NSP 11 of SARS-CoV-2 is an Intrinsically Disordered Protein

Kundlik Gadhave[1], Prateek Kumar[1], Ankur Kumar[1], Taniya Bhardwaj[1], Neha Garg[2], Rajanish Giri[1*]

[1]*School of Basic Sciences, Indian Institute of Technology Mandi, Kamand, Himachal Pradesh, 175005, India.*

[2]*Department of Medicinal Chemistry, Faculty of Ayurveda, Institute of Medical Sciences, Banaras Hindu University, Varanasi, Uttar Pradesh, 221005, India.*

***Correspondence to:** Dr. Rajanish Giri, School of Basic Sciences, Indian Institute of Technology Mandi, Himachal Pradesh 175005, India. Email: rajanishgiri@iitmandi.ac.in.

Abstract:

The intrinsically disordered proteins/regions (IDPs/IDPRs) are known to be responsible for multiple cellular processes and are associated with many chronic diseases. In viruses, the existence of disordered proteome is also proven and are related with its conformational dynamics inside the host. The SARS-CoV-2 virus has a large proteome, in which, structure and functions of many proteins are not known as of yet. Previously, we have investigated the dark proteome of SARS-CoV-2. However, the disorder status of non-structural protein 11 (nsp11) was not possible because of very small in size, just 13 amino acid long, and for most of the IDP predictors, the protein size should be at least 30 amino acid long. Also, the structural dynamics and function status of nsp11 was not known. Hence, we have performed extensive experimentation on nsp11. Our results, based on the Circular dichroism spectroscopy gives characteristic disordered spectrum for IDPs. Further, we investigated the conformational behaviour of nsp11 in the presence of membrane mimetic environment, alpha helix inducer, and natural osmolyte. In the presence of negatively charged and neutral liposomes, nsp11 remains disordered. However, with SDS micelle, it adopted an α-helical conformation, suggesting the helical propensity of nsp11. At the end, we again confirmed the IDP behaviour of nsp11 using molecular dynamics simulations.

Keywords: COVID-19, SARS-CoV-2, nsp11, CD Spectroscopy, Intrinsically disordered proteins

1. Introduction:

The existing pandemic caused by a severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is an enveloped RNA virus. Its genome organization shows the presence of a ~30 kbp positive-sense single-stranded RNA (+ssRNA) buried inside a layer of viral proteins involved in protecting and shaping the virion particle [1–3]. Once inside the host cell, genomic RNA is translated into three different types of viral proteins: structural, accessory, and non-structural proteins (nsps). Structural proteins (spike, envelope, membrane, and nucleocapsid) together forms the outer cover of virion [1,4]. Accessory proteins aid in virus survival by evading host immune system, inducing cell death, and modulating viral pathogenicity. The third category of coronavirus proteins comprises of the nsps which are cleaved from two translated polyproteins (pp1a and pp1ab) by two proteases (PLpro and 3CLpro) [4–6]. These virus-encoded nsps serves as an essential component for replication of genome and transcriptional activity. In fact, several nsps are involved throughout initial and intermediate phases of the viral life cycle [7]. The recently emerged SARS-CoV-2 has a short protein, nsp11, of only 13 amino acids [5]. Generally, depending on the CoVs species, the nsp11 comprises 13–23 residues [7]. The nsp11 protein is the cleavage product of pp1a polyprotein by 3CLpro/Mpro protease at the nsp10/11 junction (putative cleavage site is-nsp10 PMLQ|SADA nsp11) [8]. Interestingly, voluntary ribosomal frameshifting occurs in coronaviruses in replicase gene (-1 frameshift) giving rise to rest of the nsp12-16 proteins of longer polyprotein 1ab [7].

In our recent study, we analysed the propensity of intrinsic disorder in SARS-CoV-2 proteins and correlated it with its evolutionary closer human SARS and Bat SARS-like coronaviruses [9]. The multiple sequence alignment of SARS-CoV-2 nsp11 with SARS-CoV and Bat SARS-like coronavirus represents difference in protein sequence at 5th and 6th positions where glutamine and serine are observed to substitute serine and threonine residues, respectively [9]. As an uncharacterised protein of potential function, published literature on nsp11 protein of murine hepatitis virus shows cleavage mutants nsp10/nsp11 are not replication viable [10]. On contrarily, cleavage mutants nsp10-nsp11/nsp12 of avian coronavirus infectious bronchitis virus are dispensable for viral replication [6]. Sun *et al.* reported that the overexpression of nsp11 of porcine reproductive and respiratory syndrome virus (PRRSV) induced a strong suppression of interferon (IFN) production [11].

The release and fate of nsp11 protein in CoV-infected cells have not been recognised so far. Moreover, its structural, biophysical, and functional information in any coronavirus

has still not been elucidated which is vital to understand its biology and pathogenesis from structural point of view. In this study, using Far-UV CD spectroscopy, we identified SARS-CoV-2 nsp11 as an intrinsically disordered protein (IDP). IDPs are unstructured proteins that lack structural constraints forming an ensemble of structures [12,13]. In contrast to classical structure-function-paradigm, IDPs works either through structure-independent interactions or by virtue of gaining a secondary structure (called as MoRFs) such as α-helix on interaction with their physiological partner [14,15].

Several proteins from organisms of all three kingdoms of life have been studied to perform structure-independent functions [16,17]. Viruses are no less behind and also have multiple flexible proteins [18–20]. The sequence-dependent characteristic of IDPs enables them to have an exceptionally large interactome [16,17]. We, therefore, as an essential step in direction of understanding nsp11 protein and its interacting ability, employed experimental as well as computational methods to analyse its 'gain-of-structure' capacity. We used various model systems such as DOPC, DOPS, SDS micelles, TMAO, and TFE to mimic the biological environments like hydrophobic-hydrophilic interface, and protein-lipids interaction. Collectively, our findings revealed the unstructured nature of SARS-CoV-2 nsp11 protein along with membrane-mediated effects of liposomes on its structure. The comprehensive structural investigation may help to deduce the function and structure-function relationship of nsp11 protein in SARS-CoV-2. As nsp11 is present at the site where ribosomal frameshift occurs, the role of disordered nature of nsp11 therein needs to be elucidated.

2. Material and Methods:

2.1. Material: The 2,2,2-trifluoroethanol (TFE) and Sodium dodecyl sulfate **(**SDS), and trimethylamine N-oxide (TMAO) were procured from Sigma-Aldrich (USA). The lipids 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) and 1,2-dioleoyl-sn-glycero-3- phospho-L-serine (DOPS) were obtained from Avanti Polar Lipids (Alabaster, Alabama, U.S.A.) and prepared using the lipid extruder provided by the same manufacturer. nsp11 peptide (NH2-SADAQSFLNGFAV-COOH; UniProt ID: P0DTC1 (residues 4393-4405)) was procured from *Gene* Script with > 80 % purity. The peptide was dissolved in 20 mM phosphate buffer (pH 7.4) containing 50 mM NaCl to prepare stock solution (3 mg/ml).

2.2. Liposome preparation: The chloroform from lipids (DOPC/DOPS) solution was removed using rotatory evaporators. Afterward, overnight incubation was done in a desiccator to remove any traces of chloroform. The dried lipid was then hydrated with 20 mM

sodium phosphate buffer (pH 7.4) and further lipid suspension (DOPC: 29 mg/ml; DOPS: 20 mg/ml) was processed five times by freeze (liquid N_2)/thaw (60° C water bath)/vortex cycles. Lipid suspension was then subjected to extrusion to prepare large unilamellar vesicles (LUVs) using 0.1 μm pore diameter polycarbonate membrane. For the extrusion process, we used Avanti mini-extruder (Avanti) and followed the manufacturer's protocol. The prepared LUVs were stored at 4^oC and used within 5 days.

***2.3. Dynamic light scattering (DLS)*:** The hydrodynamic radius/size of the prepared LUVs was measured by using DLS (Zetasizer Nano S from Malvern Instruments Ltd., UK). The prepared LUVs were diluted in water (1:100 ratio of LUV and water) before measurements. The observed size of the DOPC and DOPS was 102 nm and 124 nm respectively.

2.4. Circular dichroism (CD) spectroscopy: The far UV-CD spectrum of nsp11 protein was recorded in J-1500 spectrophotometer (Jasco) at 25ºC. We used 1 mm optical path length quartz cuvette to record spectra in various concentrations of SDS, LUVs, TFE, and varying temperature condition. The scan range spanned 190–240 nm, with a 0.5 nm data interval. The 0.1 mm optical path length cuvette was used to record spectra in case of TMAO. 20 mM phosphate buffer (pH 7.4, 50 mM NaCl), was used to prepare the nsp11 sample (100 μM) for spectral scan. 500 μM of nsp11 were used in case of TMAO. The CD spectrum was recorded as an average of 3 consecutive scans. Each spectrum was fitted by Savitsky-Golay smoothing (15 points of smoothing window and second polynomial order). The secondary structure contents (α-helix, β-sheet, disordered, and turn) of the samples were calculated by deconvolution of CD spectroscopic data in DICHROWEB webserver (CONTIN algorithm with a reference set 7) [21].

2.5. Structure modelling and molecular dynamics (MD) simulation: The long timescale computer simulations have been very useful to comprehend the atomic-level dynamics of a peptide like nsp11 of SARS-CoV-2. There is very less information available about its intrinsic properties. Thus, to gain atomic insight on dynamics of nsp11, we have performed all-atom molecular dynamics simulations and successfully correlated with experimental observations. Firstly, a 3D model was constructed for the 13 amino acids long peptide using PEP-FOLD peptide structure prediction server [22]. It applies a coarse-grained (CG) forcefield and performs up to 200ns simulation runs to build an energy minimized structure. The detailed methodology has been given in our previous reports [23,24]. The resultant model was then prepared using Chimera by addition of missing hydrogens and proper parameterization of asymmetrical residues. We used Gromacs v5, where simulation setup was built by placing the protein structure in a cubic box along with SPC water model, 0.15M

NaCl salt concentration. After solvation, the system was charge neutralized with counterions. To attain an energy minimized simulation system, the steepest descent method was used until the system was converged within 1000 kJ/mol. Further, the equilibration of system was done to optimize solvent in the environment. Using NVT and NPT ensembles within periodic boundary conditions for 100ps each, the system was equilibrated. The average temperature at 300K and pressure at 1 bar were maintained using Nose-Hoover and Parrinello-Rahman coupling methods during simulation. All bond-related constraints were solved using SHAKE algorithm. The final production run was performed for 500ns in our high performing cluster at IIT Mandi. After analyzing the trajectory, the last frame was chosen for further studies in different conditions. Next, the structural conformations of NSP11 were analyzed in two different conditions of solvents: TFE (8M) and SDS (60 molecules) mixed with water in separate simulation runs. By using above described forcefield parameters and addition of TFE and SDS molecules into the system, the simulations were performed for 200ns and 500ns for TFE and SDS based systems, respectively.

Further, Replica Exchange MD was performed for 100 ns of last frame using our previous protocol in Desmond simulation package [25]. Here, we have chosen 8 temperatures for replica-exchange based on experiments. The temperatures were calculated by linear mode in Desmond in 8 replicas viz. 298K (25°C), 307.29K (34.29°C), 316.57K (43.57°C), 325.86K (52.86°C), 335.14K (62.14°C), 344.43K (71.43°C), 353.71K (80.71°C), and 363K (90°C). All trajectory analysis, calculations were performed using Chimera, Maestro and Gromacs commands for calculating root mean square deviation (RMSD), root mean square fluctuation (RMSF), and radius of gyration (Rg) for protein structure compactness for C-α atoms.

3. Result and Discussion:

3.1. SARS-CoV-2 nsp11 is intrinsically disordered under physiological buffer conditions:

The Far-UV CD spectrum of nsp11 protein shows a signature minima at 200 nm (**Figure 1A**), suggest the spectrum of typical disordered type protein.

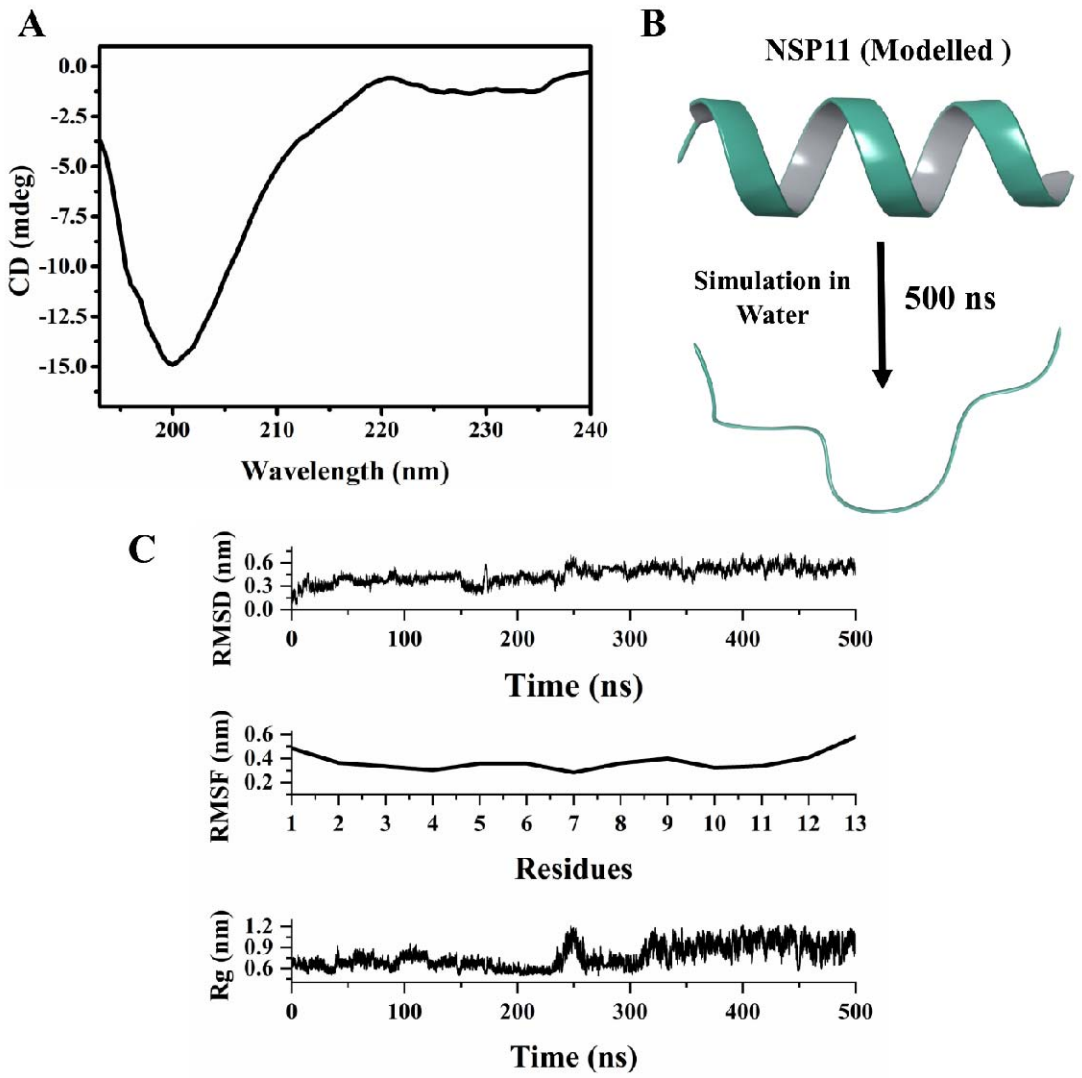

Figure 1: Secondary structure analysis and MD simulation of nsp11. (A) Far-UV CD spectra of nsp11 represents characteristic disordered spectra (B) Structure of nsp11 obtained from molecular modelling in PEP-FOLD server and further MD simulation performed at 500 ns in water. (C) RMSD, RMSF, and Rg analysis of nsp11 of MD trajectory.

The structural insight into the nsp11 were further investigated through all-atom MD simulations up to 500 ns. The structure of nsp11 was entirely in helical conformation after modelling by PEP-FOLD webserver. After preparing the structure for its proper symmetry and hydrogen addition, the structure was simulated in presence of SPC water model for 500 ns. As observed in simulations, the structure loses its helicity during simulation period (**Figure 1B**). In terms of mean fluctuations and deviations at the atomic scale, high values for RMSD, RMSF and Rg were detected (**Figure 1C**). After 250 ns, the RMSD values were increased from approximately 0.25 nm to 0.6 nm and remained stable till 500 ns. Similarly, the Rg increased 0.65 nm to 1.1 nm in radius of gyration that suggest lower compactness as

compared to the initial structure. Also, RMSF is showing higher fluctuation at the terminal residues between 0.4 to 0.6 nm as compared to the middle segment. Higher fluctuation and deviation may be due to presence of disordered structure in nsp11. The disordered proteins are highly dynamic in nature where the simulation and CD spectra revealed nsp11 as disordered protein. However, the prediction and homology modelling suggest that the nsp11 also has some helical structure indicating that it has an intrinsic ability to form helix when environmental condition demands.

3.2. Disordered nsp11 shows upsurge in α-helical content in TFE

The lipids/membrane mimetics, natural osmolytes, and organic solvents (α-helix inducer) are generally used to explore the folding propensities of disordered proteins [26]. As nsp11 is a disordered system it is wise to use TFE, SDS, Lipids, and osmolytes to investigate the fold propensity and structural transition from disordered to ordered one. The TFE is well-characterised to initiate folding and induce secondary structures in protein chains. It indirectly stabilizes the intra H-bonds of an α-helix by destabilizing or weakening the H-bonds between CO and NH groups of protein backbone with water [27]. With increasing TFE concentration it is evident that the negative ellipticity at 200 nm become lesser and negative ellipticities at 208 nm and 222 nm become towards higher value (**Figure 2A, B**). The change in the ellipticity at the corresponding wavelength indicate that nsp11 acquires an α-helical structure in presence of TFE. Additionally, the transition from disorder to order structure is accompanied by formation of an intermediate structure near 20 % of TFE (Figure 2B) representing a partial folded conformation. Simulations in presence of TFE produces remarkable change in structure of nsp11 which is consistent with the result obtained from Far-UV CD. A disorder to order (α-helix) transition is seen at residues $_{4}$AQSFLNGF$_{11}$ (**Figure 2C**) which shows stable RMSF. But both the terminals are showing higher fluctuation due to disordered terminal. RMSD and Rg shows lesser fluctuation due to higher structured content (**Figure 2D**). Further, to know α-helix promoting regions of the nsp11, we performed sequence-based secondary structure prediction by SSpro3 (from SCRATCH protein predictor) [28] and PEP2D [29] servers (**Figure 2E**). SSpro3 server predicted residues 4th -8th as α-helix (%) and residues 1st -3rd and 9th -13th as random coil (%) whereas PEP2D server predicted residues 1-4th and 10-13th as random coil (61.54 %) and residues 5-9th as α-helix (38.46 %) (**Figure 2E)**. This is in correlation with MD results. The structural changes in TFE may arise due to disordered proteins lacking a stable three-dimensional conformation at physiological condition and easily adopts a secondary structure in the presence of favourable environment or suitable binding partner [30].

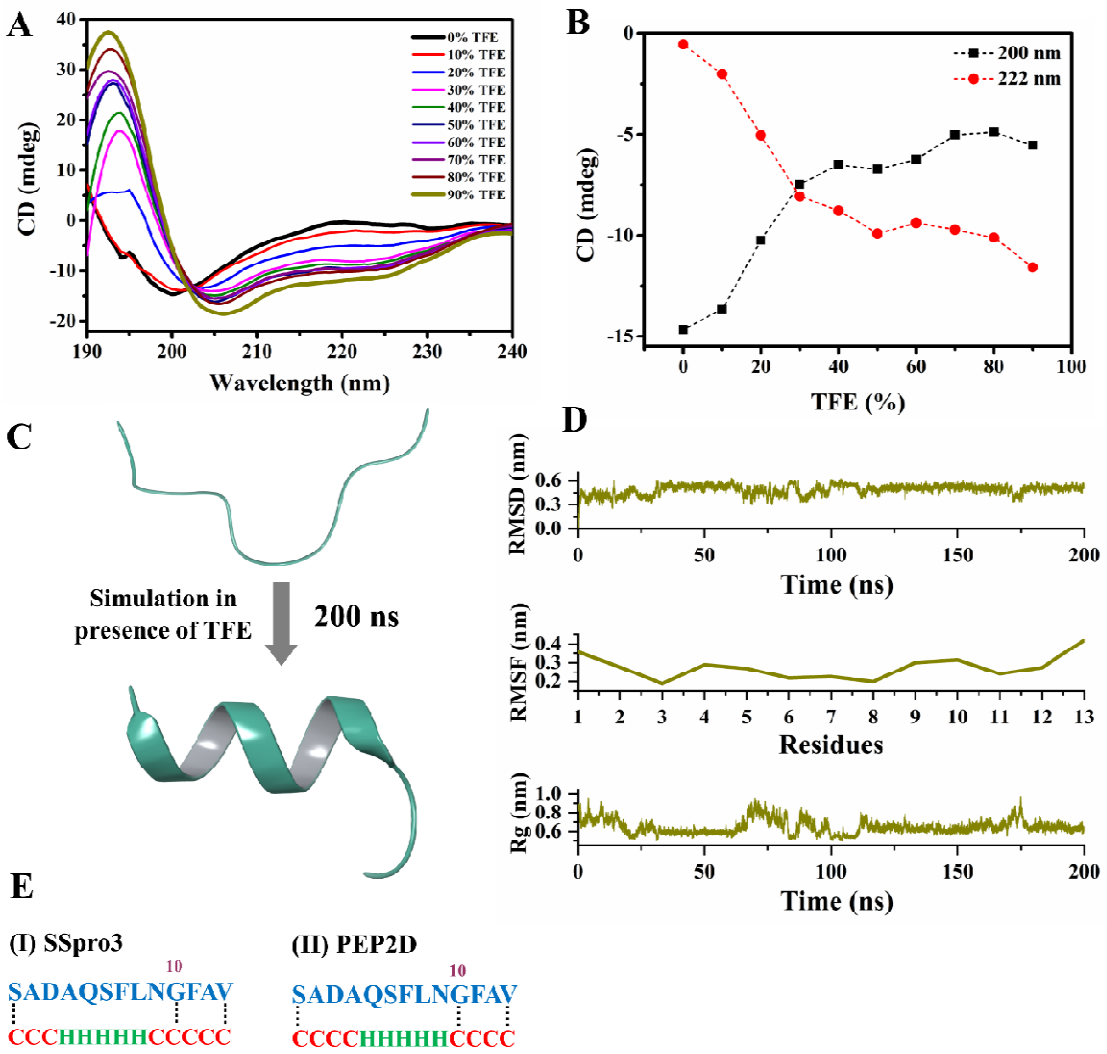

Figure 2: Secondary structure analysis of nsp11 in TFE. (A) Effect of TFE (0-90%) concentration on conformation of disordered nsp11 protein monitored by Far-UV CD. Higher TFE concentration induces α- helix in nsp11. Plot (B) signifies the change of ellipticity at 200 nm (black squares with dashed line) and 222 nm (red circles with dashed line) as the function of an increasing TFE concentration. (C) Disordered to ordered transition of nsp11 in presence of TFE at 200 ns. 3D structure of nsp11 in presence of TFE at 200 ns showing helical structure in the middle segment ($_{4}$AQSFLNGF$_{11}$). (D) The fluctuating trend in RMSD, RMSF, Rg depicts the occurring changes for nsp11 in TFE. (E) Secondary structure perdition for nsp11 using (I) SSPro3 and (II) PEP2D (H-Helix (green) and C- Coil (red)) to predict helix promoting region in nsp11.

3.3. nsp11 adopts α-helix structure in SDS but remain unaffected in negatively charged and neutral lipids

The membrane inside cells are very dynamic and therefore, artificially prepared models can be used to study protein-membrane or protein-protein interactions [31]. Here, LUVs with neutral DOPC and anionic DOPS were prepared and used for CD experiments. Several IDPs are reported to bind efficiently to the DOPC and DOPS lipids and this interaction is accompanied by enhanced level of their α-helical structure [23,26,32]. In case of nsp11, the presence of DOPS and DOPC does not show any change in secondary structure conformation (**Figure 3E and 3F**). The nsp11 has one negatively charged residue and does not contain positively charged residues that may have role in its interaction with lipids. Analogous to lipid membranes, the SDS micelles used to mimic the interface between hydrophobic and hydrophilic environment such as plasma membrane and the cytosol [33]. SDS is a denaturing detergent with folding inducing properties that forms micelles (CMC is 1.99 mM in 50 mM phosphate buffer [34]) and interacts with hydrophobic parts of proteins [35]. At lower SDS concentration (below CMC) where SDS is in the monomeric forms do not shows any structural change and the shape of the spectra remains similar as of nsp11 in the absence of SDS. Whereas at 5mM (above CMC) the shape of spectrum changes and exhibits a shift in negative ellipticity from 200 nm to 203 nm. At higher concentration (25, 50, and 100 mM) the CD spectra of nsp11 exhibited two minima at 206 and 222 nm suggesting the structural transition from the disordered to ordered conformation (**Figure 3A, B**). Further, CD deconvolution confirms the reduction in disordered structure content and increase in ordered structure (α-helix) content (**Table 1**) in nsp11 above CMC.

Table 1: Estimation of secondary structure content of nsp11 in presence SDS, TFE, and varying temperature from DICHROWEB CD analysis.

		Secondary structure content (%)			
		α-helix	**β-sheet**	**Turn**	**Disordered**
SDS	0 mM	8.85%	11.75%	12.4%	66.95%
	0.5 mM	8.25%	14.95%	12.05%	64.75%
	100 mM	98.85%	1.15%	0%	0%
TFE	10%	8.05%	18.95%	11.75%	61.25%
	50%	87.4%	12.55%	0%	0%
	90%	98.65%	1.35%	0%	0%

Like CD spectroscopy, MD simulations of nsp11 in presence of SDS molecules is also showing some noticeable change in structure as shown in **Figure 3C**. A small proportion of helix was formed by four residues $_4$AQSF$_7$ at the last frame (500 ns) simulation which contributes to nearly 30 % of ordered structure. A little fluctuation was observed in case of

RMSD and Rg may be due to disordered terminal region which shows high RMSF throughout the simulation (**Figure 3D**). The RMSF of residues at the middle segment is stable that suggest the gain of ordered structure during simulation in SDS. Collectively, in-depth insight of structural transition from disordered to ordered one suggest the possible interaction of nsp11 with membrane mimetics which may explains the host specificity to membrane disruption implicated in the viral pathogenesis.

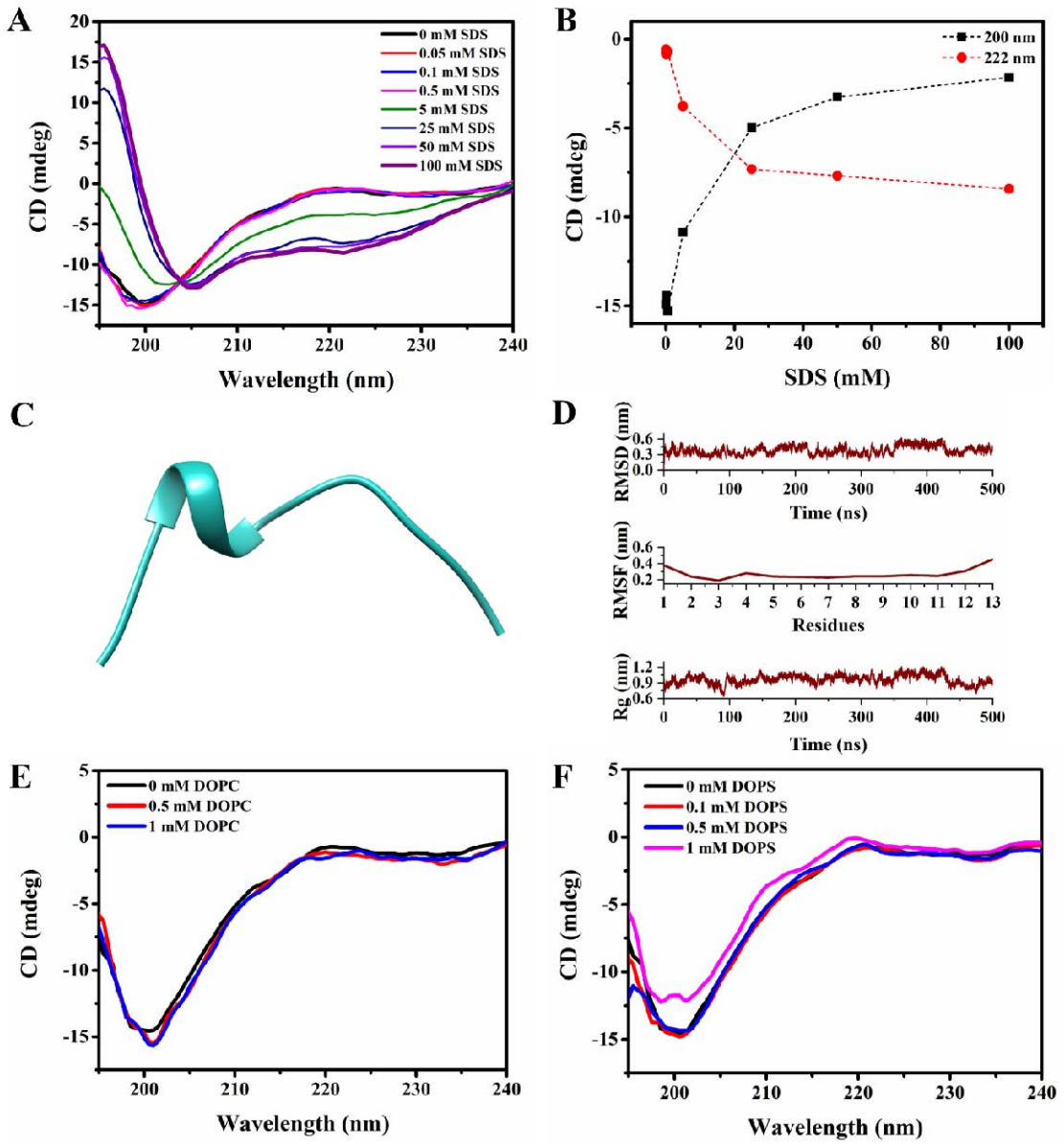

Figure 3: Effect of Membrane-Mimetic Environment on nsp11 secondary structure. (A) Far-UV CD spectra of nsp11 in presence of SDS. Increasing concentration of SDS induces helical content in nsp11. (B) representation of ellipticity at 200 nm (black squares with dashed line) and 222 nm (red circles with dashed line) as the function of SDS concentration. (C) MD simulation trajectory of nsp11

in presence of SDS. The last frame (at 500 ns) is showing a little propensity of helix ($_4$AQSF$_7$). (D) RMSD, RMSF, and Rg analysis of nsp11 in presence of SDS molecules. **(E)** Far-UV CD spectra of nsp11 measured in presence of neutral lipid (DOPC) and (F) anionic lipid DOPS. Both lipids are unable to induce any structural change in nsp11.

3.4. Osmolytes show no effect on secondary structure of nsp11

In the cell, various classes of organic osmolytes are present and may have several roles therein [26,36]. Thus, natural osmolytes such as TMAO were used on structural properties of nsp11 protein. The osmophobic effect of TMAO force thermodynamically unstable proteins or disordered proteins to fold and regain high functional activity [26]. In disordered state, the peptide backbone is largely exposed to its surrounding, as an osmolyte, TMAO exerts a solvophobic thermodynamic force which raises the free energy of disordered state preferentially shifting the equilibrium towards the folded form. This osmophobic effect drives the folding of natively unstructured proteins [37,38]. In the case of nsp11, TMAO is not exerting any effect on its secondary structure. The shape of the far-UV CD spectra remains same irrespective of increased TMAO concentration (up to 2 M). Only a slight shift in signature minimum was observed (shift from 200nm to 203 nm) at 3.25 M TMAO but the shape of the spectra retains its disordered type protein (**Figure 4**).

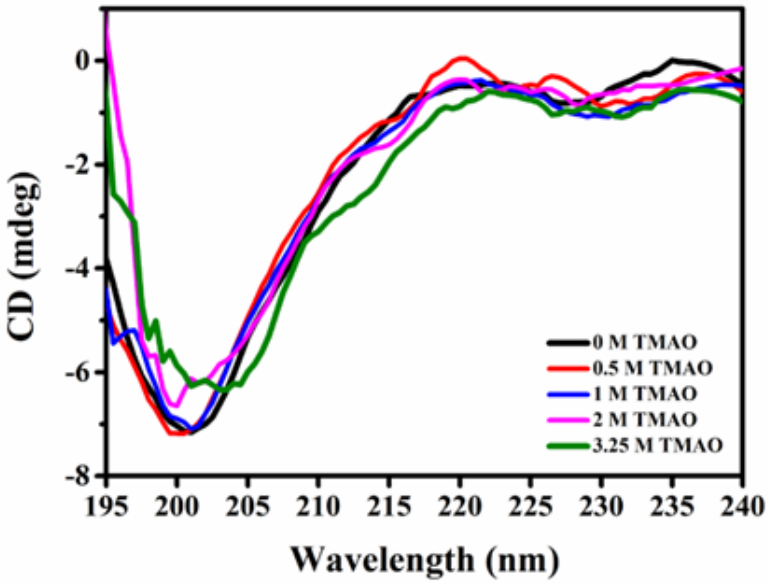

Figure 4: Far-UV CD spectra of nsp11 in presence of TMAO. Irrespective of the TMAO concentration the shape of CD spectra represents the typical disordered type protein suggesting TMAO unable to induce secondary structure change in nsp11.

3.5. Temperature-induced structural change in nsp11 appeared by contraction in structure:

Further to investigate the temperature-induced conformational changes in nsp11 we acquired CD spectra at 10 ºC to 90 ºC with 5 ºC interval as shown in **Figure 5A**. At lower temperatures, the far-UV CD spectrum is typical of an unfolded protein. As the temperature rises, the signature minima remain nearly at 200 nm, but the shape of the spectrum shows some changes at higher temperature. With increasing temperature, the ellipticity (negative) corresponding to signature minima at 200 nm becomes lower and higher at 222 nm (**Figure 5B**).

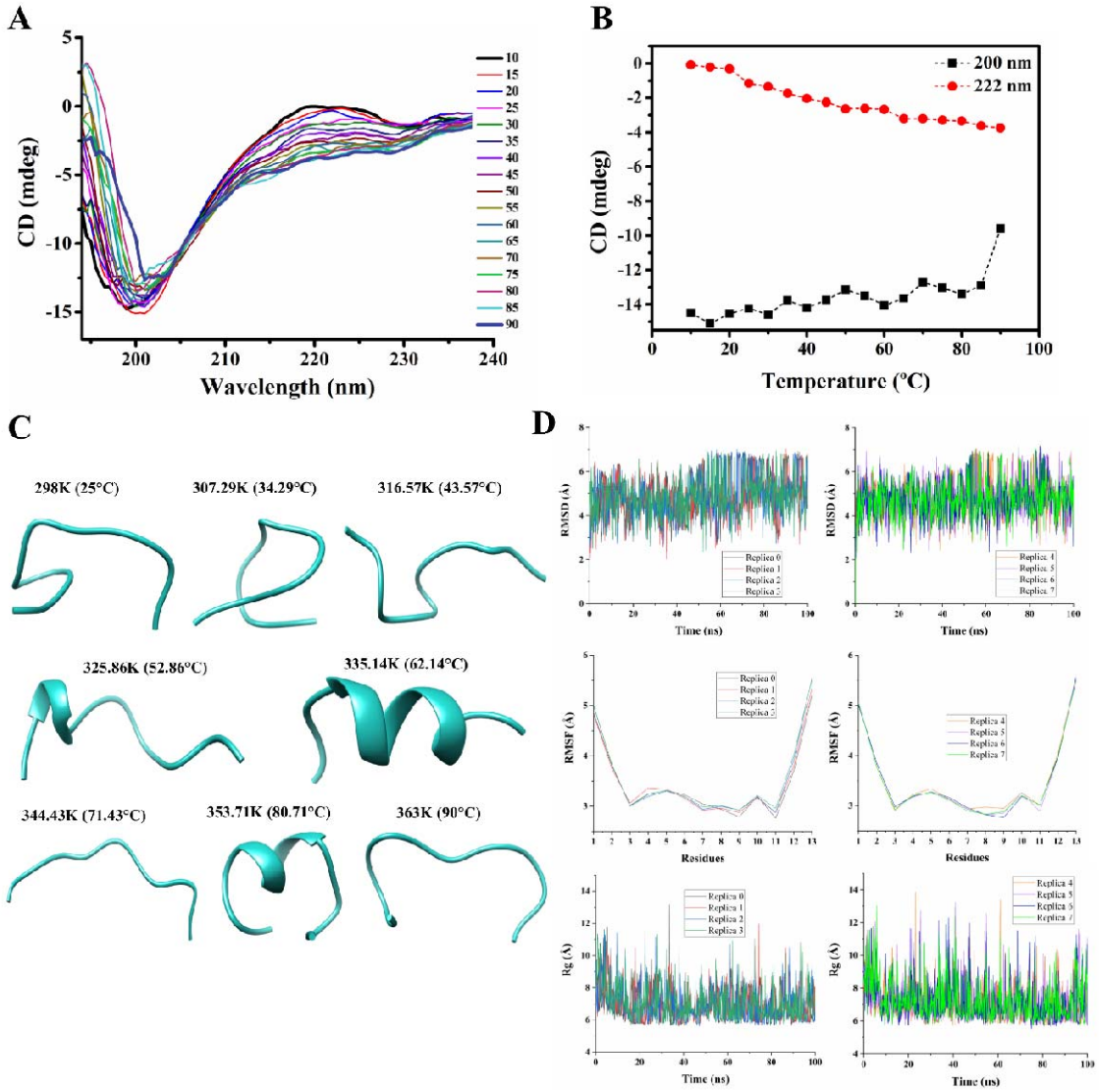

Figure 5: Effect of temperature on secondary structure of nsp11 protein. (A) Thermal stability of nsp11 secondary structure monitored by far-UV CD spectra at temperatures range from 10°C to 90°C with 5°C interval. (B) Analysis of CD spectra obtained in fig A. black line represents change of

ellipticity at 198 nm and red line represents change of ellipticity at 222 nm. (C) Last frame of each replica (0-7) of REMD simulation with increasing temperature. (D) RMSD, RMSF, and Rg analysis of each trajectory.

The change in the shape of the spectra at higher temperature, indicating temperature-induced partial folding or partial formation of ordered secondary structure. This type of change is in line with the other disordered proteins [23,32,39]. The structural changes in IDPs at higher temperature is the common phenomenon which attended by the contraction of the conformational ensemble [40]. The contraction of structure occur in nsp11 may be due to change in secondary or tertiary structure [40]. Moreover, reports suggested that the folding in the disordered proteins at higher temperatures may occur due to enhancement in the strength of hydrophobic interaction by elevated temperature [26]. The hydrophobicity value of a nsp11 protein were calculated according to ProtParam [41] analysis in shows grand average of hydropathicity (GRAVY) score is 0.500 represents its hydrophobic nature (increasing positive score indicates a greater hydrophobicity). However, the Nettels *et al* reported that that the contraction of disordered proteins is independent of the hydrophobicity of the protein, and therefore may not be due to enhanced hydrophobic interactions [42]. Some reports also suggested that the structural changes at higher temperature are independent of long-range interactions and are occurred with the contraction of the conformational ensemble [40].

Further to gain insights into temperature-induced contraction in nsp11 we employed MD simulation where temperature-based replica exchange MD were performed to see the conformational change. Total 8 replicas were run and allow them to exchange conformations during MD simulation (**Figure 5C**). The structural change is not consistent during the simulation where nsp11 exists mostly in disordered state. But at higher temperature (316.57K, 325.86K and 353.71K) nsp11 gains helical structure. The helical structure gain is restricted mostly at the middle segment of the protein whereas both terminal residues remain disordered. It is also evident from RMSF where residues fluctuation for 3rd to 11th is stable but the fluctuation is very high for residues 1-2 and 12-13 (**Figure 5D**). Additionally, more fluctuating peak of RMSD and Rg (**Figure 5D**) is suggesting that the nsp11 is mostly disordered irrespective of temperature rise. Overall, increasing temperature endorsed substantial changes in the shape of CD spectra that suggest a partial change in secondary structure conformation and this change may be due to partial helical structure gain at higher temperatures revealed by MD simulation.

Conclusion:

This study accentuates the structural conformation of SARS-CoV-2 nsp11 protein. Our study signifies that the nsp11 is an intrinsically disordered protein. The model conditions mimicking the effect of the membrane field (TFE) is effective to induce helical structure in nsp11. It shows the high specificity towards the SDS micelles, compared to neutral and anionic lipids, and readily undergo a conformational transition (disorder to order) in the presence SDS micelle. Importantly, the consequent structural transformations at different membrane environment could be dependent on the type of solvent used. Thus, our study emphasizes that the nsp11 may play functional role in the host cytosolic membrane affinity/interaction.

Future perspectives:

In the wake of ongoing pandemic, understanding the biology and structure-function paradigm is essential to understand the pathogenesis of the virus. Especially, the identification of role of such proteins which results into important events like ribosomal frameshift needs to be identified. In last two decades, it has been established that proteins with no proper three-dimensional structure has great amount of significance in most cellular processes. Herein, the identified intrinsic disordered nature of the SARS-CoV-2 nsp11 protein will surely help to investigate its role in the ribosomal frameshifting which can give some promising targets for identifying contagious nature of SARS-CoV-2 virus.

Author Contributions: RG and NG: Conception, design, review, and writing of the manuscript. KG and AK performed experiments and PK performed the computational study. KG, AK, PK, TB analyzed data and wrote the manuscript. All authors have read and approved the manuscript.

Acknowledgments: Authors are thankful to Indian Institute of Technology Mandi (BioX AMRC center, and HPC facility) for all the facilities. KG is thankful to Science and Engineering Research Board (SERB), India (Grant Number: CRG/2019/005603) for fellowship. AK were supported by fellowship from Ministry of Human Resource Development (MHRD). TB is grateful to Department of Science and Technology (DST) for Inspire Fellowship. RG is thankful to IYBA Award (Grant Number: BT/11/IYBA/2018/06) by Department of Biotechnology (DBT), India.

Conflict of Interest: The authors declare that there is no conflict of interest.

References:

[1] A.A.T. Naqvi, K. Fatima, T. Mohammad, U. Fatima, I.K. Singh, A. Singh, S.M. Atif, G. Hariprasad, G.M. Hasan, M.I. Hassan, Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach, Biochim. Biophys. Acta - Mol. Basis Dis. 1866 (2020). doi:10.1016/j.bbadis.2020.165878.

[2] I. Astuti, Ysrafil, Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2): An overview of viral structure and host response, Diabetes Metab. Syndr. Clin. Res. Rev. 14 (2020) 407–412. doi:10.1016/j.dsx.2020.04.020.

[3] S. Kumar, R. Nyodu, V.K. Maurya, S.K. Saxena, Morphology, Genome Organization, Replication, and Pathogenesis of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), in: Coronavirus Dis. 2019, Nature Publishing Group, 2020: pp. 23–31. doi:10.1007/978-981-15-4814-7_3.

[4] A. Wu, Y. Peng, B. Huang, X. Ding, X. Wang, P. Niu, J. Meng, Z. Zhu, Z. Zhang, J. Wang, J. Sheng, L. Quan, Z. Xia, W. Tan, G. Cheng, T. Jiang, Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China, Cell Host Microbe. 27 (2020) 325–328. doi:10.1016/j.chom.2020.02.001.

[5] Y. Chen, Q. Liu, D. Guo, Emerging coronaviruses: Genome structure, replication, and pathogenesis, J. Med. Virol. 92 (2020) 418–423. doi:10.1002/jmv.25681.

[6] S.G. Fang, H. Shen, J. Wang, F.P.L. Tay, D.X. Liu, Proteolytic processing of polyproteins 1a and 1ab between non-structural proteins 10 and 11/12 of Coronavirus infectious bronchitis virus is dispensable for viral replication in cultured cells, Virology. 379 (2008) 175–180. doi:10.1016/j.virol.2008.06.038.

[7] E.J. Snijder, E. Decroly, J. Ziebuhr, The Nonstructural Proteins Directing Coronavirus RNA Synthesis and Processing, in: Adv. Virus Res., Academic Press Inc., 2016: pp. 59–126. doi:10.1016/bs.aivir.2016.08.008.

[8] J.F.W. Chan, K.H. Kok, Z. Zhu, H. Chu, K.K.W. To, S. Yuan, K.Y. Yuen, Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan, Emerg. Microbes Infect. 9 (2020) 221–236. doi:10.1080/22221751.2020.1719902.

[9] R. Giri, T. Bhardwaj, M. Shegane, B.R. Gehi, P. Kumar, K. Gadhave, C.J. Oldfield, V.N. Uversky, Understanding COVID-19 via comparative analysis of dark proteomes of SARS-CoV-2, human SARS and bat SARS-like coronaviruses, Cell. Mol. Life Sci. (2020). doi:10.1007/s00018-020-03603-x.

[10] D.J. Deming, R.L. Graham, M.R. Denison, R.S. Baric, Processing of Open Reading Frame 1a Replicase Proteins nsp7 to nsp10 in Murine Hepatitis Virus Strain A59 Replication, J. Virol. 81 (2007) 10280–10291. doi:10.1128/jvi.00017-07.

[11] Y. Sun, H. Ke, M. Han, N. Chen, W. Fang, D. Yoo, Nonstructural protein 11 of porcine reproductive and respiratory syndrome virus suppresses both MAVS and RIG-I expression as

one of the mechanisms to antagonize Type I interferon production, PLoS One. 11 (2016). doi:10.1371/journal.pone.0168314.

[12] V.N. Uversky, J.R. Gillespie, A.L. Fink, Why are “natively unfolded” proteins unstructured under physiologic conditions?, Proteins Struct. Funct. Genet. 41 (2000) 415–427. doi:10.1002/1097-0134(20001115)41:3<415::AID-PROT130>3.0.CO;2-7.

[13] R. Van Der Lee, M. Buljan, B. Lang, R.J. Weatheritt, G.W. Daughdrill, A.K. Dunker, M. Fuxreiter, J. Gough, J. Gsponer, D.T. Jones, P.M. Kim, R.W. Kriwacki, C.J. Oldfield, R. V. Pappu, P. Tompa, V.N. Uversky, P.E. Wright, M.M. Babu, Classification of intrinsically disordered regions and proteins, Chem. Rev. 114 (2014) 6589–6631. doi:10.1021/cr400525m.

[14] V. Vacic, C.J. Oldfield, A. Mohan, P. Radivojac, M.S. Cortese, V.N. Uversky, A.K. Dunker, Characterization of molecular recognition features, MoRFs, and their binding partners, J. Proteome Res. 6 (2007) 2351–2366. doi:10.1021/pr0701411.

[15] C.J. Oldfield, Y. Cheng, M.S. Cortese, P. Romero, V.N. Uversky, A.K. Dunker, Coupled folding and binding with α-helix-forming molecular recognition elements, Biochemistry. 44 (2005) 12454–12470. doi:10.1021/bi050736e.

[16] C.J. Oldfield, A.K. Dunker, Intrinsically disordered proteins and intrinsically disordered protein regions, Annu. Rev. Biochem. 83 (2014) 553–584. doi:10.1146/annurev-biochem-072711-164947.

[17] P. Kulkarni, V.N. Uversky, Intrinsically Disordered Proteins: The Dark Horse of the Dark Proteome, Proteomics. 18 (2018). doi:10.1002/pmic.201800061.

[18] V.N. Uversky, Intrinsically disordered proteins and their “Mysterious” (meta)physics, Front. Phys. 7 (2019) 10. doi:10.3389/fphy.2019.00010.

[19] R. Giri, D. Kumar, N. Sharma, V.N. Uversky, Intrinsically disordered side of the Zika virus proteome, Front. Cell. Infect. Microbiol. 6 (2016). doi:10.3389/fcimb.2016.00144.

[20] D. Kumar, A. Singh, P. Kumar, V.N. Uversky, C.D. Rao, R. Giri, Understanding the penetrance of intrinsic protein disorder in rotavirus proteome, Int. J. Biol. Macromol. 144 (2020) 892–908. doi:10.1016/j.ijbiomac.2019.09.166.

[21] L. Whitmore, B.A. Wallace, DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data, Nucleic Acids Res. 32 (2004) W668. doi:10.1093/nar/gkh371.

[22] P. Thévenet, Y. Shen, J. Maupetit, F. Guyon, P. Derreumaux, P. Tufféry, PEP-FOLD: An updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides, Nucleic Acids Res. 40 (2012). doi:10.1093/nar/gks419.

[23] K.U. Saumya, D. Kumar, P. Kumar, R. Giri, Unlike dengue virus, the conserved 14–23 residues in N-terminal region of Zika virus capsid is not involved in lipid interactions, Biochim. Biophys. Acta - Biomembr. 1862 (2020) 183440. doi:10.1016/j.bbamem.2020.183440.

[24] A. Kumar, A. Kumar, P. Kumar, N. Garg, R. Giri, SARS-CoV-2 NSP1 C-terminal region (residues 130-180) is an intrinsically disordered region, BioRxiv. (2020) 2020.09.10.290932. doi:10.1101/2020.09.10.290932.

[25] A. Kumar, P. Kumar, S. Kumari, V.N. Uversky, R. Giri, Folding and structural polymorphism of p53 C-terminal domain: One peptide with many conformations, Arch. Biochem. Biophys. 684 (2020) 108342. doi:10.1016/j.abb.2020.108342.

[26] V.N. Uversky, Intrinsically disordered proteins and their environment: Effects of strong denaturants, temperature, pH, Counter ions, membranes, binding partners, osmolytes, and macromolecular crowding, Protein J. 28 (2009) 305–325. doi:10.1007/s10930-009-9201-4.

[27] P. Luo, R.L. Baldwin, Mechanism of helix induction by trifluoroethanol: A framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water, Biochemistry. 36 (1997) 8413–8421. doi:10.1021/bi9707133.

[28] C.N. Magnan, P. Baldi, SSpro/ACCpro 5: Almost perfect prediction of protein secondary structure and relative solvent accessibility using profiles, machine learning and structural similarity, Bioinformatics. 30 (2014) 2592–2597. doi:10.1093/bioinformatics/btu352.

[29] H. Singh, S. Singh, G.P. Singh Raghava, Peptide Secondary Structure Prediction using Evolutionary Information, BioRxiv. (2019) 558791. doi:10.1101/558791.

[30] R. Van Der Lee, M. Buljan, B. Lang, R.J. Weatheritt, G.W. Daughdrill, A.K. Dunker, M. Fuxreiter, J. Gough, J. Gsponer, D.T. Jones, P.M. Kim, R.W. Kriwacki, C.J. Oldfield, R. V. Pappu, P. Tompa, V.N. Uversky, P.E. Wright, M.M. Babu, Classification of intrinsically disordered regions and proteins, Chem. Rev. 114 (2014) 6589–6631. doi:10.1021/cr400525m.

[31] C.T. Höfer, S. Di Lella, I. Dahmani, N. Jungnick, N. Bordag, S. Bobone, Q. Huang, S. Keller, A. Herrmann, S. Chiantia, Structural determinants of the interaction between influenza A virus matrix protein M1 and lipid membranes, Biochim. Biophys. Acta - Biomembr. 1861 (2019) 1123–1134. doi:10.1016/j.bbamem.2019.03.013.

[32] A. Kumar, P. Kumar, R. Giri, Zika virus NS4A cytosolic region (residues 1–48) is an intrinsically disordered domain and folds upon binding to lipids, Virology. 550 (2020) 27–36. doi:10.1016/j.virol.2020.07.017.

[33] M. Popovic, A. De Biasio, A. Pintar, S. Pongor, The intracellular region of the Notch ligand Jagged-1 gains partial structure upon binding to synthetic membranes, FEBS J. 274 (2007) 5325–5336. doi:10.1111/j.1742-4658.2007.06053.x.

[34] E. Fuguet, C. Ràfols, M. Rosés, E. Bosch, Critical micelle concentration of surfactants in aqueous buffered and unbuffered systems, Anal. Chim. Acta. 548 (2005) 95–100. doi:10.1016/j.aca.2005.05.069.

[35] D. V. Tulumello, C.M. Deber, SDS micelles as a membrane-mimetic environment for transmembrane segments, Biochemistry. 48 (2009) 12096–12103. doi:10.1021/bi9013819.

[36] M.B. Burg, J.D. Ferraris, Intracellular organic osmolytes: Function and regulation, J. Biol.

Chem. 283 (2008) 7309–7313. doi:10.1074/jbc.R700042200.

[37] D.W. Bolen, I. V. Baskakov, The osmophobic effect: Natural selection of a thermodynamic force in protein folding, J. Mol. Biol. 310 (2001) 955–963. doi:10.1006/jmbi.2001.4819.

[38] I. Baskakov, D.W. Bolen, Forcing thermodynamically unfolded proteins to fold, J. Biol. Chem. 273 (1998) 4831–4834. doi:10.1074/jbc.273.9.4831.

[39] D. Kumar, P.M. Mishra, K. Gadhave, R. Giri, Conformational dynamics of p53 N-terminal TAD2 region under different solvent conditions, Arch. Biochem. Biophys. 689 (2020) 108459. doi:10.1016/j.abb.2020.108459.

[40] M. Kjaergaard, A.B. Nørholm, R. Hendus-Altenburger, S.F. Pedersen, F.M. Poulsen, B.B. Kragelund, Temperature-dependent structural changes in intrinsically disordered proteins: Formation of α-helices or loss of polyproline II?, Protein Sci. 19 (2010) 1555–1564. doi:10.1002/pro.435.

[41] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, A. Bairoch, Protein Identification and Analysis Tools on the ExPASy Server, in: Proteomics Protoc. Handb., Humana Press, 2005: pp. 571–607. doi:10.1385/1-59259-890-0:571.

[42] D. Nettels, S. Müller-Späth, F. Küster, H. Hofmann, D. Haenni, S. Rüegger, L. Reymond, A. Hoffmann, J. Kubelka, B. Heinz, K. Gast, R.B. Best, B. Schuler, Single-molecule spectroscopy of the temperature-induced collapse of unfolded proteins, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 20740–20745. doi:10.1073/pnas.0900622106.