystal structure of the N-protein from the RABV strain ERA (which is 99.11% identical to the N-protein from the RABV strain PV) in complex with RNA was solved (PDB ID: 2GTT; [122]). **Figure 4B** shows that in this RABV nucleoprotein-RNA complex, the N-proteins are organized in an undecameric ring. In such a complex, the two core domains of the nucleoprotein clamp around the RNA at their interface and shield it from the environment [122]. Polymerization of the nucleoprotein is achieved by domain exchange between protomers, with flexible hinges allowing nucleocapsid formation. The wide curvature of the C terminus of the protein, which is highly intrinsically disordered, makes it able to more tightly cover bound RNA. The nucleoprotein and the RNP core are able to adopt different confirmations as a result at different periods of the viral cycle [123]. This important observation is illustrated by **Figure 4C** showing structure of the N-protein protomer and demonstrating the presence of two "arms" in the structure (residues 6-28 and 349-414). Importantly, the C-terminal arm contains MoRF. In addition to these functional arms, each protomer has several regions of missing electron density (residues 1-5, 104-117, 186-188, 374-397, and 449-450).

Although the N-protein is expected to have low probability of the spontaneous LLPS (p_{LLPS} = 0.1405) and does not include any DPRs, one can find there four regions with the content-dependent interactions (residues 105-115, 284-318, 367-396, and 398-411), which overlap, include or are included in disordered/flexible regions of this protein (residues 103-111, 294-303, 361-400, and 403-428) (see **Figure 4A**).

3.4. Functional disorder in the G-protein

The G-protein is a 524-residue-long type I transmembrane protein with a long extravirion region (residues 20-459), a transmembrane helix (residues 460-480), and an intravirion domain (residues 481-524). Being located on the surface of RABV particles, the G-protein controls the receptor binding and the release of the viral ribonucleoprotein (RNP) in the cytoplasm via the pH-dependent membrane fusion, thereby playing a crucial role in the cell entry and *in vivo* spread [124]. Furthermore, it was shown that the G-protein (in particular, it's ectodomain) accumulates the adaptive mutations improving the release of infectious viral particles [125].

The G-protein is synthesized as a precursor with the N-terminal signal peptide (residues 1-19), which is removed during maturation of this protein. Similar to the proteins found in the envelopes of other viruses, the RABV G-protein forms homotrimers on the surface of the virion that are responsible for the attachment of the virus to the host cellular receptors, such as the muscular form of the nicotinic acetylcholine receptor (nAChR), the neuronal cell adhesion molecule (NCAM), and the p7 5 neurotrophin receptor (p75NTR). There are approximately 400 such trimeric spikes, which are tightly arranged on the surface of the virus. The C-terminal domain of the G protein is essential for trimer stability [9]. **Figure 5A** shows that the G-protein is mostly ordered and contains relatively few IDPRs, which is a characteristic feature of spike/glycoproteins of many other viruses. The most promising IDPR is the C-terminal IDPR (residues 486-524), which corresponds to the intravirion domain of this protein that is engaged in interaction with the matrix protein. Several flexible regions located within the extravirion domain serve as aid in the interaction of the G-protein with surface molecules of the host cell [24]. There are three glycosylation cites in this protein (asparagine residues 56, 266, and 338), and the C-terminally located lipidation site, S-palmitoyl cysteine 479.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2022 doi:10.20944/preprints202208.0531.v1

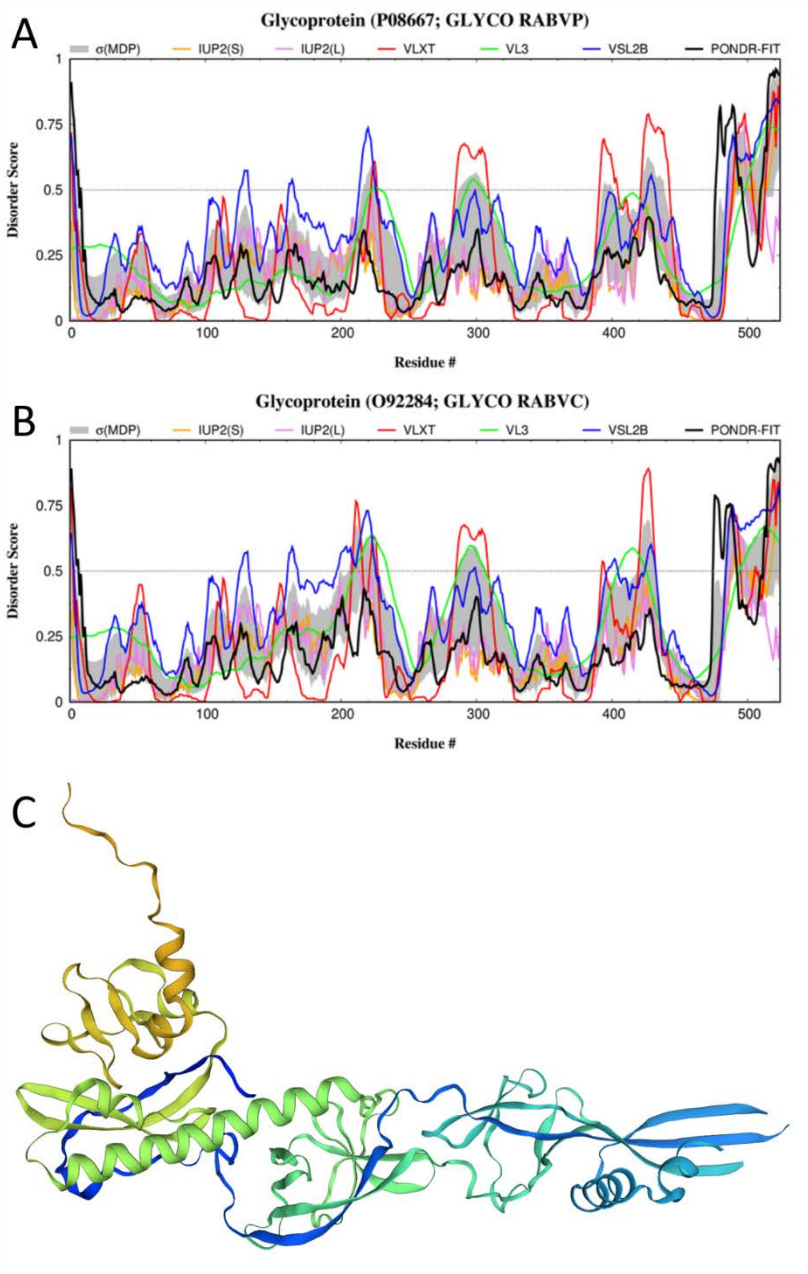

Figure 5. Structure and disorder in the G-protein from RABV. **A**. Intrinsic disorder profile generated the G-protein from RABV strain VP (UniProt ID: P08667) by RIDAO. **B**. Intrinsic disorder profile generated for the G-protein from the RABV strain CVS-11 by RIDAO (UniProt ID: O92284). **C**. A structural model for the 20-424 fragment of the G-protein from RABV strain VP generated by SWISS-MODEL [97] using as a template known structure of the G-protein from the RABV strain CVS-11 (PDB ID: 6LGW [126]; UniProt ID: O92284) with sequence identity to the query G-protein of 91.48%.

There is currently no structural information for the G-protein from the RABV strain PV. Therefore, we used SWISS-MODEL to create a homology model for this protein. **Figure 5C** shows a model for the 20-424 fragment of this protein that was generated using the known structure of the G-protein from the RABV strain CVS-11 (PDB ID: 6LGW [126]; UniProt ID: O92284) as a template with sequence identity to the query G-protein of 91.48%. This structure is characterized by a highly elongated form and a presence of a C-terminal "arm" (residues 400-416). In the original structure of the G-protein from the RABV strain CVS-11, there are several regions of missing electron density (residues 21-24, 95-101, 133-143, 202-204, and 414-429).

 doi:10.20944/preprints202208.0531.v1

Furthermore, the authors of this study represent a structure for another form of this protein (PDB ID: 6LGX), which shows different pattern of missing electron density regions (residues 21-25, 91-105, 131-147, 274-294, and 427-463). Since these two structures were resolved at different conditions (at ~pH-6.5 in the complex with a neutralizing antibody 523-11 (PDB ID: 6LGW) and at ~pH-8.0 in free form (PDB ID: 6LGX), these observations suggest that this G-protein structure is characterized by the noticeable sensitivity to the environmental conditions. The authors pointed out that the basic-to-acidic pH change results in the large re-orientations of the three domains found in the G-protein from the RABV strain CVS-11 leading to the concomitant domain-linker reconstructions that switches from a bent hairpin conformation into an extended conformation [126]. These low-pH-induced structural transitions within the domain-linker region are related to the functionality of the G-protein playing important roles in the G-protein-mediated membrane fusion [126]. **Figures 5B** and **5B** show that the G-proteins from the RABV strains PV and CVS-11 are characterized by very similar disorder profiles, suggesting that the observations made for the structural peculiarities of the RABV G-protein from the CVS-11 strain are applicable to the G-protein from the RABV strain PV as well. The IDPR in the G-protein likely serves to facilitate endocytosis into the cell through interacting with cell recognition molecules [24].

Similar to the N-protein, the glycoprotein of RABV is characterized by low probability of the spontaneous LLPS (p_{LLPS} = 0.1351) and does not include any DPRs, but contains five regions with content-dependent interactions (residues 18-29, 218-223, 229-239, 434-441, and 502-508), which overlap, include or are included in disordered/flexible regions of this protein (residues 214-225, 422-442, and 486-524) (see **Figure 5A**).

3.5. Functional disorder in the L-protein

With the amino acid sequence of 2,124 residues, the L-protein is the longest protein in the RABV proteome. This protein is an RNA-directed RNA polymerase (RdRp) that catalyzes the transcription of viral mRNAs as well as their polyadenylation and capping. It has several functional regions, such as a RdRp catalytic domain (residues 611-799), a mononegavirus-type SAM-dependent 2'-O-MTase domain (residues 1,674-1,871) included into the C-terminal region that is involved in interaction with the P-protein (residues 1,562-2,127) and containing a disorder-based interaction site, and a MoRF (residues 1,631-1,638). In addition to interactions with the P-protein, L-protein may form homodimers. Although the L-protein includes many disordered or flexible regions, its overall intrinsic disorder level is relatively low (see **Figure 6A**).

The need for interplay between ordered and disordered structures in this protein reflects its purpose to serve as a low fidelity viral polymerase. The IDPRs likely define the lower fidelity of the polymerase, which results in a high mutation rate and therefore higher flexibility in the ability of the virus to adapt to host defenses [7]. The polymerase activity of the L protein depends on the identity of residues upstream to the protein as well as the identity of C-terminal residues [127]. **Figure 6C** shows the structural model for the L-protein from RABV strain PV built using the cryo-EM structure of the large structural protein from the RABV strain SAD B19 (sequence identity: 98.68%) complexed with the fragment of the P-protein (PDB ID: 6UEB, [128]; UniProt ID: P16289) as a template. Due to a high sequence similarity, disorder profiles of the L-proteins from RABV strains SAD B19 and PV are almost identical (cf. **Figures 6A** and **6B**). Since in the aforementioned cryo-EM structure of the L-P complex, residues 1-27 of the L-protein from the RABV strain SAD B19 constitute a region of missing electron density, it is likely that this N-terminal region is disordered in the L-protein from the RABV strain PV as well.

FuzDrop indicated that the L-protein has a low p_{LLPS} value of 0.11860 and does not contain any DPRs. However, this protein possesses 24 regions with content-dependent interactions (residues 5-12, 104-109, 416-434, 456-480, 507-514, 566-571, 672-677, 821-833, 852-859, 878-899, 903-911, 924-929, 982-987, 1014-1022, 1152-1157, 1232-1238, 1383-1388,

 doi:10.20944/preprints202208.0531.v1

1390-1399, 1562-1570, 1655-1667, 1731-1753, 1980-1985, 2086-2092, and 2127-2132). Many of these regions are related to the IDPRs identified in the L-protein (e.g., residues 1-42, 108-114, 466-480, 504-512, 827-839, 1555-1599, 1642-1656, 1735-1750, 2093-2096, and 2137-2142).

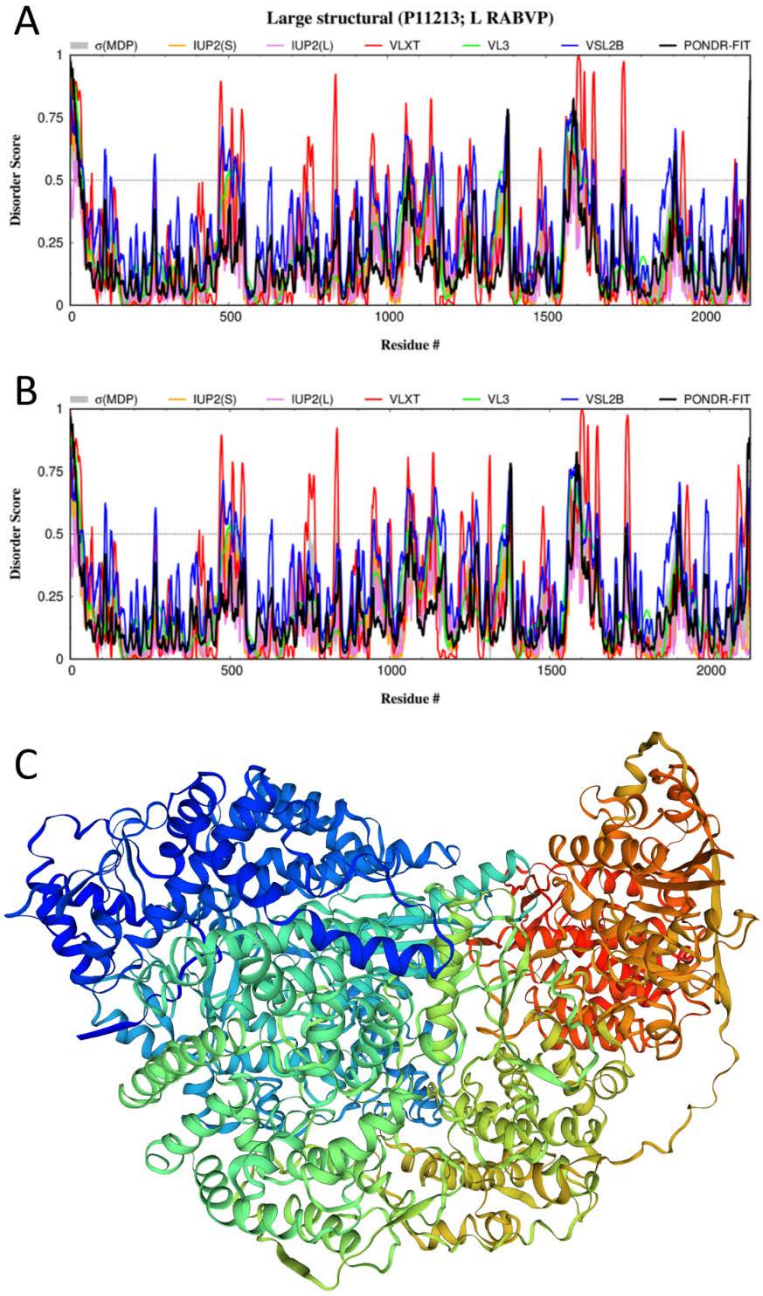

Figure 6. Structure and disorder in the L-protein from RABV. **A**. Intrinsic disorder profile generated the L-protein from RABV strain VP (UniProt ID: P11213) by RIDAO. **B**. Intrinsic disorder profile generated for the L-protein from the RABV strain SAD B19 by RIDAO (UniProt ID: P16289). **C**. A structural model for the L-protein from RABV strain PV built by SWISS-MODEL [97] using as a template the cryo-EM structure of the large structural protein from the RABV strain SAD B19 (sequence identity: 98.68%) complexed with the fragment of the P-protein (PDB ID: 6UEB, [128]; UniProt ID: P16289).

3.7. Intrinsic disorder in human proteins interacting with the RABV proteins

3.7.1. Host interactors of the P-protein

 doi:10.20944/preprints202208.0531.v1

In addition to be involved in interaction with the RABV L- and N-proteins and the viral ribonucleocapsids, the P-protein is known to bind to the multitude of host proteins, such as dynein light chain 1 and 2 (DYNLL1 (UniProt ID: P63167) and DYNLL2 (UniProt ID: Q96FJ2), respectively) [129], as well as host signal transducer and activator of transcription 1-alpha/beta and signal transducer and activator of transcription 2 (STAT1 (UniProt ID: P42224) and STAT2 proteins (UniProt ID: P52630), respectively) [16,17,100-103], promyelocytic leukemia (PML) protein (UniProt ID: P29590) [109], the ribosomal protein L9 (UniProt ID: Q02878) [15], STAT3 (UniProt ID: P40763), nucleolin (NCL; UniProt ID: P19338), focal adhesion kinase (FAK; UniProt ID: Q05397), Janus kinase 1 (JAK1; UniProt ID: P23458), inhibitor of nuclear factor kappa-B kinase subunit epsilon (IKKε; UniProt ID: Q14164), Beclin-1 (BECN1; UniProt ID: Q14457), tubulin alpha (TUB-α; UniProt ID: Q71U36 for tubulin alpha-1A chain), tubulin beta (TUB-β; UniProt ID: Q9H4B7 for tubulin beta-1 chain), ATP-binding cassette sub-family E member 1 (ABCE1; UniProt ID: P61221), T-complex protein 1 subunit gamma (CCTγ; UniProt ID: P49368), Hsp90 co-chaperone Cdc37 (CDC37; UniProt ID: Q16543), and heat shock protein HSP 90-alpha (Hsp90AA1; UniProt ID: P07900) [71]. Furthermore, P-protein can interact with the complex I in mitochondria leading to mitochondrial dysfunction, increased generation of reactive oxygen species (ROS), and oxidative stress [130]. Being the largest and most complicated component of the respiratory chain, complex I contains 45 subunits [131]. Unfortunately, the exact targets of the P-protein within this complex are unknown. Therefore, the proteins forming complex I were not included in analysis.

3.7.2. Host interactors of the M-protein

Several human proteins interacting who the RABV M-protein where established. The list includes RelAp43 (which is a splicing variant of RelA (UniProt ID: Q04206) and other proteins in the NF-κB pathway [23], as well as V-type proton ATPase catalytic subunit A (ATP6V1A; UniProt ID: P38606), E3 ubiquitin-protein ligase NEDD4 (NEDD4; UniProt ID: P46934), Transcriptional coactivator YAP1 (UniProt ID; P46937), Eukaryotic translation initiation factor 3 subunit H (EIF3H, UniProt ID: O15372), JAK1 (UniProt ID: P23458), and STAT1 (UniProt ID: P63167) [71]. Note that the M-protein shares two human interactors (JAK1 and STAT1) with the P-protein [71].

3.7.3. Host interactors of the N-protein

Very few host proteins were shown to act as physical partners of the RABV N-protein. These include heat shock 70 kDa proteins 1A (HSPA1A, UniProt ID: P0DMV8) and 1B (HSPA1B, UniProt ID: P0DMV9), prefoldin subunit 1 (PFDN1, UniProt ID: O60925), and CCTγ (UniProt ID: P49368), with the CCTγ being a shared partner for the RABV P- and N-proteins [71].

3.7.4. Host interactors of the G-protein

The major biological function of the G-protein is cell entry via interaction with RABV entry receptors, such as nicotinic acetylcholine receptors (nAChR), the neuronal cell adhesion molecule (NCAM1; UniProt ID: P13591), the low affinity p75 neurotrophin receptor (p75NTR, also known as tumor necrosis factor receptor superfamily member 16; UniProt ID: P08138), and metabotropic glutamate receptor subtype 2 (mGluR2; UniProt ID: Q14416) [132-135]. nAChRs are assemblies of five subunits, which are arranged symmetrically around a central pore. It is known that the neuronal subtypes of nAChRs (which serve as the receptors of the RABV G-protein) exist as various homomeric or heteromeric (at least one α and one β) combinations of twelve different nicotinic receptor subunits: α_2–α_{10} and β_2–β_4 (with some of the neuronal nAChR subtypes being $(\alpha_4)_3(\beta_2)_2$, $(\alpha_4)_2(\beta_2)_3$, $(\alpha_3)_2(\beta_4)_3$, $\alpha_4\alpha_6\beta_3(\beta_2)_2$, $(\alpha_7)_5$). Therefore, there are multiple possibilities for binding of the G-protein to nAChR. Physical interactions of the nicotinic acetylcholine receptor alpha 1

 doi:10.20944/preprints202208.0531.v1

(nAChR α1 or CHRNA1; UniProt ID: P02708) [136] and nAChr α7 (CHRNA7, UniProt ID: P36544) [137] with the RABV G-protein were demonstrated. Furthermore, the G-protein can bind host microtubule-associated serine/threonine-protein kinases 1 and 2 (MAST1 (UniProt ID: Q9Y2H9) and MAST2 (UniProt ID: Q6P0Q8), respectively), tyrosine-protein phosphatase non-receptor type 4 (PTPN4; UniProt ID: P29074), disks large homolog 2 (DLG2; UniProt ID: Q15700), multiple PDZ domain protein (MPDZ; UniProt ID: O75970), and synaptosomal-associated protein 25 (SNAP25; UniProt ID: P60880) [71].

3.7.5. Host interactors of the L-protein

The only human partner of the L-protein is DYNLL1, which is shared with the P-protein [71].

3.7.6. Prevalence of intrinsic disorder in human proteins interacting with RABV

Detailed characterization of the prevalence of functional disorder in each of the human proteins interacting with the RABV P-, M-, N,- G-, and L-proteins are systemized in **Supplementary Figures S1**, **S2**, **S3**, **S4**, and **S5**, respectively. The corresponding analyses were conducted using a set of computational tools, such as RIDAO, STRING, D^2P^2 and AlphaFold. This revealed a very high level of disorder in the majority of the proteins from this dataset, with the entire set being characterized by the mean PPIDR of 41.6±20.9 % (as evaluated using the outputs of the PONDR® VSL2 predictor, which is one of the most accurate stand-alone disorder predictors). **Supplementary Figures S1**, **S2**, **S3**, **S4**, and **S5** show that all these proteins contain multiple IDPRs of various length. Many proteins contain multiple MoRFs, and almost all human proteins in this dataset are densely decorated by a multitude of different PTMs. These observations indicate that intrinsic disorder in this these proteins is related to their functionality, playing a role in their binding promiscuity, as evidenced by dense PPI networks centered at these proteins.

These observations are further illustrated by **Figure 7** presenting global intrinsic disorder characteristics of 37 human proteins. In fact, **Figure 7A** shows there is no a single protein in this dataset that could be classified as mostly ordered, whereas 25 proteins (67.6%) are expected to be mostly disordered (i.e., their PPIDR exceeds 30%). Furthermore, PPIDR values for 13 proteins (NCL (P19338; 86.20%), YAP1 (P46937; 82.94%), SNAP25 (P60880; 78.16%), MAST2 (Q6P0Q8; 73.75%), MAST1 (Q9Y2H9; 68.60%), CDC37 (Q16543; 65.08%), RelAp43 (Q04206; 64.61%), NEDD4 (P46934; 62.17%), L9 (Q02878; 61.46%), BECN1 (Q14457; 58.00%), p75NTR (P08138; 57.14%), PML (P29590; 54.06%), and PFDN1 (O60925; 50%)) exceed 50%, defining them as highly disordered.

Figure 7B shows the ΔCH-ΔCDF plot (a combined binary predictor of intrinsic disorder) that verifies a global prevalence of intrinsic disorder in 37 human proteins interacting with the RABV proteins. The ΔCH-ΔCDF plot provides means for evaluation of the flavors of intrinsic disorder. **Figure 7B** shows that quadrant Q1 (bottom left corner) contains 21 proteins that are predicted to be ordered by both predictors; quadrant Q2 (bottom right corner) includes 10 proteins, which are predicted to be ordered/compact by CH-plot and disordered by CDF (i.e., it contains either molten globular proteins, which are compact, but without unique 3D structures, or hybrid proteins containing comparable levels of ordered and disordered residues); quadrant Q3 (top right corner) includes 4 highly disordered (native coils or native pre-molten globules), which are predicted as disordered by both predictors. Finally, one proteins in quadrant Q4 (top left corner) is predicted to be disordered by CH-plot and ordered by CDF-plot. Therefore, 16 human proteins interacting with the RABV proteins are predicted as containing very noticeable levels of disorder.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2022 doi:10.20944/preprints202208.0531.v1

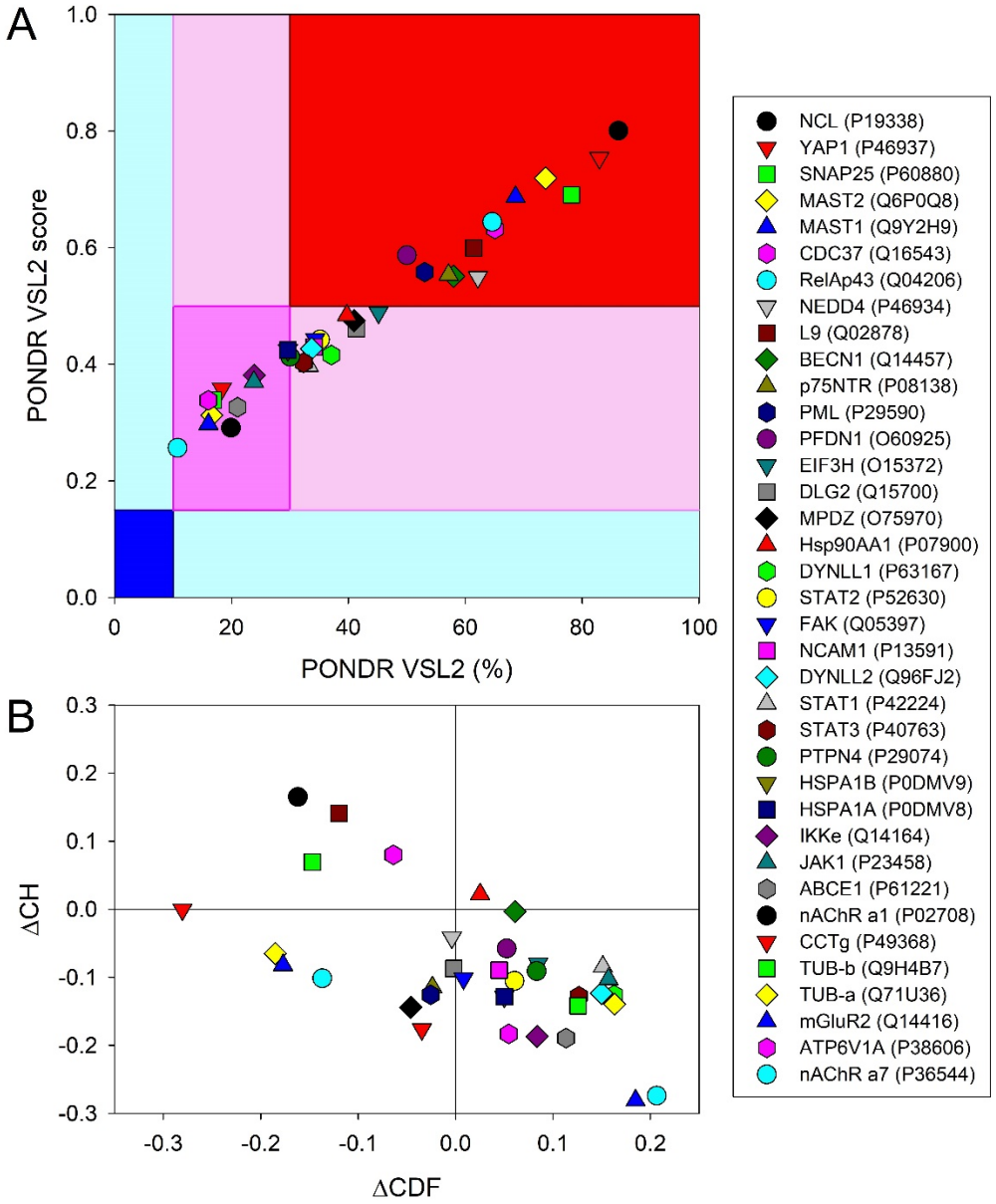

Figure 7. Evaluation of global disorder in 37 human proteins interacting with the RABV proteins. A. PONDR® VSL2 output for 37 human proteins. PONDR® VSL2 score is the average disorder score (ADS) for a protein. PONDR® VSL2 (%) is a percent of predicted disordered residues (PPDR); i.e., residues with disorder scores above 0.5. Color blocks indicate regions in which proteins are mostly ordered (blue and light blue), moderately disordered (pink and light pink), or mostly disordered (red). If the two parameters agree, the corresponding part of background is dark (blue or pink), whereas light blue and light pink reflect areas in which only one of these criteria applies. **B**. Charge-hydropathy and cumulative distribution function (CH-CDF) plot. The Y-coordinate is calculated as the distance of the corresponding protein from the boundary in the CH plot. The X-coordinate is calculated as the average distance of the corresponding protein's CDF curve from the CDF boundary. The quadrant that the protein is located determines its classification. Q1, protein predicted to be disordered by CH-plot and ordered by CDF. Q2, protein predicted to be ordered by CH-plot and CDF. Q3, protein predicted to be ordered to by CH-plot and disordered by CDF-plot. Q4, protein predicted to be disordered by CH-plot and CDF.

Not surprisingly, all 37 proteins in the analyzed set were found to form a rather dense PPI network (see **Figure 8A**), in which, on average, each protein interacts with at least 11 partners.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2022 doi:10.20944/preprints202208.0531.v1

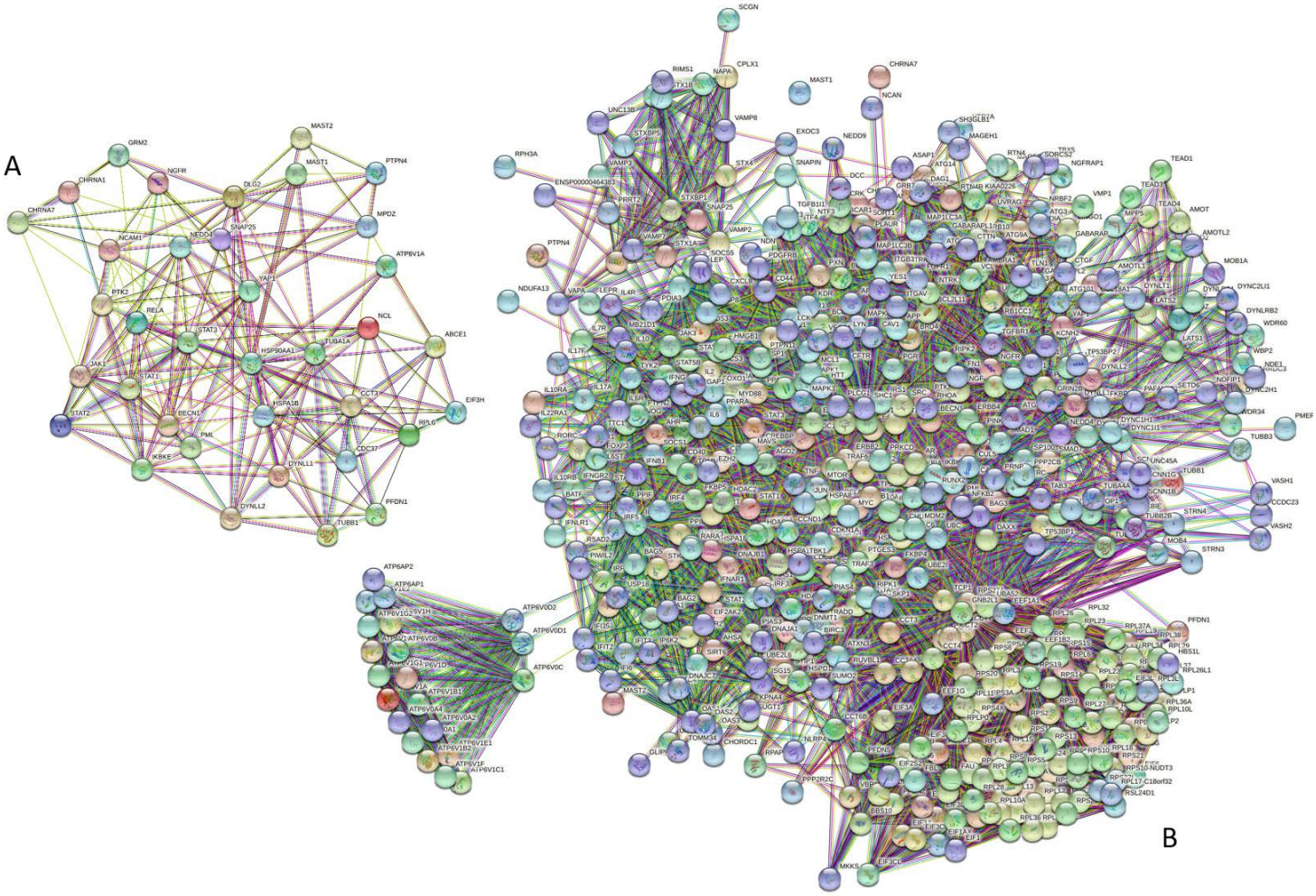

Figure 8. Intraset and set-based ineractivity of human proteins engaged in intraction with the RABV proteins. **A**. STRING-generated PPI network within the analyzed set of human prteins. To include all the proteins into the network, a low confidence of 0.15 as a minimum required interaction score was used in this case. This network includes 36 proteins linked by 204 intections. The resulting average node degree of this network is 11.3, and its average local clustering coefficient (which defines how close its neighbors are to being a complete clique; the local clustering coefficient is equal to 1 if every neighbor connected to a given node Ni is also connected to every other node within the neighborhood, and it is equal to 0 if no node that is connected to a given node Ni connects to any other node that is connected to Ni) is 0.578. Since the expected number of edges for a network of same size for proteins randomly selected from human proteome is 128; this network is characterized by the PPI enrichment p-value of 4.7×10^{-10}. **B**. STRING-generated PPI nework centered at the human proteins interacting with the RABV proteins. Note that the number of interactors in STRING is limited to 500. This network generated with high confidence score of 0.7 includes 536 proteins connected by 11,358 interactions. The average node degree average local clustering coefficient of this PPI are 42.4 and 0.631, respectively; its PPI enrichment p-value is $<10^{-16}$.

In this intraset PPI network, five the most significantly enriched biological processes were Viral process (GO:0016032; $p = 3.49\times10^{-8}$), Immune effector process (GO:0002252; p = 0.00030), Interspecies interaction between organisms (GO:0044419; p = 0.00030), Intracellular transport (GO:0046907; p = 0.00057), and Cellular component organization (GO:0016043; p = 0.00057). Among the molecular functions, five the most significantly enriched were Ubiquitin-like protein ligase binding (GO:0044389; $p = 7.02\times10^{-7}$), Protein binding (GO:0005515; $p = 1.22\times10^{-6}$), Ubiquitin protein ligase binding (GO:0031625; $p = 6.83\times10^{-5}$), Enzyme binding (GO:0019899; $p = 8.27\times10^{-5}$), and Binding (GO:0005488; $p = 8.27\times10^{-5}$). Finally, five the most significantly enriched cellular components were Postsynapse (GO:0098794; $p = 1.79\times10^{-6}$), Cytosol (GO:0005829; $p = 1.79\times10^{-6}$), Synapse

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2022 doi:10.20944/preprints202208.0531.v1

(GO:0045202; p = 2.96×10^{-5}), Cell junction (GO:0030054; p = 0.00010), and Neuron projection (GO:0043005; p = 0.00017).

We also analyzed the set-based interactivity of these human proteins. To this end, STRING was used to generate a PPI network that includes first shell interactors (see **Figure 8B**). The resulting highly interlinked interactome includes 536 proteins connected by 11,358 interactions. Therefore, this interactome is characterized by an average node degree of 42.4, and it shows an average local clustering coefficient of 0.631. The expected number of interactions for the set of proteins of its size is 5.089, indicating this PPI network centered at human proteins interacting with the RABV proteins has significantly more interactions than expected (PPI enrichment p-value is < 10^{-16}). Looking at the functional enrichment of this network based on the Gene Ontology (GO) terms revealed that five the most significantly enriched biological processes were Interspecies interaction between organisms (GO:0044419; p = 6.95×10^{-102}), Viral process (GO:0016032; p = 1.60×10^{-99}), Symbiotic process (GO:0044403; p = 1.60×10^{-99}), Translational initiation (GO:0006413; p = 7.45×10^{-88}) and Cellular response to organic substance (GO:0071310; p = 6.37×10^{-76}). Five the most significantly enriched functions were Structural constituent of ribosome (GO:0003735; p = 5.85×10^{-66}), Protein binding (GO:0005515; p = 3.38×10^{-57}), Binding (GO:0005488; p = 2.02×10^{-52}), Enzyme binding (GO:0019899; p = 3.08×10^{-46}), and RNA binding (GO:0003723; p = 2.09×10^{-37}). Among the cellular components, five the most significantly enriched were Cytosol (GO:0005829; p = 6.45×10^{-106}), Protein-containing complex (GO:0032991; p = 2.61×10^{-82}), Cytosolic ribosome (GO:0022626; p = 5.33×10^{-81}), Cytoplasm (GO:0005737; p = 6.02×10^{-68}), and Ribosomal subunit (GO:0044391; p = 5.65×10^{-65}).

Although detailed description of the potential roles of intrinsic disorder in functionality of human proteins shown to interact with the RABV proteins is outside the scopes of this study, it is clear that all the proteins analyzed here (i.e., RABV N-, L-, P-, M- and G-proteins and 37 human proteins) contain very significant levels of intrinsic disorder. This is further illustrated by **Figure 9**, showing experimentally identified and validated interactions between five RABV proteins, G (glycoprotein), N (nucleoprotein), L (RNA-dependent polymerase), P (phosphoprotein), and M (matrix protein) and 37 human proteins. This diagram clearly shows that most of the proteins (viral and human) in this network are "red" (highly disordered), and there are no "blue" (mostly ordered) proteins, suggesting importance of intrinsic disorder for RABV infection.

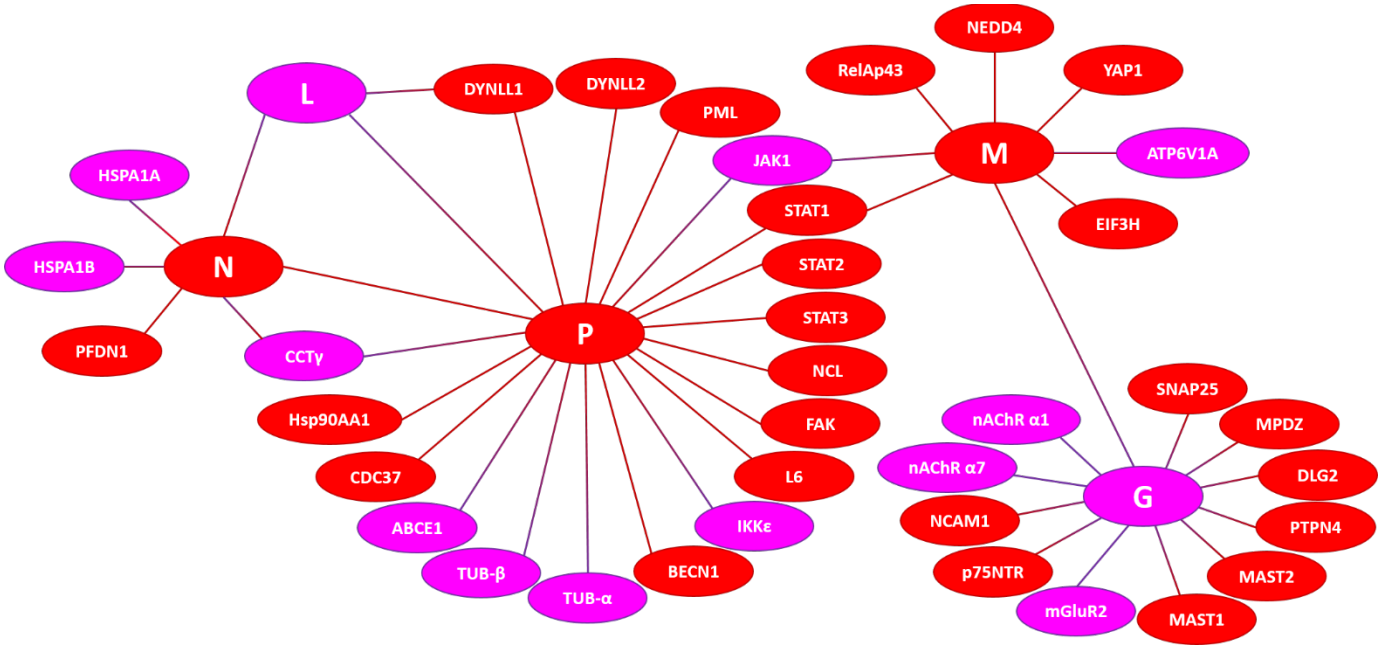

Figure 9. Disordered interactome of the RABV N-, L-, P-, M-, and G-proteins. Proteins are colored based on their PPIDR values, with highly and moderately disordered proteins being shown by red and pink color, respectively.

 doi:10.20944/preprints202208.0531.v1

4. Conclusions

All Rabies Lyssavirus PV proteins contain IDPRs, most of which aid in the flexibility of the virus and its ability to evade host antiviral defenses. All human proteins found as the RABV interactors also contain high levels of intrinsic disorder, with most of these proteins being highly disordered. This disorder-centric layer of complexity in the RABV and its interaction with host proteins adds a new angle to the search for the potential target for anti-Rabies drugs. Once the virus has infected a host cell, there are virtually no effective treatments available that can be used to destroy the virus. Although the modern RABV vaccine has largely eradicated the virus in the developed countries, many people around the world are unable to seek treatment until the virus has already crossed the blood-brain barrier. Without immediate usage of the vaccine upon infection, the RABV-related mortality rate is close to 100%. The advent of bioinformatics approaches in a clinical setting has inspired many developments that have led to the creation of new drugs. It is likely that targeting regions of intrinsic disorder within the viral proteins or host proteins directly interacting with the RABV proteins will help creating novel drugs that can target the virus once it has infected the central nervous system.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, Figure S1: Functional disorder in human proteins interacting with the RABV P-protein. For each protein, an amino acid sequence in FASTA format is shown followed by the disorder profile generated by RIDAO, D^2P^2-generated functional disorder profile, modeled 3D structure generated by AlphaFold, and STRING-based protein-protein interaction network; Figure S2. Functional disorder in human proteins interacting with the RABV M-protein. For each protein, an amino acid sequence in FASTA format is shown followed by the disorder profile generated by RIDAO, D^2P^2-generated functional disorder profile, modeled 3D structure generated by AlphaFold, and STRING-based protein-protein interaction network; Figure S3. Functional disorder in human proteins interacting with the RABV N-protein. For each protein, an amino acid sequence in FASTA format is shown followed by the disorder profile generated by RIDAO, D^2P^2-generated functional disorder profile, modeled 3D structure generated by AlphaFold, and STRING-based protein-protein interaction network; Figure S4. Functional disorder in human proteins interacting with the RABV G-protein. For each protein, an amino acid sequence in FASTA format is shown followed by the disorder profile generated by RIDAO, D^2P^2-generated functional disorder profile, modeled 3D structure generated by AlphaFold, and STRING-based protein-protein interaction network; Figure S5. Functional disorder in human proteins interacting with the RABV L-protein. For each protein, an amino acid sequence in FASTA format is shown followed by the disorder profile generated by RIDAO, D^2P^2-generated functional disorder profile, modeled 3D structure generated by AlphaFold, and STRING-based protein-protein interaction network.

Author Contributions: Conceptualization, V.N.U.; methodology, V.N.U.; validation, S.D. and V.N.U.; formal analysis, S.D. and V.N.U.; investigation, S.D. and V.N.U.; writing—original draft preparation, S.D. and V.N.U.; writing—review and editing, S.D. and V.N.U.; visualization, S.D. and V.N.U.; supervision, V.U. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article or supplementary material.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rupprecht, C.E. Rhabdoviruses: Rabies Virus. In *Medical Microbiology*, th, Baron, S., Eds. Galveston (TX), 1996.
2. Pieracci, E.G.; Pearson, C.M.; Wallace, R.M.; Blanton, J.D.; Whitehouse, E.R.; Ma, X.; Stauffer, K.; Chipman, R.B.; Olson, V. Vital Signs: Trends in Human Rabies Deaths and Exposures - United States, 1938-2018. *MMWR Morb Mortal Wkly Rep* **2019**, *68*, 524-528, doi:10.15585/mmwr.mm6823e1.

3. Dietzschold, B.; Li, J.; Faber, M.; Schnell, M. Concepts in the pathogenesis of rabies. *Future Virol* **2008**, *3*, 481-490, doi:10.2217/17460794.3.5.481.
4. Gluska, S.; Zahavi, E.E.; Chein, M.; Gradus, T.; Bauer, A.; Finke, S.; Perlson, E. Rabies Virus Hijacks and accelerates the p75NTR retrograde axonal transport machinery. *PLoS Pathog* **2014**, *10*, e1004348, doi:10.1371/journal.ppat.1004348.
5. ViralZone. Lyssavirus. Availabe online: https://viralzone.expasy.org/resources/Rhabdoviridae_virion.jpg (accessed on 07/20/2020).
6. Okada, K.; Ito, N.; Yamaoka, S.; Masatani, T.; Ebihara, H.; Goto, H.; Nakagawa, K.; Mitake, H.; Okadera, K.; Sugiyama, M. Roles of the Rabies Virus Phosphoprotein Isoforms in Pathogenesis. *J Virol* **2016**, *90*, 8226-8237, doi:10.1128/JVI.00809-16.
7. Kouznetzoff, A.; Buckle, M.; Tordo, N. Identification of a region of the rabies virus N protein involved in direct binding to the viral RNA. *J Gen Virol* **1998**, *79 (Pt 5)*, 1005-1013, doi:10.1099/0022-1317-79-5-1005.
8. Nakagawa, K.; Kobayashi, Y.; Ito, N.; Suzuki, Y.; Okada, K.; Makino, M.; Goto, H.; Takahashi, T.; Sugiyama, M. Molecular Function Analysis of Rabies Virus RNA Polymerase L Protein by Using an L Gene-Deficient Virus. *J Virol* **2017**, *91*, doi:10.1128/JVI.00826-17.
9. Desmezieres, E.; Maillard, A.P.; Gaudin, Y.; Tordo, N.; Perrin, P. Differential stability and fusion activity of Lyssavirus glycoprotein trimers. *Virus Res* **2003**, *91*, 181-187, doi:10.1016/s0168-1702(02)00267-8.
10. Gupta, P.K.; Sharma, S.; Walunj, S.S.; Chaturvedi, V.K.; Raut, A.A.; Patial, S.; Rai, A.; Pandey, K.D.; Saini, M. Immunogenic and antigenic properties of recombinant soluble glycoprotein of rabies virus. *Vet Microbiol* **2005**, *108*, 207-214, doi:10.1016/j.vetmic.2005.04.007.
11. Chenik, M.; Schnell, M.; Conzelmann, K.K.; Blondel, D. Mapping the interacting domains between the rabies virus polymerase and phosphoprotein. *J Virol* **1998**, *72*, 1925-1930.
12. Pulmanausahakul, R.; Li, J.; Schnell, M.J.; Dietzschold, B. The glycoprotein and the matrix protein of rabies virus affect pathogenicity by regulating viral replication and facilitating cell-to-cell spread. *J Virol* **2008**, *82*, 2330-2338, doi:10.1128/JVI.02327-07.
13. Mebatsion, T.; Weiland, F.; Conzelmann, K.K. Matrix protein of rabies virus is responsible for the assembly and budding of bullet-shaped particles and interacts with the transmembrane spike glycoprotein G. *J Virol* **1999**, *73*, 242-250.
14. Chenik, M.; Chebli, K.; Blondel, D. Translation initiation at alternate in-frame AUG codons in the rabies virus phosphoprotein mRNA is mediated by a ribosomal leaky scanning mechanism. *J Virol* **1995**, *69*, 707-712.
15. Li, Y.; Dong, W.; Shi, Y.; Deng, F.; Chen, X.; Wan, C.; Zhou, M.; Zhao, L.; Fu, Z.F.; Peng, G. Rabies virus phosphoprotein interacts with ribosomal protein L9 and affects rabies virus replication. *Virology* **2016**, *488*, 216-224, doi:10.1016/j.virol.2015.11.018.
16. Wiltzer, L.; Okada, K.; Yamaoka, S.; Larrous, F.; Kuusisto, H.V.; Sugiyama, M.; Blondel, D.; Bourhy, H.; Jans, D.A.; Ito, N., et al. Interaction of rabies virus P-protein with STAT proteins is critical to lethal rabies disease. *J Infect Dis* **2014**, *209*, 1744-1753, doi:10.1093/infdis/jit829.
17. Vidy, A.; El Bougrini, J.; Chelbi-Alix, M.K.; Blondel, D. The nucleocytoplasmic rabies virus P protein counteracts interferon signaling by inhibiting both nuclear accumulation and DNA binding of STAT1. *J Virol* **2007**, *81*, 4255-4263, doi:10.1128/JVI.01930-06.
18. Moseley, G.W.; Lahaye, X.; Roth, D.M.; Oksayan, S.; Filmer, R.P.; Rowe, C.L.; Blondel, D.; Jans, D.A. Dual modes of rabies P-protein association with microtubules: a novel strategy to suppress the antiviral response. *J Cell Sci* **2009**, *122*, 3652-3662, doi:10.1242/jcs.045542.
19. Masatani, T.; Ito, N.; Ito, Y.; Nakagawa, K.; Abe, M.; Yamaoka, S.; Okadera, K.; Sugiyama, M. Importance of rabies virus nucleoprotein in viral evasion of interferon response in the brain. *Microbiol Immunol* **2013**, *57*, 511-517, doi:10.1111/1348-0421.12058.

20. Masatani, T.; Ito, N.; Shimizu, K.; Ito, Y.; Nakagawa, K.; Sawaki, Y.; Koyama, H.; Sugiyama, M. Rabies virus nucleoprotein functions to evade activation of the RIG-I-mediated antiviral response. *J Virol* **2010**, *84*, 4002-4012, doi:10.1128/JVI.02220-09.
21. Ogino, M.; Ito, N.; Sugiyama, M.; Ogino, T. The Rabies Virus L Protein Catalyzes mRNA Capping with GDP Polyribonucleotidyltransferase Activity. *Viruses* **2016**, *8*, doi:10.3390/v8050144.
22. Finke, S.; Mueller-Waldeck, R.; Conzelmann, K.K. Rabies virus matrix protein regulates the balance of virus transcription and replication. *J Gen Virol* **2003**, *84*, 1613-1621, doi:10.1099/vir.0.19128-0.
23. Ben Khalifa, Y.; Luco, S.; Besson, B.; Sonthonnax, F.; Archambaud, M.; Grimes, J.M.; Larrous, F.; Bourhy, H. The matrix protein of rabies virus binds to RelAp43 to modulate NF-kappaB-dependent gene expression related to innate immunity. *Sci Rep* **2016**, *6*, 39420, doi:10.1038/srep39420.
24. Zhang, G.; Wang, H.; Mahmood, F.; Fu, Z.F. Rabies virus glycoprotein is an important determinant for the induction of innate immune responses and the pathogenic mechanisms. *Vet Microbiol* **2013**, *162*, 601-613, doi:10.1016/j.vetmic.2012.11.031.
25. Dunker, A.K.; Lawson, J.D.; Brown, C.J.; Williams, R.M.; Romero, P.; Oh, J.S.; Oldfield, C.J.; Campen, A.M.; Ratliff, C.M.; Hipps, K.W., et al. Intrinsically disordered protein. *J Mol Graph Model* **2001**, *19*, 26-59.
26. Dunker, A.K.; Obradovic, Z.; Romero, P.; Garner, E.C.; Brown, C.J. Intrinsic protein disorder in complete genomes. *Genome Inform Ser Workshop Genome Inform* **2000**, *11*, 161-171.
27. Tompa, P. Intrinsically unstructured proteins. *Trends Biochem Sci* **2002**, *27*, 527-533.
28. Uversky, V.N. Natively unfolded proteins: a point where biology waits for physics. *Protein Sci* **2002**, *11*, 739-756.
29. Uversky, V.N. The mysterious unfoldome: structureless, underappreciated, yet vital part of any given proteome. *J Biomed Biotechnol* **2010**, *2010*, 568068, doi:10.1155/2010/568068.
30. Uversky, V.N.; Dunker, A.K. Understanding protein non-folding. *Biochim Biophys Acta* **2010**, *1804*, 1231-1264, doi:10.1016/j.bbapap.2010.01.017.
31. Uversky, V.N.; Gillespie, J.R.; Fink, A.L. Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* **2000**, *41*, 415-427, doi:10.1002/1097-0134(20001115)41:3<415::AID-PROT130>3.0.CO;2-7 [pii].
32. Dunker, A.K.; Cortese, M.S.; Romero, P.; Iakoucheva, L.M.; Uversky, V.N. Flexible nets. The roles of intrinsic disorder in protein interaction networks. *The FEBS journal* **2005**, *272*, 5129-5148, doi:10.1111/j.1742-4658.2005.04948.x.
33. Ward, J.J.; Sodhi, J.S.; McGuffin, L.J.; Buxton, B.F.; Jones, D.T. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J Mol Biol* **2004**, *337*, 635-645.
34. Dunker, A.K.; Obradovic, Z.; Romero, P.; Garner, E.C.; Brown, C.J. Intrinsic protein disorder in complete genomes. *Genome Inform. Ser. Workshop Genome Inform.* **2000**, *11*, 161-171.
35. Ward, J.J.; Sodhi, J.S.; McGuffin, L.J.; Buxton, B.F.; Jones, D.T. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* **2004**, *337*, 635-645.
36. Xue, B.; Dunker, A.K.; Uversky, V.N. Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life. *J. Biomol. Struct. Dyn.* **2012**, *30*, 137-149, doi:10.1080/07391102.2012.675145.
37. Peng, Z.; Yan, J.; Fan, X.; Mizianty, M.J.; Xue, B.; Wang, K.; Hu, G.; Uversky, V.N.; Kurgan, L. Exceptionally abundant exceptions: comprehensive characterization of intrinsic disorder in all domains of life. *Cell Mol Life Sci* **2015**, *72*, 137-151, doi:10.1007/s00018-014-1661-9.
38. Xue, B.; Dunker, A.K.; Uversky, V.N. Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life. *J Biomol Struct Dyn* **2012**, *30*, 137-149, doi:10.1080/07391102.2012.675145.
39. Iakoucheva, L.M.; Brown, C.J.; Lawson, J.D.; Obradovic, Z.; Dunker, A.K. Intrinsic disorder in cell-signaling and cancer-associated proteins. *J Mol Biol* **2002**, *323*, 573-584.
40. Uversky, V.N. Unusual biophysics of intrinsically disordered proteins. *Biochim Biophys Acta* **2013**, *1834*, 932-951, doi:10.1016/j.bbapap.2012.12.008.

41. Dyson, H.J.; Wright, P.E. Intrinsically unstructured proteins and their functions. *Nature reviews. Molecular cell biology* **2005**, *6*, 197-208, doi:10.1038/nrm1589.
42. Wright, P.E.; Dyson, H.J. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J Mol Biol* **1999**, *293*, 321-331.
43. Tompa, P. The interplay between structure and function in intrinsically unstructured proteins. *FEBS letters* **2005**, *579*, 3346-3354, doi:10.1016/j.febslet.2005.03.072.
44. Radivojac, P.; Iakoucheva, L.M.; Oldfield, C.J.; Obradovic, Z.; Uversky, V.N.; Dunker, A.K. Intrinsic disorder and functional proteomics. *Biophys J* **2007**, *92*, 1439-1456, doi:10.1529/biophysj.106.094045.
45. Vucetic, S.; Xie, H.; Iakoucheva, L.M.; Oldfield, C.J.; Dunker, A.K.; Obradovic, Z.; Uversky, V.N. Functional anthology of intrinsic disorder. 2. Cellular components, domains, technical terms, developmental processes, and coding sequence diversities correlated with long disordered regions. *Journal of proteome research* **2007**, *6*, 1899-1916, doi:10.1021/pr060393m.
46. Xie, H.; Vucetic, S.; Iakoucheva, L.M.; Oldfield, C.J.; Dunker, A.K.; Obradovic, Z.; Uversky, V.N. Functional anthology of intrinsic disorder. 3. Ligands, post-translational modifications, and diseases associated with intrinsically disordered proteins. *Journal of proteome research* **2007**, *6*, 1917-1932, doi:10.1021/pr060394e.
47. Xie, H.; Vucetic, S.; Iakoucheva, L.M.; Oldfield, C.J.; Dunker, A.K.; Uversky, V.N.; Obradovic, Z. Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. *Journal of proteome research* **2007**, *6*, 1882-1898, doi:10.1021/pr060392u.
48. Uversky, V.N.; Oldfield, C.J.; Dunker, A.K. Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annu Rev Biophys* **2008**, *37*, 215-246, doi:10.1146/annurev.biophys.37.032807.125924.
49. Vacic, V.; Markwick, P.R.; Oldfield, C.J.; Zhao, X.; Haynes, C.; Uversky, V.N.; Iakoucheva, L.M. Disease-associated mutations disrupt functionally important regions of intrinsic protein disorder. *PLoS computational biology* **2012**, *8*, e1002709, doi:10.1371/journal.pcbi.1002709.
50. Dunker, A.K.; Garner, E.; Guilliot, S.; Romero, P.; Albrecht, K.; Hart, J.; Obradovic, Z.; Kissinger, C.; Villafranca, J.E. Protein disorder and the evolution of molecular recognition: theory, predictions and observations. *Pac Symp Biocomput* **1998**, 473-484.
51. Daughdrill, G.W.; Pielak, G.J.; Uversky, V.N.; Cortese, M.S.; Dunker, A.K. Natively disordered proteins. In *Handbook of Protein Folding*, Buchner, J., Kiefhaber, T., Eds. Wiley-VCH, Verlag GmbH & Co. KGaA: Weinheim, Germany, 2005; pp. 271-353.
52. Uversky, V.N. Intrinsic disorder-based protein interactions and their modulators. *Curr Pharm Des* **2013**, *19*, 4191-4213.
53. Uversky, V.N. Functional roles of transiently and intrinsically disordered regions within proteins. *The FEBS journal* **2015**, *282*, 1182-1189, doi:10.1111/febs.13202.
54. Uversky, V.N. p53 Proteoforms and Intrinsic Disorder: An Illustration of the Protein Structure-Function Continuum Concept. *Int J Mol Sci* **2016**, *17*, 1874, doi:10.3390/ijms17111874.
55. Smith, L.M.; Kelleher, N.L.; Consortium for Top Down, P. Proteoform: a single term describing protein complexity. *Nat Methods* **2013**, *10*, 186-187, doi:10.1038/nmeth.2369.
56. Xue, B.; Ganti, K.; Rabionet, A.; Banks, L.; Uversky, V.N. Disordered interactome of human papillomavirus. *Curr Pharm Des* **2014**, *20*, 1274-1292, doi:10.2174/13816128113199990072.
57. Uversky, V.N.; Roman, A.; Oldfield, C.J.; Dunker, A.K. Protein intrinsic disorder and human papillomaviruses: increased amount of disorder in E6 and E7 oncoproteins from high risk HPVs. *Journal of proteome research* **2006**, *5*, 1829-1842, doi:10.1021/pr0602388.
58. Xue, B.; Mizianty, M.J.; Kurgan, L.; Uversky, V.N. Protein intrinsic disorder as a flexible armor and a weapon of HIV-1. *Cell Mol Life Sci* **2012**, *69*, 1211-1259, doi:10.1007/s00018-011-0859-3.
59. Dolan, P.T.; Roth, A.P.; Xue, B.; Sun, R.; Dunker, A.K.; Uversky, V.N.; LaCount, D.J. Intrinsic disorder mediates hepatitis C virus core-host cell protein interactions. *Protein Sci* **2015**, *24*, 221-235, doi:10.1002/pro.2608.

60. Fan, X.; Xue, B.; Dolan, P.T.; LaCount, D.J.; Kurgan, L.; Uversky, V.N. The intrinsic disorder status of the human hepatitis C virus proteome. *Mol Biosyst* **2014**, *10*, 1345-1363, doi:10.1039/c4mb00027g.
61. Meng, F.; Badierah, R.A.; Almehdar, H.A.; Redwan, E.M.; Kurgan, L.; Uversky, V.N. Unstructural biology of the Dengue virus proteins. *The FEBS journal* **2015**, *282*, 3368-3394, doi:10.1111/febs.13349.
62. Kumar, D.; Singh, A.; Kumar, P.; Uversky, V.N.; Rao, C.D.; Giri, R. Understanding the penetrance of intrinsic protein disorder in rotavirus proteome. *Int J Biol Macromol* **2020**, *144*, 892-908, doi:10.1016/j.ijbiomac.2019.09.166.
63. Whelan, J.N.; Reddy, K.D.; Uversky, V.N.; Teng, M.N. Functional correlations of respiratory syncytial virus proteins to intrinsic disorder. *Mol Biosyst* **2016**, *12*, 1507-1526, doi:10.1039/c6mb00122j.
64. Mishra, P.M.; Uversky, V.N.; Giri, R. Molecular Recognition Features in Zika Virus Proteome. *J Mol Biol* **2018**, *430*, 2372-2388, doi:10.1016/j.jmb.2017.10.018.
65. Giri, R.; Kumar, D.; Sharma, N.; Uversky, V.N. Intrinsically Disordered Side of the Zika Virus Proteome. *Front Cell Infect Microbiol* **2016**, *6*, 144, doi:10.3389/fcimb.2016.00144.
66. Singh, A.; Kumar, A.; Yadav, R.; Uversky, V.N.; Giri, R. Deciphering the dark proteome of Chikungunya virus. *Sci Rep* **2018**, *8*, 5822, doi:10.1038/s41598-018-23969-0.
67. Redwan, E.M.; AlJaddawi, A.A.; Uversky, V.N. Structural disorder in the proteome and interactome of Alkhurma virus (ALKV). *Cell Mol Life Sci* **2019**, *76*, 577-608, doi:10.1007/s00018-018-2968-8.
68. Bhardwaj, T.; Saumya, K.U.; Kumar, P.; Sharma, N.; Gadhave, K.; Uversky, V.N.; Giri, R. Japanese encephalitis virus - exploring the dark proteome and disorder-function paradigm. *The FEBS journal* **2020**, 10.1111/febs.15427, doi:10.1111/febs.15427.
69. Giri, R.; Bhardwaj, T.; Shegane, M.; Gehi, B.R.; Kumar, P.; Gadhave, K.; Oldfield, C.J.; Uversky, V.N. Understanding the COVID-19 via Comparative Analysis of Dark Proteomes of SARS-CoV-2, Human SARS and Bat SARS-Like Coronaviruses. *Cell Mol Life Sci* **2020**, *In press*.
70. UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* **2019**, *47*, D506-D515, doi:10.1093/nar/gky1049.
71. Zandi, F.; Goshadrou, F.; Meyfour, A.; Vaziri, B. Rabies Infection: An Overview of Lyssavirus-Host Protein Interactions. *Iran Biomed J* **2021**, *25*, 226-242, doi:10.52547/ibj.25.4.226.
72. Chen, J.; Kriwacki, R.W. Intrinsically Disordered Proteins: Structure, Function and Therapeutics. *J Mol Biol* **2018**, *430*, 2275-2277, doi:10.1016/j.jmb.2018.06.012.
73. Xue, B.; Dunbrack, R.L.; Williams, R.W.; Dunker, A.K.; Uversky, V.N. PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. *Biochim Biophys Acta* **2010**, *1804*, 996-1010, doi:10.1016/j.bbapap.2010.01.011.
74. Dosztanyi, Z.; Csizmok, V.; Tompa, P.; Simon, I. The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. *J Mol Biol* **2005**, *347*, 827-839, doi:10.1016/j.jmb.2005.01.071.
75. Meszaros, B.; Erdos, G.; Dosztanyi, Z. IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res* **2018**, *46*, W329-W337, doi:10.1093/nar/gky384.
76. Meszaros, B.; Simon, I.; Dosztanyi, Z. Prediction of protein binding regions in disordered proteins. *PLoS computational biology* **2009**, *5*, e1000376, doi:10.1371/journal.pcbi.1000376.
77. Dosztanyi, Z.; Meszaros, B.; Simon, I. ANCHOR: web server for predicting protein binding regions in disordered proteins. *Bioinformatics* **2009**, *25*, 2745-2746, 10.1093/bioinformatics/btp518.
78. Dayhoff, G.W.I.; Uversky, V.N. Rapid prediction and analysis of protein intrinsic disorder. *Protein Science* **2022**, In press.
79. Romero, P.; Obradovic, Z.; Li, X.; Garner, E.C.; Brown, C.J.; Dunker, A.K. Sequence complexity of disordered protein. *Proteins* **2001**, *42*, 38-48.
80. Peng, K.; Radivojac, P.; Vucetic, S.; Dunker, A.K.; Obradovic, Z. Length-dependent prediction of protein intrinsic disorder. *BMC bioinformatics* **2006**, *7*, 208, doi:10.1186/1471-2105-7-208.

81. Peng, K.; Vucetic, S.; Radivojac, P.; Brown, C.J.; Dunker, A.K.; Obradovic, Z. Optimizing long intrinsic disorder predictors with protein evolutionary information. *Journal of bioinformatics and computational biology* **2005**, *3*, 35-60.
82. Dosztanyi, Z.; Csizmok, V.; Tompa, P.; Simon, I. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* **2005**, *21*, 3433-3434, doi:10.1093/bioinformatics/bti541.
83. Mohan, A.; Sullivan, W.J., Jr.; Radivojac, P.; Dunker, A.K.; Uversky, V.N. Intrinsic disorder in pathogenic and non-pathogenic microbes: discovering and analyzing the unfoldomes of early-branching eukaryotes. *Mol Biosyst* **2008**, *4*, 328-340, doi:10.1039/b719168e.
84. Sun, X.; Xue, B.; Jones, W.T.; Rikkerink, E.; Dunker, A.K.; Uversky, V.N. A functionally required unfoldome from the plant kingdom: intrinsically disordered N-terminal domains of GRAS proteins are involved in molecular recognition during plant development. *Plant Mol Biol* **2011**, *77*, 205-223, doi:10.1007/s11103-011-9803-z.
85. Xue, B.; Oldfield, C.J.; Van, Y.Y.; Dunker, A.K.; Uversky, V.N. Protein intrinsic disorder and induced pluripotent stem cells. *Mol Biosyst* **2012**, *8*, 134-150, doi:10.1039/c1mb05163f.
86. Huang, F.; Oldfield, C.; Meng, J.; Hsu, W.L.; Xue, B.; Uversky, V.N.; Romero, P.; Dunker, A.K. Subclassifying disordered proteins by the CH-CDF plot method. *Pac Symp Biocomput* **2012**, 128-139.
87. Rajagopalan, K.; Mooney, S.M.; Parekh, N.; Getzenberg, R.H.; Kulkarni, P. A majority of the cancer/testis antigens are intrinsically disordered proteins. *J Cell Biochem* **2011**, *112*, 3256-3267, doi:10.1002/jcb.23252.
88. Oates, M.E.; Romero, P.; Ishida, T.; Ghalwash, M.; Mizianty, M.J.; Xue, B.; Dosztanyi, Z.; Uversky, V.N.; Obradovic, Z.; Kurgan, L., et al. D(2)P(2): database of disordered protein predictions. *Nucleic Acids Res* **2013**, *41*, D508-516, doi:10.1093/nar/gks1226.
89. Ishida, T.; Kinoshita, K. PrDOS: prediction of disordered protein regions from amino acid sequence. *Nucleic Acids Res* **2007**, *35*, W460-464, doi:10.1093/nar/gkm363.
90. Walsh, I.; Martin, A.J.; Di Domenico, T.; Tosatto, S.C. ESpritz: accurate and fast prediction of protein disorder. *Bioinformatics* **2012**, *28*, 503-509, doi:10.1093/bioinformatics/btr682.
91. Andreeva, A.; Howorth, D.; Brenner, S.E.; Hubbard, T.J.; Chothia, C.; Murzin, A.G. SCOP database in 2004: refinements integrate structure and sequence family data. *Nucleic Acids Res* **2004**, *32*, D226-229, doi:10.1093/nar/gkh039.
92. Murzin, A.G.; Brenner, S.E.; Hubbard, T.; Chothia, C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* **1995**, *247*, 536-540, doi:10.1006/jmbi.1995.0159.
93. de Lima Morais, D.A.; Fang, H.; Rackham, O.J.; Wilson, D.; Pethica, R.; Chothia, C.; Gough, J. SUPERFAMILY 1.75 including a domain-centric gene ontology method. *Nucleic Acids Res* **2011**, *39*, D427-434, doi:10.1093/nar/gkq1130.
94. Hornbeck, P.V.; Kornhauser, J.M.; Tkachev, S.; Zhang, B.; Skrzypek, E.; Murray, B.; Latham, V.; Sullivan, M. PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res* **2012**, *40*, D261-270, doi:10.1093/nar/gkr1122.
95. Szklarczyk, D.; Franceschini, A.; Kuhn, M.; Simonovic, M.; Roth, A.; Minguez, P.; Doerks, T.; Stark, M.; Muller, J.; Bork, P., et al. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res* **2011**, *39*, D561-568, doi:10.1093/nar/gkq973.
96. Hardenberg, M.; Horvath, A.; Ambrus, V.; Fuxreiter, M.; Vendruscolo, M. Widespread occurrence of the droplet state of proteins in the human proteome. *Proceedings of the National Academy of Sciences* **2020**, *117*, 33254-33262, doi:10.1073/pnas.2007670117.
97. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; de Beer, T.A.P.; Rempfer, C.; Bordoli, L., et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **2018**, *46*, W296-W303, doi:10.1093/nar/gky427.

98. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A., et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583-589, doi:10.1038/s41586-021-03819-2.
99. Nevers, Q.; Albertini, A.A.; Lagaudriere-Gesbert, C.; Gaudin, Y. Negri bodies and other virus membrane-less replication compartments. *Biochim Biophys Acta Mol Cell Res* **2020**, *1867*, 118831, doi:10.1016/j.bbamcr.2020.118831.
100. Brzozka, K.; Finke, S.; Conzelmann, K.K. Identification of the rabies virus alpha/beta interferon antagonist: phosphoprotein P interferes with phosphorylation of interferon regulatory factor 3. *J Virol* **2005**, *79*, 7673-7681, doi:10.1128/JVI.79.12.7673-7681.2005.
101. Vidy, A.; Chelbi-Alix, M.; Blondel, D. Rabies virus P protein interacts with STAT1 and inhibits interferon signal transduction pathways. *J Virol* **2005**, *79*, 14411-14420, doi:10.1128/JVI.79.22.14411-14420.2005.
102. Lieu, K.G.; Brice, A.; Wiltzer, L.; Hirst, B.; Jans, D.A.; Blondel, D.; Moseley, G.W. The rabies virus interferon antagonist P protein interacts with activated STAT3 and inhibits Gp130 receptor signaling. *J Virol* **2013**, *87*, 8261-8265, doi:10.1128/JVI.00989-13.
103. Niu, X.; Tang, L.; Tseggai, T.; Guo, Y.; Fu, Z.F. Wild-type rabies virus phosphoprotein is associated with viral sensitivity to type I interferon treatment. *Arch Virol* **2013**, *158*, 2297-2305, doi:10.1007/s00705-013-1743-2.
104. Brice, A.; Whelan, D.R.; Ito, N.; Shimizu, K.; Wiltzer-Bach, L.; Lo, C.Y.; Blondel, D.; Jans, D.A.; Bell, T.D.; Moseley, G.W. Quantitative Analysis of the Microtubule Interaction of Rabies Virus P3 Protein: Roles in Immune Evasion and Pathogenesis. *Sci Rep* **2016**, *6*, 33493, doi:10.1038/srep33493.
105. Rowe, C.L.; Wagstaff, K.M.; Oksayan, S.; Glover, D.J.; Jans, D.A.; Moseley, G.W. Nuclear Trafficking of the Rabies Virus Interferon Antagonist P-Protein Is Regulated by an Importin-Binding Nuclear Localization Sequence in the C-Terminal Domain. *PLoS One* **2016**, *11*, e0150477, doi:10.1371/journal.pone.0150477.
106. Pasdeloup, D.; Poisson, N.; Raux, H.; Gaudin, Y.; Ruigrok, R.W.; Blondel, D. Nucleocytoplasmic shuttling of the rabies virus P protein requires a nuclear localization signal and a CRM1-dependent nuclear export signal. *Virology* **2005**, *334*, 284-293, doi:10.1016/j.virol.2005.02.005.
107. Gupta, A.K.; Blondel, D.; Choudhary, S.; Banerjee, A.K. The phosphoprotein of rabies virus is phosphorylated by a unique cellular protein kinase and specific isomers of protein kinase C. *J Virol* **2000**, *74*, 91-98, doi:10.1128/jvi.74.1.91-98.2000.
108. Zhan, J.; Watts, E.; Brice, A.M.; Metcalfe, R.D.; Rozario, A.M.; Sethi, A.; Yan, F.; Bell, T.D.M.; Griffin, M.D.W.; Moseley, G.W., et al. Molecular Basis of Functional Effects of Phosphorylation of the C-Terminal Domain of the Rabies Virus P Protein. *J Virol* **2022**, *96*, e0011122, doi:10.1128/jvi.00111-22.
109. Blondel, D.; Regad, T.; Poisson, N.; Pavie, B.; Harper, F.; Pandolfi, P.P.; De The, H.; Chelbi-Alix, M.K. Rabies virus P and small P products interact directly with PML and reorganize PML nuclear bodies. *Oncogene* **2002**, *21*, 7957-7970, doi:10.1038/sj.onc.1205931.
110. Horvath, A.; Miskei, M.; Ambrus, V.; Vendruscolo, M.; Fuxreiter, M. Sequence-based prediction of protein binding mode landscapes. *PLoS computational biology* **2020**, *16*, e1007864, doi:10.1371/journal.pcbi.1007864.
111. Goh, G.K.; Dunker, A.K.; Foster, J.A.; Uversky, V.N. Shell disorder analysis predicts greater resilience of the SARS-CoV-2 (COVID-19) outside the body and in body fluids. *Microb Pathog* **2020**, *144*, 104177, doi:10.1016/j.micpath.2020.104177.
112. Goh, G.K.; Dunker, A.K.; Foster, J.A.; Uversky, V.N. Rigidity of the Outer Shell Predicted by a Protein Intrinsic Disorder Model Sheds Light on the COVID-19 (Wuhan-2019-nCoV) Infectivity. *Biomolecules* **2020**, *10*, doi:10.3390/biom10020331.
113. Goh, G.K.; Dunker, A.K.; Foster, J.A.; Uversky, V.N. Nipah shell disorder, modes of infection, and virulence. *Microb Pathog* **2020**, *141*, 103976, doi:10.1016/j.micpath.2020.103976.
114. Goh, G.K.; Dunker, A.K.; Foster, J.A.; Uversky, V.N. Zika and Flavivirus Shell Disorder: Virulence and Fetal Morbidity. *Biomolecules* **2019**, *9*, doi:10.3390/biom9110710.

115. Goh, G.K.; Dunker, A.K.; Foster, J.A.; Uversky, V.N. HIV Vaccine Mystery and Viral Shell Disorder. *Biomolecules* **2019**, *9*, doi:10.3390/biom9050178.
116. Goh, G.K.; Dunker, A.K.; Uversky, V.N. Correlating Flavivirus virulence and levels of intrinsic disorder in shell proteins: protective roles vs. immune evasion. *Mol Biosyst* **2016**, *12*, 1881-1891, doi:10.1039/c6mb00228e.
117. Goh, G.K.; Dunker, A.K.; Uversky, V.N. Shell disorder, immune evasion and transmission behaviors among human and animal retroviruses. *Mol Biosyst* **2015**, *11*, 2312-2323, doi:10.1039/c5mb00277j.
118. Goh, G.K.; Dunker, A.K.; Uversky, V. Prediction of Intrinsic Disorder in MERS-CoV/HCoV-EMC Supports a High Oral-Fecal Transmission. *PLoS Curr* **2013**, *5*, doi:10.1371/currents.outbreaks.22254b58675cdebc256dbe3c5aa6498b.
119. Goh, G.K.; Dunker, A.K.; Uversky, V.N. Understanding Viral Transmission Behavior via Protein Intrinsic Disorder Prediction: Coronaviruses. *J Pathog* **2012**, *2012*, 738590, doi:10.1155/2012/738590.
120. Graham, S.C.; Assenberg, R.; Delmas, O.; Verma, A.; Gholami, A.; Talbi, C.; Owens, R.J.; Stuart, D.I.; Grimes, J.M.; Bourhy, H. Rhabdovirus matrix protein structures reveal a novel mode of self-association. *PLoS Pathog* **2008**, *4*, e1000251, doi:10.1371/journal.ppat.1000251.
121. Montespan, C.; Wiethoff, C.M.; Wodrich, H. A Small Viral PPxY Peptide Motif To Control Antiviral Autophagy. *J Virol* **2017**, *91*, doi:10.1128/JVI.00581-17.
122. Albertini, A.A.; Wernimont, A.K.; Muziol, T.; Ravelli, R.B.; Clapier, C.R.; Schoehn, G.; Weissenhorn, W.; Ruigrok, R.W. Crystal structure of the rabies virus nucleoprotein-RNA complex. *Science* **2006**, *313*, 360-363, doi:10.1126/science.1125280.
123. Luo, M.; Green, T.J.; Zhang, X.; Tsao, J.; Qiu, S. Conserved characteristics of the rhabdovirus nucleoprotein. *Virus Res* **2007**, *129*, 246-251, doi:10.1016/j.virusres.2007.07.011.
124. Etessami, R.; Conzelmann, K.K.; Fadai-Ghotbi, B.; Natelson, B.; Tsiang, H.; Ceccaldi, P.E. Spread and pathogenic characteristics of a G-deficient rabies virus recombinant: an in vitro and in vivo study. *J Gen Virol* **2000**, *81*, 2147-2153, doi:10.1099/0022-1317-81-9-2147.
125. Nitschel, S.; Zaeck, L.M.; Potratz, M.; Nolden, T.; Te Kamp, V.; Franzke, K.; Hoper, D.; Pfaff, F.; Finke, S. Point Mutations in the Glycoprotein Ectodomain of Field Rabies Viruses Mediate Cell Culture Adaptation through Improved Virus Release in a Host Cell Dependent and Independent Manner. *Viruses* **2021**, *13*, doi:10.3390/v13101989.
126. Yang, F.; Lin, S.; Ye, F.; Yang, J.; Qi, J.; Chen, Z.; Lin, X.; Wang, J.; Yue, D.; Cheng, Y., et al. Structural Analysis of Rabies Virus Glycoprotein Reveals pH-Dependent Conformational Changes and Interactions with a Neutralizing Antibody. *Cell Host Microbe* **2020**, *27*, 441-453 e447, doi:10.1016/j.chom.2019.12.012.
127. Schnell, M.J.; Conzelmann, K.K. Polymerase activity of in vitro mutated rabies virus L protein. *Virology* **1995**, *214*, 522-530, doi:10.1006/viro.1995.0063.
128. Horwitz, J.A.; Jenni, S.; Harrison, S.C.; Whelan, S.P.J. Structure of a rabies virus polymerase complex from electron cryo-microscopy. *Proc Natl Acad Sci U S A* **2020**, *117*, 2099-2107, doi:10.1073/pnas.1918809117.
129. Liu, X.; Nawaz, Z.; Guo, C.; Ali, S.; Naeem, M.A.; Jamil, T.; Ahmad, W.; Siddiq, M.U.; Ahmed, S.; Asif Idrees, M., et al. Rabies Virus Exploits Cytoskeleton Network to Cause Early Disease Progression and Cellular Dysfunction. *Front Vet Sci* **2022**, *9*, 889873, doi:10.3389/fvets.2022.889873.
130. Kammouni, W.; Wood, H.; Saleh, A.; Appolinario, C.M.; Fernyhough, P.; Jackson, A.C. Rabies virus phosphoprotein interacts with mitochondrial Complex I and induces mitochondrial dysfunction and oxidative stress. *J Neurovirol* **2015**, *21*, 370-382, doi:10.1007/s13365-015-0320-8.
131. Sharma, L.K.; Lu, J.; Bai, Y. Mitochondrial respiratory complex I: structure, function and implication in human diseases. *Curr Med Chem* **2009**, *16*, 1266-1277, doi:10.2174/092986709787846578.
132. Wang, J.; Wang, Z.; Liu, R.; Shuai, L.; Wang, X.; Luo, J.; Wang, C.; Chen, W.; Wang, X.; Ge, J., et al. Metabotropic glutamate receptor subtype 2 is a cellular receptor for rabies virus. *PLoS Pathog* **2018**, *14*, e1007189, doi:10.1371/journal.ppat.1007189.

133. Tuffereau, C.; Benejean, J.; Blondel, D.; Kieffer, B.; Flamand, A. Low-affinity nerve-growth factor receptor (P75NTR) can serve as a receptor for rabies virus. *EMBO J* **1998**, *17*, 7250-7259, doi:10.1093/emboj/17.24.7250.
134. Thoulouze, M.I.; Lafage, M.; Schachner, M.; Hartmann, U.; Cremer, H.; Lafon, M. The neural cell adhesion molecule is a receptor for rabies virus. *J Virol* **1998**, *72*, 7181-7190, doi:10.1128/JVI.72.9.7181-7190.1998.
135. Lentz, T.L.; Burrage, T.G.; Smith, A.L.; Crick, J.; Tignor, G.H. Is the acetylcholine receptor a rabies virus receptor? *Science* **1982**, *215*, 182-184, doi:10.1126/science.7053569.
136. Sajjanar, B.; Dhusia, K.; Saxena, S.; Joshi, V.; Bisht, D.; Thakuria, D.; Manjunathareddy, G.B.; Ramteke, P.W.; Kumar, S. Nicotinic acetylcholine receptor alpha 1(nAChRalpha1) subunit peptides as potential antiviral agents against rabies virus. *Int J Biol Macromol* **2017**, *104*, 180-188, doi:10.1016/j.ijbiomac.2017.05.179.
137. Embregts, C.W.E.; Begeman, L.; Voesenek, C.J.; Martina, B.E.E.; Koopmans, M.P.G.; Kuiken, T.; GeurtsvanKessel, C.H. Street RABV Induces the Cholinergic Anti-inflammatory Pathway in Human Monocyte-Derived Macrophages by Binding to nAChr alpha7. *Front Immunol* **2021**, *12*, 622516, doi:10.3389/fimmu.2021.622516.