Journal of proteome research

pubs.acs.org/jpr

Article

Ebola Virus VP35 Protein: Modeling of the Tetrameric Structure and an Analysis of Its Interaction with Human PKR

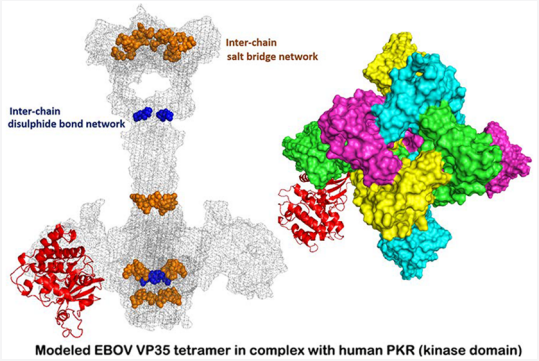
Anupam Banerjee and Pralay Mitra*

Cite This: *J. Proteome Res.* 2020, 19, 4533−4542

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The Viral Protein 35 (VP35), a crucial protein of the Zaire Ebolavirus (EBOV), interacts with a plethora of human proteins to cripple the human immune system. Despite its importance, the entire structure of the tetrameric assembly of EBOV VP35 and the means by which it antagonizes the autophosphorylation of the kinase domain of human protein kinase R (PKR^K) is still elusive. We consult existing structural information to model a tetrameric assembly of the VP35 protein where 93% of the protein is modeled using crystal structure templates. We analyze our modeled tetrameric structure to identify interchain bonding networks and use molecular dynamics simulations and normal-mode analysis to unravel the flexibility and deformability of the different regions of the VP35 protein. We establish that the C-terminal of VP35 ($VP35^C$) directly interacts with PKR^K to prevent it from autophosphorylation. Further, we identify three plausible $VP35^C$−PKR^K complexes with better affinity than the PKR^K dimer formed during autophosphorylation and use protein design to establish a new stretch in $VP35^C$ that interacts with PKR^K. The proposed tetrameric assembly will aid in better understanding of the VP35 protein, and the reported $VP35^C$−PKR^K complexes along with their interacting sites will help in the shortlisting of small molecule inhibitors.



Modeled EBOV VP35 tetramer in complex with human PKR (kinase domain)

KEYWORDS: *Zaire Ebolavirus, VP35 protein, tetrameric assembly, human PKR, autophosphorylation, protein design*

INTRODUCTION

The World Health Organization (WHO) Ebola situation report declares[1] the Zaire Ebolavirus (hereafter referred to as EBOV) belonging to the Ebolavirus genera of the Filoviridae family[2] as one of the most deadly and potent health threats to the global community even in the current scenario.[3] The EBOV is a filamentous encased negative-sense single-stranded RNA virus whose ∼19 kb genome encodes three nonstructural and seven structural proteins.[4,5] Out of the seven structural proteins, the viral proteins (VPs) are mainly responsible for antagonizing the Interferon (IFN) proteins (and the signaling pathways involving the IFNs and IFN-stimulated genes, ISGs), thereby crippling the immune responses in the human body. Among the VPs, the EBOV-VP35 (hereafter referred to as VP35) antagonize the maximum number of human proteins involved in different IFN and ISG signaling pathways.[6] The VP35 protein, along with other structural proteins (NP and L in the presence of VP30), acts as a crucial component of the viral replication complex. VP35 interacts with a host of human proteins to specifically antagonize the RIG-I (Retinoic acid-Inducible Gene-I), TRIF (Toll/interleukin-1 Receptor domain-containing adapter Inducing IFN-β), and IRF7 (Interferon Response Factor 7) pathways. Particularly, the RNA binding domain (also known as Interferon Inhibitory Domain, IID) of VP35 is responsible for the inhibition of signaling involving the Protein Activator of interferon-induced protein kinase (PACT) and Protein Kinase R (PKR). Considering its significance, the structural assembly of the entire VP35 protein will considerably help in unraveling the arrangement, bonding landscape, and dynamics of and between the different subunits of the viral protein.

Exploring VP35, we find that the crystal structure of the IID domain and a small stretch at the N-terminal (residues 21−57) are available in the Protein Data Bank (PDB) with PDB ID: 4ZTA (Chain A) and 5BPV (Chain A), respectively. Existing in vitro experiments confirm that the VP35 protein exists in a tetrameric state.[7,8] In 2016, Edwards et al. established the tetrameric structure for residues 80−340 of the protein with the help of size exclusion chromatography coupled with multiangle light scattering experiments.[7] Separately, in 2017 Bruhn et al. reported that the entire VP35 protein exists as a tetramer.[8] Recently in 2019, Zinzula et al. determined the trimeric crystal structure of the coiled-coil oligomerization

Special Issue: Proteomics in Pandemic Disease

Received: June 27, 2020
Published: September 1, 2020

domain (residues 83−145) and reported that the oligomerization domain of the VP35 protein can exist both in a trimeric and tetrameric state.[9] However, as the trimeric state is only reported for the oligomerization domain, we rely on the multiple experiments in favor of the tetrameric state for larger stretches of the VP35 protein. Nevertheless, experimentally it is challenging to determine the entire structure of the tetrameric VP35 (with 340 residues in each chain) due to the presence of a sizable coiled-coil oligomerization domain and lack of regular secondary structures (loop region) in large stretches of the protein. For brevity, now onward, we denote VP35 consisting of three disjoint regions viz., $VP35^{N}$ (N-terminal region: residue 1−82), $VP35^{O}$ (extended oligomerization region: residues 83−185), and $VP35^{C}$ (C-terminal region: residues 186−340) or a combination of any two of them viz., $VP35^{N-O}$, $VP35^{O-C}$, and $VP35^{N-C}$. Additionally, we indicate $VP35^{X}_{y-z}$ as a stretch that varies from residue y to residue z in the X region, where X can be N, O, C, N−O, O−C, or N−C.

We rely on the available experimentally determined crystal structures of major parts of VP35, and consult the predicted secondary structures for the remaining intermittent parts and combine them to construct the tertiary structure of the monomeric (and subsequently tetrameric) unit of the VP35 protein. The predicted secondary structures inform us regarding the residues that belong to the loop region in the remaining intermittent parts and guides the computational loop modeling in those regions. Zinzula et al. (in 2019) used secondary structure prediction to demarcate the boundary of the oligomerization domain of the VP35 protein before opting for in vitro experimentation.[9] Similarly, in our endeavor we consider the secondary structure information to guide and ratify the tertiary structure prediction of the residues whose structure is unavailable to date.

The modeled tetramer adheres to all structural constraints and provides us novel insights into the plausible arrangement of the four subunits of the VP35 protein. We carry out Molecular Dynamics (MD) simulation as well as Normal Mode Analysis (NMA) of the predicted structure to establish, analyze, and quantify the structural uniqueness of the viral protein. Along with interchain hydrogen bonds, several symmetric interchain salt bridge and disulfide bond networks in both the N-terminal domain and the oligomerization domain ensure the structural integrity of the tetramer. We note that the adjacent protein chains interact with each other in the N-terminal and oligomerization domain of the protein. But, the presence of a highly flexible $VP35^{O-C}_{168-221}$ joining the oligomerization domain to the C-terminal ensure that the DNA binding/IID domain remains available to interact with human proteins.

Next, we unravel the mechanism by which VP35 impedes the autophosphorylation of the human PKR protein. The PKR protein on binding with a double-stranded RNA (dsRNA) or with the PACT protein gets activated, and post autophosphorylation phosphorylates the eukaryotic translation initiation factor eIF2α. Phosphorylation of the eIF2α inhibits the translation of both viral as well as cellular mRNAs.[10] It is established that VP35 inhibits the autophosphorylation of PKR and consequently inhibits the phosphorylation of eIF2α.[11] However, the exact mechanism by which VP35 inhibits the autophosphorylation of PKR is not known to date. The fact that Lys309Ala mutation results in the inability of VP35 to bind with dsRNA but does not hinder its ability to antagonize PKR indicates that RNA sequestration is not responsible for inhibiting autophosphorylation of the kinase.

We hypothesize that the C-terminal of VP35 competitively binds with the kinase domain of the PKR protein and inhibits it from autophosphorylation. Toward this endeavor, we identify three VP35−PKR complexes with the help of protein−protein docking and establish that their interaction affinity is higher in comparison to a dimer of the PKR kinase domain necessary for autophosphorylation. Additionally, we validate the shortlisted complexes against the experimentally established Alanine replacements of Lys309 and Arg312 in VP35 that prevent it from antagonizing PKR autophosphorylation.[11] The recent success of protein design in identifying critical interacting residues for EBOV VP24 with human karyopherin alpha 5 (KPNA5) proteins[6] and other similar studies[12,13] inspired us to design the VP35 interface in search of novel hotspots, which when mutated will lead to a diminished affinity between the VP35 and PKR proteins. Toward this direction, we design the interface of VP35 protein with PKR in two residue stretches of the C-terminal ($VP35^{C}_{230-239}$ and $VP35^{C}_{267-279}$). The design suggests mutations, which significantly disrupt the VP35−PKR complex and hence can be inferred as interacting hotspots. We believe that the interacting hotspots can be considered for small molecule discovery. The quaternary structures discussed in this work and the computational methods elucidated can be extended to the analysis of similar viral assemblies and host−pathogen interactions for advancing the development of novel therapeutics. The tetrameric complex of VP35, the shortlisted VP35−PKR complexes, along with results from all MD simulations can be accessed at https://cosmos.iitkgp.ac.in/EbolaVP35.

MATERIALS AND METHODS

Modeling of the Tetrameric Structure of VP35

First, we model the individual subunit structures of VP35 utilizing existing partial coordinate information. Throughout the modeling of the tetramer, we accommodate the PSIPRED[14] and PSSPRED[15] predicted secondary structure of the VP35 monomer and assemble the segments whose crystallographic information is not available to us. We consider the available N-terminal coordinate information of VP35 from the crystal structure of the Ebola virus nucleoprotein bound to VP35 chaperoning peptide (PDB ID: 4ZTA, Chain A). We model the $VP35^{N}_{15-60}$ subunit of the protein by fixing the broken backbone using MODELLER.[16] In the absence of homologous templates, we model $VP35^{N-O}_{71-185}$ with the help of the ab initio protein structure prediction software viz, QUARK.[17] We rely on I-TASSER[18] to predict the structure of $VP35^{C}_{217-340}$ considering the crystal structure of the VP35 RNA binding domain (PDB ID: 5BPV, Chain A) as a template. Now, we integrate all these structures along with the crystal structure from the PDB ID: 1CLH, Chain A (as a homolog to model the loop region of the $VP35^{N}_{52-70}$) as templates and again use MODELLER to computationally model a monomer for the entire stretch of the VP35 protein. To ensure that $VP35^{N-O}_{71-185}$ containing the large helical coiled-coil oligomerization domain retains its structural integrity, we place distance constraints between each residue of the stretch during template-based modeling of the monomer. We consult the discrete optimized protein energy (DOPE) score[19] to shortlist the monomer from a set of 500 modeled structures.

https://dx.doi.org/10.1021/acs.jproteome.0c00473
J. Proteome Res. 2020, 19, 4533−4542

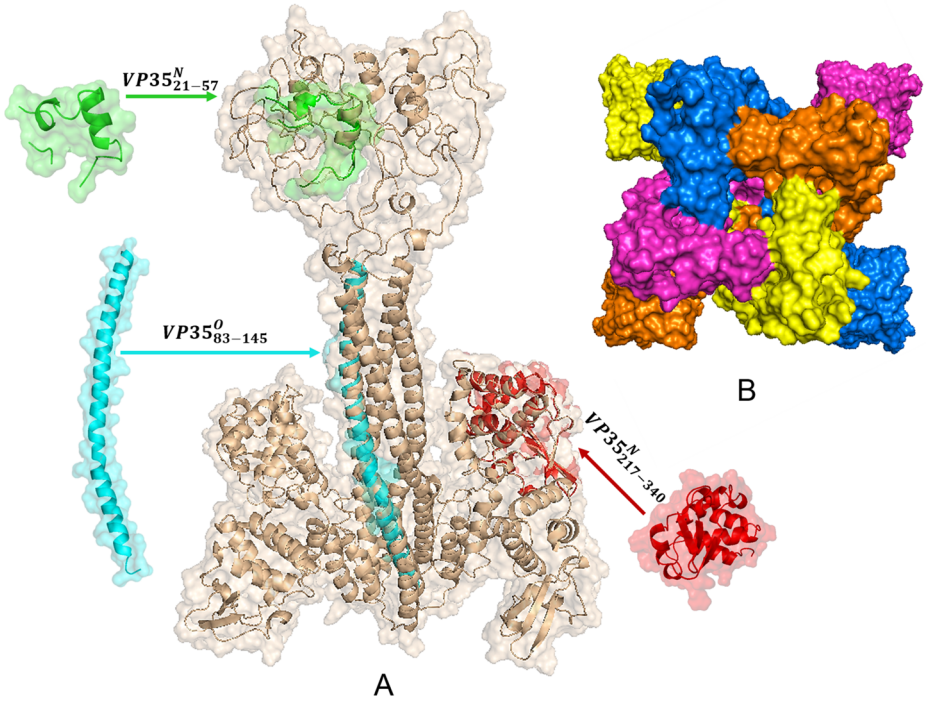


Figure 1. (A) Front view of the modeled homotetrameric structure of VP35 (in brown). The piecewise crystal structures of $VP35^{N}_{21-57}$ in green (extracted from PDB ID: 4ZTA, Chain A), of $VP35^{C}_{217-340}$ in red (extracted from PDB ID: 5BPV, Chain A), and of $VP35^{O}_{83-145}$ in cyan (extracted from PDB ID: 6GBO, Chain A) that are used to model the homooligomeric assembly. (B) Top view of the modeled VP35 tetrameric structure (Chain A in orange, Chain B in marine blue, Chain C in magenta, and Chain D in yellow).

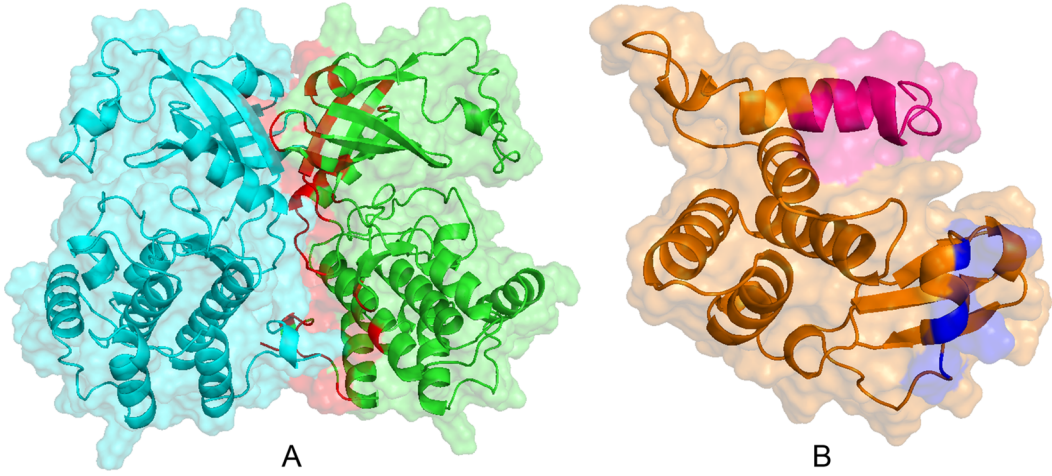


Figure 2. (A) The I-TASSER extended the modeled dimer of PKR's kinase domain (PKR^{K}) based on the crystal structure with PDB ID: 2A19, Chain B/C. The red-colored region refers to the interface between the dimer. (B) The I-TASSER extended modeled IID of VP35 in yellow comprising of $VP35^{C}$ based on the crystal structure with PDB ID: 5BPV, Chain A. The magenta-colored residue corresponds to $VP35^{C}_{186-200}$, and the blue colored residues correspond to residues Arg305, Lys309, and Arg312 of VP35.

Next, we use SymmDock[20] to prepare tetrameric assemblies from the VP35 monomer and, based on experimental information regarding the interaction of individual subunits, shortlist a plausible assembly for further modeling. We will refer to this plausible assembly as initial assembly in further discussions.

We use the coordinate information of a single chain from the newly discovered crystal structure of the oligomerization domain ($VP35^{O}_{83-145}$, PDB ID: 6GBO, Chain A) of VP35 (in its trimeric state) to construct the tetrameric oligomerization domain of VP35. As suggested by Zinzula et al.,[9] we superimpose $VP35^{O}_{83-145}$ on the crystal structure of the oligomerization domain of VP35 from the Reston virus (PDB ID: 6GBQ) to obtain the tetrameric model for the said stretch. Next, we consult the secondary structure for the next 40 residues ($VP35^{O}_{146-185}$) of VP35 and accordingly rely on the tetramer from the crystal structure of Nipah Virus (PDB ID: 4N5B) to model this stretch using the multichain modeling module of MODELLER. Considering $VP35^{O}$, we again use the multichain modeling module to combine the tetrameric structure from $VP35^{O}_{83-145}$ and $VP35^{O}_{146-185}$. Relying on the DOPE score of the models, we finalize the tetrameric assembly for $VP35^{O}$ of the tetramer. Finally, we consider the tetrameric unit for $VP35^{N}$ and $VP35^{C}$ from the initial assembly

https://dx.doi.org/10.1021/acs.jproteome.0c00473
J. Proteome Res. 2020, 19, 4533−4542

and together with the finalized assembly for $VP35^{O}$, rely on the multichain modeling of MODELLER to get the tetrameric structure of VP35. Note that although we use ab initio structure prediction to construct the initial assembly, the final tetrameric structure consists of I-TASSER or MODELLER modeled subparts with crystal structure templates provided for homology modeling. Hence the modeled structure can be thought of as acceptably reliable since 93% of the VP35 protein is modeled using crystal structure templates. Figure 1A contains the crystal structure of the N-terminal subunit (in green), the C-terminal subunit (in red), and the oligomerization domain (in cyan) used for modeling the tetramer.

Computational Docking of VP35 and PKR Subunits

Figure 2A corresponds to the dimer of PKR's kinase domain (PKR^{K}) responsible for its activation and further activation of eIF2α. We model the stretch 258−551 of the kinase domain adhering to the crystal structure of PKR (PDB ID: 2A19, Chain B/C) using I-TASSER to get the entire interface shared in the dimer. While it is experimentally determined that the deletion of $VP35^{N-C}_{1-200}$ robs VP35 of its ability to antagonize PKR activation,[21] it is also reported that mutation of any two of Arg305 or Lys309 or Arg312 (blue stretch in Figure 2B) by alanine inhibits VP35 from antagonizing PKR activation.[11] With the objective of identifying docked conformations that best represent a plausible interaction between VP35 and PKR, we dock the I-TASSER modeled structures of $VP35^{C}$ with PKR^{K} using well established protein−protein docking protocols. We rely on the docking software HADDOCK,[22] and FireDock[23] to form an initial pool of 136 docking decoys and further shortlist complexes that best represent a VP35−PKR interaction either in $VP35^{C}_{186-200}$ or in $VP35^{C}_{305-312}$. The initial shortlisting of complexes is done on the basis of the presence of either a hydrogen bond and/or salt bridge involving at least two out of Arg305, Lys309, and Arg312 and the presence of at least one hydrogen bond/salt bridge in $VP35^{C}_{186-200}$.

Additionally, the number of interface residues of PKR in the complexes in common with the interface residues of PKR in its dimeric form is also considered. The initial shortlisting results in 26 protein complexes that reduce to 10 when we apply a constraint that necessitates at least three hydrogen bonds and/or salt bridge involving Arg305, Lys309, and Arg312. Since alanine mutation in any two out of these three residues prevents VP35 antagonism, we perceive that each of the three residues participates in at least one interprotein bonding interaction. Following that, two more protein complexes are discarded as the VP35 protein in the complex interacts with less than 50% of the interface residues (residues in red in Figure 2A) of PKR. This criterion is applied to weed out complexes in which the VP35 protein will not antagonize the PKR dimerization. To recognize stable interaction we empirically consider three shortlisted complexes (SCs) viz., SC1, SC2, and SC3 with FoldX interaction energy <2 kcal/mol (as per FoldX software; higher the value lesser is the interaction), interface area >1000 $Å^2$ (computed using NIP-NSc software[24]), and at least 10 hydrogen bonds between the VP35 and PKR proteins. Please note that a higher interface area with a greater number of interprotein hydrogen bonds along with lower interaction energy ensures that the chosen complexes are sufficiently stable.

Considering SC1, SC2, and SC3, we explore three possible conformations of PKR binding to an almost identical interface of VP35. Exploiting available experimental evidence regarding the importance of the residues in VP35's interface, we allow minimum variation in its binding site, but at the same time, consider three different binding regions for PKR for a detailed analysis of plausible binding conformations. Following our hypothesis on competitive VP35 binding with PKR to prevent PKR from dimerization, we ensure that the residues in PKR's interface in all the SCs intersect with the residues in its dimeric interface. The minimum root-mean-square deviation (RMSD) between the three SCs is 25.72 Å (between SC1 and SC3), whereas the maximum RMSD between them is 33.26 Å (between SC2 and SC3).

Computational Protein Design of VP35 Interface Residues

We perform computational protein design on the Midstretch of VP35 that interacts with PKR in the three shortlisted VP35−PKR protein complexes. Earlier, we relied on the EvoDesign protein design framework to design EBOV VP24 residues that interact with the human KPNA5 protein.[6] EvoDesign is a structural homology driven computational protein design method, and in the present case, the structural homologs of $VP35^{C}$ are considered to guide the algorithm. A residue profile based on their frequency in the structural homologs is constructed to suggest mutations during the Monte Carlo (MC) search of optimal design sequences. MC simulation, guided by a hybrid energy function (evolutionary energy and physics-based energy), is performed for 30 000 steps in ten simulation trajectories. The energetically fit sequences are finally clustered, and the cluster centers from the ten most significant sequence clusters are considered as design sequences. In the present context, residues in selected stretches of VP35 are only mutated during the MC simulation. Lastly, one design sequence is chosen manually for further analysis.

Molecular Dynamics Simulation

We perform MD simulation using GROMACS 5.1.5[25] with the OPLS-AA force field. The protein structures are solvated in a cubic box, ensuring that all the protein atoms are at a distance of 1.0 nm from their respective box edges. The energy of the solvated systems is minimized for 1000 steps using the steepest descent minimization algorithm. Further, the energy-reduced systems are equilibrated for 100 ps at 298 K to maintain the desired temperature of 298 K. Subsequently, the systems are equilibrated for 1 ns at 298 K to arrive at the desired pressure of 1 bar. The solvated systems are then subjected to final MD simulation for 100 ns at 298 K. All the complexes considered for the VP35−PKR interaction study are first refined using GalaxyRefine[26] before running MD simulations. The MD simulations are carried out in two NVIDIA Tesla P100-PCIE-16GB GPUs where a representative VP35−PKR complex with 449 residues takes 41 h and 25 min to complete 100 ns of simulation.

Normal Mode Analysis

We perform NMA using the Bio3D normal-mode analysis function[27] provided by the DynaMut web server.[28] Herein we use the C-alpha force field[29] to analyze the interactions between constituent atoms of the protein. We carry out deformation and fluctuation analysis of the VP35 tetramer based on the coordinate information from the first ten nontrivial normal modes. However, we consider all nontrivial normal modes for the cross-correlation analysis between all residues in the tetrameric assembly of VP35.

https://dx.doi.org/10.1021/acs.jproteome.0c00473
J. Proteome Res. 2020, 19, 4533−4542

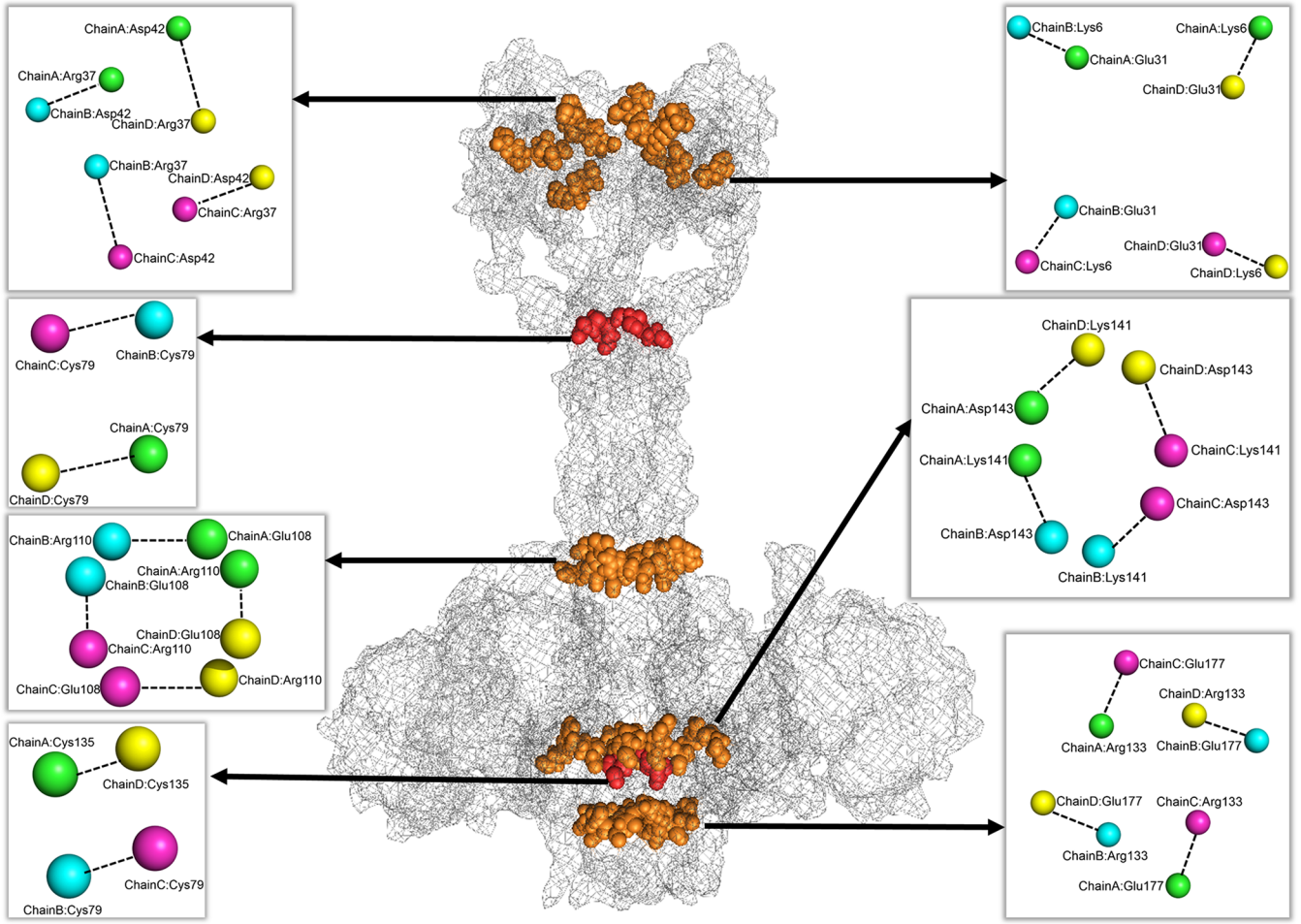


Figure 3. Interchain salt bridges (in orange) and disulfide bonds (in red) in the modeled tetrameric assembly of VP35. Top view of individual salt bridge networks and disulfide bond networks between residues (from top left anticlockwise) Arg37 and Asp42, Cys79 and Cys79, Glu108, and Arg110, Cys135 and Cys135, Arg133 and Glu177, Lys141 and Asp143 and between Lys6 and Glu31 in adjacent chains of the VP35 tetramer.

RESULTS AND DISCUSSION

Analysis of the Modeled Tetrameric Structure of VP35

The top view of the computationally modeled tetrameric structure of VP35 reveals that the assembly follows C4 point group symmetry (Figure 1B). The modeled structure is in well accordance with the crystal structure of the N-terminal domain (RMSD of 1.20 Å with PDB ID: 4ZTA, Chain A), with the oligomerization domain (RMSD of 1.97 Å with PDB ID: 6GBO, Chain A), and with the C-terminal domain (RMSD of 0.74 Å with PDB ID: 5BPV, Chain A) of the viral protein (Figure 1A). The RMSD is computed when the three crystal structures are aligned with Chain A (can be aligned with any chain since it is a homomer) of the modeled VP35 tetramer. The low RMSD values establish that the chosen segments are well aligned and in accordance with the structural constraints of the crystal structures. Considering $VP35^{N}_{1-82}$ of the modeled structure, we find that the entire subunit has no regular secondary structure except for two small helical segments spanning $VP35^{N}_{28-35}$ and $VP35^{N}_{40-43}$ in the protein. Despite a few missing residues, the presence of significant bends along the two small helices in the crystal structure (PDB ID: 4ZTA, Chain A) made us think against the free-flowing nature of the N-terminal domain. Hence, we focus on possible interchain interactions in this region. Observing the tetrameric assembly, we find residues from $VP35^{N}_{6-64}$ share an interface with residues from the $VP35^{N}_{24-37}$ in the adjacent chain. Following the coiled-coil oligomerization domain, we note that residues $VP35^{O}_{83-145}$ adhere to the known[9] helical structure and residues from the same stretch interact with the adjacent chains to give rise to the intertwined tetramer.

Following the predicted secondary structure, we model $VP35^{O-C}_{146-221}$ and observe that helices from $VP35^{O}_{152-164}$, $VP35^{O}_{174-183}$, and $VP35^{C}_{195-204}$ span the region intermediate to the oligomerization domain and the C-terminal domain. $VP35^{C}_{222-340}$ follows the exact known structural assembly of the crystal structure, and shares no interface with adjacent chains. Considering the residues of VP35 sharing an interface with residues of an adjacent chain, we find that a network of symmetric interchain salt bridges and interchain disulfide bonds acts as cages to maintain the integrity of the tetrameric structure (Figure 3). Eight such interchain salt bridges between Lys6-Glu31 and Arg37-Asp42 of each adjacent chains help to maintain the integrity of the assembly despite no predominant regular secondary structure in the N-terminal of the protein. Additionally, 12 such interchain salt bridges (four each between Glu108-Arg110, Lys141-Asp143, and between Arg133-Glu177) and four interchain disulfide bonds (Cys79A-Cys79D, Cys79B-Cys79C, Cys135A-Cys135D, Cys135B-Cys135C) add to the much-needed rigidity of the oligomerization domain. While the three (one) interchain salt bridge (disulfide bond) networks are quite evident from the

https://dx.doi.org/10.1021/acs.jproteome.0c00473
J. Proteome Res. 2020, 19, 4533−4542

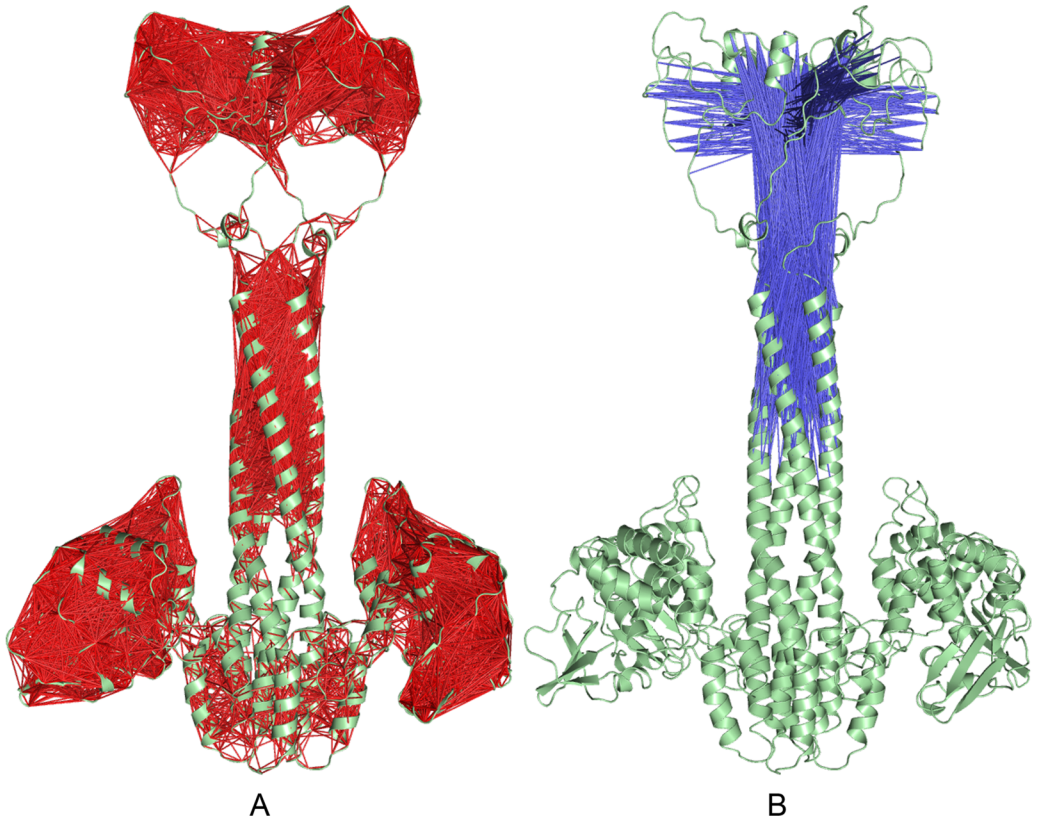


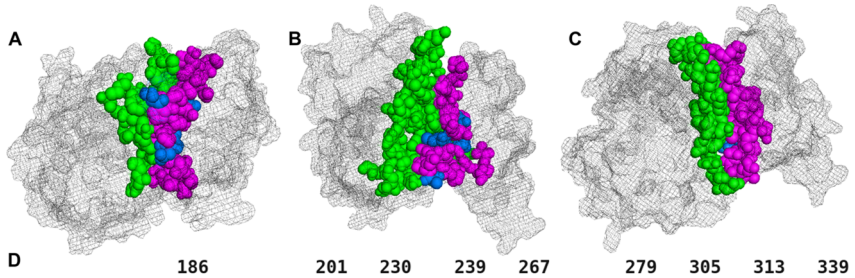
Figure 4. (A) Linkers between residues that fluctuate with a cross-correlational coefficient between 0.8 and 1 across the four chains of the tetramer (in red). (B) Linkers between residues that fluctuate with a cross-correlational coefficient between −0.6 and −0.8 across the four chains of the tetramer (in blue). Figures are generated using the DynaMut web server.

plausible assembly of the known oligomerization domain,[9] the entire modeled structure of the tetramer informs us about novel salt bridge networks between Lys6-Glu31 and Arg133-Glu177 along with a disulfide bond between Cys135 in adjacent chains.

We perform Molecular Dynamics (MD) simulation using GROMACS 5.1.5 to carry out a comprehensive analysis of the arrangement, dynamics, and stability of the modeled VP35 tetramer. We construct the entire tetrameric assembly in three steps (i) modeling of the extended oligomerization domain (four chains each containing $VP35^{O}$), (ii) modeling the tetramer containing both the N-terminal and the oligomerization domain (four chains each containing$VP35^{N-O}_{1-185}$), and (iii) building the entire VP35 tetramer (four chains each containing $VP35^{N-C}_{1-340}$). We perform 100 ns MD simulations for the tetrameric assembly generated in each of these steps. We analyze the RMSD, root-mean-square fluctuation (RMSF), and the number of hydrogen bonds for each MD simulation (Figure S1), where the RMSD corresponds to the combined deviation of the four chains of the tetrameric assemblies. The detailed analysis of the MD simulation for the first two steps is provided in the Supporting Information.

Considering the MD simulation of the entire tetrameric assembly, we find the RMSD peaks at 2.1 nm and stabilize within 20 ns of the simulation (Figure S1G). The RMSD plot captures the RMSD of each protein structure in the sampled frames (across the MD simulation) with respect to the initial protein structure. As a protein undergoes MD simulation, the kinetic energy of the individual atoms along with many other interactions result in the fluctuations of the RMSD value which stabilizes once the structure attains an equilibrium state. Although the extent of conformational change is not a direct measure of a protein's stability, a nonplateauing RMSD plot indicates poor structural integrity and possible unfolding. The RMSF plot (that measures the fluctuation of each residue in uniformly sampled protein structures across the MD simulation) of the entire assembly in Figure S1H indicates sizable fluctuation of residues in $VP35^{C}_{195-230}$ and suggests a highly flexible intermediate region between the C-terminal and the oligomerization domain of the VP35 protein. The flexibility of the intermediate region and the flexibility of the highly unstructured N-terminal domain contribute to the high RMSD of the modeled tetramer. An excellent average Pearson correlation coefficient (PCC) of 0.83 between the RMSF of residues in individual chains suggest highly correlated and symmetric dynamics in the tetrameric assembly. The number of hydrogen bonds after a steady increase saturates at around 60 ns and averages at 895 for the entire duration of the simulation (Figure S1I). We find that the number of interchain hydrogen bonds (computed for 100 frames using HBPLUS[30]) steadily increases to a maximum of 100 and averages at 75 for the entire duration of the simulation (Figure S2A). A very similar trend is also observed for the interchain salt bridges (a steady increase to a maximum of 27 with an average of 21) (Figure S2B). We find that residues participating in symmetric interchain salt bridges with alternative residues bearing opposite polarity in adjacent chains support the modeled tetrameric assembly to remain intact during the 100 ns of MD simulation. Interestingly, the residues of the N-terminal interacting with adjacent chains continue similar interactions for the 100 ns of the simulation. The stability imparted to the interacting residues due to the two interchain salt-bridge

https://dx.doi.org/10.1021/acs.jproteome.0c00473
J. Proteome Res. 2020, 19, 4533−4542



D

	186 201 230 239 267 279 305 313 339
Native	ESRDETVPQSVREAFN---DHLPGFGTAF---LAEGDSPQCALIQ---RACQKSLRP---KI
SC1	------VP---R--------H-PGF-TAF---LAEGDSPQC--I----RA--KS-RP---KI
SC2	-------PQ--R--FN---DHLPGF-TAF--------SPQ--------RA--KSLRP---KI
SC3	------VP---R--F-----HLPGF-TAF---LAEGDSPQC--IQ---RA-QKS-R----KI
For_Design	------------------DHLPGF-TAF---LAEGDSPQC----------
Middes_SC1	------------------T----N--P-----E-----EA----------
Middes_SC2	------------------T----H--P-----S--EN--A----------
Middes_SC3	----------N----------

Figure 5. (A) SC1, (B) SC2, (C) SC3 with the interface residues of PKR^{K} shown in green color. The interface residues of $VP35^{C}$ are shown in magenta color, and the residues mutated in the Midstretch of $VP35^{C}$ while performing protein design (Middes) using EvoDesign are shown in marine color. (D) The interface residues of VP35 in the 3 SCs along with their respective mutations, according to Middes as determined using EvoDesign.

networks is sufficient to hold the interacting stretches of the N-terminal together. Despite the presence of highly flexible loop regions in the N-terminal, the interacting stretches continue to interact for the entire duration of the MD simulation.

Figure S3 presents the variation in secondary structure for the 340 residues across the four chains of the tetramer computed using the STRIDE[31] software. Observing the distribution of the secondary structures, we note that 83% of protein reports a consistent secondary structure during 75% of the simulation time. We find that regions like $VP35^{O}_{152-164}$, $VP35^{O}_{174-183}$, and $VP35^{C}_{195-204}$ are transient helices and, at times, partly convert to loop. These transient parts along with the existing loop provide the much necessary flexibility to the intermediate region joining the oligomerization domain to the C-terminal of the protein. The flexibility in this region allows the C-terminal to interact with multiple human proteins and in the process facilitate inhibition of interferon activation.

The symmetry in a protein's oligomeric assembly influences its flexibility.[32] Therefore, it is important that we observe the same for the C-terminal domain of the modeled VP35 tetramer. Although the modeled tetramer has a symmetrical structure, no additional constraints are imposed to accommodate the same during the MD simulation. Instead, we intend to analyze the flexibility and symmetry of the modeled structure in the context of its unconstrained motion. Considering $VP35^{C}_{186-340}$, we find that the average RMSF of the residue stretch 186−340 (0.99 nm) is much higher than the average RMSF for the entire tetramer (0.78 nm). The observation suggests that the modeled VP35 tetramer accommodates the much needed flexibility of the C-terminal domain. Well correlated fluctuations of residues among different chains of an oligomer indicate symmetric movement of the constituent parts. We observe an average PCC of 0.75 of the RMSF of the residue stretch 186−340 among all the four chains of the modeled tetramer. Our analysis of the C-terminal domain indicates symmetry and the necessary flexibility needed for interacting with different human proteins.

We use the Bio3D package[27] provided by the DynaMut web server[28] for NMA. The deformation energy (Figure S4A,B) and fluctuations (Figure S4C,D) of individual residues in the tetrameric assembly computed using the first ten nontrivial normal modes reestablish that residues in the intermediate region of the N-terminal and oligomerization domain and residues in the intermediate of oligomerization domain and the C-terminal domain are prone to maximum deformation. Analyzing the cross-correlation coefficient between all residues in the entire tetrameric assembly (computed by considering all normal modes), we find that the fluctuation of identical residues across the four chains is highly correlated, and the fluctuation in residues in the N-terminal sharing an interface with the residues of the adjacent chain reports a high negative cross-correlation with the fluctuation of residues in the upper half of the oligomerization domain (Figure 4).

Similar to MD simulation, no separate symmetry based constraints are set for the NMA. Considering the C-terminal domain of the modeled tetramer, we find that the residues 186−230 have an average deformation energy of 0.32 (the energy term has no unit[33]) which is much higher than the average deformation energy of 0.12 for the entire tetramer. These sets of residues fluctuate heavily according to the NMA and they provide the much needed flexibility to the C-terminal region. In addition, we find that the fluctuations in $VP35^{C}_{186-340}$ are well correlated (average PCC about 1.00) and indicate toward symmetric motion of the residues. Hence, we conclude that the C-terminal domain demonstrates symmetry and is also flexible for necessary interactions with different human proteins.

Establishing VP35's Interaction with Human PKR

Adhering to our hypothesis that the IID of VP35 directly binds with PKR to prevent it from autophosphorylation, we analyze plausible interactions between them. We dock[22,23,34−37] the I-TASSER modeled PKR^{K} with $VP35^{C}$ (as modeled earlier by us) to shortlist SC1, SC2, and SC3 based on physicochemical specifications (detailed in Material and Methods section) and manual observations regarding the $VP35^{C}$−PKR^{K} interaction (Figure 5A−C). In all the three SCs residues 305, 309, and 312 along with a subset of residues from $VP35^{C}_{186-200}$ share an interface with residues responsible for the dimerization of PKR^{K}.

Although we shortlist three complexes, we find that the $VP35^{C}$ interface is almost identical while allowing variation in the way PKR^{K} binds to the viral protein. The VP35 interface being almost identical for the three SCs is used to identify hot spots, which when mutated lead to reduced binding affinity

https://dx.doi.org/10.1021/acs.jproteome.0c00473
J. Proteome Res. 2020, 19, 4533−4542

Table 1. Summary of VP35−PKR Interaction for the Shortlisted Complexes Averaged over 100 ns of MD Simulation

	SC1			SC2			SC3		
complex	IA[a]	NCB[b]	IE[c]	IA[a]	NCB[b]	IE[c]	IA[a]	NCB[b]	IE[c]
native	986	11	0.21	737	10	−3.81	871	11	−2.75
PKR_Dimer	1071	11	−0.11	1071	11	−0.11	1071	11	−0.11
Ala_mut	703	6	2.45	783	7	−0.70	956	14	0.97
Middes	705	10	2.00	865	10	−2.60	763	10	0.42

[a]IA: Interface Area (Å^2) rounded off to the nearest integer. [b]NCB: Number of noncovalent bonds (hydrogen bonds and salt bridges) rounded off to the nearest integer. [c]IE: Interaction energy (kcal/mol), higher the value lower is the interaction.

with PKR^K. In Figure S5, we superimpose all SCs with the predicted tetrameric structure (chain A) of VP35. In Figure S5A and Figure S5C, we find that on superposing the SCs, the constituent PKR^K does not interfere with the remaining tetrameric structure of VP35, although, for SC2 (Figure S5B), the PKR^K does interfere marginally with the tetrameric structure. However, considering the fact that the VP35^C is highly flexible owing to the high deformability of VP35$^{O-C}_{168-221}$, the PKR^K in SC2 can be accommodated without possible interference in the VP35 tetramer. Considering the three SCs, the residues of VP35^C sharing an interface with PKR^K spans over three stretches: VP35$^C_{192-201}$, VP35$^C_{305-313}$, and VP35$^C_{230-239}$ along with the patch VP35$^C_{267-279}$ (further discussed as Midstretch).

Literature informs that any two out of three alanine replacements at position 305, 309, and 312 will inhibit VP35 from antagonizing PKR autophosphorylation. Therefore, we consider position 309 and 312 to examine the binding affinity of the SC1, SC2, and SC3 with the PKR^K dimer by accommodating the Lys309Ala and Arg312Ala mutations (denoted as Ala_mut). We perform 100 ns MD simulation for the PKR^K dimer, for each of SC1, SC2, SC3, and in the corresponding conformations accommodate the two alanine replacements. Figure S6 contains the RMSD and Radius of gyration (R_g) plots for each of the protein complexes while Table S1 includes the peak RMSD values and the average R_g values corresponding to their simulations. The RMSD values of all the complexes are more or less comparable with a slightly higher R_g value for the PKR^K dimer possibly due to its bigger size. We consider 100 frames of coordinate information of the protein complexes sampled at regular intervals during the 100 ns of simulation for our analysis. We compute the average of the noncovalent interactions (consisting of the number of interprotein hydrogen bonds and salt bridges) and interaction energy (computed using the AnalyseComplex module of FoldX[38]) between the interacting proteins in the chosen complexes for our comparison (Table 1).

We find that in SC1, the average noncovalent interactions are the same as that of the PKR^K dimer and the interaction energy of the kinase dimer is only marginally lower than that of SC1. However, the number of noncovalent interactions drops drastically to six and the interaction energy increases to 2.45 kcal/mol for Ala_mut of SC1, although geometric fitting[24] for Ala_mut of SC1 is better than that of the PKR^K dimer (Table S2). For SC2, the number of noncovalent interactions is one less than that of the PKR^K dimer. However, the interacting energy (−3.81 kcal/mol) for the SC2 complex is much lower than that of the kinase dimer. Again, Ala_mut of SC2 results in much lower noncovalent interactions (only seven) but with higher interaction energy (−0.7 kcal/mol) than native SC2. The interface packing and surface complementarity of SC2 are higher than both the PKR^K dimer and SC2 with Ala_mut. Analyzing SC3, we find that the noncovalent interaction of the complex is comparable with that of the PKR^K dimer, although, with interaction energy of −2.75 kcal/mol, the SC3 is much more stable than the PKR^K dimer. SC3 with Ala_mut, report three more noncovalent interactions as compared to SC3 but again report higher interaction energy (0.97 kcal/mol). The interface packing and surface complementarity of SC3 are again better than that of the kinase dimer and SC3 with Ala_mut. Summarizing three SCs, we observe an overall higher binding affinity between VP35^C and PKR^K than in the PKR^K dimer, indicating the possibility of direct interaction with VP35^C, which in turn prevents the PKR from dimerization. The fact that the noncovalent interactions and interaction energy of the PKR^K dimer primarily lies in between that of the native SCs and the Ala_mut SCs further validates the reason why the alanine replacements inhibit VP35^C from antagonizing PKR^K autophosphorylation.

Protein Design to Identify Interacting Hotspots in VP35

Truncation of VP35$^{N-C}_{1-200}$ and the alanine replacements in residues 305, 309, 312 of VP35$^C_{305-312}$ both inhibit VP35 from antagonizing PKR autophosphorylation. While the alanine replacements experiment at 305, 309, and 312 ensure that VP35 does not antagonize the autophosphorylation of PKR, it is difficult to comment on the fold level integrity of the protein due to such replacements. We perform computational protein design[39−41] in the Midstretch residues to evaluate the binding potential of an unexplored stretch in the VP35 protein. The protein design algorithm[6,13,40,41] explores the possible mutations in the VP35 protein that will not compromise its structural integrity and therefore provide us with a better opportunity to study the altered binding affinity in the chosen complex. We design the Midstretch of VP35 to witness an altered binding affinity among the two proteins. The efficacy of computational protein design to predict EBOV VP24 mutations in disrupting its interaction with Human KPNA5[6] is already established. Computational protein design is also successfully exploited in designing protein sequences (or sequence patches) for advancing the development of novel therapeutics.[12,13] Protein design ensures the generation of novel amino acid sequences following the homology information of the target protein while ensuring the fold level fitness of the entire protein. During protein design, we consider three design sequences each one from SC1, SC2, and SC3 (hereafter called Middes), in which certain residues in the Midstretch (Figure 5D) of the VP35 protein are mutated. We tabulate the details on the mutations reported in Middes of the three SCs in Table S3.

Adhering to the same kind of analysis as mentioned in the previous section, we compare the binding affinity of all the modeled SCs with the SCs accommodating the Middes mutations. Figure S6 and Table S1 contain the RMSD and

https://dx.doi.org/10.1021/acs.jproteome.0c00473
J. Proteome Res. 2020, 19, 4533−4542

R_g data corresponding to these mutations for the respective SCs. We again consider the average of the noncovalent interactions and interaction energy between the interacting proteins in the SCs and their corresponding Middes mutations across the 100 frames sampled across 100 ns of MD simulations (Table 1). We note that the Middes mutations of SC1 (Middes_SC1 in Figure 5D) result in one lesser noncovalent interaction and increased interaction energy compared to SC1 (and the PKR^K dimer). In SC2, we find that the Middes mutations (Middes_SC2 in Figure 5D) result in a similar number of noncovalent interactions, higher interaction energy than native SC2, but lower interaction energy than the PKR^K dimer. Interestingly in SC3, only one mutation (Ser272Asn) in the Midstretch is sufficient to result in a lower number of noncovalent interactions and higher interaction energy than both native SC3 and the PKR^K dimer. Although both the alanine mutations and the Middes mutations lead to decreased binding affinity than the native protein complexes, a direct comparison of the two is difficult as we are unaware of the structural changes induced by the alanine mutations in the VP35 protein. It is important to note that the output of protein design on the Midstretch of $VP35^C$ provides alternative residue mutations that will ensure the fold level integrity of $VP35^C$ but in no way determines its interaction affinity with its binding partner. Many of the alternative mutations in the Midstretch lead to an increased affinity with PKR^K, but we carried our analysis only on those mutations that lead to a decreased affinity with PKR^K. The analysis of the Middes mutations indicates that the Midstretch can be alternatively explored in wet-lab experiments to plausibly inhibit VP35 from interacting with PKR.

CONCLUSION

Considering the overarching importance of the VP35 protein in the EBOV−human protein interaction landscape, we model the tetrameric assembly (structure) of the viral protein. In a piece-wise manner, the structural information of about 67% of only one subunit out of four subunits of the VP35 protein is known. Nevertheless, an entire tetrameric assembly of the 340 residues long VP35 protein is of immense importance to better comprehend the structural characteristics of the viral protein. The stability and dynamics of the modeled tetramer are assessed and analyzed with the help of MD simulations and NMA. The simulations indicate highly symmetric behavior among identical residues across all the four chains of the tetramer. In the tetrameric assembly, we identify five symmetrical interchain salt bridge networks and two interchain disulfide bond networks between the adjacent chains that are responsible for the integrity of the entire assembly of the viral protein. We note that the flexibility of the $VP35^C$ in engaging human IFN antagonism is due to the highly flexible $VP35^{O-C}_{168-221}$ region (present in the intermediate region between the oligomerization domain and the C-terminal DNA binding domain/IID). We predict transient helical subunits in $VP35^{O-C}_{168-221}$ that along with adjacent loop regions provide the much-needed flexibility to the C-terminal. From a computational modeling point of view, the construction of the VP35 tetramer posed unique challenges owing to the large stretches of residues with no regular secondary structure and the presence of a large coiled-coil oligomerization domain. Hence, the protocol followed to construct the assembly that can also be extended to model a similar assembly of proteins.

In addition to the computational modeling of the tetrameric assembly, we propose interactions between the $VP35^C$ and PKR^K molecules plausibly responsible for the VP35 in inhibiting the autophosphorylation of the PKR. While establishing our hypothesis, we present three shortlisted docked complexes that best represent the interaction between the VP35 and PKR proteins. MD simulation studies suggest that in the shortlisted complexes, PKR^K has a greater affinity toward $VP35^C$ as compared with its affinity for another PKR^K molecule in a PKR^K dimer. The findings suggest the possibility that VP35 directly interacts with PKR through its dimerization interface and prevents it from autophosphorylation. The MD simulation of the shortlisted complexes subject to alanine replacements in Lys309 and Arg312 of the VP35 protein indicate reduced stability than a PKR^K dimer. The observation is in line with the experimental fact that alanine replacements of any two residues among Arg305, Lys309, and Arg312 of the VP35 protein resulted in its inability to stop PKR's autophosphorylation. Through a protein design study, we locate mutations in $VP35^C_{230-239}$ and $VP35^C_{267-279}$ which noticeably weaken the interaction between the VP35 and PKR proteins. For experimental validation, the stretch we consider for protein design can prove to be a specific target for selected small molecule inhibitors.

Computational alternatives (and augmentation) to wet-lab experimentation are particularly important as the EBOV is highly contagious. In the present work, we present a computationally predicted tetrameric structure of the VP35 protein to augment our understanding regarding the assembly, dynamics, and bonding patterns of the viral protein. Further, we show that VP35 competitively binds with PKR and prevents it from autophosphorylation. Toward this direction, we provide extensive analysis from docking and design studies to establish the hypothesis and identify novel interacting hotspots in the VP35 protein that can be considered for drug discovery.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.0c00473.

MD simulation analysis of the modeled intermediate tetrameric assemblies; Figures S1−S6; Tables S1−S3 (PDF)

AUTHOR INFORMATION

Corresponding Author

Pralay Mitra − *Department of Computer Science and Engineering, Indian Institute of Technology Kharagpur, West Bengal 721302, India;* orcid.org/0000-0003-4119-3788; Phone: +91-3222-282344; Email: pralay@cse.iitkgp.ac.in

Author

Anupam Banerjee − *Advanced Technology Development Centre, Indian Institute of Technology Kharagpur, West Bengal 721302, India;* orcid.org/0000-0002-2859-7705

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.0c00473

Notes

The authors declare no competing financial interest.

https://dx.doi.org/10.1021/acs.jproteome.0c00473
J. Proteome Res. 2020, 19, 4533−4542

■ ACKNOWLEDGMENTS

We are thankful to Abantika Pal, a research scholar from our group, for initial discussions.

■ REFERENCES

(1) Ebola situation report; World Health Organization, 2016.
(2) Ye, Y. Z.; Tang, H. X.; Ding, D. F. Design of Protein Cores by Screening Combinatorial Sequence Library. *Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai)* **1998**, *30* (5), 525−539.
(3) *Ebola Virus Disease Democratic Republic of Congo: External Situation Report*; World Health Organization, Oct 22, 2019; p 64.
(4) Messaoudi, I.; Amarasinghe, G. K.; Basler, C. F. Filovirus pathogenesis and immune evasion: insights from Ebola virus and Marburg virus. *Nat. Rev. Microbiol.* **2015**, *13* (11), 663−76.
(5) Kuhl, A.; Pohlmann, S. How Ebola virus counters the interferon system. *Zoonoses Public Health* **2012**, *59* (Suppl 2), 116−31.
(6) Banerjee, A.; Pal, A.; Pal, D.; Mitra, P. Ebolavirus interferon antagonists-protein interaction perspectives to combat pathogenesis. *Briefings Funct. Genomics* **2017**, *17* (6), 392−401.
(7) Edwards, M. R.; Liu, G.; Mire, C. E.; Sureshchandra, S.; Luthra, P.; Yen, B.; Shabman, R. S.; Leung, D. W.; Messaoudi, I.; Geisbert, T. W.; et al. Differential regulation of interferon responses by Ebola and Marburg virus VP35 proteins. *Cell Rep.* **2016**, *14* (7), 1632−1640.
(8) Bruhn, J. F.; Kirchdoerfer, R. N.; Urata, S. M.; Li, S.; Tickle, I. J.; Bricogne, G.; Saphire, E. O. Crystal structure of the Marburg virus VP35 oligomerization domain. *J. Virol.* **2017**, *91* (2), No. e01085-16.
(9) Zinzula, L.; Nagy, I.; Orsini, M.; Weyher-Stingl, E.; Bracher, A.; Baumeister, W. Structures of Ebola and Reston Virus VP35 Oligomerization Domains and Comparative Biophysical Characterization in All Ebolavirus Species. *Structure* **2019**, *27* (1), 39−54.
(10) Garcia, M. A.; Meurs, E. F.; Esteban, M. The dsRNA protein kinase PKR: virus and cell control. *Biochimie* **2007**, *89* (6−7), 799−811.
(11) Schumann, M.; Gantke, T.; Muhlberger, E. Ebola virus VP35 antagonizes PKR activity through its C-terminal interferon inhibitory domain. *J. Virol.* **2009**, *83* (17), 8993−7.
(12) Rajesh, Y.; Banerjee, A.; Pal, I.; Biswas, A.; Das, S.; Dey, K. K.; Kapoor, N.; Ghosh, A. K.; Mitra, P.; Mandal, M. Delineation of crosstalk between HSP27 and MMP-2/MMP-9: A synergistic therapeutic avenue for glioblastoma management. *Biochim. Biophys. Acta, Gen. Subj.* **2019**, *1863* (7), 1196−1209.
(13) Shultis, D.; Mitra, P.; Huang, X.; Johnson, J.; Khattak, N. A.; Gray, F.; Piper, C.; Czajka, J.; Hansen, L.; Wan, B.; Chinnaswamy, K.; Liu, L.; Wang, M.; Pan, J.; Stuckey, J.; Cierpicki, T.; Borchers, C. H.; Wang, S.; Lei, M.; Zhang, Y. Changing the Apoptosis Pathway through Evolutionary Protein Design. *J. Mol. Biol.* **2019**, *431* (4), 825−841.
(14) McGuffin, L. J.; Bryson, K.; Jones, D. T. The PSIPRED protein structure prediction server. *Bioinformatics* **2000**, *16* (4), 404−405.
(15) Yan, R.; Xu, D.; Yang, J.; Walker, S.; Zhang, Y. A comparative assessment and analysis of 20 representative sequence alignment methods for protein structure prediction. *Sci. Rep.* **2013**, *3*, 2619.
(16) Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinf.* **2016**, *54*, 5.6.1−5.6.37.
(17) Xu, D.; Zhang, Y. Toward optimal fragment generations for ab initio protein structure assembly. *Proteins: Struct., Funct., Genet.* **2013**, *81* (2), 229−39.
(18) Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y. The I-TASSER Suite: protein structure and function prediction. *Nat. Methods* **2015**, *12* (1), 7−8.
(19) Shen, M. Y.; Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein Sci.* **2006**, *15* (11), 2507−24.
(20) Schneidman-Duhovny, D.; Inbar, Y.; Nussinov, R.; Wolfson, H. J. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.* **2005**, *33*, W363−W367.
(21) Feng, Z.; Cerveny, M.; Yan, Z.; He, B. The VP35 protein of Ebola virus inhibits the antiviral effect mediated by double-stranded RNA-dependent protein kinase PKR. *J. Virol.* **2007**, *81* (1), 182−92.
(22) van Zundert, G. C. P.; Rodrigues, J.; Trellet, M.; Schmitz, C.; Kastritis, P. L.; Karaca, E.; Melquiond, A. S. J.; van Dijk, M.; de Vries, S. J.; Bonvin, A. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.* **2016**, *428* (4), 720−725.
(23) Mashiach, E.; Schneidman-Duhovny, D.; Andrusier, N.; Nussinov, R.; Wolfson, H. J. FireDock: a web server for fast interaction refinement in molecular docking. *Nucleic Acids Res.* **2008**, *36*, W229−W232.
(24) Mitra, P.; Pal, D. New measures for estimating surface complementarity and packing at protein−protein interfaces. *FEBS Lett.* **2010**, *584* (6), 1163−1168.
(25) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29* (7), 845−854.
(26) Heo, L.; Park, H.; Seok, C. Galaxy Refine: Protein structure refinement driven by side-chain repacking. *Nucleic Acids Res.* **2013**, *41*, W384−W388.
(27) Skjaerven, L.; Yao, X. Q.; Scarabelli, G.; Grant, B. J. Integrating protein structural dynamics and evolutionary analysis with Bio3D. *BMC Bioinf.* **2014**, *15*, 399.
(28) Rodrigues, C. H.; Pires, D. E.; Ascher, D. B. DynaMut: Predicting the Impact of Mutations on Protein Conformation, Flexibility and Stability. *Nucleic Acids Res.* **2018**, *46* (W1), W350−w355.
(29) Hinsen, K.; Petrescu, A.-J.; Dellerue, S.; Bellissent-Funel, M.-C.; Kneller, G. R. Harmonicity in slow protein dynamics. *Chem. Phys.* **2000**, *261* (1−2), 25−37.
(30) McDonald, I. K.; Thornton, J. M. Satisfying Hydrogen Bonding Potential in Proteins. *J. Mol. Biol.* **1994**, *238* (5), 777−93.
(31) Heinig, M.; Frishman, D. STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res.* **2004**, *32*, W500−W502.
(32) Schulze, B.; Sljoka, A.; Whiteley, W. How does symmetry impact the flexibility of proteins? *Philos. Trans. R. Soc., A* **2014**, *372* (2008), 20120041.
(33) Hollup, S. M.; Salensminde, G.; Reuter, N. WEBnm@: a web application for normal mode analyses of proteins. *BMC Bioinf.* **2005**, *6*, 52.
(34) Kozakov, D.; Hall, D. R.; Xia, B.; Porter, K. A.; Padhorny, D.; Yueh, C.; Beglov, D.; Vajda, S. The ClusPro web server for protein-protein docking. *Nat. Protoc.* **2017**, *12* (2), 255−278.
(35) Pierce, B. G.; Wiehe, K.; Hwang, H.; Kim, B. H.; Vreven, T.; Weng, Z. ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics* **2014**, *30* (12), 1771−3.
(36) Mitra, P.; Pal, D. PRUNE and PROBE–two modular web services for protein-protein docking. *Nucleic Acids Res.* **2011**, *39*, W229−W234.
(37) Mitra, P.; Pal, D. Using correlated parameters for improved ranking of protein-protein docking decoys. *J. Comput. Chem.* **2011**, *32* (5), 787−96.
(38) Schymkowitz, J.; Borg, J.; Stricher, F.; Nys, R.; Rousseau, F.; Serrano, L. The FoldX web server: an Online Force Field. *Nucleic Acids Res.* **2005**, *33*, W382−W388.
(39) Mitra, P.; Shultis, D.; Zhang, Y. EvoDesign: De novo protein design based on structural and evolutionary profiles. *Nucleic Acids Res.* **2013**, *41*, W273−W280.
(40) Mitra, P.; Shultis, D.; Brender, J. R.; Czajka, J.; Marsh, D.; Gray, F.; Cierpicki, T.; Zhang, Y. An evolution-based approach to De Novo protein design and case study on Mycobacterium tuberculosis. *PLoS Comput. Biol.* **2013**, *9* (10), No. e1003298.
(41) Banerjee, A.; Pal, K.; Mitra, P. An evolutionary profile guided greedy parallel replica-exchange Monte Carlo search algorithm for rapid convergence in protein design. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **2019**, 1.

https://dx.doi.org/10.1021/acs.jproteome.0c00473
J. Proteome Res. 2020, 19, 4533−4542