



Research article

Mutational analysis of structural proteins of SARS-CoV-2



Shweta Jakhmola^{a,1}, Omkar Indari^{a,1}, Dharmendra Kashyap^{a,1}, Nidhi Varshney^a, Ayan Das^a,
Elangovan Manivannan^b, Hem Chandra Jha^{a,*}

^a Department of Biosciences and Biomedical Engineering, Indian Institute of Technology Indore, Indore, India

^b School of Pharmacy, Devi Ahilya Vishwavidyalaya, Indore, India

ARTICLE INFO

Keywords:

SARS-CoV-2
Spike protein
Mutations
D614G
Integrin
Signaling pathways

ABSTRACT

SARS-CoV-2 transmissibility is higher than that of other human coronaviruses; therefore, it poses a threat to the populated communities. We investigated mutations among envelope (E), membrane (M), and spike (S) proteins from different isolates of SARS-CoV-2 and plausible signaling influenced by mutated virus in a host. We procured updated protein sequences from the NCBI virus database. Mutations were analyzed in the retrieved sequences of the viral proteins through multiple sequence alignment. Additionally, the data was subjected to ScanPROSITE to analyse if the mutations generated a relevant sequence for host signaling. Unique mutations in E, M, and S proteins resulted in modification sites like PKC phosphorylation and N-myristoylation sites. Based on structural analysis, our study revealed that the D614G mutation in the S protein diminished the interaction with T859 and K854 of adjacent chains. Moreover, the S protein of SARS-CoV-2 consists of an Arg-Gly-Asp (RGD) tripeptide sequence, which could potentially interact with various members of integrin family receptors. RGD sequence in S protein might aid in the initial virus attachment. We speculated crucial host pathways which the mutated isolates of SARS-CoV-2 may alter like PKC, Src, and integrin mediated signaling pathways. PKC signaling is known to influence the caveosome/raft pathway which is critical for virus entry. Additionally, the myristoylated proteins might activate NF- κ B, a master molecule of inflammation. Thus the mutations may contribute to the disease pathogenesis and distinct lung pathophysiological changes. Further the frequently occurring mutations in the protein can be studied for possible therapeutic interventions.

1. Introduction

In December 2019, a series of pneumonia cases with clinical symptoms much resembling a viral infection emerged at the Wuhan city of Hubei province, China, with no clear indication of the cause [1, 2]. A while after, the whole genome sequencing of respiratory tract samples from the infected individuals, indicated an unfamiliar beta-coronavirus, which was initially called the 2019 novel coronavirus (2019-nCoV) [1, 3]. On further phylogenetic investigation the International Committee on taxonomy of viruses renamed the virus as severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) [4]. Moreover, a recently published data revealed that the human SARS-CoV-2 shared the highest nucleotide sequence similarity with the bat CoV: RaTG13 [5]. The illness caused by SARS-CoV-2 is referred to as the coronavirus disease of 2019 (COVID-19) [6,7]. Recognized as a global health emergency, COVID-19 was declared a pandemic on 11th March 2020 by the World Health Organisation

(WHO). Since then, the virus has spread to more than 223 countries and territories [8]. As on 27th of January there are 98,280,844 total new SARS-CoV-2 cases, and about 2,112,759 deaths were reported worldwide [8]. Since then researchers are relentlessly trying to address this problem of high transmissibility of the virus from one person to another [9].

Importantly the SARS-CoV-2 proteins that help these viruses to infect the host and propagate inside the host cells include the spike (S), membrane (M) and envelope (E) proteins [10]. Viral attachment is the first step towards establishment of a plausible successful infection in the host. Several *in silico* studies have demonstrated that SARS-CoV-2 utilises its S protein to bind to the angiotensin-converting enzyme 2 (ACE-2) receptor and the outcome was further strengthened by various *in vitro* and *in vivo* studies [11, 12]. Yet another *in silico* study revealed the key residues involved in establishing the stable interaction between S and ACE-2 proteins [13]. The information could be useful in determining the binding affinity of the viral protein to different allelic variants of ACE-2

* Corresponding author.

E-mail address: hemcjha@iiti.ac.in (H.C. Jha).

¹ Equal contributors.

found in COVID-19 patients. Also, the establishment of stable interaction between the viral and host proteins is as important as their binding. A computational study identified mutations in the S protein that increased the affinity of the virus towards human ACE-2 thus maintaining stable connections [14]. Furthermore, scrutiny of drugs from the available databases can aid in procurement of FDA approved therapeutics. Several antivirals have already been investigated for the treatment of SARS-CoV-2 infection [15]. Computation docking studies have been conducted to target 3C like protease (3CLpro/M^{pro}). Drugs like Remdesivir, zanamivir, saquinavir and indinavir are known to potentially target the protease (PDB ID: 6LU7) [16]. Further a docking study concluded that chloroquine preferred binding to the S protein rather than the M^{pro} [16, 17]. An in depth molecular docking and dynamic simulation study demonstrated that a protein inhibitor Δ ABP-D25Y exclusively bound to the S protein ACE-2 binding site, thus the inhibitor could be a potential blocker of S protein and receptor binding domain (RBD) attachment [16, 17, 18].

Numerous investigations have elucidated similarities and differences among the genomes and proteins of different human CoVs using various bioinformatics tools [19]. These studies have helped in understanding the relatedness of the virus with other human CoVs. In the current study we took advantage of the available database containing the information of SARS-CoV-2 proteins and focused on the analysis of various mutations in these proteins from different isolates from multiple geographical regions. Knowing the mutations in various proteins of SARS-CoV-2 may unravel the mystery of higher transmission rates of COVID-19 leading to such a pandemic and might provide a helping hand to target the virus specifically. Studies have also shown that often the drugs used to target the virus may not be equally effective in all the infected individuals [20]. The reason may be the varying host conditions like the immune status, age, gender etc., however it may also be due to variants of the virus that infect the individuals. Thus, variants of SARS-CoV-2 may influence the host differently. Therapeutics targeting a specific variant or a combination of drugs effective against a broad number of viral isolates may be prepared upon having knowledge regarding the mutation in the viruses. In addition, our study will help in the evaluation of the effects of essential mutations in SARS-CoV-2 on the host.

Besides we investigated relevant mutations in the sequences of SARS-CoV-2 structural proteins. We observed that single amino acid substitutions generated unique sites like myristoylation and protein kinase C (PKC) phosphorylation sites. Notably, some modifications in viral proteins are vital to the viruses life cycle, for example, myristoylation and phosphorylation site of S and M protein respectively, enhance the arenavirus and human immunodeficiency virus (HIV) entry into the host cell and further help in the replication and the hijacking of the host cell machinery [21, 22]. Phosphorylation of membrane protein at PKC sites mediates budding, pathogenesis, overall assembly of viral particles, and viral envelope formation in Herpes simplex virus (HSV), HIV, and influenza virus [23, 24, 25]. Various recently published data have established mutation in the RBD of SARS-CoV-2 S protein, which contributed to the enhanced infectivity of the virus [26, 27]. Briefly, the S protein of SARS-CoV-2, a trimeric multifunctional molecule, has two subunits S1 and S2, while the S1 subunit helps to bind with the host receptor, the S2 subunit provides entry into the host cells through the fusion process [26, 28, 29]. Moreover, the S1 subunit has two domains that recognize ACE-2 and ACE-1, and integrin molecules [30, 31, 32]. It is noteworthy to mention that we have identified a conserved Arg-Gly-Asp (RGD) sequence in S protein of SARS-CoV-2 capable of interacting with an array of integrin family receptors [33]. This could contribute to the enhanced ability of the virus to transmit from one person to another. Conclusively, mutations, and modifications in protein sequences could explain the virus adaptability or infectivity pattern [34]. It will also help us understand the host response to the virus and thereby cater to the infection in a much better way. The interactions may alter the host microenvironment and intracellular signaling triggering the

inflammatory pathways. The alterations brought in by the virus may contribute to the COVID-19 associated disease pathogenesis.

2. Materials and methods

2.1. Sequence retrieval and multiple sequence alignment (MSA)

Amino acid sequences of three proteins, namely E, M, and S protein from different isolates of SARS-CoV-2, were retrieved from the NCBI Virus database [35]. The sequences generated from 24th January 2020 to 17th July 2020 were utilized. The isolates with partial or missing amino acid sequences for respective proteins were excluded. Overall for E, M, and S proteins, the sequences from 9463, 6970, and 6476 isolates were considered respectively. Multiple sequence alignments (MSA) of amino acid sequences of all three proteins were performed separately in MEGA X software using clustalW algorithm [36].

2.2. Protein motif and domain/modification site analysis

Protein motifs and modification sites of E, M, and S proteins were analyzed using the online public database tool ScanPROSITE [37, 38].

2.3. Structure retrieval and processing

The crystal structure of the SARS-CoV-2 S protein RBD bound to the ACE-2 receptor (PDB: 6MOJ) and the crystal structure of trimeric S protein with a single RBD in the up conformation (PDB: 6VSB) were retrieved from the RCSB Protein Data Bank (PDB). The amino acid sequences of RBD in 6MOJ and 6VSB structures were identical. Hence utilized in the docking studies with ACE-2 or integrin molecules as per requirement. The missing amino acid residues, domains, and disulfide bonds in the S protein structure were resolved as per previously reported protocols [39, 40]. The RBD from the 6MOJ structure (ACE: RBD assemble) was separated and edited using Chimera 1.14 to create a mutation in the protein structure. Three separate structures of RBD with a point mutation, each namely S477N, V483A, and N501Y were created. Similarly, the protein sequence of 6VSB was edited at 614 position to create D614G mutation. Further, all the proteins were prepared with the protein preparation wizard in MAESTRO v10.6 [41]. During the protein preparation process, hydrogen atoms and the bond orders were assigned. Further formal charges were added to hetero groups in the amino acid residues. The hydrogen-bonding network in amino acid residues was optimized by selecting the histidine tautomers and predicting the ionization states. The optimized protein structure was then subjected to all-atom constrained energy minimization using MAESTRO v10.6 with the OPLS3 force field.

2.4. Molecular docking and interaction analysis

We performed the molecular docking of ACE-2 with wild type or mutated RBD structures using the HDock server [42, 43]. For ACE-2: RBD, a template free and targeted docking was performed considering 24–42, 79–83, and 353–357 residue range of ACE-2 while 446–456 and 486–505 residue range of RBD predicted to participate in the interaction [44]. The docked structures with the lowest root mean square deviation (RMSD) and docking scores were considered for further study. RMSD is the average distance between the atoms of superimposed proteins (with and without bound ligand) and provides information if the two conformers are intrinsically similar or not. Similarly, a template free targeted docking of the SARS-CoV-2 S protein (6VSB) was performed with various integrin molecules using the HDock server. Top 10 models proposed by the software with minimum RMSD and docking score values were analyzed for the interaction of the 403–405 RGD tripeptide sequence in the S protein with the integrin molecules. The protein-protein interactions in the selected docked structures/models were inferred using

interacting residue plots generated by the DIMPLOT tool of Ligplot + v.2.2 [45,46].

2.5. Study of plausible host cell signaling pathway influenced by mutated virus proteins

The information regarding the host cell response for shortlisted protein modification sites obtained as a result of mutations in various virus proteins was acquired through existing literature [47, 48, 49]. The schematic representation and elucidation of possible signaling triggered due to mutated virus proteins were interpreted (Figure 9).

3. Results

3.1. Multiple sequence alignment of SARS-CoV-2 E, M and S proteins revealed unique mutations

We compared three proteins- E, M, and S of various SARS-CoV-2 isolates through MSA. Total 42, 156, and 5449 isolates showed at least one mutation in the E, M, and S protein sequences respectively. Each variation and its proportion are displayed in Figure 1 A-D. A significant mutation observed in nine sequences of E protein was P71L. However, the substitution of the amino acid resulted in a similar type of amino acid, i.e., a nonpolar amino acid. Further, a polar to nonpolar amino acid mutation S68F was observed in five isolates of E protein. L73F mutation, a nonpolar to nonpolar amino acid substitution occurred in four isolates. Remaining one or two mutations per sequence were seen in less than 20

isolates. In M protein, D209Y, T175M, and K15R mutations with all polar to polar amino acid substitution were seen in 20, 19, and 18 isolates respectively. Additionally, mutations that resulted in similar kinds of amino acids were V70F (nonpolar to nonpolar) and H125Y (polar to polar) obtained in more than ten isolates. Nonetheless, mutation with a change in polar to nonpolar amino acid was also observed, D3G, in various isolates. Other modifications were prevalent in less than ten isolates. Compared to E and M proteins, many isolates showed variations in the S protein sequences. The most prominent mutation D614G was observed in 4555 isolates, out of 6476 total isolates. The mutation resulted in the replacement of a hydrophilic and negatively charged D to hydrophobic and neutral G. The second and third prominent mutation was nonpolar aliphatic to an aromatic amino acid, L5F and L54F, each observed in 68 isolates. The S477N (polar to polar) and V483A (nonpolar to nonpolar) mutation was observed in 39 and 22 isolates. Interestingly, these mutations were observed in the RBD of the S protein. All other mutations were found in less than 20 isolates. In addition, we have shown the appearance of particular mutations in E, M and S proteins over time along with the location in which these isolates emerged (Figure 2). The exact date and month of identification of these SARS-CoV-2 mutants is given in table S1.

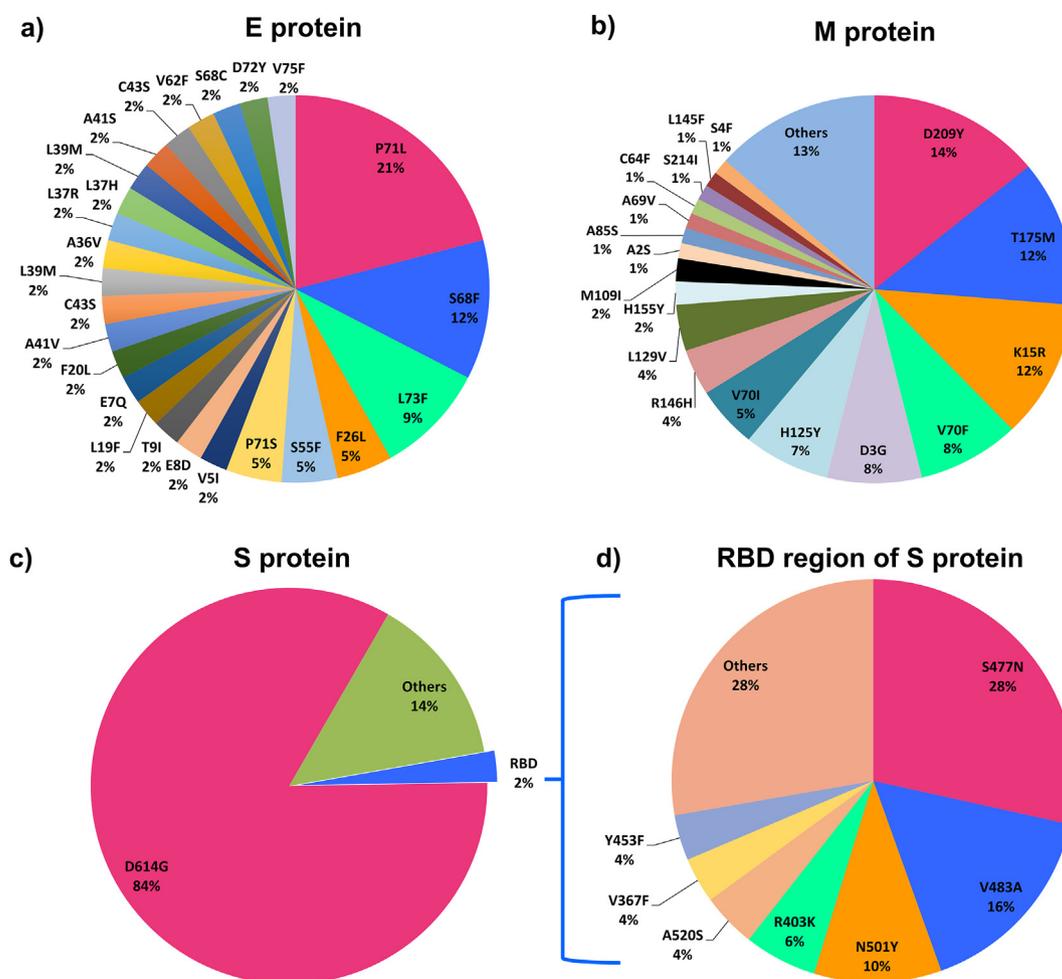


Figure 1. Pie diagram representing the proportions of mutations in a) E protein b) M protein c) S protein and d) RBD region of the S protein. For each protein, the total isolates with at least one mutation considered here are 42, 156 and 5449 for E, M, and S protein respectively.

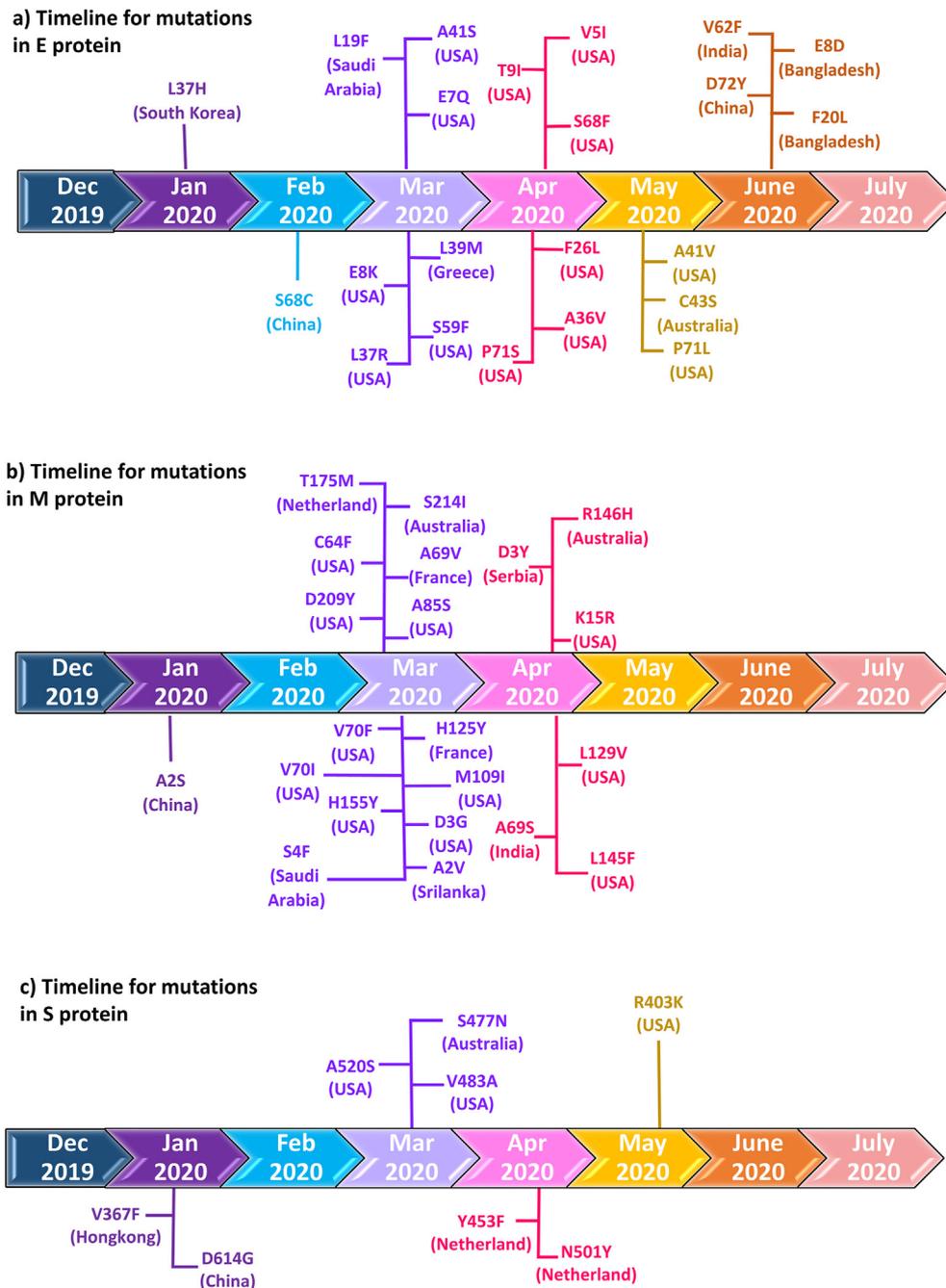


Figure 2. Timeline for mutations in a) E protein b) M protein and c) S protein, appearing for the first time along with country where the mutation was first reported. The top mutations observed in each of the mentioned protein are considered in timeline preparation.

3.2. Mutations in E, M and S protein of SARS-CoV-2 resulted in unique sequences recognised by the host factors

3.2.1. Glutamic acid (E) to lysine (K) mutation forms a PKC phosphorylation site in E protein of SARS-CoV-2

We identified a single mutation, E8K, in an isolate QKO24093.1 from the USA in E protein of SARS-CoV-2 (Figure 3a). The mutation resulted in a PKC phosphorylation site unique to the sequence. Other domains observed in all the protein sequences under investigation included N glycosylation sites at 48–51 and 66–69 positions, and a PKC phosphorylation site at 67–69 position (Figure 3b).

3.2.2. Unique mutations in M protein result in N-myristoylation sites

The D3G mutation was observed in 12 isolates of M protein resulting in N-myristoylation site at 3–8 position (Figure 3d). In addition, a single amino acid substitution V10A in isolate QLF97810.1 at the 10th position of the M protein sequence resulted in N-myristoylation site (Figure 3e). Also, P132S mutation caused the occurrence of a casein kinase II phosphorylation site (Figure 3f). Additionally, all the sequences under consideration included modifications sites namely, N-glycosylation site (NGTI; 5–8 amino acids), casein kinase II phosphorylation site (TVVE; 9–12 and SSSD; 212–215 amino acids), N-myristoylation site (GIAIAM; 79–84 and GTILTR; 126–131), and protein kinase C phosphorylation site

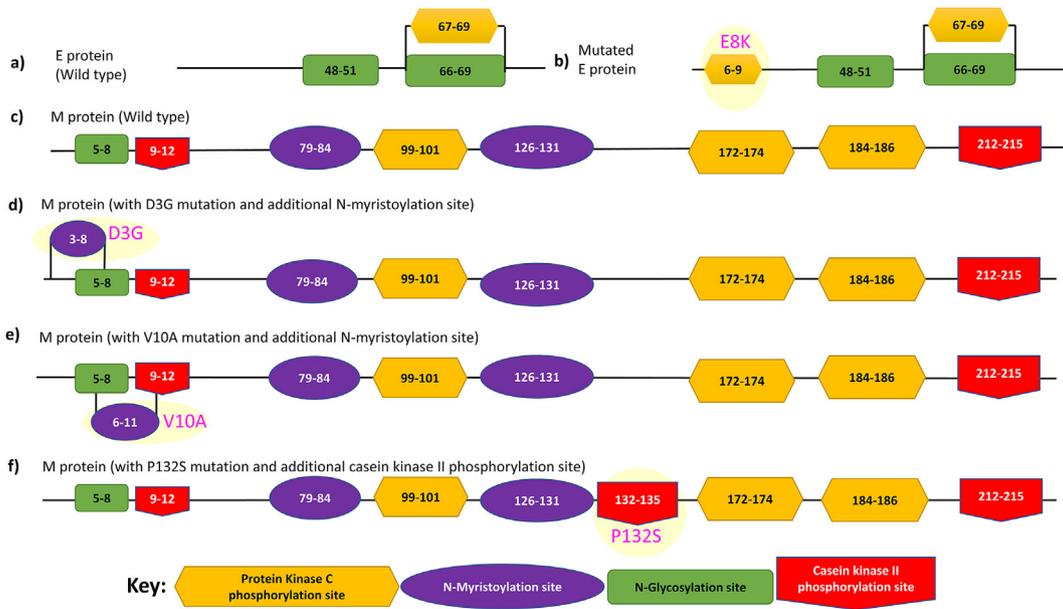


Figure 3. Schematic representation of modification sites in the envelope (E) and membrane (M) protein sequences of SARS-CoV-2 generated due to amino acid substitution. a) In wild type E protein, two N glycosylation sites at 48–52 and 66–69 positions were present. b) Single amino acid substitution E8K in the E protein (QKO24093.1 isolate) resulted in a PKC site at position 6–9 position. c) Wild type M protein contained an N glycosylation site at 5–8, two casein kinase II (CK II) phosphorylation sites at 9–12 and 212–215, two N-myristoylation sites at 79–84 and 126–131, and three PKC phosphorylation sites at 99–101 172–174 and 184–186 positions. d) In addition to the above mentioned modification sites in the M protein, a single point mutation D3G in QIZ16332.1 isolate resulted in a N myristoylation site at 3–8 position. e) Similarly, a single amino acid substitution V10A in the M protein sequence of QLF97810.1 isolate resulted in an N myristoylation site at 6–11 position. f) Furthermore a CK II phosphorylation site was observed at 9–12 position due to a single amino acid substitution P132S in the sequence of QKG90089.1 isolate.

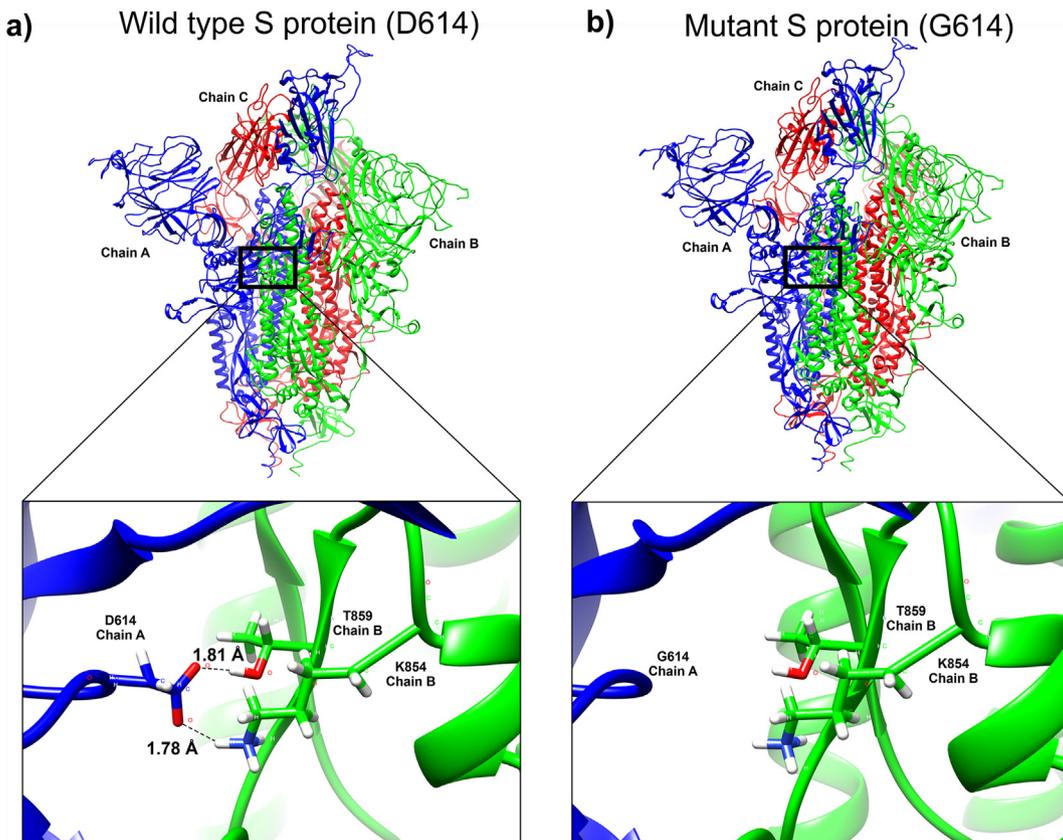


Figure 4. Intramolecular interactions of a) D614 (wild type) and b) G614 (mutant) residue of chain A with chain B residues of the S protein. D614 amino acid in the chain A of SARS-CoV-2 S protein (with up-RBD confirmation), interacted with T869 and K854 of adjacent chains while in the D614G mutant, G614 in chain A did not establish interaction with any nearby residue.

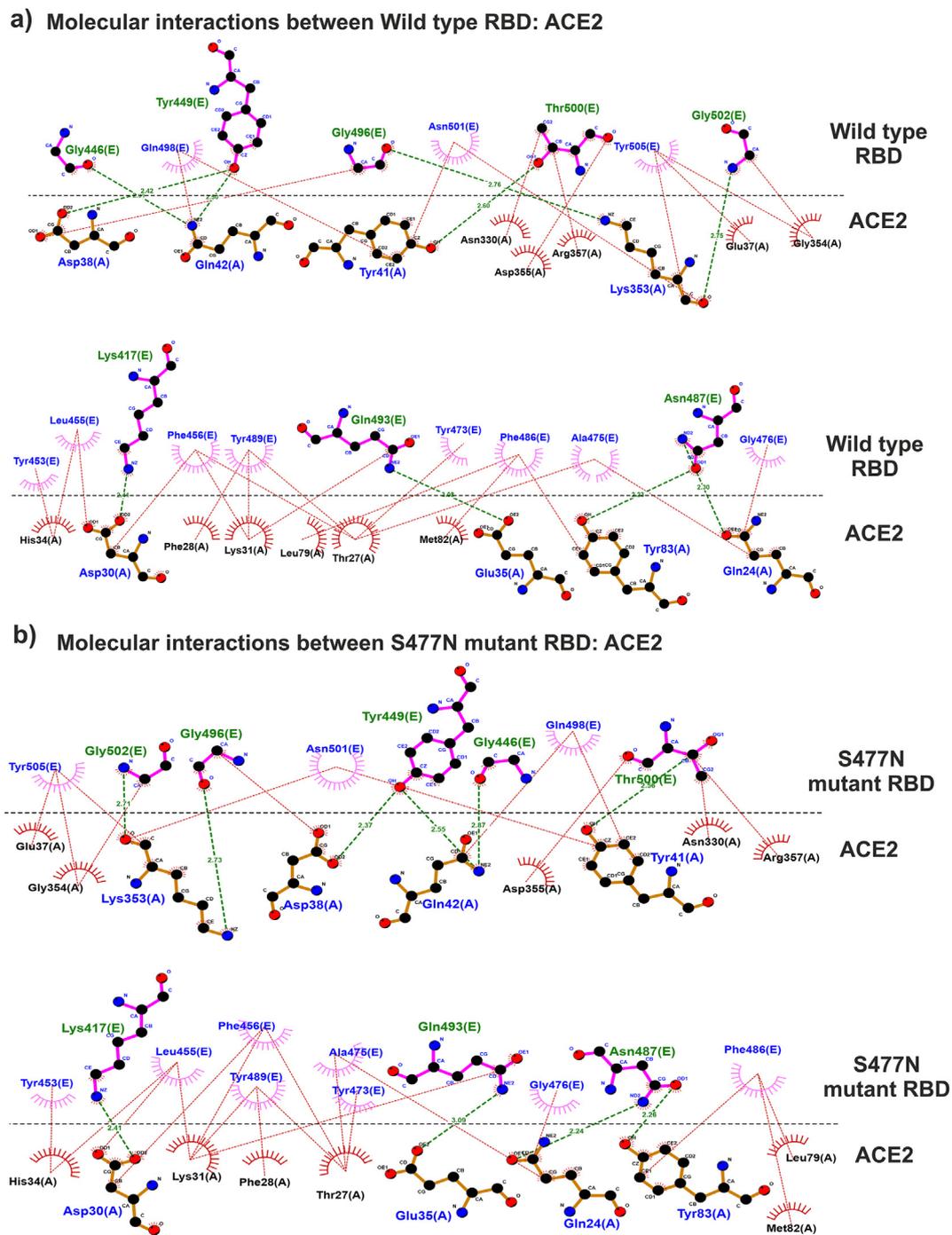


Figure 5. Interacting amino acids between a) Wild Type-RBD: ACE-2; b) S477N-RBD: ACE-2. A template free targeted molecular docking between wild type RBD (PDB: 6MOJ) and ACE-2, and RBD mutant with ACE-2 was performed using HDOCK server. The models with lowest RMSD values and highest docking scores were analysed for interactions using Ligplot + v.2.2 servers.

(SFR; 99–101, TSR; 172–174 and SQR; 184–186 amino acids) (Figure 3c).

3.2.3. Presence of RGD sequence in S protein and D614G mutation resulted in an additional N-myristoylation site in S protein of SARS-CoV-2

In this study, we procured 6476 S protein sequences from different countries. To name a few, we included sequences from Australia, Brazil, China, India, Italy, Japan, Pakistan, South Korea, Spain, Sweden, Taiwan, USA, Finland, Vietnam, etc. The 1273 aa long S protein sequences of the SARS-CoV-2 were subjected to functional domain

analysis using ScanPROSITE. The protein was observed to have a cysteine-rich profile (score = 14.622), with the cysteine-rich region lying between 1235–1254. We further obtained various frequently occurring post translation modification (PTM) sites like the N-glycosylation, PKC phosphorylation, casein kinase II phosphorylation, N-myristoylation, and cAMP cGMP-dependent protein kinase phosphorylation sites along with an RGD tripeptide region. 84% of the mutated S protein sequences possessed D614G mutation. Interestingly this mutation contributed to an additional N-myristoylation site (614–619). Nonetheless, we identified the RGD site at 403–405 amino

acid position in all the isolated protein sequences. The RGD tripeptide sequence obtained in all the S protein sequences might interact with the host integrin and enhance the virus's ability to infect. The virus has, therefore, evolved itself to manipulate the host differently. Various mutations away from the modification sites were also observed.

3.3. Impact of D614G mutation on its intramolecular interaction

In the wild type trimeric structure of S protein, the Aspartate residue at 614 position of A chain with an up RBD confirmation formed a hydrogen bond with T859. Also, it created a salt bridge with K854 of the

adjacent chain, chain B (Figure 4a). However, these interactions were not seen in the mutated structure (D614G mutation) of the S protein (Figure 4b). The lack of onsite acetyl group in the mutated structure prevented such interactions.

3.4. Molecular interactions between ACE-2 and wild type or mutated RBD

Different H-bond and hydrophobic interactions between ACE-2 and wild type or mutated RBD were analyzed using the DIMPLOTT tool of Ligplot + v.2.2 and were displayed in Figures 5 and 6. The critical residues of S protein involved in hydrogen bond formation in wild type RBD: ACE-2 docking was K417, G446, Y449, N487, Q493, G496, T500, and

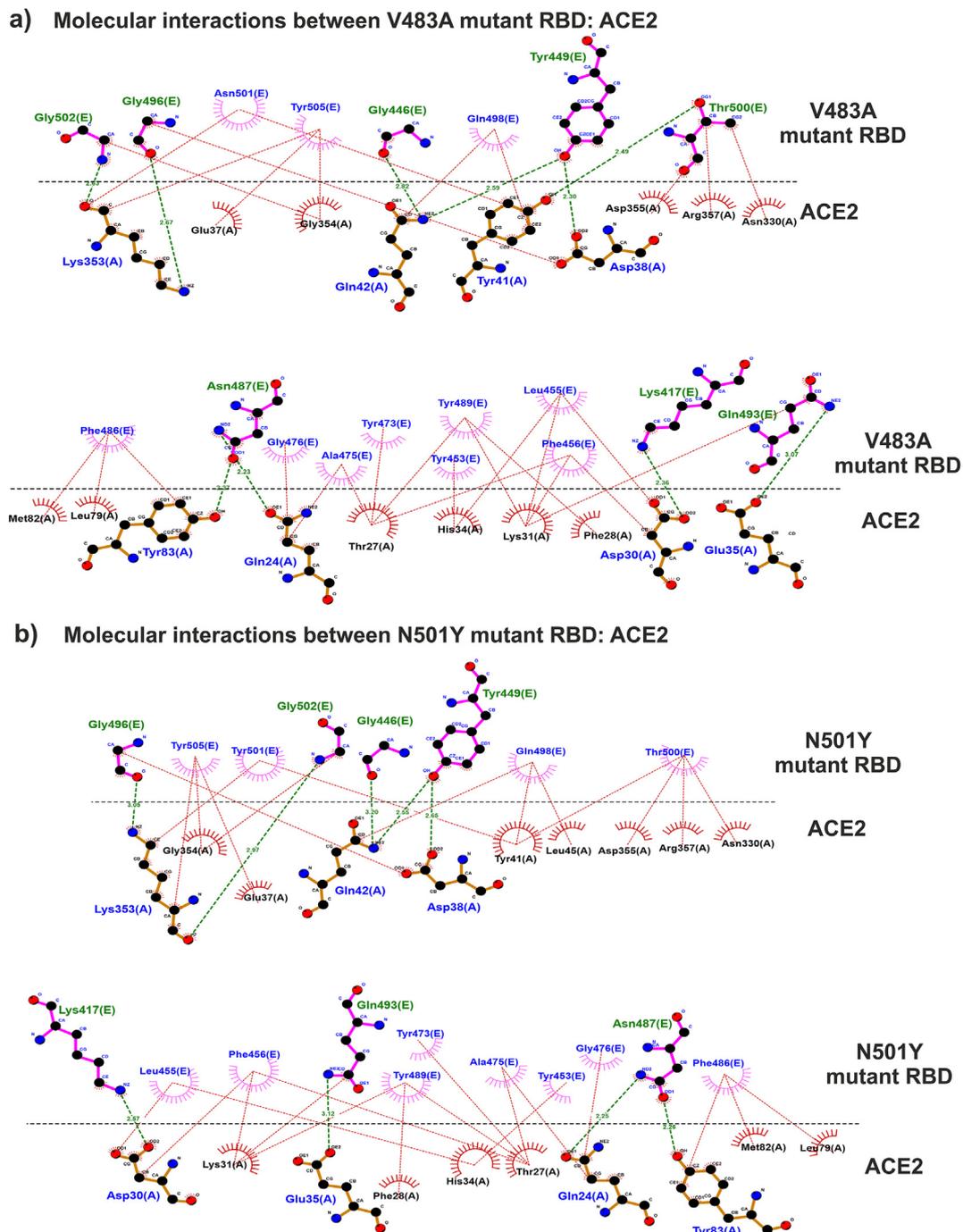


Figure 6. Interacting amino acids between a) V483A-RBD: ACE-2; b) N501Y-RBD: ACE-2. A template free targeted molecular docking between the mutated RBD (prepared from wild type RBD from PDB: 6M0J) with ACE-2 (retrieved from PDB: 6M0J) was performed using HDOCK server. The models with lowest RMSD values and highest docking scores were analysed for interactions using Ligplot + v.2.2 servers.

G502. Similarly, the key residues of ACE-2 protein were Q24, D30, E35, D38, Y41, Q42, Y83, and K353. Briefly, G502 (S) and Y505 (S) interacted with L353 (ACE-2); Y449 (S) with D39 (ACE-2) and Q42 (ACE-2); G446 (S) interacted with Q42 (ACE-2); T500 (S) bonded to Y41 (ACE-2), N487 (S) hydrogen-bonded with Y84 (ACE-2) and Q24 (ACE-2); lastly, L417 (S) and Q493 (S) interacted with D30 (ACE-2) and E35 (ACE-2) respectively. Eleven residues of the S protein and ACE-2 were involved in hydrophobic interactions (Figure 5a). Upon creating mutations within the RBD domain and docking with ACE-2, the same overall residues were involved in the interaction with three exceptions (Figures 5b, 6a and 6b). In S477A and V483A RBD structure docked with ACE-2, additional hydrophobic interaction between L455 (S) with K31 (ACE-2) was observed. Another was in the case of the docked structure of N501Y-RBD: ACE-2 interaction, T500 of S protein failed to form H bond with Y41 of ACE-2 and showed

only hydrophobic interaction. Moreover, Y at the 501th position formed a hydrophobic bond as of N with Y41 (ACE-2). Further, Q498 (S) generated an additional hydrophobic interaction with L45 (ACE-2) in the docked structure of N501Y-RBD: ACE-2.

3.5. Protein-protein docking of SARS-CoV-2 RBD containing RGD tripeptide sequence with integrin molecules

On performing targeted protein-protein docking of the SARS-CoV-2 RBD (6VSB) containing RGD tripeptide sequence with α IIb β 3, we obtained five different models involving the interaction of virus RGD with the integrin molecule. D405 in the virus RBD bound through a hydrogen bond to R139 in Model 1 (Figure 7 a, d), Q451 in model 4 (Figure 7c) and R281 in model 7 (Figure 7e). Additionally, R403 interacted with N227 in

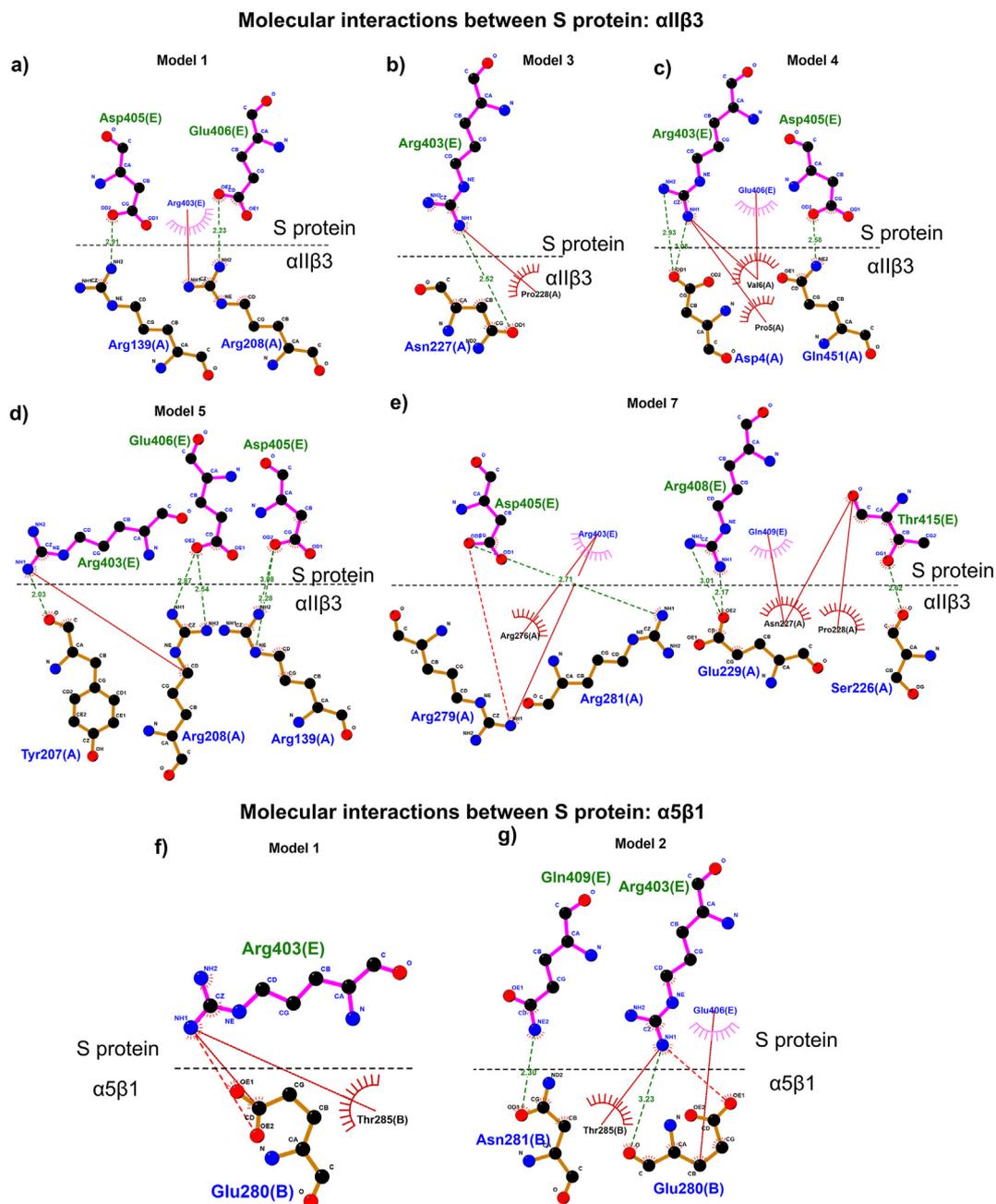


Figure 7. Molecular interactions between RBD of S-protein with integrins [α IIb β 3 (a–e), α 5 β 1 (f–g)]. A template free targeted molecular docking between wild type RBD (PDB: 6VSB) with various integrins was performed using HDock server. The models showing possible interactions between RBD and integrins were analysed for interactions using Ligplot + v.2.2 servers.

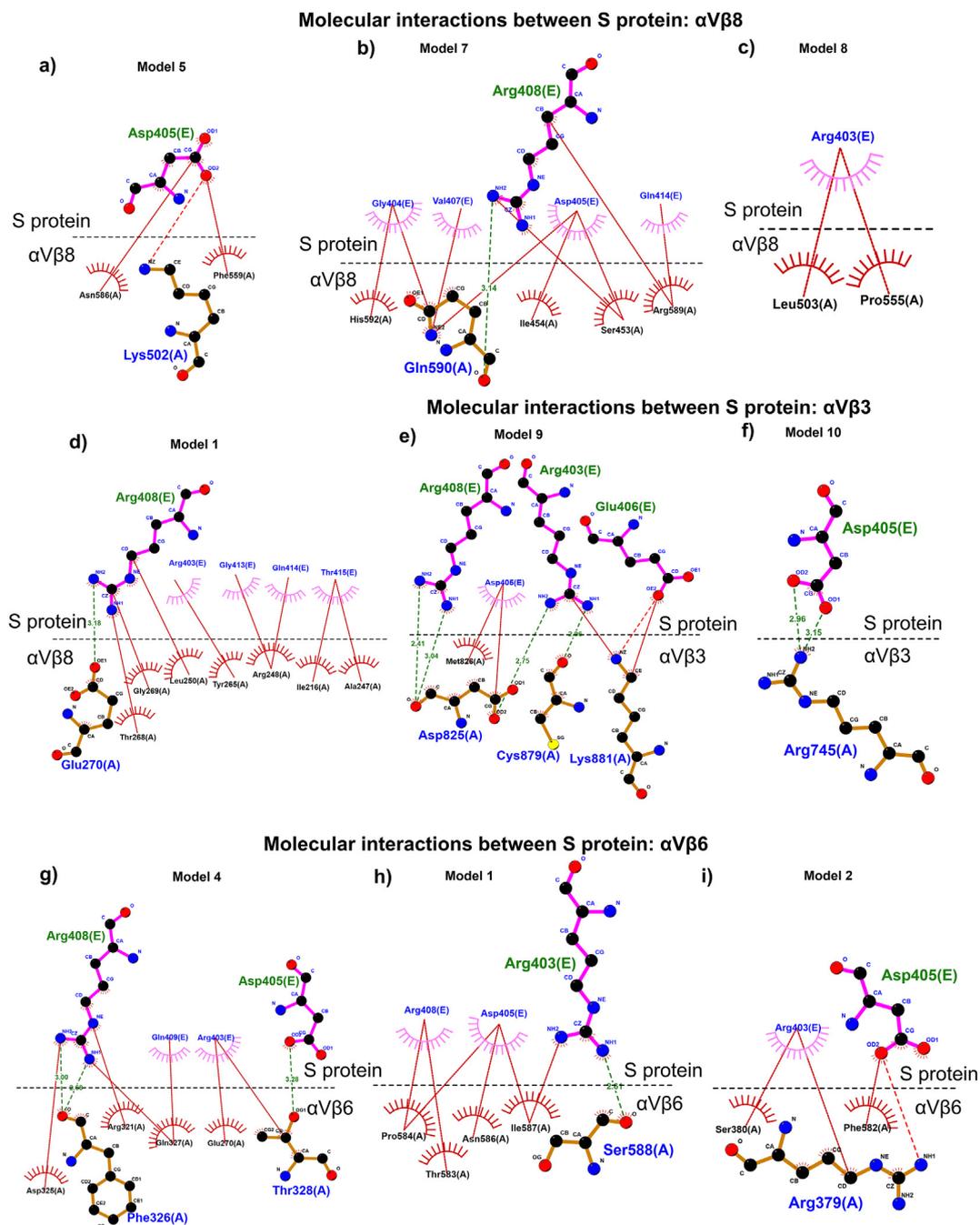


Figure 8. Molecular interactions between RBD of S-protein with integrins [α v β 8 (a–d), α v β 3 (e–f), and α v β 6 (g–i)]. The template free and targeted molecular docking between wild type RBD (retrieved from PDB: 6VSB) with various integrins was performed using HDOCK server. The numerous models showing possible interactions between RBD and integrins were analysed for interactions using Ligplot + v.2.2 servers.

model 3 (Figure 7B), D4 in model 4 (Figure 7C) and Y207 in model 5 (Figure 7d). Weak bonds further stabilized the overall interactions. Specifically, R403 showed hydrophobic interactions with R208 in model 1 (Figure 7A), P228 in model 3 (Figure 7B), P5 in model 4 (Figure 7c), R208 in model 5 (Figure 7D) and R276 and R279 in model 7 (Figure 7e). Similarly, D405 interacted through weak binding with R279 in model 7 (Figure 7e). The interaction of α v β 1 with the SARS-CoV-2 RBD domain mainly involved weak interactions between the RGD of the virus and the integrin molecule. The main hydrophobic interactions included R403 with E280 and T285 in model 1 (Figure 6 f, g). R403 also hydrogen bonded to E280 in model 2 (Figure 7g). Similar to α v β 1 interaction with SARS-CoV-2, α v β 8 interactions with integrin RGD majorly included hydrophobic contacts. R403 interacted with L503 and P555 in model 8

(Figure 8c) and Y265 (Figure 8d) in model 1. D405 showed hydrophobic contacts with F559, N586, and K502 in model 5 (Figure 8A), Q590, I454, and S453 in model 7 (Figure 8b). Interestingly R408 of virus RBD interacted through a hydrogen bond with the integrin molecule in model 1 (Figure 8B) and 7 (Figure 8d) in case of α v β 8 interaction with SARS-CoV-2 RBD. Interestingly, α v β 3 interacted with the SARS-CoV-2 RGD tripeptide region through hydrogen and weak hydrophobic interactions. R403 hydrogen bonded to C879 and D825 in model 9 (Figure 8e). Besides, D405 is hydrogen-bonded to R745 in model 10 (Figure 8f). The hydrophobic contacts included D405 with M826 and D825 and R403 with K881 in model 9 (Figure 8e). Interestingly, D405 of SARS-CoV-2 interacted via hydrogen bond with T328 in model 4 (Figure 8g) and through hydrophobic interactions with P584, N586, I587 in model 1

(Figure 8H), and R379 in model 2 (Figure 8i) of $\alpha\text{v}\beta 6$. Moreover, R403 established hydrogen bonds with S588 in model 1 (Figure 8h). The weak interactions of R403 were observed with E270, T328 in model 4 (Figure 8G), I587 in model 1 (Figure 8H), and S380 and R379 in model 2 (Figure 8i).

3.6. Screening of plausible host signaling influenced by the mutant SARS-CoV-2

Based on the findings of this study regarding the modification sites generated due to mutations, we screened the host's possible signaling pathways. Intriguingly, our results regarding modification sites of the viral E and M proteins enabled us to correlate with possible signaling pathways mediated for SARS-CoV-2 rapid internalization and hijacking of host cellular machinery. Some significant downstream signaling pathways affected by these mutations include; protein phosphorylation, cytoskeleton remodelling, and inflammatory responses (Figure 9).

Briefly, the virus might utilize the clathrin-mediated endocytosis (CME) pathway (Figure 9a) for its entry into the cells and mediate the host intracellular responses through the PKC pathway [50]. Nonetheless, the newly generated PKC sites on the virus E and M protein might get phosphorylated through the PKC pathway. It may facilitate the budding process of the virus as in the case of other viruses. The virus may also gain entry into the host cell using the PKC caveolae/raft mediated endocytosis pathway (Figure 9b). We also looked into the plausible functional aspects regulated by the binding of the virus's RGD to integrin molecules on the host surface. RGD is a conserved motif on SARS-CoV-2 S protein that has a potential to bind to the integrin molecule, which may enhance the binding potential of the protein to ACE-2 [38,51,6,7,52]. Integrins can mediate signaling through the cell membrane by forming adhesome complex, which includes talin, paxillin, vinculin, focal adhesion kinase (FAK), Src, integrin-linked kinase (ILK), p130cas, and GTPase of the Rho family (Figure 9c) [53]. Upon integrin ligation, FAK may get activated. FAK, a tyrosine kinase, possesses an Src Homology 2 (SH2) domain binding site. FAK may bind to the phosphorylated site of integrin and recruit Src. Once the FAK-Src complex is activated, it may lead to the phosphorylation of downstream targets such as Rho-GTPases (RAC1, RhoA, and Cdc42), which primarily aid in cytoskeleton reorganization facilitating virus entry [54]. Additionally, FAK-Src complexes may integrate into the Ras-Raf-MEK-ERK pathway, which regulate critical cellular functions such as cell proliferation and migration [55]. On the other hand, FAK-Src complexes also may phosphorylate downstream targets such as interferon regulatory transcription factor (IRF3), which upon entry into the nucleus activate NF- κ B that could lead to Interferon alpha/beta (IFN α/β), Interleukin 2 (IL2) and IL10 secretions [56]. Viruses utilize their RGD motif to interact with the integrin molecules and initiate these wide varieties of host responses [38, 48, 57]. Intriguingly, integrins are also involved in the innate immune response of the pathogen.

Moreover, we obtained sites of myristoylation on the S and M protein. The virus with myristoylated S protein bound to the ACE-2 receptor may activate the $G\alpha$, which further enables membrane protein phosphatidylinositol bisphosphate (PIP2) (Figure 9d) [22, 48, 57]. Furthermore, PIP2 may activate the phosphoinositide-3 kinase (PIP3) and DAG. Subsequently, the activated PIP3 may interact with the smooth endoplasmic reticulum (SER) and enable the Ca^{2+} channel. The burst of Ca^{2+} ions may subsequently activate the calmodulin. Additionally, activated calmodulin may signal calcineurin, which in turn results in recruitment of the transcription factor, i.e., nuclear factor of activated T cells (NFAT). NFAT may transcribe various genes that help in cell proliferation, differentiation, and inflammation [58]. The virus coils also potentially activate another signaling pathway through binding to the catalytic domain of Src. HCK upon recognition of myristoylated group may activate microtubule-associated protein kinase (MAPK) and PI3K.

Furthermore, MAPK may phosphorylate ERK and PI3K, which sequentially could activate protein kinase B (Akt). Together ERK and Akt may recruit the P70/S6K complex to activate nuclear transcription factor GATA-binding factor 1 (GATA-1). GATA-1 is known to transcribe various genes involved in cell differentiation and lung cell respiratory burst [59]. On the other hand, binding of the virus to the host can stimulate lymphocyte-specific protein tyrosine kinase (LCK) directly through $G\alpha$. Activated LCK kinase may further activate the downstream molecules, namely ZAP-70, ITK, and Grb2 complex. Nonetheless, this complex may cause conversion of Ras-GDP to Ras-GTP through the son of sevenless (SOS), and thus C-Raf. C-Raf may induce MAPK, which may move to the nucleus and activate TF AP1. TF AP1 may help in the transcription of various genes that could help in cell differentiation, proliferation, and apoptosis [47].

4. Discussion

The SARS-CoV-2 pandemic has brought the *coronaviridae* family back in the limelight. Such a case of virus spread necessarily arises from mutations in the viral genome and functional proteins that aid in the adaptation of the virus to a new host [60]. Several efforts have been made to tackle SARS-CoV-2 infection and its spread worldwide. Mutations include the variations upon which natural selection acts, and often result in novelty [61]. The data reported for the age, and gender influenced by the infection, along with the time of recovery of an individual from COVID-19 varied for different countries [52, 62, 63]. This depicted the virus's ability to be ever-evolving making it a better fit to the host. However, the type and rate of mutations in the virus remain enigmatic. In the present study, we have correlated the impact of these mutations with host cell mechanistic, though further *in-vitro* and *in-vivo* studies need to be conducted.

Protein modification is an important phenomenon in the world of the living and could closely be related to the disease pathogenesis. It often makes an organism more suited to an adverse environment [64]. Interestingly, the presence of such sites enables additional modifications in the proteins. The human genome encodes a wide variety of enzymes, namely; histone acetyltransferase, N-myristoylase, and Casein II phosphorylase, which help in protein modification post-translation [64, 65]. Thus, the cells regulate processes like growth, replication, transcription, and translation through these modifications [22, 48]. Furthermore, pathogens like bacteria and viruses also possess such sites, which helps them better adapt to their host environment [66]. Some viruses even code for enzymes that help in virus modification [67, 68]. In our study, we identified several distinct mutations in the E, M, and S protein of SARS-CoV-2. We further speculated and investigated the repercussions of these mutations on the virus driven host processes. Consequently these insights would help to relate viral mutations and thereby generated modification sites, with subsequent SARS-CoV-2 disease pathogenesis. Notably, 9463 isolates of E and 6970 isolates of M proteins were compared to analyze the mutation sites.

E protein, a minor component of the virus membrane, is involved in virus replication and infection processes [69]. All the isolates under investigation possessed a PKC phosphorylation site at 66–69 aa position. In addition E8L mutation in the E protein resulted in a PKC site at 6–8 aa position. In the case of several RNA viruses it is already reported that PKC phosphorylations are crucial in various steps of viral replication [70]. Further, M protein is more prevalent within the virus membrane, and is imperative for coronavirus budding and determining the shape of the viral envelope. During the assembly of the virus particle, M protein interacts with nucleocapsid, E, S, and M protein. Mutations occurring in the M protein could influence the host cell interaction. In the current study, we have observed several PKC sites in the M protein of SARS-CoV-2. Hui and Nayak reported that the PKC phosphorylation sites in viral membranes may change their dynamic ruffling and endocytosis [71]. For example respiratory syncytial virus (RSV) is known to induce PKC

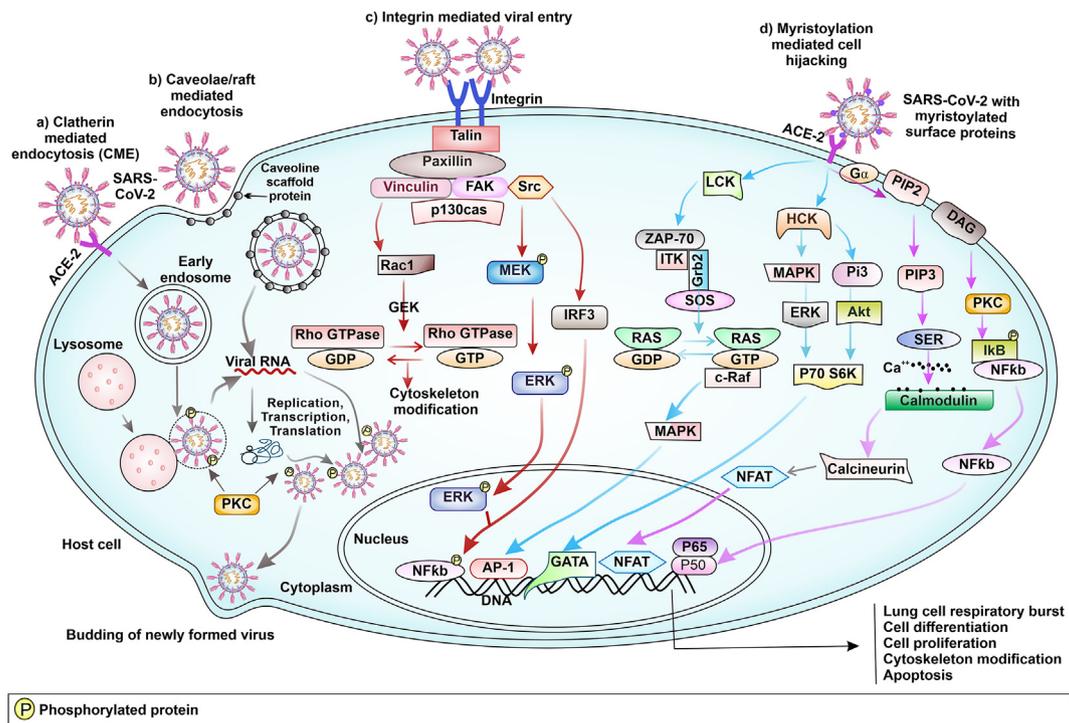


Figure 9. Possible signaling pathways utilized by SARS-CoV-2 for cellular hijacking. a) SARS-CoV-2 may bind to the angiotensin-converting enzyme-2 (ACE-2) receptor and enter the cell through CME mediated pathway. b) Caveolae/raft dependent endocytosis: SARS-CoV-2 may function by attaching to the higher cholesterol regions of the cell membrane. It may facilitate invagination of the cell membrane, and viral complex. c) SARS-CoV-2 may also bind to the integrin through the RGD tripeptide domain and activate the downstream signaling complexes. The activated complex may thus influence the RAS-related C3 botulinum toxin substrate-1 (RAC-1), Mitogen-activated protein kinase kinase (MEK), and interferon regulatory transcription factor (IRF3). d) SARS-CoV-2 may also bind to ACE-2 receptors and activate the downstream $G\alpha$ subunit of ACE-2 receptors and transcription factor NF- κ B leading to transcription of various inflammatory genes.

activation and its cytoplasmic translocation. Also the RSV particles are colocalized with PKC enzyme, required for the fusion of viruses to host cell membranes [25]. PKC helps viruses evade the phagolysosomal pathway through caveolae formation; further allowing viruses such as HIV type 1, filoviruses (Ebola), and simian virus 40 evade the immune system [72]. HSV-1 recruits the PKC, where it phosphorylates lamin B to aid in the nuclear lamina modification and enhance budding at the nuclear membrane [73]. Availability of PKC phosphorylation sites in the viral proteins may facilitate the recruitment of PKC thus mediating the aforementioned processes. To further elaborate on the PKC site's importance, it is noteworthy to state the central role of phosphoprotein P in Borna disease virus (BDV) pathogenesis. A study has demonstrated this phenomenon by using a recombinant BDV with a mutated PKC phosphorylation site. PKC plays a pivotal role in adenovirus infection of corneal fibroblasts and regulation of downstream molecules, including the important lipid raft component caveolin-1 [23].

Nevertheless, we identified myristoylation sites in the M protein. Also D614G mutation in the S protein resulted in a myristoylation site. Myristoylation involves attachment of a myristoyl group to the N-terminal of a glycine from myristic acid [74]. Myristoylation is a critical PTM that happens in a host cell and viruses which lack the transmembrane domain [74]. Viruses like the arena and HIV use this modification to promote the viral budding process [24]. Myristoylation of the virus protein may promote viral infectivity and replication [22, 75, 76]. Besides, Thomas Strecker *et al.*, showed that inhibition of myristoylation prevented the lymphocytic choriomeningitis virus and Lassa viruses [77]. Also, a study found that myristoylation played a crucial role in poliovirus infectivity and processing of structural precursor protein [48]. Moreover, myristoylation of the hypothetical protein of African cassava virus suppressed the host RNA interference. Thus, a mutation on N-terminal glycine will affect virus binding and pathogenicity [78, 79]. Further, Johnson *et al.* demonstrated the importance of Ser/Thr residue at the fifth

position of the consensus G1X2X3X4S/T5X6X7X8 apart from the presence of a glycine at the N-terminal end. This consensus sequence is present in almost all the myristoylated proteins [80].

Moreover, the virus may enter the host cell through a receptor-mediated mechanism that requires the adherence of the virus with the host receptor. The attachment process is a crucial first step towards establishing a successful infection. The severity of COVID-19 and an enhanced transmission ability of SARS-CoV-2 is related to its increased potential to bind to host ACE-2 [81,82]. Moreover, the virus's cross-species migration involves the ability to engage ACE-2, varying among the animal species [83]. We analyzed the binding of wild type and mutant SARS-CoV-2 S protein to ACE-2 compared to the wild type. Also, unlike previously reported studies [44], we observed two extra residues 473 and 475 of S proteins involved, while R393 of ACE-2 was missing from the interaction. However, the influence of these interactions on virus internalization and pathogenesis needs to be evaluated. Additionally, numerous anti-SARS-CoV-2 vaccines are currently under different phases of clinical trials. Most of the vaccines are designed targeting the S-protein to inhibit the viral entry and further transmission in the body. To check the effect of distinct mutations against the vaccine efficacy is of crucial importance. The widespread mutation D614G lies at the inner side at the S1-S2 junction of S protein and don't lead to any change on epitopes or surface structure of S protein and hence may not get impacted by vaccines [84, 85, 86]. However, the mutations in the RBD region of S protein could affect the efficacy of vaccines. A study by Starr *et al.* [86], mapped viral mutations that enabled the virus to escape the antibodies against SARS-CoV-2. However, the study utilized only two types of antibodies in a cocktail i.e., Regeneron's REGN-COV2 cocktail and Eli Lilly's LY-CoV016 antibody. Therefore the question still remains regarding the efficacy of vaccines under preparation against mutated viruses. Recently two new variants of SARS-CoV-2 have been identified

to have greater transmission in the reported areas. The UK variant with 23 distinct mutations known as Variant of Concern, year 2020, month 12, variant 01 (VOC 202012/01) was reported in December 2020. The South Africa variant named 501Y.V2 was also reported shortly after the UK variant in December 2020. Both the variants carry N501Y mutation lying in the RBD region of S protein. Our study has demonstrated that this mutation has maintained the interaction with ACE-2 with involvement of an additional residue (Q498) of S protein. Currently various studies are under investigation to check efficacy of the vaccines under trials against variants carrying this mutation. Further, the crystal structure of S protein generated by Wrapp *et al* [87] depicted that D614 residue interacted with T859 of an adjacent chain when any one of the chains has RBD in up conformation. Similar observation has been reported in an earlier study as well [85]. Another recent report [88] has shown that D614 forms salt bridges with K854 of the fusion peptide proximal region (FPPR). The report suggested that the D614-K854 interaction supported the role of FPPR in membrane fusion. Our study denoted the interaction of D614 with T859 and K854, which further led us to conclude D614 interaction with both residues might be reinforcing the role of FPPR. The results also depicted the elimination of this intramolecular interaction on D614G mutation. However, relevant structure-based studies are required for further validation.

In addition, Wrapp *et al* demonstrated a significantly higher affinity of S protein of SARS-CoV-2 towards ACE-2 [89]. Studies using cryo-EM have elucidated the interaction at the Angstrom level [63, 87]. Furthermore, a study documented the importance of the RGD sequence in SARS-CoV-2 transmission and pathology. We have validated the presence of the mutation in all of the isolates. The RGD sequence is located at 403–405 neighbouring to the ACE-2 interaction site. Integrin's, the multifunctional heterodimeric cell surface receptor molecules, serve as potent receptors utilizing the RGD thus facilitating the entry process of numerous viruses. Viruses like the adenoviruses and herpesviruses, exhibit an RGD (Arg-Gly-Asp) sequence on their virus surfaces [90, 91]. Notably, integrin's namely $\alpha 5\beta 1$, $\alpha 8\beta 1$, $\alpha \nu\beta 1$, $\alpha \nu\beta 3$, $\alpha \nu\beta 8$, $\alpha \nu\beta 5$, $\alpha \nu\beta 6$, may recognize the RGD region on the host cell or the SARS-CoV-2 [92, 93, 94, 95]. We found that various members of the integrin family react differently with the SARS-CoV-2 RBD domain-containing RGD tripeptide sequence. Integrin heterodimers such as $\alpha \text{IIb}\beta 3$, $\alpha \nu\beta 3$, $\alpha \nu\beta 6$ interact with strong hydrogen bonds with R403 and D405 of SARS-CoV-2 RBD. While some integrins like $\alpha 5\beta 1$, $\alpha 5\beta 1$ are weakly bonded mainly involving hydrophobic interactions with R403 and D405 of SARS-CoV-2. Therefore, the expression of a particular type of integrin on the cell may contribute to strong or weak attachment of SARS-CoV-2 to its surface. Furthermore, integrin's favoured the cell to cell and cell to pathogen attachment and may be necessary to induce leukocyte migration aggravating the inflammatory reactions [96]. Thus, the RGD sequence of the S protein of SARS-CoV-2 may also be recognized and recruited by integrin's in alveolar epithelial cells to accelerate the infection process [38, 51, 52, 53, 97]. Earlier studies have shown that blocking or deleting the RGD reduced the viral DNA uptake [98]. Thus, RGD interactions with the integrin's seem to regulate not only the virus entry process but also the associated pathogenesis [38]. Therefore, identification of potential blockers that prevent the S protein binding to the ACE-2 and integrin molecules can be performed. Targeting this interaction might provide possible therapeutic interventions for COVID-19 treatment. The receptor-ligand interaction can thereby alter a plethora of intracellular signaling pathways like PI-3K, FAK, Rho GTPases, Src, and diaphanous 2 (Dia2)-associated signaling, which is necessary for the internalization of the virus [55]. Nonetheless, phosphorylated integrin's can activate the FAK-Src complex, which further phosphorylates downstream targets such as Rho-GTPases (RAC1, RhoA, and Cdc42) that ultimately lead to cytoskeleton reorganization to facilitate virus entry inside the cell [54]. Additionally, integrin's have a distinctive property of shifting their ligand binding conformations from high to low, followed by signal transduction [33]. In some cases, the inactive integrin's can bind to the ligand. In other cases, the extracellular domain of the protein needs to undergo

conformational changes brought in by the intracellular signaling modulations. Possibly, we hypothesize that the binding of the S protein to ACE-2 might provide this inside out signaling to the integrin that might help stabilize the interaction. The sequence might further facilitate SARS-CoV-2 infection and be responsible for its increased ability to transmit from person to person. However, a detailed study needs to be conducted.

5. Conclusion

We hypothesize that this rapidly evolving RNA virus is capable of undergoing mutations that could impart the virus with an ability to manipulate the host and thereby establish itself successfully. We have obtained some significant differences in E, M, and S protein, which may contribute to an enhanced ability of SARS-CoV-2 to infect epithelial cells better than SARS-CoV. We observed aspartic acid (D) to glycine (G) mutation at 614th position of S protein of 84% of the isolates, forming a myristoylation site [22]. Also, we noted the presence of a conserved tripeptide sequence RGD. Both the mutations may significantly contribute to viruses' enhanced potential to transmit among the population. Furthermore, the virus can regulate the critical intracellular signaling pathways like PKC signaling, myristoylation signaling and immune regulatory pathways, affecting the disease pathogenesis [22]. What needs to be seen is if these mutations contribute to enhanced pathogenicity or reduced capability of the virus to infect the host cells.

Declarations

Author contribution statement

Shweta Jakhmola, Omkar Indari, Dharmendra Kashyap: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Nidhi Varshney: Performed the experiments; Contributed reagents, materials, analysis tools or data.

Ayan Das: Performed the experiments.

Elangovan Manivannan: Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Hem Chandra Jha: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.

Funding statement

This work was supported by the Council for Scientific and Industrial Research, India grant no 37 (1693)/17/EMR-II and Department of Science and Technology, India (DST) as Ramanujan fellowship grant no. SB/S2/RJN-132/20/5. We are thankful to the Ministry of Human Resource, Department of Biotechnology, Council of Scientific & Industrial Research, DST-Inspire of Government of India for fellowship to Shweta Jakhmola, Omkar Indari, Dharmendra Kashyap, Nidhi Varshney respectively in the form of research stipend.

Data availability statement

Data included in article/supp. material/referenced in article.

Competing interest statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at <https://doi.org/10.1016/j.heliyon.2021.e06572>.

Acknowledgements

We appreciate our lab colleagues for insightful discussions and advice. We gratefully acknowledge the Indian Institute of Technology Indore for providing facilities and support.

References

- [1] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Z. Cheng, T. Yu, J. Xia, Y. Wei, W. Wu, X. Xie, W. Yin, H. Li, M. Liu, Y. Xiao, H. Gao, L. Guo, J. Xie, G. Wang, R. Jiang, Z. Gao, Q. Jin, J. Wang, B. Cao, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, *Lancet* 395 (2020) 497–506.
- [2] K. Bavarisetti, Review of “A Pneumonia Outbreak Associated with a New Coronavirus of Probable Bat Origin,” 2020.
- [3] P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li, C.-L. Huang, H.-D. Chen, J. Chen, Y. Luo, H. Guo, R.-D. Jiang, M.-Q. Liu, Y. Chen, X.-R. Shen, X. Wang, X.-S. Zheng, K. Zhao, Q.-J. Chen, F. Deng, L.-L. Liu, B. Yan, F.-X. Zhan, Y.-Y. Wang, G.-F. Xiao, Z.-L. Shi, A pneumonia outbreak associated with a new coronavirus of probable bat origin, *Nature* 579 (2020) 270–273.
- [4] Y.-R. Guo, Q.-D. Cao, Z.-S. Hong, Y.-Y. Tan, S.-D. Chen, H.-J. Jin, K.-S. Tan, D.-Y. Wang, Y. Yan, The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak - an update on the status, *Mil. Med. Res.* 7 (2020) 11.
- [5] K.G. Andersen, A. Rambaut, W. Ian Lipkin, E.C. Holmes, R.F. Garry, The proximal origin of SARS-CoV-2, *Nat. Med.* 26 (2020) 450–452.
- [6] H. Li, S.-M. Liu, X.-H. Yu, S.-L. Tang, C.-K. Tang, Coronavirus disease 2019 (COVID-19): current status and future perspectives, *Int. J. Antimicrob. Agents* 55 (2020) 105951.
- [7] S. Jakhmola, O. Indari, D. Kashyap, N. Varshney, A. Rani, C. Sonkar, B. Baral, S. Chatterjee, A. Das, R. Kumar, H.C. Jha, Recent updates on COVID-19: a holistic review, *Heliyon* 6 (2020), e05706.
- [8] Novel Coronavirus, Situation Reports, nCoV, 2019, <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>. (Accessed 3 February 2021).
- [9] J.L. Geoghegan, A.M. Senior, F. Di Giallonardo, E.C. Holmes, Virological factors that increase the transmissibility of emerging human viruses, *Proc. Natl. Acad. Sci. Unit. States Am.* 113 (2016) 4170–4175.
- [10] I. Astuti, Ysrafil, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2): an overview of viral structure and host response, *Diabetes Metab. Syndr.* 14 (2020) 407–412.
- [11] A. Choudhury, S. Mukherjee, In silico studies on the comparative characterization of the interactions of SARS-CoV-2 spike glycoprotein with ACE-2 receptor homologs and human TLRs, *J. Med. Virol.* 92 (2020) 2105–2113.
- [12] P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li, C.-L. Huang, H.-D. Chen, J. Chen, Y. Luo, H. Guo, R.-D. Jiang, M.-Q. Liu, Y. Chen, X.-R. Shen, X. Wang, X.-S. Zheng, K. Zhao, Q.-J. Chen, F. Deng, L.-L. Liu, B. Yan, F.-X. Zhan, Y.-Y. Wang, G.-F. Xiao, Z.-L. Shi, Addendum: a pneumonia outbreak associated with a new coronavirus of probable bat origin, *Nature* 588 (2020) E6.
- [13] E. Laurini, D. Marson, S. Attilic, M. Fermoeglia, S. Pricl, Computational alanine scanning and structural analysis of the SARS-CoV-2 spike protein/angiotensin-converting enzyme 2 complex, *ACS Nano* 14 (2020) 11821–11830.
- [14] J.D. Blanco, X. Hernandez-Alias, D. Cianferoni, L. Serrano, In silico mutagenesis of human ACE2 with S protein and translational efficiency explain SARS-CoV-2 infectivity in different species, *PLoS Comput. Biol.* 16 (2020), e1008450.
- [15] O. Indari, S. Jakhmola, M. Elangovan, H.C. Jha, An update on antiviral therapy against SARS-CoV-2: how far have we come? Review, *Front. Pharmacol. - Exp. Pharmacol. Drug Discov.* (2020).
- [16] D.C. Hall Jr., H.-F. Ji, A search for medications to treat COVID-19 via in silico molecular docking models of the SARS-CoV-2 spike glycoprotein and 3CL protease, *Trav. Med. Infect. Dis.* 35 (2020) 101646.
- [17] R.S. Cheke, The molecular docking study of potential drug candidates showing anti-COVID-19 activity by exploring of therapeutic targets of SARS-CoV-2, *Eurasian J. Med. Oncol.* 4 (2020) 185–195.
- [18] G. Jaiswal, V. Kumar, In-silico design of a potential inhibitor of SARS-CoV-2 S protein, *PLoS One* 15 (2020), e0240004.
- [19] S. Srinivasan, H. Cui, Z. Gao, M. Liu, S. Lu, W. Mkandawire, O. Narykov, M. Sun, D. Korin, Structural genomics of SARS-CoV-2 indicates evolutionary conserved functional regions of viral proteins, *Viruses* 12 (2020) 360.
- [20] WHO Solidarity Trial Consortium, H. Pan, R. Peto, A.-M. Henao-Restrepo, M.-P. Preziosi, V. Sathiyamoorthy, Q. Abdool Karim, M.M. Alejandria, C. Hernández García, M.-P. Kiemy, R. Malekzadeh, S. Murthy, K.S. Reddy, M. Roses Periago, P. Abi Hanna, F. Ader, A.M. Al-Bader, A. Alhasawi, E. Allum, A. Alotaibi, C.A. Alvarez-Moreno, S. Appadoo, A. Asiri, P. Aukrust, A. Barratt-Due, S. Bellani, M. Branca, H.B.C. Cappel-Porter, N. Cerrato, T.S. Chow, N. Como, J. Eustace, P.J. García, S. Godbole, E. Gotuzzo, L. Grisevicius, R. Hamra, M. Hassan, M. Hassany, D. Hutton, I. Irmansyah, L. Jancoriene, J. Kirwan, S. Kumar, P. Lennon, G. Lopardo, P. Lydon, N. Magrini, T. Maguire, S. Manevska, O. Manuel, S. McGinty, M.T. Medina, M.L. Mesa Rubio, M.C. Miranda-Montoya, J. Nel, E.P. Nunes, M. Perola, A. Portolés, M.R. Rasmín, A. Raza, H. Rees, P.P.S. Reges, C.A. Rogers, K. Salami, M.I. Salvadori, N. Sinani, J.A.C. Sterne, M. Stevanovikj, E. Tacconelli, K.A.O. Tikkinen, S. Trelle, H. Zaid, J.-A. Røttingen, S. Swaminathan, Repurposed antiviral drugs for covid-19 - interim WHO solidarity trial results, *N. Engl. J. Med.* (2020).
- [21] M. Benthams, Role of myristoylation and N-terminal basic residues in membrane association of the human immunodeficiency virus type 1 Nef protein, *J. Gen. Virol.* 87 (2006) 563–571.
- [22] S. Maurer-Stroh, F. Eisenhaber, Myristoylation of viral and bacterial proteins, *Trends Microbiol.* 12 (2004) 178–185.
- [23] M.A. Yousuf, J.S. Lee, X. Zhou, M. Ramke, J.Y. Lee, J. Chodosh, J. Rajaiya, Protein kinase C signaling in adenoviral infection, *Biochemistry* 55 (2016) 5938–5946.
- [24] L. Chen, O.T. Keppler, C. Schölz, Post-translational modification-based regulation of HIV replication, *Front. Microbiol.* 9 (2018) 2131.
- [25] H. San-Juan-Vergara, M.E. Peebles, R.F. Lockey, S.S. Mohapatra, Protein kinase C- α activity is required for respiratory syncytial virus fusion to human bronchial epithelial cells, *J. Virol.* 78 (2004) 13717–13726.
- [26] R.N. Kirchdoerfer, C.A. Cottrell, N. Wang, J. Pallesen, H.M. Yassine, H.L. Turner, K.S. Corbett, B.S. Graham, J.S. McLellan, A.B. Ward, Pre-fusion structure of a human coronavirus spike protein, *Nature* 531 (2016) 118–121.
- [27] M. Letko, A. Marzi, V. Munster, Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses, *Nat. Microbiol.* 5 (2020) 562–569.
- [28] S. Belouzard, J.K. Millet, B.N. Licitra, G.R. Whittaker, Mechanisms of coronavirus cell entry mediated by the viral spike protein, *Viruses* 4 (2012) 1011–1033.
- [29] K. Wu, G. Peng, M. Wilken, R.J. Geraghty, F. Li, Mechanisms of host receptor adaptation by severe acute respiratory syndrome coronavirus, *J. Biol. Chem.* 287 (2012) 8904–8911.
- [30] F. Li, Structural analysis of major species barriers between humans and palm civets for severe acute respiratory syndrome coronavirus infections, *J. Virol.* 82 (2008) 6984–6991.
- [31] R. Yan, Y. Zhang, Y. Li, L. Xia, Y. Guo, Q. Zhou, Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2, *Science* 367 (2020) 1444–1448.
- [32] C. Tikellis, M.C. Thomas, Angiotensin-converting enzyme 2 (ACE2) is a key modulator of the renin angiotensin system in health and disease, *Int. J. Pept. 2012* (2012) 256294.
- [33] H.A.M. Hussein, L.R. Walker, U.M. Abdel-Raouf, S.A. Desouky, A.K.M. Montasser, S.M. Akula, Beyond RGD: virus interactions with integrins, *Arch. Virol.* 160 (2015) 2669–2681.
- [34] R. Sanjuán, P. Domingo-Calap, Mechanisms of viral mutation, *Cell. Mol. Life Sci.* 73 (2016) 4433–4448.
- [35] NCBI virus. <https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/>. (Accessed 3 February 2021).
- [36] S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, X. Mega, Molecular evolutionary genetics analysis across computing platforms, *Mol. Biol. Evol.* 35 (2018) 1547–1549.
- [37] E. de Castro, C.J.A. Sigrist, A. Gattiker, V. Bulliard, P.S. Langendijk-Genevaux, E. Gasteiger, A. Bairoch, N. Hulo, ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins, *Nucleic Acids Res.* 34 (2006) W362–W365.
- [38] C.J. Sigrist, A. Bridge, P. Le Mercier, A potential role for integrins in host cell entry by SARS-CoV-2, *Antivir. Res.* 177 (2020) 104759.
- [39] O.C. Grant, D. Montgomery, K. Ito, R.J. Woods, Analysis of the SARS-CoV-2 spike protein glycan shield reveals implications for immune recognition, *Sci. Rep.* 10 (2020) 14991.
- [40] H. Woo, S.-J. Park, Y.K. Choi, T. Park, M. Tanveer, Y. Cao, N.R. Kern, J. Lee, M.S. Yeom, T.I. Croll, C. Seok, W. Im, Developing a fully glycosylated full-length SARS-CoV-2 spike protein model in a viral membrane, *J. Phys. Chem. B* 124 (2020) 7128–7137.
- [41] S.P. Niinivehmas, E. Manivannan, S. Rauhamäki, J. Huuskonen, O.T. Pentikäinen, Identification of estrogen receptor α ligands with virtual screening techniques, *J. Mol. Graph. Model.* 64 (2016) 30–39.
- [42] Y. Yan, H. Tao, J. He, S.-Y. Huang, The HDock server for integrated protein–protein docking, *Nat. Protoc.* 15 (2020) 1829–1852.
- [43] Y. Yan, Z. Wen, X. Wang, S.-Y. Huang, Addressing recent docking challenges: a hybrid strategy to integrate template-based and free protein–protein docking, *Proteins* 85 (2017) 497–512.
- [44] J. Lan, J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, Q. Zhang, X. Shi, Q. Wang, L. Zhang, X. Wang, Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor, *Nature* 581 (2020) 215–220.
- [45] R.A. Laskowski, M.B. Swindells, LigPlot+: multiple ligand–protein interaction diagrams for drug discovery, *J. Chem. Inf. Model.* 51 (2011) 2778–2786.
- [46] A.C. Wallace, R.A. Laskowski, J.M. Thornton, LIGPLOT: a program to generate schematic diagrams of protein–ligand interactions, *Protein Eng.* 8 (1995) 127–134.
- [47] L.D. Notarangelo, Combined immunodeficiencies with nonfunctional T lymphocytes, *Adv. Immunol.* 121 (2014) 121–190.
- [48] I. Corbic Ramljak, J. Stanger, A. Real-Hohn, D. Dreier, L. Wimmer, M. Redlberger-Fritz, W. Fischl, K. Klingel, M.D. Mihovilovic, D. Blaas, H. Kowalski, Cellular N-myristoyltransferases play a crucial picornavirus genus-specific role in viral assembly, virion maturation, and infectivity, *PLoS Pathog.* 14 (2018), e1007203.
- [49] S.E. Byeon, Y.-S. Yi, J. Oh, B.C. Yoo, S. Hong, J.Y. Cho, The role of Src kinase in macrophage-mediated inflammatory responses, *Mediat. Inflamm.* 2012 (2012) 512926.

- [50] P. Cossart, A. Helenius, Endocytosis of viruses and bacteria, *Cold Spring Harb. Perspect. Biol.* 6 (2014) a016972, a016972.
- [51] J. Shang, Y. Wan, C. Luo, G. Ye, Q. Geng, A. Auerbach, F. Li, Cell entry mechanisms of SARS-CoV-2, *Proc. Natl. Acad. Sci. U.S.A.* 117 (2020) 11727–11734.
- [52] S. Jakhmola, O. Indari, B. Baral, D. Kashyap, N. Varshney, A. Das, S. Chatterjee, H.C. Jha, Comorbidity assessment is essential during COVID-19 treatment, *Front. Physiol.* 11 (2020) 984.
- [53] J.T. Parsons, A.R. Horwitz, M.A. Schwartz, Cell adhesion: integrating cytoskeletal dynamics and cellular tension, *Nat. Rev. Mol. Cell Biol.* 11 (2010) 633–643.
- [54] C.M. Teoh, S.S.L. Tan, T. Tran, Integrins as therapeutic targets for respiratory diseases, *Curr. Mol. Med.* 15 (2015) 714–734.
- [55] K. Ley, C. Laudanna, M.I. Cybulsky, S. Nourshargh, Getting to the site of inflammation: the leukocyte adhesion cascade updated, *Nat. Rev. Immunol.* 7 (2007) 678–689.
- [56] T. Gianni, G. Campadelli-Fiume, The epithelial $\alpha v \beta 3$ -integrin boosts the MYD88-dependent TLR2 signaling in response to viral and bacterial components, *PLoS Pathog.* 10 (2014), e1004477.
- [57] S. Yan, H. Sun, X. Bu, G. Wan, New strategy for COVID-19: an evolutionary role for RGD motif in SARS-CoV-2 and potential inhibitors for virus infection, *Front. Pharmacol.* 11 (2020) 912.
- [58] J. Zhang, H. Feng, S. Xu, P. Feng, Hijacking GPCRs by viral pathogens and tumor, *Biochem. Pharmacol.* 114 (2016) 69–81.
- [59] F.M. Roversi, F.V. Pericole, J.A. Machado-Neto, A. da Silva Santos Duarte, A.L. Longhini, F.A. Corrocher, B. Palodetto, K.P. Ferro, R.G. Rosa, M.O. Baratti, S. Verjovski-Almeida, F. Traina, A. Molinari, M. Botta, S.T.O. Saad, Hematopoietic cell kinase (HCK) is a potential therapeutic target for dysplastic and leukemic cells due to integration of erythropoietin/PI3K pathway and regulation of erythropoiesis: HCK in erythropoietin/PI3K pathway, *Biochim. Biophys. Acta Mol. Basis Dis.* 1863 (2017) 450–461.
- [60] P. Agudelo-Romero, P. Carbonell, M.A. Perez-Amador, S.F. Elena, Virus adaptation by manipulation of host's gene expression, *PLoS One* 3 (2008) e2397.
- [61] C.F. Baer, Does mutation rate depend on itself, *PLoS Biol.* 6 (2008) e52.
- [62] S. Jakhmola, B. Baral, H.C. Jha, A comparative analysis of COVID-19 outbreak on age groups and both the sexes of population from India and other countries, *J. Infect. Develop. Count.* (2020).
- [63] W. Song, M. Gui, X. Wang, Y. Xiang, Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2, *PLoS Pathog.* 14 (2018), e1007236.
- [64] P. Eriksson, Avian Influenza Virus : deciphering receptor interactions and their role in interspecies transmission, *Acta Univ. Upsal.* (2019). <http://uu.diva-portal.org/smash/get/diva2:1365237/FULLTEXT01>. (Accessed 25 September 2020).
- [65] M. Narasumani, P.M. Harrison, Discerning evolutionary trends in post-translational modification and the effect of intrinsic disorder: analysis of methylation, acetylation and ubiquitination sites in human proteins, *PLoS Comput. Biol.* 14 (2018), e1006349.
- [66] B. Macek, K. Forchhammer, J. Hardouin, E. Weber-Ban, C. Grangeasse, I. Mijakovic, Protein post-translational modifications in bacteria, *Nat. Rev. Microbiol.* 17 (2019) 651–664.
- [67] P. Wimmer, S. Schreiner, Viral mimicry to usurp ubiquitin and SUMO host pathways, *Viruses* 7 (2015) 4854–4872.
- [68] Y. Yanku, A. Ori, Regulation of protein-protein interactions by the SUMO and ubiquitin pathways, *Protein Interact.* (2012). (Accessed 25 September 2020).
- [69] M. Bianchi, D. Benvenuto, M. Giovanetti, S. Angeletti, M. Ciccozzi, S. Pascarella, Sars-CoV-2 envelope and membrane proteins: structural differences linked to virus characteristics? *BioMed Res. Int.* 2020 (2020) 4389089.
- [70] A.B. Blázquez, Á. Vázquez-Calvo, M.A. Martín-Acebes, J.-C. Saiz, Pharmacological inhibition of protein kinase C reduces west nile virus replication, *Viruses* 10 (2018) 91.
- [71] E.K.-W. Hui, D.P. Nayak, Role of G protein and protein kinase signalling in influenza virus budding in MDCK cells, *J. Gen. Virol.* 83 (2002) 3055–3066.
- [72] M.J. Duncan, J.-S. Shin, S.N. Abraham, Microbial entry through caveolae: variations on a theme, *Cell Microbiol.* 4 (2002) 783–791.
- [73] R. Park, J.D. Baines, Herpes simplex virus type 1 infection induces activation and recruitment of protein kinase C to the nuclear membrane and increased phosphorylation of lamin B, *J. Virol.* 80 (2006) 494–504.
- [74] D.I. Udenwobe, R.-C. Su, S.V. Good, T.B. Ball, S. Varma Shrivastav, A. Shrivastav, Myristoylation: an important protein modification in the immune response, *Front. Immunol.* 8 (2017) 751.
- [75] D. Schoeman, B.C. Fielding, Coronavirus envelope protein: current knowledge, *Virol. J.* 16 (2019) 69.
- [76] P. Gripon, I. Cannie, S. Urban, Efficient inhibition of hepatitis B virus infection by acylated peptides derived from the large viral surface protein, *J. Virol.* 79 (2005) 1613–1622. Philippe Gripon, Isabelle Cannie, Stephan Urban.
- [77] T. Strecker, A. Maisa, S. Daffis, R. Eichler, O. Lenz, W. Garten, The role of myristoylation in the membrane association of the Lassa virus matrix protein Z, *Virol. J.* 3 (2006) 93.
- [78] A.V. Carluccio, M.I. Prigigallo, T. Rosas-Diaz, R. Lozano-Duran, L. Stavolone, S-acylation mediates Mungbean yellow mosaic virus AC4 localization to the plasma membrane and in turns gene silencing suppression, *PLoS Pathog.* 14 (2018), e1007207.
- [79] R.V. Chowda-Reddy, F. Achenjang, C. Felton, M.T. Etarock, M.-T. Anangfac, P. Nugent, V.N. Fondong, Role of a geminivirus AV2 protein putative protein kinase C motif on subcellular localization and pathogenicity, *Virus Res.* 135 (2008) 115–124.
- [80] D.R. Johnson, R.S. Bhatnagar, L.J. Knoll, J.I. Gordon, Genetic and biochemical studies of protein N-myristoylation, *Annu. Rev. Biochem.* 63 (1994) 869–914.
- [81] A.C. Walls, Y.-J. Park, M.A. Tortorici, A. Wall, A.T. McGuire, D. Velesler, Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein, *Cell* 181 (2020) 281–292, e6.
- [82] S. Jakhmola, O. Indari, S. Chatterjee, H.C. Jha, SARS-CoV-2, an underestimated pathogen of the nervous system, *SN Compr. Clin. Med.* (2020) 1–10.
- [83] B.J. Bosch, W. Bartelink, P.J.M. Rottier, Cathepsin L functionally cleaves the severe acute respiratory syndrome coronavirus class I fusion protein upstream of rather than adjacent to the fusion peptide, *J. Virol.* 82 (2008) 8887–8890.
- [84] D.C. Groves, S.L. Rowland-Jones, A. Angyal, The D614G mutations in the SARS-CoV-2 spike protein: implications for viral infectivity, disease severity and vaccine design, *Biochem. Biophys. Res. Commun.* S0006–291X (2020), 32038–6.
- [85] A.J. McAuley, M.J. Kuiper, P.A. Durr, M.P. Bruce, J. Barr, S. Todd, G.G. Au, K. Blasdell, M. Tachedjian, S. Lowther, G.A. Marsh, S. Edwards, T. Poole, R. Layton, S.-J. Riddell, T.W. Drew, J.D. Druce, T.R.F. Smith, K.E. Broderick, S.S. Vasan, Experimental and in silico evidence suggests vaccines are unlikely to be affected by D614G mutation in SARS-CoV-2 spike protein, *Npj Vaccines* 5 (2020) 1–5.
- [86] T.N. Starr, A.J. Greaney, A. Addetia, W.W. Hannon, M.C. Choudhary, A.S. Dingens, J.Z. Li, J.D. Bloom, Prospective mapping of viral mutations that escape antibodies used to treat COVID-19, *Science* (2021) eabf9302.
- [87] D. Wrapp, N. Wang, K.S. Corbett, J.A. Goldsmith, C.-L. Hsieh, O. Abiona, B.S. Graham, J.S. McLellan, Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation, *Science* 367 (2020) 1260–1263.
- [88] Y. Cai, J. Zhang, T. Xiao, H. Peng, S.M. Sterling, R.M. Walsh Jr., S. Rawson, S. Rits-Volloch, B. Chen, Distinct conformational states of SARS-CoV-2 spike protein, *Science* 369 (2020) 1586–1592.
- [89] C. Sonkar, D. Kashyap, N. Varshney, B. Baral, H.C. Jha, Impact of gastrointestinal symptoms in COVID-19: a molecular approach, *SN Compr. Clin. Med.* (2020) 1–12.
- [90] D. Velesler, K. Cupelli, M. Burger, P. Gräber, T. Stehle, J.E. Johnson, Single-particle EM reveals plasticity of interactions between the adenovirus penton base and integrin $\alpha v \beta 3$, *Proc. Natl. Acad. Sci. U.S.A.* 111 (2014) 8815–8819.
- [91] M.V. Veetil, S. Sadagopan, N. Sharma-Walia, F.-Z. Wang, H. Raghu, L. Varga, B. Chandran, Kaposi's sarcoma-associated herpesvirus forms a multimeric complex of integrins ($\alpha v \beta 5$, $\alpha v \beta 3$, and $\alpha 3 \beta 1$) and CD98-xCT during infection of human dermal microvascular endothelial cells, and CD98-xCT is essential for the postentry stage of infection, *J. Virol.* 82 (2008) 12126–12144.
- [92] B. Chandran, Early events in kaposi's sarcoma-associated herpesvirus infection of target cells, *J. Virol.* 84 (2010) 2188–2199.
- [93] T. Raymond, E. Gorbunova, I.N. Gavrilovskaya, E.R. Mackow, Pathogenic hantaviruses bind plexin-semaphorin-integrin domains present at the apex of inactive, bent v 3 integrin conformers, *Proc. Natl. Acad. Sci. Unit. States Am.* 102 (2005) 1163–1168.
- [94] E. Ruoslahti, M.D. Pierschbacher, Arg-Gly-Asp: a versatile cell recognition signal, *Cell* 44 (1986) 517–518.
- [95] D.M. Shayakhmetov, A.M. Eberly, Z.-Y. Li, A. Lieber, Deletion of penton RGD motifs affects the efficiency of both the internalization and the endosome escape of viral particles containing adenovirus serotype 5 or 35 fiber knobs, *J. Virol.* 79 (2005) 1053–1061.
- [96] B.-H. Luo, C.V. Carman, T.A. Springer, Structural basis of integrin regulation and signaling, *Annu. Rev. Immunol.* 25 (2007) 619–647.
- [97] E.J. Plosa, J.T. Benjamin, J.M. Sucre, P.M. Gulleman, L.A. Gleaves, W. Han, S. Kook, V.V. Polosukhin, S.M. Haake, S.H. Guttentag, L.R. Young, A. Pozzi, T.S. Blackwell, R. Zent, $\beta 1$ Integrin regulates adult lung alveolar epithelial cell inflammation, *JCI Insight* 5 (2020), e129259.
- [98] G. Fox, N.R. Parry, P.V. Barnett, B. McGinn, D.J. Rowlands, F. Brown, The cell attachment site on foot-and-mouth disease virus includes the amino acid sequence RGD (arginine-glycine-aspartic acid), *J. Gen. Virol.* 70 (Pt 3) (1989) 625–637.