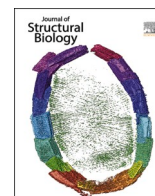




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Spike protein fusion loop controls SARS-CoV-2 fusogenicity and infectivity

Debnath Pal

Indian Institute of Science, Bengaluru 560012, India

ARTICLE INFO

Edited by Bauke Dijkstra

Keywords:

Coronavirus
COVID-19
MHV-A59
Two consecutive prolines
Structural rigidity

ABSTRACT

The high SARS-CoV-2 reproductive number driving the COVID-19 pandemic has been a mystery. Our recent *in vitro*, and *in vivo* coronaviral pathogenesis studies involving Mouse Hepatitis Virus (MHV-A59) suggest a crucial role for a small host membrane-virus contact initiator region of the Spike protein, called the fusion peptide that enhances the virus fusogenicity and infectivity. Here I study the Spike from five human β -coronaviruses (HCoV) including the SARS-CoV-2, and MHV-A59 for comparison. The structural and dynamics analyses of the Spike show that its fusion loop spatially organizes three fusion peptides contiguous to each other to synergistically trigger the virus-host membrane fusion process. I propose a Contact Initiation Model based on the architecture of the Spike quaternary structure that explains the obligatory participation of the fusion loop in the initiation of the host membrane contact for the virus fusion process. Among all the HCoV Spikes in this study, SARS-CoV-2 has the most hydrophobic surface and the extent of hydrophobicity correlates with the reproductive number and infectivity of the other HCoV. Comparison between results from standard and replica exchange molecular dynamics reveal the unique physicochemical properties of the SARS-CoV-2 fusion peptides, accrued in part from the presence of consecutive prolines that impart backbone rigidity which aids the virus fusogenicity. The priming of the Spike by its cleavage and subsequent fusogenic conformational transition steered by the fusion loop may be critical for the SARS-CoV-2 spread. The importance of the fusion loop makes it an apt target for anti-virals and vaccine candidates.

1. Introduction

The novel Coronavirus Disease 2019 (COVID-19) caused by SARS-CoV-2, Severe Acute Respiratory Syndrome (SARS) by SARS-CoV, and Middle Eastern Respiratory Syndrome (MERS) by MERS-CoV, induce severe acute respiratory distress in patients. Though these diseases share similar clinical and pathological features, COVID-19 differs in overlapping yet distinct phases of infection (Shi et al., 2020; Zheng, 2020). The degree of infectivity is significantly high in SARS-CoV-2, and far more aggressive, as evidenced by the current global pandemic. This can be quantified by the preliminary reproductive number (R_0) of COVID-19 (2.0–2.5), which is higher than the R_0 of SARS (1.7–1.9) and far higher than that of MERS (<1) (Petrossillo et al., 2020). The significant difference in R_0 may accrue due to environmental, immunological, or molecular reasons. COVID-19 transmission has been attributed to the long life of SARS-CoV-2 outside the host as it increases the chances of infection through cross-contamination by contact in the population (van Doremalen et al., 2020). The large distance distribution of the SARS-CoV-2 particles from the infected person due to activities like sneezing and coughing (Morawska and Cao, 2020), and the tiny size of the virus

droplets may be more efficient in penetrating deeply into the pulmonary system to allow the rapid spread of the disease (Pedersen and Ho, 2020). However, the SARS-CoV also has high genomic similarity with the SARS-CoV-2, and one would have expected it to have similar transmission behavior and R_0 , which is evidently not the case. For that matter, the environmental spread of other viruses should have been far more widespread than coronaviruses, given that the coronaviruses have the largest RNA viral genomes and therefore the largest particle size and consequently higher aerosol size compared to many other viruses. This is again not that we observe in practice. Another possibility of high viral spread may accrue from intense viral shedding, where SARS-CoV-2 has succeeded early on in rapid viral replication and cell-to-cell spread before the onset of acute inflammatory response. Here the extent of the viral replication is dependent on the immune containment-response, but given that MERS has shown 34% case fatality compared to 9.5% for SARS-CoV and only 2.3% for SARS-CoV-2 thus far (Petrossillo et al., 2020), one can argue that the R_0 values should also have been in the same order, which is evidently not the case. This suggests that neither environmental nor immunological response suitably explains the higher R_0 of SARS-CoV-2, suggesting that the key reason may be molecular. The

E-mail address: dpal@iisc.ac.in.

<https://doi.org/10.1016/j.jsb.2021.107713>

Received 15 September 2020; Received in revised form 23 February 2021; Accepted 25 February 2021

Available online 1 March 2021

1047-8477/© 2021 Elsevier Inc. All rights reserved.

Spike (S) glycoprotein that protrudes ~150 Å out of the 500–2000 Å diameter coronavirus envelope is the most suitable molecule for making the first contact with the host cell, and is, therefore, a key molecular factor that determines virus fusion, entry and spread in the host, and thus holds clues for the rapid spread of SARS-CoV-2.

Densely glycosylated Spike protein has been suggested as the prime reason for the high SARS-CoV-2 infectivity (Wrapp et al., 2020). This extensive glycosylation of the S2 domain is driven by an intracellular C-terminal signaling peptide for transport and retention in the endoplasmic reticulum (see Fig. S1A Multiple Sequence Alignment bottom panel). However, this signal sequence is absent in a fellow murine β -coronavirus MHV-A59 Spike protein (Sadasivan et al., 2017), indicating limited glycosylation. Yet, MHV-A59 aggressively infects the mouse liver and brain. Upon intracranial inoculation in mice, it can cause acute stage meningoencephalitis and myelitis, chronic stage demyelination, and axonal loss (Das Sarma et al., 2009, 2002, 2008). It infects the neurons profusely and can spread from neuron to neuron. Its propagation from grey matter neuron to white matter and release at the nerve ends to infect the oligodendrocytes by cell-to-cell fusion (Das Sarma et al., 2009) are robust mechanisms to evade immune responses and induce chronic stage progressive neuroinflammatory demyelination concurrent with axonal loss in the absence of functional virions (Das Sarma et al., 2009; Kenyon et al., 2015). Therefore, high infectivity of the MHV-A59 Spike does not appear to be contingent on glycosylation and one can argue the same for SARS-CoV-2 Spike, where its glycosylation may only marginally raise the basal fusion efficiency. Surface glycosylation thus may not be a contributing factor to host cell binding, although the successive virus-to-cell and cell-to-cell fusion may all together play an important role in higher virus infectivity.

It has been suggested that enhanced virus-to-cell infection can be propelled by the increased number of hydrogen-bonded contacts between SARS-CoV-2 Receptor Binding Domain (RBD) and ACE2 receptor leading to higher affinity and improved host targeting compared to the SARS-CoV (Tai et al., 2020; Wrapp et al., 2020). However, a significantly higher affinity between ACE2 and SARS-CoV-2 has not been experimentally corroborated (Walls et al., 2020). Besides, such a proposition is weak because the RBDs in all HCoVs are diverse, including SARS-CoV, where the minimal RBD (318 to 510 residues) (Hofmann and Pöhlmann, 2004) shares only 74% sequence identity with SARS-CoV-2 (Fig. S1A). Also, the SARS-CoV-2 Spike may interact with other receptors such as DC-SIGN and DC-SIGNR as in SARS-CoV to increase tropism (Marzi et al., 2004) and viral spread. Therefore, there is no direct consequence of ACE2 recognition with infectivity unless a virus entry can be realized; however, when RBDs interact with ACE2 in large numbers during the acute stage of the infection, they may modulate the host immune response by downregulating hydrolysis of the pro-inflammatory angiotensin II to anti-inflammatory angiotensin 1–7 in the renin-angiotensin signaling pathway (Vaduganathan et al., 2020). This can alter the immune response and increase infectivity. But such effects can manifest only beyond the early stage of the infection, and for that to happen the efficiency of viral entry is the rate-limiting step.

The cleavage of the Spike protein is said to prime it for the efficient virus-host membrane fusion process. How essential is this for virus fusogenicity and infectivity is an important consideration. The Spike cleavage potentially removes any *in situ* covalent and noncovalent constraints that the S1 domain may impose on the S2 domain impeding its conformational transition that facilitates the virus entry. It has been proposed that SARS-CoV-2 Spike is preactivated by cleavage at the S1/S2 site when it is packaged inside the host, and the S2' site is cleaved when the Spike gets attached to the host receptor, which makes the priming process very efficient (Hoffmann et al., 2020). However, a comparison of the S1/S2 cleavage signal sequence ...RXXR... shows that SARS-CoV-2 "...RRARS..." Furin recognition site is similar to MHV-A59 Spike's "...RRAHR...", and others like MERS, HCoV-OC43 and HCoV-HKU1 Spike have a conserved motif sequence as well (Fig. S1A). The cleavage site signal at S2' embedding a single Arginine is highly

conserved across all HCoVs. Therefore, the efficient priming advantage available to SAR-CoV-2 Spike is equally present for MHV-A59, MERS, HCoV-OC43, and HCoV-HKU1 Spike. In contrast, the canonical S1/S2 cleavage recognition sequence is missing in SARS-CoV with only a single Arginine present there. A regular cleavage at this site has not been reported, and cleavage by trypsin has been shown to activate the virus independent of the pH due to the presence of a single Arginine. The importance of this region has been aptly corroborated by S2' site cleavage studies in SARS-CoV (Belouzard et al., 2010). Besides, it is also possible for Spike to be activated by the low pH environment through protonation of residues if it internalizes in the endosome post interaction with the host receptor. In contrast, fusion processes are known to happen in MHV-A59 Spike without cleavage as well (Hingley et al., 2002). Therefore, based on the similarity of the SARS-CoV-2 Spike with others, one can argue that it is competent to access multiple pathways for priming that enhances its infection capability, including the possibility that it can infect without a cleavage as well – though these advantages are not unique.

Among all the components of the fusion apparatus in the Spike S2 domain, the FPs are the least studied although they have been suggested to contribute to the trigger that drives the virus-host fusion process by initiating the protein-host membrane contacts. Limited experimental information available shows mutation in FP of SARS-CoV Spike can significantly perturb the fusion efficiency (Broer et al., 2006) as much as >70% (Petit et al., 2005). Most studies of the FP regions have used synthetic peptides in a fusion assay system to understand their membrane perturbing capabilities (Alsaadi et al., 2019; Guillen et al., 2008a, 2008b; Sainz et al., 2005), and how Ca^{+2} ions may interact with these peptides to modulate fusion (Tang et al., 2020). Interestingly, the FPs also contains a central proline (White, 1990) in several viruses such as the Avian Sarcoma/Leucosis virus (Delos et al., 2000), Ebola virus (Gómara et al., 2004), Vesicular Stomatitis virus (Fredericksen and Whitt, 1995), and Hepatitis C virus (Drummer and Pombourios, 2004), where its important role has been investigated through mutation studies. The location of coronavirus FPs proximal to the N-terminal of the S2 domain is reminiscent of FPs from HIV-1, influenza virus, and paramyxoviruses. They have been suggested to be located at the head of a pre-hairpin intermediate structure (Harrison, 2008) predicted for the current model of class I viral fusion proteins.

Although the FPs are believed to be the early initiators of protein and host membrane contact, there is still no consensus on their location. For example, the fusion peptide for SARS-CoV-2 Spike has been cited at 788–806 position by Xia et al. (2020) compared to 816–833 by Wrapp et al. (2020). When inferred from alignment to SARS-CoV Spike, two additional FP segments at 875–902 position and 1203–1220 can be proposed based on experimental studies by Ou et al. (2016) and Guillen et al. (2008), respectively. In reality, all four Spike fusion domain segments (FP-I to FP-IV; Fig. S1A) mentioned can be identified by a simple window-based analysis using interfacial hydrophobicity scales such as from Wimley and White (1996). Given that FP-I to FP-III are contiguous to each other in the sequence, the whole segment spanning the beginning of FP-I to the end of FP-III can be termed together as the “fusion loop” (Fig. S1B). But how these FPs in the loop can act synergistically to rapidly trigger the membrane fusion process is an important point for study.

Recently we have seen from *in silico*, *in vitro*, and *in vivo* studies (Singh et al., 2019) that additional “rigidity” accrued from the presence of proline in FP-III from MHV-A59 Spike is critical for virus fusogenicity, infectivity, and pathogenicity. The two consecutive prolines at 938–939 of Spike from MHV-A59 (S-MHV-A59(PP)) in the FP-III and its proline deletion (Δ 938) mutant S-MHV-A59(P) from an isogenic recombinant strain of MHV-A59, RSA59 (Singh et al., 2019) reveal slower trafficking of the latter to the cell surface and significantly less fusogenicity. The proline deleted targeted recombinant mutant strain RSA59(P-) which contains S-MHV-A59(P) when compared to S-MHV-A59(PP) containing parental isogenic strain RSA59 for infection in neuronal cell line

demonstrate less aggressive and fewer syncytia formation, and one order lower viral titers post-infection *in vitro*. The *in vivo* studies in mice parallels the *in vitro* studies demonstrating significantly reduced viral replication and consecutive disease pathologies, like less severity in meningitis, encephalitis, and demyelination, and inability to infect the retina nor induce loss of retinal ganglion cells (Rout et al., 2020). The non-neurotropic strain MHV-2 sharing 91% genome identity with MHV-A59, and 83% pairwise Spike sequence identity, with a single central proline in FP-III, cause only meningitis and is unable to invade the brain parenchyma. Computational studies of S2 fusion domains of S-MHV-A59 (PP) and MHV-2 Spike involving molecular dynamics confirm the former to be more rigid and containing more residues in the regular secondary structure (Singh et al., 2019).

The above studies point to the critical role of proline in the fusion loop, a point of further study in this paper. I perform a comprehensive computational study of Spike protein fusion peptides from five human β -coronaviruses and MHV-A59 for comparison. In line with the observations on Spike from MHV-A59, the high SARS-CoV-2 infectivity makes a particular case for further studies due to the presence of double proline in FP-I. The understanding of the intrinsic properties of the fusion peptides in the context of Spike trimeric architecture offers an insight into how they contribute to the host-membrane interaction to enhance virus fusogenicity and shed light on the distinctive reproductive number observed in SARS-CoV-2.

2. Materials and methods

The sequences used in this study (Fig. S1) were downloaded from the NCBI database (URL: <http://www.ncbi.nlm.nih.gov>). The multiple sequence alignments were performed using the T-coffee webserver (<http://tcoffee.crg.cat/>). The default parameter values for alignment available in the server were used. The server combines several methods to come up with an optimal multiple sequence alignment (Di Tommaso et al., 2011).

All protein three-dimensional structures were downloaded from the Protein Data Bank (<http://www.rcsb.org>). The PDB IDs for the downloaded structures are HCoV-HKU1: 5I08, MHV-A59: 3JCL, 6VSJ, HCoV-OC43: 6NZK, MERS-CoV: 6Q04, SARS-CoV-2: 6VXX, 6VSB, and SARS-CoV: 5XLR. Structures with the highest resolution were preferred when more than one model was available. Coordinates from these files were extracted for obtaining starting models of the FP-I, FP-II, FP-III, used in our molecular dynamics simulations. Whenever there were missing coordinates, they were modeled as an extended structure in the FP. The Solvent Accessible Surface Area (SASA) was calculated by the program NACCESS (<http://wolf.bms.umist.ac.uk/naccess/>); residues for which no atom coordinates are present in the PDB file was considered as fully exposed to solvent while calculating the relative SASA values. The secondary structure was calculated by the PROCHECK program from the PROCHECK suite (<https://www.ebi.ac.uk/thornton-srv/software/PROCHECK/>). The electrostatic surface potential was calculated using the APBS plugin (Jurris et al., 2018) inside the PyMol software (<http://pymol.org>). The default parameters were used for the calculations. All cartoon diagrams and surfaces were rendered by the PyMol software. For a more accurate calculation of the surface electrostatic potential, the surface was defined by the MSMS program (Sanner et al., 1996). MSMS computes the reduced surface of a set of spheres using a defined probe in which an analytical description of the solvent excluded surface is computed from the reduced surface. The electrostatic potential calculated by the DELPHI program (Li et al., 2012) is mapped to the MSMS computed surface and then partitioned per residue to estimate the mean electrostatic potential at the individual solvent-exposed residue surface. Both APBS and DELPHI implements the Poisson-Boltzmann method for electrostatic potential calculation.

The molecular dynamics simulations were performed with GROMACS 2019 version software (Abraham et al., 2015) (<http://www.gromacs.org>). The simulations were performed using the CHARMM 27

forcefield with cmap (Bjellmar et al., 2010). Each FP was placed in a cubic box solvated with water (TIP3P model) (Miyamoto and Kollman, 1992; Smith and van Gunsteren, 1993). Equal dimensions (nm) of the cubic box used for simulation as per the protein order: 3JCL, 5I08, 5XLR, 6NZK, 6Q04, 6VXX are FP-I: 11.16, 9.07, 6.94, 6.95, 7.04, 6.75, FP-II: 6.65, 6.66, 6.69, 6.68, 6.83, 6.61, and FP-III: 6.76, 7.05, 7.05, 6.77, 6.92, 8.51. Solvent molecules were randomly replaced with Na^+ and Cl^- to neutralize the system and additional ions were added to bring the final concentration of NaCl to 0.1 M. Periodic boundary conditions were enforced in all three directions. The system was minimized until the maximum force in the system reached below $1000 \text{ kJ mol}^{-1} \text{ nm}^{-1}$. Thereafter, the MD simulations were run in two modes, one for standard dynamics and another in replica-exchange mode. For standard dynamics, the fusion peptide systems were equilibrated for 2 ns under the NVT ensemble. In this step, a modified Berendsen thermostat was used without pressure coupling (Berendsen, 1991). This was followed by equilibration for 2 ns in the NPT ensemble at 1 atm and 300 K. In this step, the modified Berendsen thermostat was coupled with the Parrinello-Rahman barostat (Parrinello and Rahman, 1981). The production run was executed for 500 ns saving the output every 100 ps yielding 5000 frames for analysis. The analysis of the trajectory was performed using GROMACS utilities. The diagrams were created using our in-house software.

The replica exchange molecular dynamics (Sugita and Okamoto, 1999) were performed using identical simulation parameters as above but at different temperatures. The lowest temperature was 300 K, the same as the simulation described above. The higher temperature steps were calculated using the protocol by Patriksson (Patriksson and van der Spoel, 2008) keeping the minimum replica-exchange probability at 20%. The highest temperature of the simulation was 320 K and 22 simulations were run parallelly at the specific temperature steps for 20 ns. The highest temperature was chosen so as to simulate a cumulative time of 440 ns. Processing of the trajectories was done using GROMACS utilities and in-house scripts.

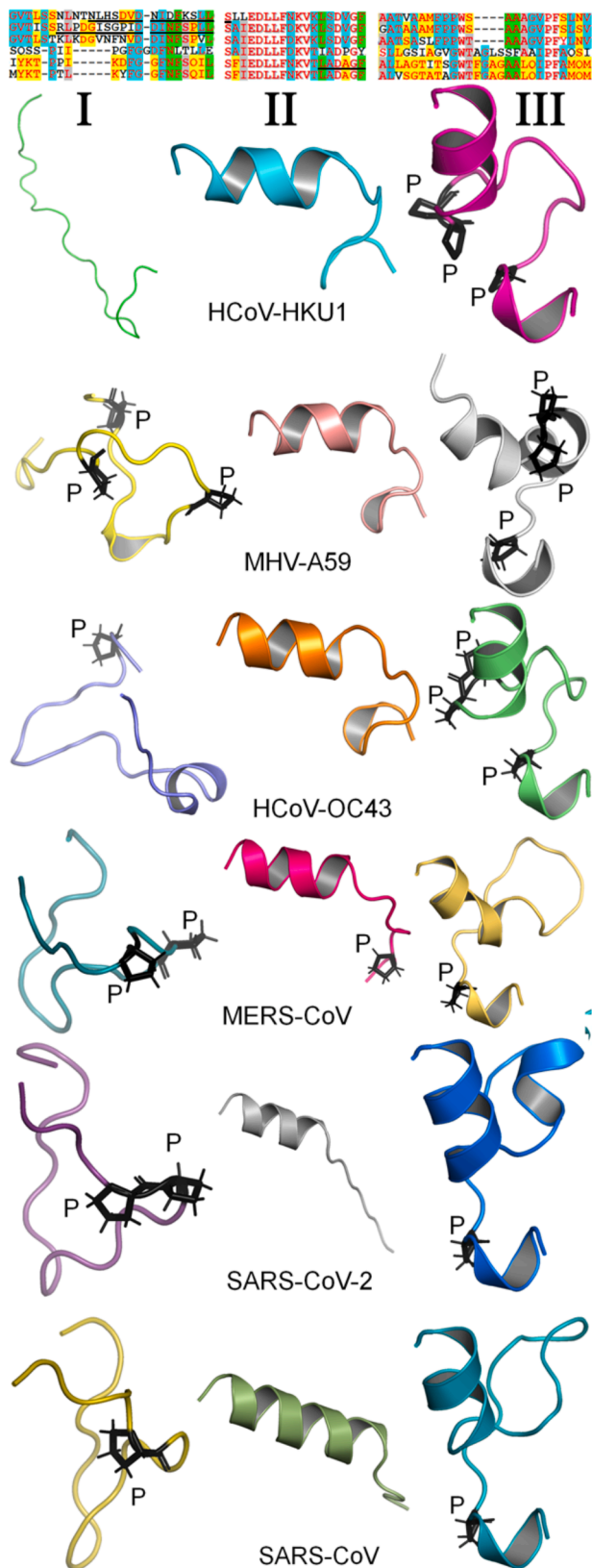
3. Results

3.1. Spike fusion peptides

The fusion peptides studies in this study have lengths in the range 19–25 for FP-I, 18 for FP-II, and 23–27 for FP-III (Fig. 1). Simple alignment suggests that FP-I is the least conserved, while FP-II is the most conserved. This is also reflected in the secondary structure where FP-I is largely in irregular or loop conformation. SARS-CoV-2 FP-I has consecutive prolines which creates a bend in the structure. The FP-II always has an α -helix in the N-terminal half that often extends into the C-terminal part. FP-II is devoid of Pro except for a single case at the C-terminal segment of MERS-CoV. Consecutive prolines are present in FP-III of HCoV-HKU1, MHV-A59, and HCoV-OC43. All the FP-III peptides are helix-loop-helix structures. In general, it can be said that none of the FP-II peptides have a central proline, while all FP-Is have a central proline except, HCoV-HKU1 and HCoV-OC43. These do have a central proline in FP-III along with MHV-A59. Interestingly, all FP segments are expected to be membranotropic; however, we do find charged and polar amino acids in them. These are as well conserved alongside the aromatic and hydrophobic residues. Although the location of FP-IV is not available from the three-dimensional structure, it can be inferred from the primary structure and the location of the transmembrane domain (Fig. S1A)). It is deeply buried and interface the virus membrane. Since it is unlikely to be involved in the contact initiation with the host membrane, it is not further studied in this work.

3.2. Spike receptor binding and the fusion loop

To understand the synergy of the FPs in the fusion trigger process, it is important to understand the structure of the Spike protein in the



(caption on next column)

Fig. 1. The sequences of the fusion peptides used in this study and their three-dimensional structures. Each row of sequence and their corresponding structures are from the Spike protein from a given virus, and the Fusion Peptides I, II, III are marked as per their location in the Spike primary sequence as indicated in Fig. S1. Proline, whenever present in the structure is shown as stick-model and labeled. The PDB structures corresponding to each virus are 5I08, 3JCL, 6NZK, 6Q04, 6VXX, and 5XLR. Note that residue coordinates for FP-I, from 5I08 and 3JCL are partly absent; the same is true from FP-II from 6VXX. The missing coordinates have been modeled. The aligned sequences for the fusion peptides are annotated as follows: amino acid at the given position in alignment is present in all six Spike proteins – RED bold with shadow, present in five – green background, present in four – grey, present in three – cyan, present in two – yellow, present once – white background, black font.

proper context. The FP-I to FP-III are surface exposed and are contiguous to each other in space as seen from the three-dimensional structures of the Spike fusion domain (Fig. 2A). FP-I to FP-III is always surface exposed in the full-length Spike, and therefore, always available for early contacts with the host membrane; whereas, FP-IV can participate in the process post the conformational transition which may expose it for interaction with the host membrane. To understand what obligates the FP I-III surfaces to make the initial protein-host membrane contact, one must look at the possible modes of virus-host attachment mediated by the Spike. For this, I propose a new Contact Initiation Model, where there is no requirement of a fusion peptide to be at the N-terminal of the conformationally transformed pre-hairpin intermediate (Tang et al., 2020). To understand the model, let us consider the different options for host receptor binding. For example, if all three RBDs in the trimeric structure find the host receptors, it can attain a tripod binding mode (Fig. 2A). A recent structure of the trimeric Spike complexed with a host receptor reveals the precise geometry of CEACAM1 binding the RBD of Spike from MCoV-A59 (Shang et al., 2020). However, a tripod binding requires receptor molecules on the host membrane to be pre-available in a specific arrangement. High expression of receptor molecules on the host cell surface is expected to increase the probability of tripod binding, but there is no existing information on whether such a precise arrangement is present on the host surface suitable for interaction with the trimeric RBD. Moreover, the membrane bilayer structure does not contain any feature that can direct such a regular host receptor arrangement. It may be noted that only in a tripod arrangement, the N-terminal segment of the S2 domain will interact early during the protein-membrane contact post intermediate structure formation, and in such a case the pre-hairpin/pre-bundle helices of the fusion domain are expected to interact head-on with the host membrane (Tang et al., 2020). However, FP-I and FP-II are located at the middle of the cylinder-like Spike S2 structure (Fig. 2A) and an intermediate formation through conformational transition is needed to place them near the head of the cylinder. Here, FP-I is expected to make the early host-membrane contact if the cleavage is at the S1/S2 site and FP-II if the cleavage is at the S2' site (see Fig. S1B). The FP-III region has limited scope to make any early contact due to its farthest position from the host membrane surface. In tripod binding mode the virus membrane is still ~ 150 Å away, and the three Heptad Repeat 2 (HR2) regions need to fold back and bind to the hydrophobic grooves of the Heptad Repeat 1 (HR1) trimer in an antiparallel manner to bridge this gap and form a hemifusion structure with the host membrane (Tang et al., 2020).

3.3. The Contact Initiation Model – spike fusion peptide trigger

If only one or two RBDs bind the receptor, the vertical anchoring of the Spike fusion domain relative to the host surface lacks the third anchor rendering the vertical orientation unstable and unfeasible. Also, a recent trimeric structure of SARS-CoV-2 Spike has shown a single RBD to be in the open conformation (Wrapp et al., 2020), where it is swiveled away from its core structure originally interacting with the N-Terminal Domain (NTD) and the fusion domain. In such a state, the interaction of

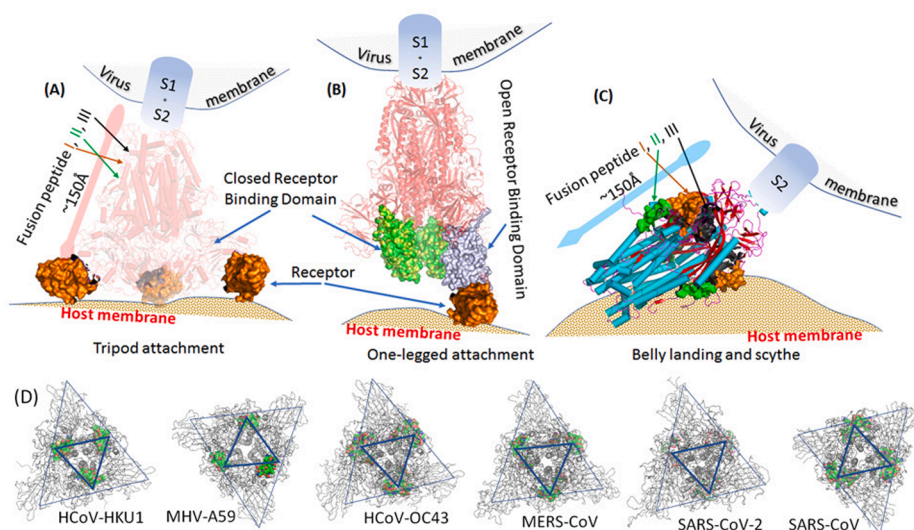


Fig. 2. Schematic diagrams explaining the mode of binding of Spike protein to the host receptor and its putative orientation relative to the host membrane surface during the virus-host attachment. (A). A cartoon diagram created using PDB ID: 6VVSJ where the Spike protein from MHV-A59 is complexed with CEACAM1 receptor from the mouse. Since all three subunits from the Spike RBD attach to the receptor, it allows the virus to anchor in a tripod mode. The approximate length of the Spike in the longer dimension is indicated along with the FP locations. (B). A cartoon diagram created using PDB ID: 6VSB, where one of the RBD is shown in an open conformation. A receptor molecule ACE2 has been drawn to show the putative attachment in one-legged mode. A two-legged mode may also be possible similarly. (C). The S2 domain of SARS-CoV-2 Spike protein is shown in a belly-landing orientation. The approximate length of the Spike in the longer dimension is indicated along with the marked FP locations. The S1 domain is likely to be loosely bound and not shown for clarity. (D) A ribbon diagram of the trimeric Spike proteins from the six coronaviruses used in this study. Two triangles are marked on each structure, where the relative locations of the

truding N-Terminal Domains (NTDs) appear near the outer triangle vertices, and the FP-I, FP-II, and FP-III are located near the inner triangle vertices. The triangles are marked to bring out the relative positions of the NTDs and the FP surfaces. The Spike structure is oriented such that the RBD appears closest to the eye, followed by the NTD and then the FPs.

the open RBD with the host (Fig. 2B) is not expected to stabilize the Spike anchoring in any specific orientation relative to the host due to the weak interaction with the fusion domain. Here a post-cleaved S2 domain is expected to interact side-on to the membrane surface through a “belly” landing to trigger the fusion process (Fig. 2C). Such a process would be sterically facile if there are no other Spike in the vicinity on the virus surface. It is to be noted that the shape of the trimeric S2 domain is not a proper cylinder, but with a bulge in the mid-segment which I call a “belly”. The FP-II and FP-I surfaces are located at the crest of this bulge, such that it is able to make the initial contact with the host membrane. The structural constraints that obligate the “belly” landing can be understood from the overall geometry of the Spike (Fig. 2D). It is to be noted that the most stable and eventual landing posture of an object having an uneven surface will be the one that guarantees the largest surface area of contact with the host landing surface. The Spike architecture is such where the relative location of the NTDs approximately form three vertices of a triangle, while the FP-I to F-III are located midpoint of the sides, forming the vertices of an inner triangle. Based on the premise of the maximum landing surface, if we consider the receptor attachment to be in two RBD locations, the S2 landing will be close to the midpoint of the two vertices of the outer triangle coincident to the FP surfaces. Even for one-legged attachment, the contact must always be directed towards the midpoint of the two NTDs because that allows maximum contact surface to be formed where the Spike can stably rest on the host surface. During the contact, the membranotropic segment of the fusion loop spanning the FPs is expected to engage the host membrane during the fusogenic conformational change. In this case, the FP sites are proximal to the virus membrane surface such that hemifusion membrane structures can be initiated early in comparison to the tripod-binding mode which requires an intermediate pre-hairpin structure to be formed. It is also to be noted that weak RBD binding to the S2 domain or disintegration of the Spike trimer post tripod binding can mimic the one- or two-legged binding mode. For additional lucidity, the structural basis of the Contact Initiation Model is further explained pictorially in a lay manner in Fig. S2.

3.4. SARS-CoV-2 fusion peptide distinctions

The physicochemical property of the fusion loop and the synergy of

the FPs therein are critical to the rapid initiation and transition to the hemifusion stage of the membrane fusion process. The proposed Contact Initiation Model ensures that the S2 domain lies on the belly contacting the host surface during the conformational transition. Since the exerted force during the conformational transition is tangential to the host surface, the orthogonal frictional forces may allow an efficient scything action based on the physicochemical nature of the contact engaged by the fusion loop. The nature of the initial surfaces of the fusion loop can be obtained from the electrostatic potential of the FP-I to FP-III surface patches (Jurus et al., 2018) (Fig. 3A, surface diagrams). Among the six Spike proteins considered in this study, SARS-CoV-2 has the most hydrophobic/neutral electrostatic FP patches (WHITE colored) most suitable for membrane disruption (Fig. 3B). It also has the least amount (10.9%, excluding MHV-A59) of highly negative (≤ -10 kT/e) electrostatic surface (RED colored patches) which repel the attachment of the S2 domain to the host membrane due to the repelling hydrophilic surface of the outer membrane composed of negatively charged fatty acid groups. On the other hand, a positive electrostatic surface (BLUE colored patches) may allow the protein surface to tightly attach to the membrane exterior through charge attraction. The presence of an interfacial ion like the Ca^{+2} can cap the negative charge at the FP surface to assist fusion trigger - consistent with the membrane charge compatibility requirements. The disruption efficiency of surface patches is, therefore, likely to be highest for the hydrophobic/neutral, followed by the positive and negative electrostatic potential, unless modulated by an interfacial cation like the Ca^{2+} . One may argue that large patches of negative electrostatic surface potential (RED color) in the HCoV-HKU1 and HCoV-OC43 Spike fusion domain may explain the mild nature of those viruses. Interestingly, the rank order of reproductive number (R_0) of the MERS, SARS-CoV, and SARS-CoV-2 appears to agree with the order of the molecular surface mean electrostatic potentials (Fig. 3B). The benign nature of HCoV-HKU1 and HCoV-OC43 fits well within this trend. It may be recalled that MHV-A59 Spike undergoes limited glycosylation due to lack of localization signal (Fig. S1A); therefore the lowest molecular surface mean electrostatic potential agrees well with its high virulence experimentally observed in mouse model systems. A membrane-bound simulation of the Spike proteins as per the Contact Initiation Model proposed in this work would be useful to understand the key membranotropic interactions that drive the virus fusogenicity.

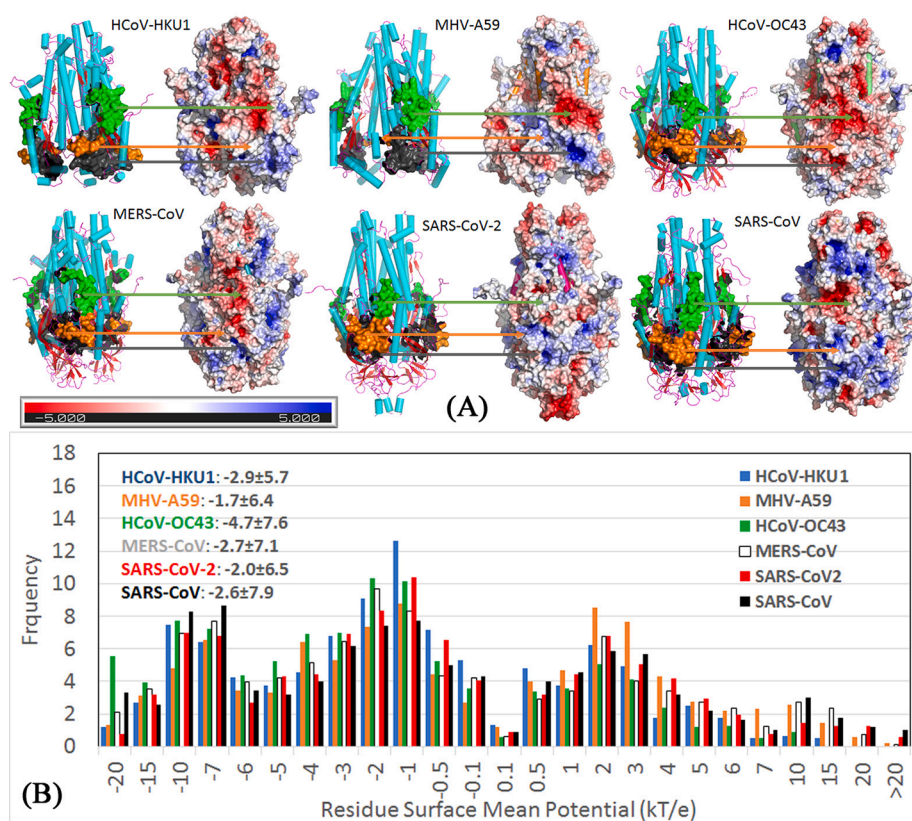


Fig. 3. Spike protein and its electrostatic potential surface. (A). Cartoon diagram showing the fusion peptide regions FP-I (orange), FP-II (green), and FP-III (dark grey) on the trimeric fusion domain structure of Spike from six viruses. The electrostatic surface of the fusion domain is shown adjacent to each structure and the corresponding fusion peptide regions are marked by an arrow. The orientation of the cartoon structure and the electrostatic surface are aligned among themselves. Note that one face of the electrostatic surface that is visible is repeated on the other side due to the symmetry arising out of the trimeric quaternary structure. The fusion domain structures are shown in cyan for helices, strands in red, and loops in magenta. Please refer to Fig. S1A for the sequence alignment and its legend for the PDB ID of files used to draw the structures. Note that a part of the FP-I surface is absent for MHV-A59 and HCoV-HKU1 due to unavailable atom coordinates in the PDB file. The same is true for a small section of FP-II from SARS-CoV-2 (see Fig. 1). All FP-I regions have at least one Asn-linked site, only SARS-CoV-2, SARS-CoV, and HCoV-HKU1 has a site in FP-II, and HCoV-OC43 and HCoV-HKU1 in FP-III. Contiguous to the FP-III are the heptad repeat regions where the glycosylation and sequence conservation among Spike is the highest. Spike from MHV-A59 shares a 64–66% overall sequence identity with HCoV-HKU1 and OC43, 71–76% identity with the corresponding S2 domains. A similar match with SARS-CoV, MERS-CoV, and SARS-CoV-2 are at <30% pairwise sequence identity. (B). A histogram showing the mean electrostatic potential at the surface of solvent-exposed residues of the Spike proteins. The values from the individual monomers were averaged to compute the histogram. The mean values for the whole molecule are also indicated.

Aside from the electrostatics, the physical rigidity of the fusion loop in Spike is of prime importance for the virus fusogenicity. This is because the Spike fusion domain being metastable, any local alteration of rigidity has global implications for the molecule. This has been alluded to by the mutation studies on the fusion peptide central prolines (Delos et al., 2000; Drummer and Pountourios, 2004; Fredericksen and Whitt, 1995; Gómara et al., 2004), but its criticality was recently revealed from our comprehensive studies on centrally located consecutive prolines in MHV-A59 Spike fusion peptide (Singh et al., 2019). Proline being an imino-acid with a unique structure has restricted torsional freedom, which in turn restricts the torsional freedom of the protein backbone where it is located (Chakrabarti and Pal, 2001). When two consecutive prolines are located, the rigidity of the protein backbone is further enhanced.

To understand the intrinsic flexibility/rigidity of the surface exposed FPs, I set up 500 ns standard molecular dynamics simulations (Fig. 4) and 440 ns enhanced sampling replica exchange molecular dynamics (REMD) simulations (Fig. 5). The intrinsic flexibility of the fusion peptides can be better assessed by simulating the molecular dynamics at a series of higher temperatures in combination with replica exchange across simulations. As seen from the plots in Fig. 5 in comparison to Fig. 4, the fraction of regular secondary structure has decreased in REMD simulations and the root-mean-square fluctuation (RMSF) has increased significantly owing to higher simulation temperature. The presence of regular secondary structure at high RMSF suggests the presence of specific joint regions in the peptide where the global alterations are principally taking place, with smaller local structures remaining stable.

If we look at FP-I from the SARS-CoV-2 it is the only one with a central double proline, where 33% of the peptide is surface exposed during the standard simulation. Leaving out the termini, the exposure is

highest in the immediate two residue neighborhood of the double proline. For all FPs in the FP-I region, although local secondary structures are induced during the standard simulation, the lowest RMSF for a residue is achieved only by the SARS-CoV-2 double proline neighborhood. This trend is preserved in the REMD simulations as well and the lowest RMSF is enforced three residues downstream of the PP dipeptide at the Asp residue. Interestingly, FP-I of SARS-CoV-2 and SARS-CoV bear high sequence similarity and their RMSF curve look similar in standard simulation, but not in REMD, which points to the importance of a double proline in the peptide segment. There are no central prolines in FP-II and FP-III in the SARS-CoV-2 Spike. FP-II is largely stabilized by the presence of helical hydrogen bonds in the N-terminal segment. SARS-CoV-2 FP-III, however, contains a unique “Thr-Ile-Thr” segment constituted of three consecutive β -branched side-chain residues that can impart substantial rigidity based on steric considerations (Chakrabarti and Pal, 2001) if not as much as the consecutive prolines. The REMD simulations confirm the same as it induces a higher fraction of helical conformation with a swivel at the nearby Gly (Fig. 5). This can be again confirmed from the energy landscape plot which shows a single smooth dominant conformational well in contrast to the other FP-IIIs (Fig. S3).

The role of single proline can also be seen from the FP-I in MERS, SARS-CoV, and MHV-A59. The starting structures of all the FP-I fragments are devoid of stable secondary structures such as helices or sheets and are dominated by turns, and irregular structures indicating that the region prefers to be in a loop conformation. This is true even for the MHV-A59 Spike, which is flexible as experimentally evidenced by the lack of coordinate from the electron density map for complete FP-I in the PDB file: 3JCL. Standard simulation induces regular secondary structure in the FP-I, but they are drastically eliminated in the REMD simulations except for HCoV-OC43. REMD shows that P is able to restrain the RMSF in its neighborhood even in absence of a secondary structure. In

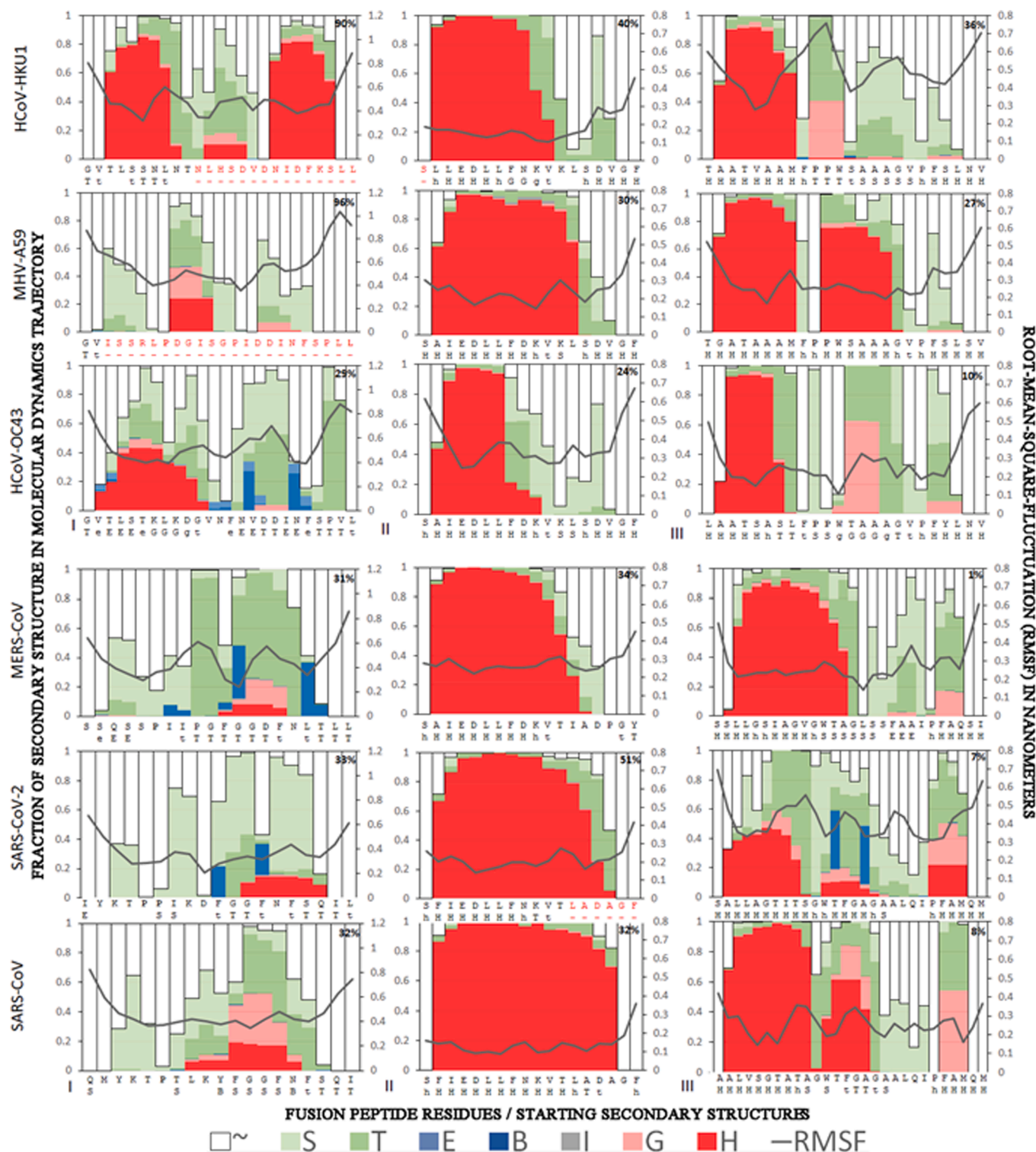


Fig. 4. Graphs showing the fraction of secondary structures and the RMSF calculated from a standard molecular dynamics simulation trajectory of the isolated fragments from FP-I, FP-II, and FP-III of six coronaviruses. The bar diagrams are drawn showing the secondary structure prevalence in 500 ns simulation and the Root-Mean-Square-Fluctuation (RMSF; Å). The X-axis indicates the residues for each FP fragment. The central proline residues are marked in BOLD and those residues for which the atom coordinates are absent in the PDB file are marked in RED. Note that the central proline is located in the central region of the fusion peptide. Below the FP residues, the corresponding secondary structures are indicated as present in the full-length protein. The symbols mean as follows, H: α -helix, h: α -helix termini, G: 3_{10} -helix, g: 3_{10} -helix termini, B: β -bridge, E: β -strand, e: β -strand termini, T: hydrogen-bonded turn, t: hydrogen-bonded turn termini, S: Bend, and ~: irregular secondary structure. The first, second, and third columns are indicated by Labels I, II, and III correspond to FP-I, FP-II, and FP-III in each Spike protein from a given virus, respectively. The relative SASA value of the whole FP fragment in the S2 domain expressed as a percentage is indicated in BOLD on the top part of each plot.

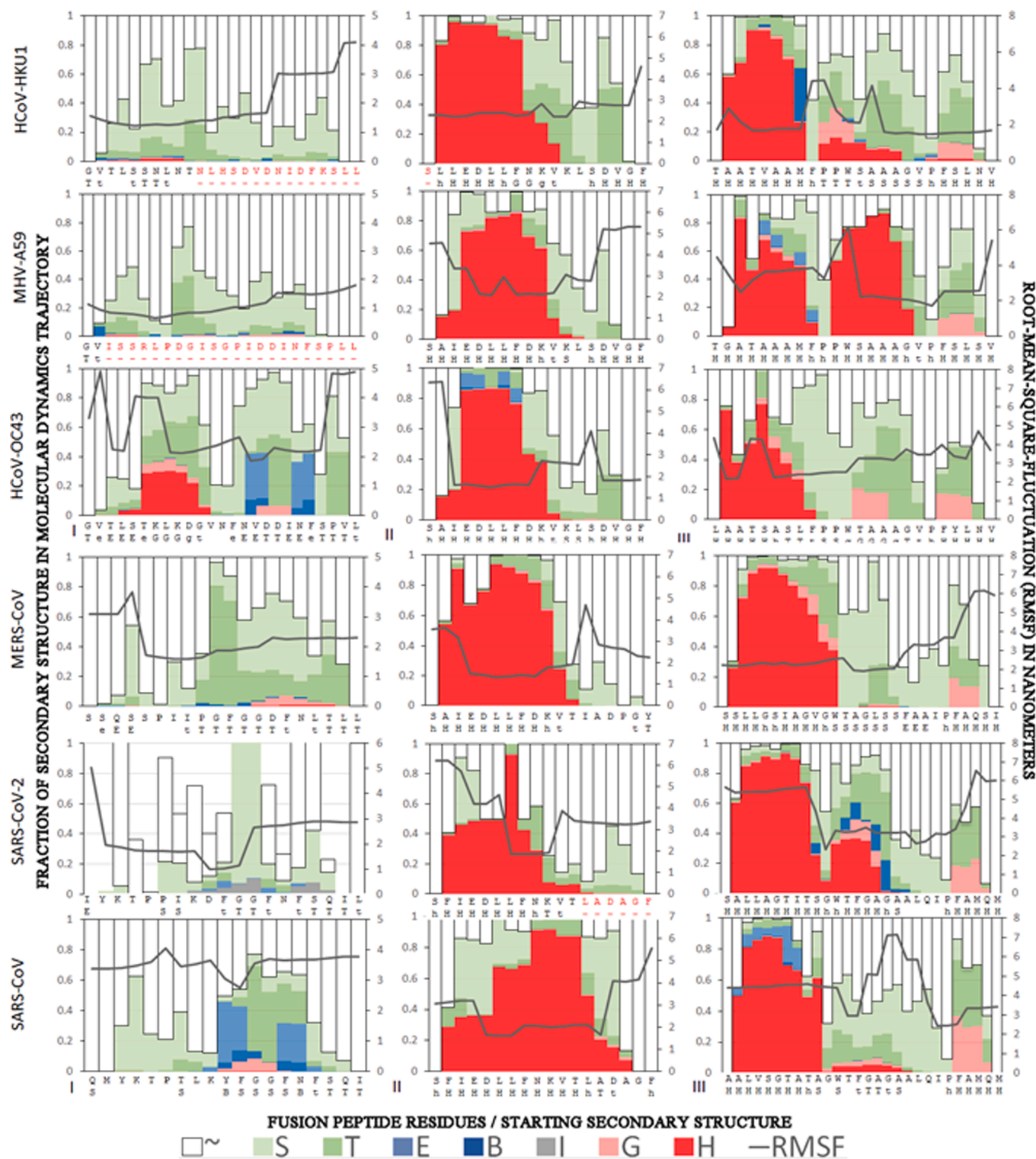


Fig. 5. Graphs showing the fraction of secondary structures and the RMSF calculated from enhanced sampling replica exchange simulation trajectory of the isolated fragments of FP-I, FP-II, and FP-III from six coronaviruses. The bar diagrams are drawn showing the secondary structure prevalence in the aggregated 440 ns simulation and the Root-Mean-Square-Fluctuation (RMSF; Å). Other details are identical to Fig. 4.

comparison, FP-IIs in all cases are dominated by helical conformation, with the N-terminal FP segments always in a helical state, which is true for FP-III region fragments as well. There are no central prolines in FP-II and a single conserved central proline in FP-IV, while consecutive prolines exist at a central location in HCoV-HKU1, HCoV-OC43, and MHV-A59 FP-III Spikes. Given an irregular structure, the effect of proline is more dramatic on the FP-I region in contrast to FPs from FP-II and FP-III

which are already stabilized by hydrogen bonds in helices. This is also reflected in the energy landscape plot created from the standard molecular dynamics simulation trajectory (Fig. S3) where the lower RMSD structure shows the most compact single conformational well among all the FP-Is. If we look at the FP-III regions where other consecutive proline containing FPs exist, HCoV-OC43 has the lowest residue RMSF at the double prolines (Fig. 4). The REMD simulations show that the double

proline contributes to sustaining the secondary structures flanking it, and this may be its steric ability to minimize fluctuations at the higher temperature.

The observations are consistent with our previous comparative molecular dynamics studies (Singh et al., 2019) on PP containing MHV-A59 FP-III and single central proline containing FP-III from another parental non-neurotropic strain of murine hepatitis virus (MHV), MHV-2, where we found the FP-III from the former became more rigid than the latter in methanolic conditions compared to water. Additionally, NMR studies on the MHV-A59 FP-III fragment revealed its unique ability to form *cis*-peptide at the central P-P peptide bond (Singh et al., 2019). A *cis*-trans isomerization during the membrane fusion process has the potential to expose the hydrophobic residues efficiently around the isomerized-peptide neighborhood, enhancing the fusion trigger potential.

While the structural role of proline cannot be discounted, other stabilizing interactions also dominate the FPs, among which the formation of the aromatic/hydrophobic clusters is of relevance to the fusion process (see Fig. S3 for examples). As observed from the packing of the loops, a combination of aromatic and Val/Ile/Leu side chains pack tightly to exclude water. But when these regions become exposed during the conformational transition, they would enhance the hydrophobic interaction of the protein surface with the membrane. In all S2 domains, however, FP-III regions are in part masked by the FP-I segment, as evident from the relative solvent accessible surface area (SASA) values indicated in Fig. 4. The FP-III region can get fully exposed post cleavage at the S2' site which dislodges the sheathing FP-I segment, or when the site gets progressively exposed during the conformational transition amidst the fusion process.

3.5. Target for therapy

Our study brings out the importance of the fusion loop region which could be a legitimate target for the design of vaccine or synthetic agents for therapy against COVID-19. The S2 domain serves as a better therapeutic target than S1 due to higher evolutionary conservation, but the few attempts made have mainly focussed attention on the heptad repeat regions (Tang et al., 2020). Two important features that also make the fusion loop an attractive therapeutic target are its accessibility due to its surface exposure in the full-length Spike, and the relatively high conservation of residues in the fusion loop, especially around the FP-II region which increases its scope as a pan-coronavirus target that can cater to future pandemic threats as well. Mimetic peptides can be designed to bind to the fusion loop to inhibit the fusogenic conformational transition of the S2 domain. Impairment of the fusion trigger would have a direct bearing on the fusogenicity of the virus and contribute to the reduction of lung invasion and damage that clinically results in acute pneumonia. Systematic studies can identify the minimal motif in the fusion loop serving as the fusogenic determinant to improve our selection of a potential therapeutic target to prevent cell-to-cell fusion and subsequent pathogenesis.

4. Discussion

The interplay of the outlined physicochemical features determines the virus-entry process to become more efficient. The local and global stability of the S2 domain is important. Since the S2 domain undergoes a conformational transition, local stability means reasonably rigid moving parts (akin to mechanical systems with joints), and global stability means a well-defined conformational transition pathway from the metastable to the stable state. This local and global stability requirement could be attributed to the physicochemical efficiency needed in disturbing the host membrane. Secondly, the electrostatic potential of the fusion loop-derived surface patches must be neutral or positive to be able to engage the host membrane. The presence of glycosylation sites adds to the hydrophobicity of the S2 domain surface, but it is a small fraction of the available surface for interaction with the host membrane.

The fusogenic conformational transition requires optimal synergy between the physical and chemical properties of the fusion loop to allow a concurrent scything action to rapidly facilitate the transition to the host-virus hemifusion membrane state. The free-energy available from the conformational transition of S2 to a more relaxed helical bundle is available to disrupt the host membrane and overcome the kinetic barrier to bring the host and virus membrane lipid bilayers together. The Contact Initiation Model ensures that the virus and host membrane are in close proximity for the formation of a hemifusion structure. The pre-hairpin S2 intermediate as suggested to exist by many researchers may be one of the many conformational states interacting with the host membrane. Priming by Spike cleavage is important for facilitating the rapid fusion process and therefore a part of the synergy at play. However, the open conformation of RBD seen in PDB ID: 6VSB for SARS-CoV-2 suggests that flexible linker segments loosely connect the RBD back to the fusion domain leaving it relatively free for an unfettered conformational transition. Therefore, multifarious options to prime and trigger appear to be available to SARS-CoV-2 for viral entry, which contributes to its increased infectivity. Preventing the trigger by inhibiting the fusion loop is therefore a suitable target for therapy. Given the importance, a more extensive study of SARS-CoV-2 Spike protein and the mechanistic hypothesis described here is therefore warranted.

5. Conclusions

The three proximal fusion peptides constituting the fusion loop in Spike protein are the membranotropic segments most suitable for engaging the host membrane surface for its disruption. Spike's unique quaternary structure architecture drives the fusion peptides to initiate the protein host membrane contact. The SARS-CoV-2 Spike trimer surface is relatively more hydrophobic among other human coronavirus Spikes, including the fusion peptides that are structurally more rigid owing to the presence of consecutive prolines, aromatic/hydrophobic clusters, a stretch of consecutive β -branched amino acids, and the hydrogen bonds. The synergy accrued from the location of the fusion peptides, their physicochemical features, and the fusogenic conformational transition appears to drive the virus fusion process and may explain the high spread of the SARS-CoV-2.

Declaration of Competing Interest

The author declares that he has no known competing financial interests or personal relationship that could have appeared to influence the work reported in this paper.

Acknowledgment

DP thanks the Department of Biotechnology, New Delhi for supporting the computational facilities, and Prof. Jayasri Das Sarma, Department of Biological Sciences, Indian Institute of Science Education and Research Kolkata for her generous support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jsb.2021.107713>.

References

- Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J.C., Hess, B., Lindahl, E., 2015. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1, 19–25.
- Alsaadi, E.A.J., Neuman, B.W., Jones, I.M., 2019. A fusion peptide in the spike protein of MERS coronavirus. *Viruses* 11 (9), 825.
- Belouzard, S., Madu, I., Whittaker, G.R., 2010. Elastase-mediated activation of the severe acute respiratory syndrome coronavirus spike protein at discrete sites within the S2 domain. *J. Biol. Chem.* 285, 22758–22763.

- Berendsen, H., 1991. Transport properties computed by linear response through weak coupling to a bath. In: *Computer Simulation in Materials Science*. Springer, pp. 139–155.
- Bjelkmar, P., Larsson, P., Cuendet, M.A., Hess, B., Lindahl, E., 2010. Implementation of the CHARMM force field in GROMACS: analysis of protein stability effects from correction maps, virtual interaction sites, and water models. *J. Chem. Theory Comput.* 6, 459–466.
- Broer, R., Boson, B., Spaan, W., Cosset, F.-L., Corver, J., 2006. Important role for the transmembrane domain of severe acute respiratory syndrome coronavirus spike protein during entry. *J. Virol.* 80, 1302–1310.
- Chakrabarti, P., Pal, D., 2001. The interrelationships of side-chain and main-chain conformations in proteins. *Prog. Biophys. Mol. Biol.* 76, 1–102.
- Das Sarma, J., Kenyon, L.C., Hingley, S.T., Shindler, K.S., 2009. Mechanisms of primary axonal damage in a viral model of multiple sclerosis. *J. Neurosci.* 29, 10272–10280.
- Das Sarma, J., Scheen, E., Seo, S.H., Koval, M., Weiss, S.R., 2002. Enhanced green fluorescent protein expression may be used to monitor murine coronavirus spread in vitro and in the mouse central nervous system. *J. Neurovirol.* 8, 381–391.
- Das Sarma, J., Iacono, K., Gard, L., Marek, R., Kenyon, L.C., Koval, M., Weiss, S.R., 2008. Demyelinating and nondemyelinating strains of mouse hepatitis virus differ in their neural cell tropism. *J. Virol.* 82, 5519–5526.
- Delos, S., Gilbert, J., White, J., 2000. The central proline of an internal viral fusion peptide serves two important roles. *J. Virol.* 74, 1686–1693.
- Di Tommaso, P., Moretti, S., Xenarios, I., Orobitg, M., Montanyola, A., Chang, J.-M., Taly, J.-F., Notredame, C., 2011. T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res.* 39, W13–W17.
- Drummer, H.E., Poubourios, P., 2004. Hepatitis C virus glycoprotein E2 contains a membrane-proximal heptad repeat sequence that is essential for E1E2 glycoprotein heterodimerization and viral entry. *J. Biol. Chem.* 279, 30066–30072.
- Fredericksen, B.L., Whitt, M.A., 1995. Vesicular stomatitis virus glycoprotein mutations that affect membrane fusion activity and abolish virus infectivity. *J. Virol.* 69, 1435–1443.
- Gómara, M.A.J., Mora, P., Mingarro, I., Nieva, J.L., 2004. Roles of a conserved proline in the internal fusion peptide of Ebola glycoprotein. *FEBS Lett.* 569, 261–266.
- Guillen, J., Kinnunen, P.K., Villalain, J., 2008. Membrane insertion of the three main membranotropic sequences from SARS-CoV S2 glycoprotein. *Biochim. Biophys. Acta* 1778, 2765–2774.
- Guillen, J., Perez-Berna, A.J., Moreno, M.R., Villalain, J., 2008. A second SARS-CoV S2 glycoprotein internal membrane-active peptide. Biophysical characterization and membrane interaction. *Biochemistry* 47, 8214–8224.
- Harrison, S.C., 2008. Viral membrane fusion. *Nat. Struct. Mol. Biol.* 15, 690–698.
- Hingley, S.T., Leparco-Goffart, I., Seo, S.H., Tsai, J.C., Weiss, S.R., 2002. The virulence of mouse hepatitis virus strain A59 is not dependent on efficient spike protein cleavage and cell-to-cell fusion. *J. Neurovirol.* 8, 400–410.
- Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T.S., Herrler, G., Wu, N.-H., Nitsche, A., 2020. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 181 (2), 271–280.
- Hofmann, H., Pöhlmann, S., 2004. Cellular entry of the SARS coronavirus. *Trends Microbiol.* 12, 466–472.
- Jurrus, E., Engel, D., Star, K., Monson, K., Brandi, J., Felberg, L.E., Brookes, D.H., Wilson, L., Chen, J., Liles, K., 2018. Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* 27, 112–128.
- Kenyon, L.C., Biswas, K., Shindler, K.S., Nabar, M., Stout, M., Hingley, S.T., Grinspan, J. B., Das Sarma, J., 2015. Gliopathy of demyelinating and non-demyelinating strains of mouse hepatitis virus. *Front. Cell Neurosci.* 9, 488.
- Li, L., Li, C., Sarkar, S., Zhang, J., Witham, S., Zhang, Z., Wang, L., Smith, N., Petukh, M., Alexov, E., 2012. DelPhi: a comprehensive suite for DelPhi software and associated resources. *BMC Biophys.* 5, 9.
- Marzi, A., Gramberg, T., Simmons, G., Moller, P., Rennekamp, A.J., Krumbiegel, M., Geier, M., Eisemann, J., Turza, N., Saunier, B., Steinkasserer, A., Becker, S., Bates, P., Hofmann, H., Pöhlmann, S., 2004. DC-SIGN and DC-SIGNR interact with the glycoprotein of Marburg virus and the S protein of severe acute respiratory syndrome coronavirus. *J. Virol.* 78, 12090–12095.
- Miyamoto, S., Kollman, P.A., 1992. Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* 13, 952–962.
- Morawska, L., Cao, J., 2020. Airborne transmission of SARS-CoV-2: The world should face the reality. *Environ. Int.* 139, 105730.
- Ou, X., Zheng, W., Shan, Y., Mu, Z., Dominguez, S.R., Holmes, K.V., Qian, Z., 2016. Identification of the fusion peptide-containing region in betacoronavirus spike glycoproteins. *J. Virol.* 90, 5586–5600.
- Parrinello, M., Rahman, A., 1981. Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* 52, 7182–7190.
- Patriksson, A., van der Spoel, D., 2008. A temperature predictor for parallel tempering simulations. *Phys Chem Chem Phys* 10, 2073–2077.
- Pedersen, S.F., Ho, Y.-C., 2020. SARS-CoV-2: a storm is raging. *J. Clin. Investig.* 130 (5), 2202–2205.
- Petit, C.M., Melancon, J.M., Chouljenko, V.N., Colgrove, R., Farzan, M., Knipe, D.M., Kousoulas, K., 2005. Genetic analysis of the SARS-coronavirus spike glycoprotein functional domains involved in cell-surface expression and cell-to-cell fusion. *Virology* 341, 215–230.
- Petrosillo, N., Viceconte, G., Ergonul, O., Ippolito, G., Petersen, E., 2020. COVID-19, SARS and MERS: are they closely related? *Clin. Microbiol. Infect.* 26, 729–734.
- Rout, S.S., Singh, M., Shindler, K.S., Das Sarma, J., 2020. One proline deletion in the fusion peptide of neurotropic mouse hepatitis virus (MHV) restricts retrograde axonal transport and neurodegeneration. *J. Biol. Chem.* 295 (20), 6926–6935.
- Sadasivan, J., Singh, M., Sarma, J.D., 2017. Cytoplasmic tail of coronavirus spike protein has intracellular targeting signals. *J. Biosci.* 42, 231–244.
- Sainz Jr., B., Rausch, J.M., Gallaher, W.R., Garry, R.F., Wimley, W.C., 2005. Identification and characterization of the putative fusion peptide of the severe acute respiratory syndrome-associated coronavirus spike protein. *J. Virol.* 79, 7195–7206.
- Sanner, M.F., Olson, A.J., Spehner, J.C., 1996. Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers* 38, 305–320.
- Shang, J., Wan, Y., Liu, C., Yount, B., Gully, K., Yang, Y., Auerbach, A., Peng, G., Baric, R., Li, F., 2020. Structure of mouse coronavirus spike protein complexed with receptor reveals mechanism for viral entry. *PLoS Pathog.* 16, e1008392.
- Shi, Y., Wang, Y., Shao, C., Huang, J., Gan, J., Huang, X., Bucci, E., Piacentini, M., Ippolito, G., Melino, G., 2020. COVID-19 infection: the perspectives on immune responses. *Cell Death Differ.* 27, 1451–1452.
- Singh, M., Kishore, A., Maity, D., Sunanda, P., Krishnarajuna, B., Vappala, S., Raghobama, S., Kenyon, L.C., Pal, D., Das Sarma, J., 2019. A proline insertion-deletion in the spike glycoprotein fusion peptide of mouse hepatitis virus strongly alters neuropathology. *J. Biol. Chem.* 294, 8064–8087.
- Smith, P.E., van Gunsteren, W.F., 1993. The viscosity of SPC and SPC/E water at 277 and 300 K. *Chem. Phys. Lett.* 215, 315–318.
- Sugita, Y., Okamoto, Y., 1999. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* 314, 141–151.
- Tai, W., He, L., Zhang, X., Pu, J., Voronin, D., Jiang, S., Zhou, Y., Du, L., 2020. Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine. *Cell. Mol. Immunol.* 17 (6), 613–620.
- Tang, T., Bidon, M., Jaimes, J.A., Whittaker, G.R., Daniel, S., 2020. Coronavirus membrane fusion mechanism offers as a potential target for antiviral development. *Antiviral Res.* 178, 104792.
- Vaduganathan, M., Vardeny, O., Michel, T., McMurray, J.J.V., Pfeffer, M.A., Solomon, S. D., 2020. Renin-angiotensin-aldosterone system inhibitors in patients with Covid-19. *N. Engl. J. Med.* 382 (17), 1653–1659.
- van Doremalen, N., Bushmaker, T., Morris, D.H., Holbrook, M.G., Gamble, A., Williamson, B.N., Tamin, A., Harcourt, J.L., Thornburg, N.J., Gerber, S.I., Lloyd-Smith, J.O., de Wit, E., Munster, V.J., 2020. Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *New Engl. J. Med.* 382, 1564–1567.
- Walls, A.C., Park, Y.-J., Tortorici, M.A., Wall, A., McGuire, A.T., Veesler, D., 2020. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* 181, 281–292.
- White, J.M., 1990. Viral and cellular membrane fusion proteins. *Annu. Rev. Physiol.* 52, 675–697.
- Wimley, W.C., White, S.H., 1996. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat. Struct. Biol.* 3, 842–848.
- Wrapp, D., Wang, N., Corbett, K.S., Goldsmith, J.A., Hsieh, C.-L., Abiona, O., Graham, B. S., McLellan, J.S., 2020. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* 367, 1260–1263.
- Xia, S., Liu, M., Wang, C., Xu, W., Lan, Q., Feng, S., Qi, F., Bao, L., Du, L., Liu, S., 2020. Inhibition of SARS-CoV-2 (previously 2019-nCoV) infection by a highly potent pan-coronavirus fusion inhibitor targeting its spike protein that harbors a high capacity to mediate membrane fusion. *Cell Res.* 30 (4), 343–355.
- Zheng, J., 2020. SARS-CoV-2: an emerging coronavirus that causes a global threat. *Int. J. Biol. Sci.* 16, 1678–1685.