Abstract| 2025

In December 2019, an outbreak of a novel respiratory disease with pneumonia-like symptoms emerged in Wuhan, China. This disease, later termed COVID-19, was identified as being caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Within a short period, COVID-19 rapidly spread across the globe, prompting the World Health Organization (WHO) to officially declare it a pandemic on March 11, 2020. The delayed implementation of preventive measures in several countries contributed to a sudden surge in cases worldwide. Despite significant advancements in clinical research that have enhanced the understanding of SARS-CoV-2, many nations continued to experience prolonged outbreaks. As with other RNA viruses, SARS-CoV-2 undergoes genetic mutations, leading to the emergence of new variants, which differs significantly from their ancestral strains. The continuous evolution of the virus, coupled with the absence of fully effective and approved antiviral treatments, underscores the urgent need for the development of efficient therapeutic strategies for COVID-19. The global impact of the COVID-19 pandemic has driven extensive efforts to develop preventive and therapeutic interventions. These initiatives led to the remarkably rapid development of multiple effective vaccines. However, despite large-scale vaccination campaigns and the unprecedented speed of vaccine production, the emergence of new SARS-CoV-2 variants continues to pose a challenge to disease control efforts. This viral evolution threatens to overturn the progress made in limiting the spread of this disease.

Two critical target proteins play a fundamental role in the development of therapeutic interventions against COVID-19. The first is the main protease (Mpro), also known as 3CLpro, which is a highly conserved cysteine hydrolase. During the viral replication process, Mpro facilitates the cleavage of polyproteins at multiple sites, generating several functional proteins necessary for viral multiplication. Given its essential role in coronavirus replication, 3CLpro has been identified as a promising target for the development of broad-spectrum antiviral agents against coronaviruses. The second key target is the spike glycoprotein (S protein), which plays a crucial role in mediating host–pathogen interactions. It binds to the angiotensin-converting enzyme 2 (ACE2) receptor on the host cell surface, facilitating viral entry through membrane fusion. Due to its crucial function in viral infection, the S protein has been extensively studied for its potential in antiviral therapeutic strategies. The detailed structural characterization of this glycoprotein has provided valuable insights, serving as the foundation for COVID-19 vaccine development. To gain a deeper understanding of these target proteins, it is essential to analyse their structural features. Therefore, the initial phase of this thesis focuses on investigating the key structural attributes of Mpro and the S protein. Additionally, the study explores the conformational accessibility of the S protein with its ACE2 receptor providing insights into its role in viral entry. As already mentioned, the main protease (Mpro) plays a pivotal role in the viral life cycle by cleaving viral polyproteins into functional components, a process essential for replication. Inhibiting Mpro disrupts

viral maturation, thereby preventing its effective proliferation. Notably, the active site of Mpro exhibits a distinct structural composition compared to human proteases, enabling the development of selective inhibitors that specifically target the viral enzyme without significantly impacting host proteins. Another critical aspect that makes Mpro an ideal drug target is its highly conserved nature across various coronaviruses. This conservation suggests that inhibitors designed against Mpro may have broad-spectrum efficacy against multiple strains of the virus. Given these attributes, the subsequent section of this thesis focuses on the design and development of small-molecule and peptide-based inhibitors targeting Mpro, with the goal of identifying potent therapeutic candidates against SARS-CoV-2.

In addition to therapeutic advancements, an essential aspect of COVID-19 research involves studying the emergence of SARS-CoV-2 variants over time. Analysing the interaction and binding profiles of different variants provides crucial insights into how specific mutations influence viral virulence. Consequently, the following section of this thesis explores the impact of mutations in the S protein on its interaction with the ACE2 receptor, with a particular focus on variants that have significantly affected global health throughout the pandemic. While it is well established that SARS-CoV-2 entry into host cells primarily depends on the binding of the S protein to the ACE2 receptor, emerging studies indicate that ACE2 expression alone does not fully correlate with infection patterns, immune responses, or clinical manifestations. This observation suggests the potential involvement of alternative receptors in viral entry. Large-scale clinical studies show that the lung, where ACE2 expression is least, is the major site of infection for COVID-19, suggesting that additional factors may be involved in viral entry. Therefore, this thesis also investigates the interaction profile of the receptor-binding domain (RBD) of the S protein with several alternative or co-receptors of SARS-CoV-2, contributing to a broader understanding of viral-host interactions.

The initial part of the thesis comprises of the study on the salient structural features of the target proteins, Mpro and S protein to get insight into some physio-chemical features of those proteins. The study involves few bioinformatics tools like *Expasy Protparam*, *GOR IV*, *Clustal Omega*, *PONDR* server in order to perform the analysis. Results depicted the thermostability of the Mpro and the S protein. The secondary structure prediction analyzed using GOR IV shows that the Mpro as well as S protein mainly exhibit three types of secondary structure namely Alpha helix, Extended strand and Random coil. The conserved regions predicted using Clustal Omega for both the SARS-CoV Mpro and SARS CoV-2 Mpro depicted that Mpro is highly conserved in the SARS-CoV and SARS-CoV-2 with a sequence similarity of 98.69%. Moreover, the conserved regions of the RBD of S protein for both the SARS-CoV and SARS CoV-2 was also analyzed and results depicts that there is a sequence identity of 72.09% and sequence similarity of 75.22 % between them. The disordered regions were also predicted using PONDR® VLXT and results shows that both the Mpro as well as S protein is highly ordered in

nature. Along with the salient structural feature analysis of the S protein, the conformational accessibility of SARS-CoV-2 S-protein with the ACE2 receptor using three different states (open, closed and Intermediate) of the S protein was performed and the extent of accessibility and inaccessibility of ACE2 receptor to different states of S protein was checked. Results depicted that the open state of the S protein is the most accessible to the ACE2 receptor and the closed state is the least accessible to the ACE2 receptor.

The next section of the study consists of the development of small molecule inhibitors and peptide inhibitors against the Mpro. Studies suggest that during the pandemic the scientific community has tried to development many potential inhibitors against Mpro and compounds like Carnosol (CAN), Arjunglucoside-I (ARJ), and Rosmanol (ROS) from Indian spices have been identified as SARS-CoV-2 Mpro inhibitors. The structural dynamics and characteristic features of binding of these small molecules to the SARS-CoV-2 Mpro are not well understood. So, the potential of mean force (PMF) has been constructed in order to analyze the binding-unbinding (or association-dissociation) pathway of the Mpro-small molecule inhibitor complexes using umbrella sampling method. Mpro-small molecule inhibitor complexes exhibited relatively higher dissociation energy values than the alpha-ketoamide (AKA)-Mpro complex (positive control) from the PMF calculations. Also, the binding affinity between protein and ligand is found to be higher in Mpro-ARJ complex [ΔG_{bind} = 19.74 kcal/mol from MM-GBSA and ΔG_{bind} = -9.13 kcal/mol from MM-PBSA] followed by the Mpro-CAN complex, Mpro-ROS complex and the Mpro-AKA complex. According to the MM-GBSA/MM-PBSA calculations, the small molecule inhibitors studied in this work exhibit a significantly higher binding affinity for Mpro. Stability profile analysis also depicted a similar stability profile for the three complexes compared to the positive control. Stability profiles depicted that the small molecules are stable in the binding pocket of the Mpro. Our findings demonstrated that those small molecules obtained from the Indian spices have higher binding affinity to the Mpro compared to a standard control, indicating potential therapeutic value.

Even though small molecules currently dominate the majority of the drug market, peptide inhibitors continue to be a class of promising candidates due to their low toxicity, high affinity, and similarity to endogenous ligand. It has been validated that therapeutic peptides can effectively and selectively inhibit the protein-protein interactions in viruses. Therefore, the development of potential peptide inhibitors is necessary to inhibit the impact of the disease. This study involves development of potential target peptides that can act against the Mpro mimicking histone deacetylase (HDAC2) which had a high-confidence interaction with Mpro. Based on the interacting residue between Mpro and HDAC2, 13 peptides were designed out of which based on toxicity, binding affinity and binding site prediction, two peptides (peptide2 and peptide4) were screened and subjected to MD simulation. Results demonstrate that the two peptides bind to the active site of the Mpro and it attains a higher stability upon binding to

the peptides. Moreover, according to the binding affinity analysis the Mpro have a strong binding affinity with both the peptides (GB_{TOT} = -72.85 kcal/mol for Mpro-peptide2 complex and GB_{TOT} = -46.36 kcal/mol for the Mpro-peptide4 complex). Apart from the inhibitory action of Mpro, the effect of mutations in the S protein on its interaction with ACE2 receptor was also analyzed. Few variants having a serious impact on the global health throughout the pandemic was studied which will provide an insight into the potential strength of each mutation on the extent of virulence on the variants. During the initial phase of the pandemic, SARS-CoV-2 strain (B.1.617 double mutant variant) has raised alarms in India and other nations. B.1.617 variant was found to contain two key mutations (L452R and E484Q) in the RBD region of the spike protein. Immediately after the B.1.617 variant, the delta variant (B.1.617.2) emerged in India leading to more challenges and deterioration to the human health. In addition to Delta, the Delta Plus sub-variants had become a new cause of global concern. Hence, the interaction profile of RBD of spike protein of the Double mutant (DM), Delta and Delta-plus variants of SARS-CoV-2 with ACE2 receptor were compared and we also focus on the effect of those mutations on the binding of the S protein to ACE2. Stability profile analysis from the MD simulation depicts that the mutated S protein-ACE2 complexes attain a higher stability upon binding of the S protein and ACE2 as compared to the wild type (WT) S protein-ACE2 complex. From the results it was also observed that the relative stability of the area in and around the mutated residues attain a lesser fluctuation leading to the more stable binding between the S protein and ACE2 in the DM, Delta and Delta-plus variant. Binding affinity analyzed using MM-GBSA/MM-PBSA algorithms shows the binding energy between S protein (DM) and ACE2 is GB_{TOT} = -47.09 kcal/mol, PB_{TOT} = -19.93 kcal/mol; S protein (Delta) and ACE2 is GB_{TOT} = -39.36 kcal/mol, PB_{TOT}= -17.52 kcal/mol and S protein (Delta-plus) is GB_{TOT} = -36.83 kcal/mol, PB_{TOT} = -16.03 kcal/mol. Moreover, the binding energy between the S protein WT and ACE2 was found to be GB_{TOT} = -31.79 kcal/mol, PB_{TOT} = -6.33 kcal/mol which suggests a stronger binding of the S protein to the ACE2 in the mutated complexes as compared to the WT complex. Hence, a stable and strong binding between the S protein and ACE2 may be a root cause for the higher virulence of the DM, Delta and Delta-plus variants. Mutations (L452R, E484Q) enhance binding to the ACE2 receptor, increasing virus stability and infectivity.

The COVID-19 pandemic was still thriving due to its continuous mutation and evolution into new strains. When the world thought of the end of the pandemic, omicron emerged out as a surprise moiety and because of the high transfer rate Omicron has affected every corner of the world and has a significant impact on India as well. Following the original B.1.1.529 variant, mutations resulted in several subvariants of Omicron including: BA.1, BA.2, BA.3, BA.4 and BA.5. Apart from the mainstream subvariants (BA.1, BA.2, BA.3,

BA.4, and BA.5), these subvariants gave rise to other sub-variants that had a significant effect on the

pandemic and the entire globe. These includes BA.2.12 and BA.2.12.1, BA.2.75 and BA.2.75.2, XBB and XBB.1 and so on. Among all the omicron sub-variants, BA.1, BA.2, BA.4, BA.2.12.1, BA.2.75 and BA.2.75.2 was considered in our study as they cause huge destruction to the world during the pandemic. Under this study we used MD simulation method to show the impact of these omicron variants on the SARS-CoV-2 RBD's tendency to connect with ACE2. From the RMSD, RMSF, and a number of intermolecular hydrogen bond analyses, we found the S protein (BA.2)-ACE2 complex to have enhanced stability than the S protein (BA.1) ACE2 complex. From the binding free energy calculations of the S protein (BA.1)-ACE2 and S protein (BA.2)-ACE2 complexes, we found that the affinity between S protein and ACE2 is higher in the BA.2 complex. The overall stability of the S protein (BA.2)-ACE2 complex and the increased affinity between S protein (BA.2) and ACE2 may be the result for higher virulence of the BA.2 strain than its BA.1 type strain.

In case of the BA.4 and BA.2.12.1 variants, significant structural changes in the mutation area of the spike protein were observed in both the complexes. From the secondary structure prediction analysis, we observed BA.2.12.1 has a higher alpha helix and extended strand structure than BA.4, which explains the increase in stability of the secondary structure in the BA.2.12.1 structure. The RMSD, RMSF, and inter-molecular hydrogen bond studies showed a similar pattern in the case of the stability of the S protein (BA.2.12.1)-ACE2 complex over the S protein (BA.4)-ACE2 complex. Based on calculations of the binding free energies of the S protein (BA.4)-ACE2 and S protein (BA.2.12.1)-ACE2 complexes, the affinity between S protein and ACE2 was found to be higher in the BA.2.12.1 complex as compared to BA.4 complex.

For the BA.2.75 and BA.2.75.2 variants, the physiochemical characteristics revealed that both omicron variants were found to exhibit secondary structural characteristics that were more or less comparable, with the secondary structure alpha helix dominating, followed by an extended strand and random coil. It was determined from the analysis of protein stability after mutation that the mutations N501Y and H655Y make these variants more stable which may contribute to the higher virulence for these variants. From the protein-protein interaction study, we observed the total number of interactions to be higher in the case of the spike (BA.2.75)-ACE2 complex in comparison with spike (BA2.75.2)-ACE2 complex. From the SIFT analysis, it was found that for the BA.2.75 variant, Y505H and N764K impaired protein function and an increased risk of disease, whereas in BA.2.75.2, the three mutations Y505H, N764K, and D1199N were observed to impair the protein function. This study also demonstrates the effect of BA.2.75 and BA.2.75.2 variants on SARS CoV-2 RBD towards its binding with the ACE2 by employing MD simulation. RMSD, RMSF, and inter-molecular hydrogen bond analyses depicted higher stability for the S protein (BA.2.75)-ACE2 complex than the S protein (BA.2.75.2)-ACE2 complex. Binding free energy calculations for both the complexes showed higher binding affinity between the S protein and

ACE2 in the S protein (BA.2.75)-ACE2 complex. Those results gave an overview on the important interactions between the S protein and ACE2 in the omicron variants which might be employed to develop new inhibitors against the SARS-CoV-2 virus. As mentioned earlier since the lungs are the organ that COVID-19 is most likely to infect, the comparatively low expression of this recognized receptor suggests that there may be alternative co-receptors or alternative SARS-CoV-2 receptors that cooperate with ACE2. In addition to ACE2, many candidate receptors of SARS-CoV-2 were reported to be specifically and highly expressed in SARS-CoV-2 affected tissues. Among these receptors, the binding affinity of CAT and L-SIGN to the S protein has been reported to be higher in one of the recent studies. Therefore, the last part of this thesis involves the study of the binding interactions between these potential receptors and the RBD region of the S protein. From the RMSD, RMSF, and a number of inter-molecular hydrogen bond analyses, we found the S protein-CAT complex to have a greater stability when compared to the S protein-L-SIGN complex and S protein-ACE2 complex. From the binding free energy calculations of the S protein-CAT and S protein-L-SIGN complexes, we found that the affinity between S protein and CAT is higher which may infer the higher stability of the S protein-CAT complex. Higher virulence of SARS-CoV-2 may derive from the S protein-CAT complex's overall stability and greater affinity of S protein for CAT.