The EuroBiotech Journal
RESEARCH ARTICLE
Bioinformatics

Insilico Screening for Identification of Hits against SARS-Cov-2 Variant of Concern B.1.617 and NSP12 Mutants by Molecular Docking and Simulation Studies

Vinuthna Vani Madishetti[1], Sudhakar Reddy[2], Sudhakar kalagara[3], Ashish Garg[4], Sreenivas Enaganti[5], Sardar Hussain[6]

Abstract

Human coronaviruses (HCoVs), including severe acute respiratory syndrome coronavirus (SARS-CoV) and 2019 novel coronavirus (2019-nCoV), also known as SARS-CoV-2, have caused global epidemics with high morbidity and mortality. Active research on finding effective drugs against 2019-nCoV/SARS-CoV-2 is going on. *In silico* screening represents the best approach for hits identification and could shorten the time and reduce cost compared to *de novo* drug discovery. Recently, CoV2 mutations have been a big concern in India, particularly on non-structural proteins (NSPs) and Spike Protein (B.1.617) which are the key targets that play a pivotal role in mediating viral replication and transcription. Herein, this study analyzed the NSPs and spike's structural aspects of mutant strains of SARS-CoV-2. The three-dimensional structures of NSPs and S Spike proteins were retrieved from the protein data bank or modeled. And a dataset of an antiviral compound library containing 490,000 drug-like ligands and structurally diverse biologically active scaffolds was used for our studies. Initially, the molecular alignment was performed for library compounds with the reference drug molecule to find targets that match the field points. Antiviral compounds having a similarity score >0.6; were selected for further docking studies with wild and mutant NSPs and S Spike protein of SARS-CoV-2 variant B.1.617. The docking studies identified a potent analog MA-11, which exhibited the highest binding affinity towards wild and mutant proteins. Further, molecular dynamics simulation studies of selected compounds confirmed their perfect fitting into NSP12 and spike active sites and offer direction for further lead optimization and rational drug design.

Keywords: Coronaviruses, COVID-19, Mutation, Non-structural proteins, SARS-CoV-2, Spike Protein

Introduction

Coronavirus disease (COVID-19) is a viral infection caused by SARS-CoV-2. It's a public emergency crisis that affected almost all countries (1-3). SARS-CoV-2 was first identified in Wuhan, Hubei province, China, following reports of severe pneumonia. The World Health Organization (WHO) officially met on 30th January 2020 and declared the outbreak a public health emergency (4).COVID-19 has been described as a significant global public health pandemic by the WHO due to its high mortality rate, rapid spread, and lack of drugs and vaccines (5,6). As of 14th July 2020, there are 13.1 million reported cases and 5,73,000 deaths of COVID-19 (7). The Coronaviruses are a family of viruses capable of causing illnesses from the common cold to severe diseases such as MERS, SARS and COVID-19 (7). The coronavirus name originated from the Latin word "corona", which means "crown." (8). The coronaviruses are enveloped viruses and have a single-stranded RNA genome (positive-sense) with the nucleocapsid in helical symmetry (9).

The cycle of coronavirus consists of three stages: Entry, Replication, and Release. In the Entry phase, when viral S (spike) glycoprotein interacts with the host receptor, the host protease cuts and activates the S protein (spike), and the infection begins (10). Once the virus enters the host cell, it gets disassembled to release the nucleocapsid and the viral genome. The coronavirus RNA genome with 5-prime Met- cap and 3-prime poly-A tail makes itself available to the host ribosome, which leads to translation. The coronavirus first overlapping ORF gets translated by the host ribosome into a long chain polypeptide

[1,5] Department of Bioinformatics, Averinbiotech laboratories, 208, 2nd Floor, Windsor Plaza, Nallakunta, Hyderabad, Telangana, India.
[2] University of Massachusetts Chan medical school, RNA therapeutics institute, Worcester, Massachusetts, USA, 01655.
[3] Department of chemistry and biochemistry, University of the Texas at El Paso, 500 W University Ave,El paso TX 79968.
[4] Department of PG Studies and Research in Chemistry and Pharmacy, Rani Durgavati University, Jabalpur, MP, India 482003.
[6] Department of Biotechnology, Government Science College, Chitradurga, Karnataka, India.

Corresponding author:
sreenivas.bioinfo@gmail.com

DOI: 10.2478/ebtj-2023-0009

which cuts itself into many non-structural proteins (NSPs) (11).

In the Replication stage, many NSP proteins are involved in Replicase-Transcriptase Complex (RTC) formation (12). The chief Replicase-Transcriptase (RT) protein is RNA-dependent-RNA-polymerase (RdRp), which is entangled in the transcription and replication of RNA from RNA strand (13). It is important to investigate the frequency, nature and likely results of mutants observed in the NSP regions of the virus due to their key roles in SARS-CoV2 replication and life cycles.

The emergence of SARS-CoV-2 variants of concern (VOC) has sabotaged strategies to combat the COVID-19 pandemic through vaccines or non-pharmaceutical interventions like face masks and social distancing (14). These variants contain mutations that increase transmissibility or immune evasion. The recent upsurge of COVID cases in India was due to the double mutant variant of the SARS-CoV-2, identified by sequence analysis of the samples (15). The COVID cases from western Maharashtra had mutations in spike protein at E484Q and L452R and developed as a separate lineage in India (16). The linage B.1.617 possesses common mutations D111D, G142D, L452R, E484Q, D614G, and P681R in RBD of spike protein (17). Among them, mutations at 452, 484, and 681 residue positions are the VOC globally.(17) In many cases, a novel variant B.1.617 carried a total of eight S protein mutations, including L452R and E484Q mutations, which changed the amino acid type known to modulate neutralization by the antibody. (17) In this study, we have performed *Insilico* computational screening for ligand hits against the mutant strains of SARS-CoV-2. A dataset containing an antiviral compound library was used for the identification of hits. Firstly, a ligand-based screening protocol using field overlay molecular alignment was used to study the structural features of the field points viz. electrostatic, hydrophobic and van der Waals field points. The molecular docking studies were carried out to predict binding interactions of ligands with S protein of SARS-CoV-2 S variant B.1.617 and NSP-12 (RNA-directed RNA polymerase) wild type and mutant structures. Further, we also studied the effect of solvation at the active site using the 3D-RISM protocol and found promising ligands against COVID19. The main aim of our work is to find the best fit ligand against the mutant strains of SARS-CoV-2, particularly against NSPs and the S Spike protein of B.1.617 VOC.

Materials-Methods

Dataset

The dataset containing antiviral compounds was retrieved from website of Life Chemicals, Canada which is the supplier of high-throughput screening (HTS) compounds. (18). It has a repository of 490,000 drug-like ligands that are structurally diverse with various biologically active scaffolds. Life Chemicals has created dedicated antiviral libraries with more than 13,700 drug-like screening compounds with antiviral activity for high-throughput screening (HTS) and high-content screening (HCS) projects. 10,000 structurally diverse compounds were selected for the identification of hits against mutant strains of SARS-CoV-2. The 2D structures of ligands were checked using Marwin software, Hungary and were minimized and optimized to 3D structures. All ligands were saved in .SDF format and full protonation and minimization were performed in the Flare software.

Molecular Alignment

Field-based molecular alignment was performed using the Align Molecule Wizard in the Forge module of Cresset Software, United Kingdom (19). The binding conformation of Remdesivir with nsp12-nsp7-nsp8 complex bound to the template-primer RNA Protein (7BV2) was used as the reference and uploaded in the wizard. The shape and field-based templates were generated for the reference ligand. The Forge calculates four types of fields: positive and negative electrostatics, van der Waals, and hydrophobicity. The 3D conformations of all 10,000compounds were generated by the Conformation Hunt module, following an 'accurate but slow' option (Max no of conformations: 200; Gradient cut-off for conformer minimization: 0.1 kcal/mol; Energy: 3 kcal/mol; RMS: 0.5). The 'normal' option with 50% field and 50% shape similarity in the Alignment module was used to derive the overall similarity between the two conformations.

Sequence analysis and Protein Preparation

The present study aimed to perform computational analysis of selected antiviral compounds against mutant strains of SARS-CoV-2 non-structural protein and spike protein targets. The spike protein sequence of SARS-CoV-2 S variant B.1.617 (Accession ID: 364 PI ISL 1360382) was obtained from the GISAID database (https://www.gisaid.org/). In the SWISS-MODEL tool, the 3D structure was generated by homology modeling using the Spike SARS-CoV-2 protein on PDB: 7CN8 (20) as a reference. The 3D structures of NSP-12 (RNA-directed RNA polymerase) wild-type and mutant proteins were retrieved from the Database of Non-Structural Proteins of India specific Novel Coronavirus (21) with mutant PDB ID's : NSP12_97_A_V and NSP12_323_P_L. The structural preparation of proteins (in PDB format) was conducted in the macromolecule protocol of Flare software with default settings. Protein structures were cleaned, missing residues and Hydrogen molecules were added, and 3D-protonation was performed on the target protein and minimized for the selected active residues. All the proteins and energies were minimized using the protein preparation wizard.

Molecular Docking Studies

The molecular docking experiments were performed in Cresset Flare Docking software (22). The molecular docking of the best-aligned ligands, having a similarity score >0.6, was conducted to predict binding interactions with the S protein of SARS-CoV-2 S variant B.1.617. Similarly, docking studies were also performed on NSP-12 (RNA-directed RNA polymerase) wild-type and mutant structures. The NSP12 (7BV2) complex

with template-primer RNA and Remdesivir in triphosphate form (RTP) was initially validated by performing the redocking experiment. Molecular docking investigated by normal mode, and results were analyzed, and protein-ligand interactions were obtained. The post docking analysis was performed to find the best poses based on the binding affinity score and protein-ligand interactions were visualized. The 2D interactions were generated from web server protein plus.

Molecular Dynamics

Molecular Dynamics (MD) simulations were performed in Flare using the Open MM package (22). MD simulation will determine the physical movements of molecules and atoms of docked protein-ligand complex. We performed MD simulations on the docked complex of compound MA-11 with the protein NSP12_323. Before performing the MD simulation, the protein is checked for any missing residues with the structure check wizard and preparation is done with the protein preparation wizard. In Flare, ligands are minimized using the AMBER force field and protein with the AMBER FF14SB force field. MD simulation is performed with default settings with a normal calculation method, the simulation length is 10 ns and Solvent Model is the Explicit method. Force fields are AMBER for small molecules, and the AM1-BCC charge method is applied with a 10 A˚ solvent box. By default, the protein-ligand complex was first minimized to 0.25 kcal/mol, and the system is equilibrated for 200ps before the production run begins. After the completion, the simulation trajectory was analyzed for RMSD (Root mean square deviation) plot and protein-ligand contacts. The dynamic analysis tab in Flare software will generate contact statistics for favorable ligand-protein interactions observed during the simulation.

3D RISM Solvation Studies

The RISM is a modern solvation approach based on the molecular Ornstein-Zernike equation (23). The 3D RISM protocol was used to explore the stability and location of water molecules in the active site of the NSP12_323 protein. A 3D-RISM calculation produces a density grid with the respective particle densities at each point, as well as a correlation grid. In our study, 3D RISM was performed using XED forcefields on MA-11 bound to mutant NSP12_323 protein complexes. We used KH closure in this study, and equation (1) shows the total correlation function h(r12) among particles as a contribution from a direct correlation (r12) and from chains of mediating particles that are responsible for the total integral sum.

$$h(r12) = c(r1.2) + \int dr_3\, C(r_{13})p(r_3)h(r_{23})$$

Results

Field-based Molecular Alignment

The molecular alignment was carried out initially for all 10,000 structurally diverse compounds using Forge Software. The binding conformation of remdesivir with the nsp12-nsp7-nsp8 complex with template-primer RNA protein (7BV2) was taken as a reference ligand. The shape and field-based templates were generated for the reference ligand and identified the field points of the highest alignment score of compounds. It illustrates that molecules from various structural classes can act at the same biological site. Thus, if two molecules with different structures interact similarly with an enzyme or receptor at the same site, their bonding conformations will contain similar surface and electrostatic characteristics. As represented in **Figure 1**, the four fields (positive electrostatic, negative electrostatic, VanderWaals and hydrophobicity) are displayed as iso surfaces at the field strength. Most of the dataset ligands have aligned with the reference molecule (Remdesivir) and shows similarity score in the range from 0.640 to 0.491. The overall field and shape similarity score of compounds MA-01 is 0.640 and MA-02 is 0.636. The compounds selected have the same pharmacophore features with reference molecule remdesivir. The positive electrostatic (red color) and negative electrostatic fields (blue color) are observed above and below part of the reference molecule. Strong hydrophobic fields (brown color) and van der Waals fields (yellow color) surrounded the aromatic ring of compounds (**Figure 1**). We selected compounds having the best similarity scores for further computational studies.

Molecular docking studies with NSP-12 Wild and Mutant Type

The molecular docking studies were carried out in Cresset Docking software. Antiviral compounds with similarity scores >0.6 were selected for docking studies. The docking studies were conducted with NSP12 (RNA-directed RNA polymerase) wild-type and mutant protein structures. NSP12 mutant structure has two mutations at position 97 (A-->V) and 323 (P-->L) with QMEAN Score of -2.19 and -1.98 and ERRAT quality factors of 73.60 % and 75.95%, respectively. Remdesivir exhibited H-bond interactions with Arg553, Lys621, Arg555, Thr556, Arg624, and hydrophobic interactions with Tyr455, Arg624, Arg555 amino acid residues of NSP12 (PDB: 7BV2). The complex showed an LF Rank score of -10.43 and a binding affinity value of LF dG -5.72 kcal/mol, a similar type of interaction for compound MA-05 with LF Rank Score -11.12 and a binding affinity value of LF dG -7.12 kcal/mol with H-bond interactions with Arg555, Arg553, Thr556 and Arg624 and hydrophobic interactions with Val557, Ser682, Phe571, and Lys574 residues were observed. While, compound MA-11 with LF Rank Score -10.98 and binding affinity value LF dG -6.58 kcal/mol exhibited H-bond interactions with Phe442, Arg553, Gln444 and hydrophobic interactions with Arg555, Ser549, Phe571 and Lys574 amino acid residues (**Figure 2** and **Table 1**).

Next, molecular docking studies with NSP12 mutants were conducted to see the variation in binding affinity. NSP12 had two mutations at position 97 (A-->V) and 323 (P-->L), and these proteins were docked with RTP and the best antiviral compounds having the highest similarity score. RTP docked with NSP12_323 mutant protein has shown an LF Rank Score of -10.93 and involves H-bond interactions with Lys551, Ser814, Ile548, and hydrophobic interactions with Ala547, Arg836

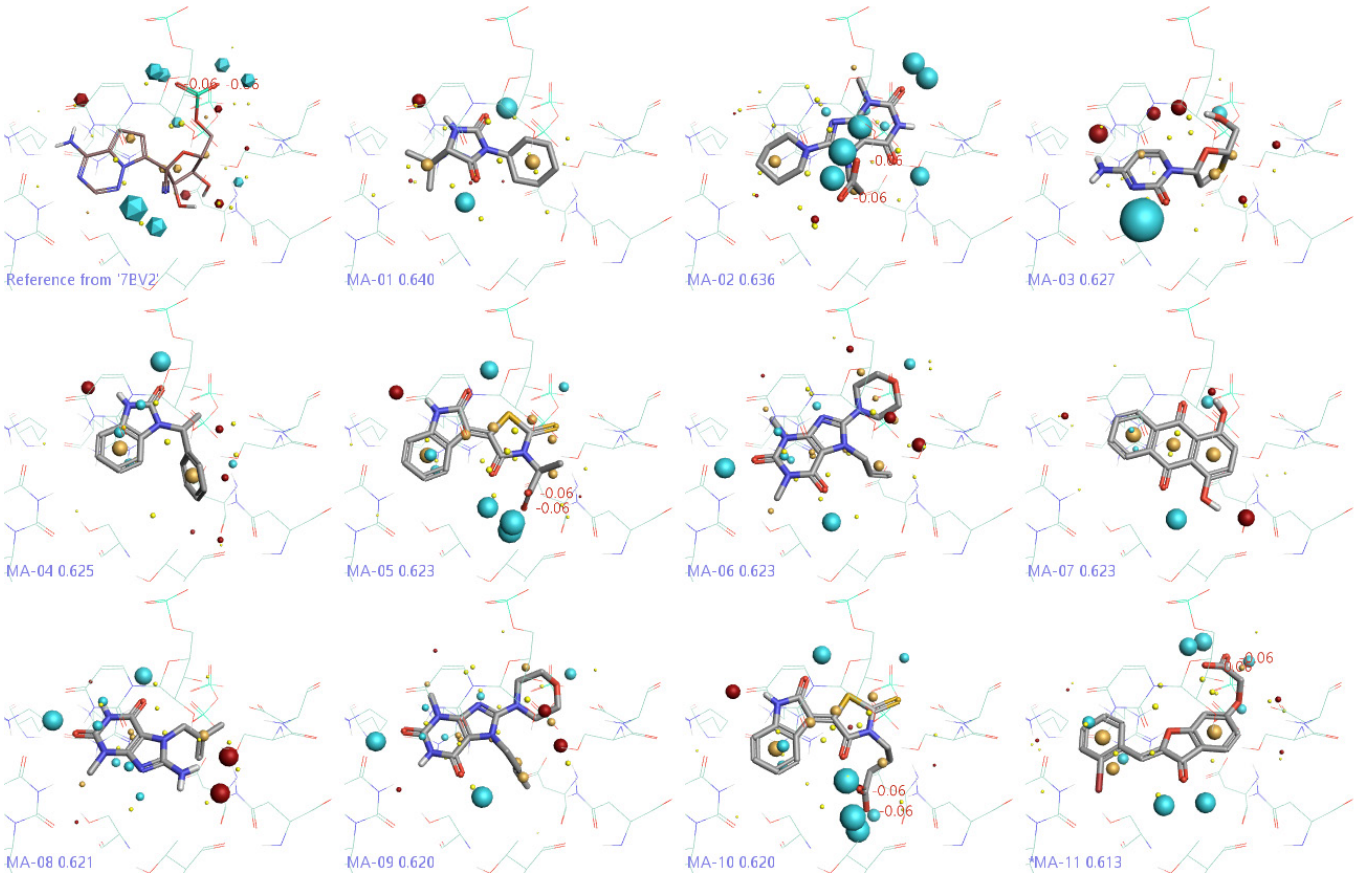

Figure 1. Alignment of antiviral compounds onto molecular field points generated from the reference compound remdesivir (7BV2). Field points for positive electrostatic (red) and negative electrostatic forces (blue) are shown as above and below parts of the molecule. The aromatic ring of compounds is surrounded by strong hydrophobic fields (brown) and van der Waals fields (yellow).

amino acid residues. Similarly, these types of interactions were observed with MA-11 having an LF Rank Score of -10.35 and binding affinity value LF dG -6.93 kcal/mol and involving H-bond interactions with Ala550, Lys551, Asp833, and hydrophobic interactions with Cys813, Ala547, Ser814 and Arg836 amino acid residues (**Figure 3**). With NSP12_97 mutant structure, remdesivir showed LF Rank Score of -10.05 and binding affinity value of LF dG -6.19 kcal/mol and MA-11 with LF Rank Score -9.87 and binding affinity value LF dG -5.97 kcal/mol involves H-bond interactions with Arg553, Arg555, Lys545, and hydrophobic interactions with Ser549, Lys551, Ala547 and Arg836 amino acid residues. From these results, we identified that Remdesivir had a strong binding affinity with the NSP12 mutant structures. The docking studies have identified potent analog MA-11, which exhibited the highest binding affinity with both wild and mutant NSP12 proteins. (**Figure 4**)

We also performed the docking studies with S Spike Protein SARS-CoV-2 S variant B.1.617. The RBD of the S protein harbors two mutations, L452R and E484Q. We built the protein structure on SWISS-MODEL using sequence accession from GISAID with ID PI_ISL_1360382 using template 6XR8 protein structure. Interestingly, the MA-11 compound has the highest binding affinity with LF Rank Score -10.25 and binding affinity value LF dG -5.57 kcal/mol and exhibits H-bond interactions with Ser349, Ala348 and hydrophobic interactions with Phe347 and Val341 residues. (**Figure 5**)

3D RISM Solvation Studies for MA-11 with Mutant Protein NSP12_323

A three-dimensional reference-site (3D-RISM) algorithm was used to study the stability of the protein in bulk water molecules generated during Flare docking studies at the active site. At the end of 3D-RISM calculations, the high-water density at the active site NSP12 was formed. **Figure 6** reveals the oxygen and hydrogen density of MA-11 at the active site of the NSP12_323 protein. It is represented in two 3D-RISM colors, the green spheres indicating favourable and the red sphere indicating unfavourable regions at the active site. Solvation studies show that ligand MA-11 bound to the COVID-19 NSP12_323 active site is stable in bulk water. The studies indicate that the green color sphere (Happy) will favor the interactions between the

MA-11 and NSP12_323 protein (**Figure 6**).

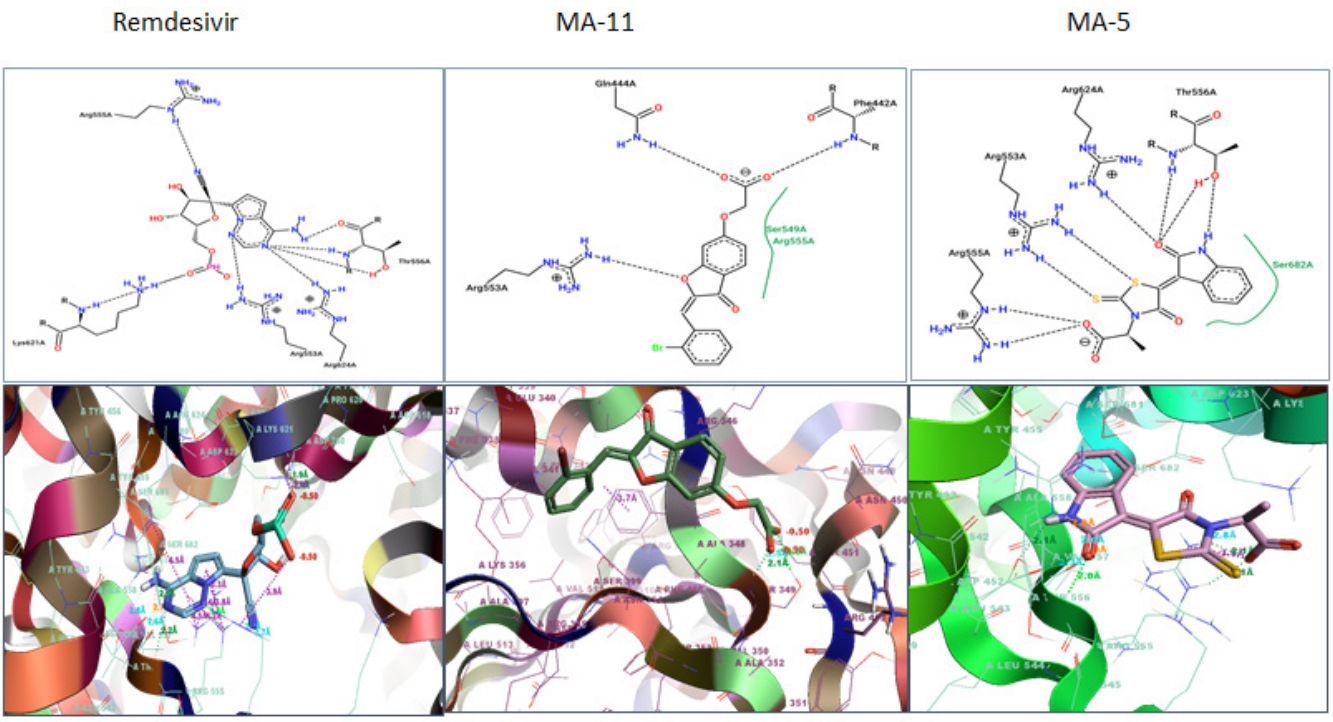

Figure 2. Molecular Docking studies of Remdesivir, MA-5 and MA-11 with Wild Type NSP12 (PDB ID: 7BV2). Key H bond and hydrophobic interactions between Wild Type NSP12 and Remdesivir, MA-5 and MA-11 are shown.

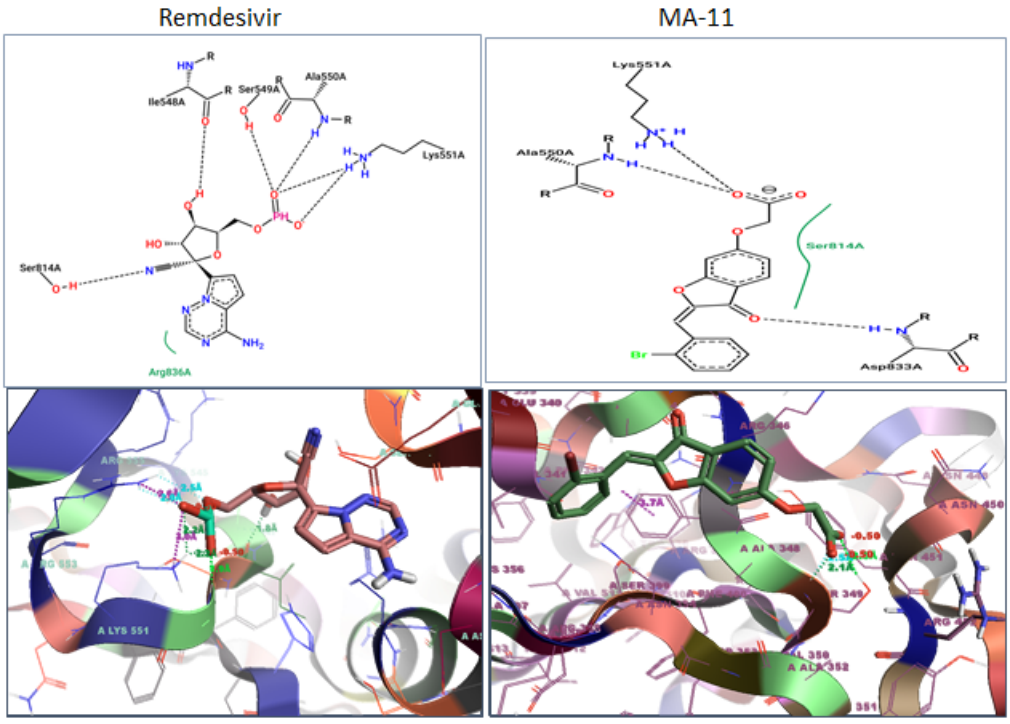

Figure 3. Molecular docking studies of Remdesivir and MA-11 with mutant type NSP12_323

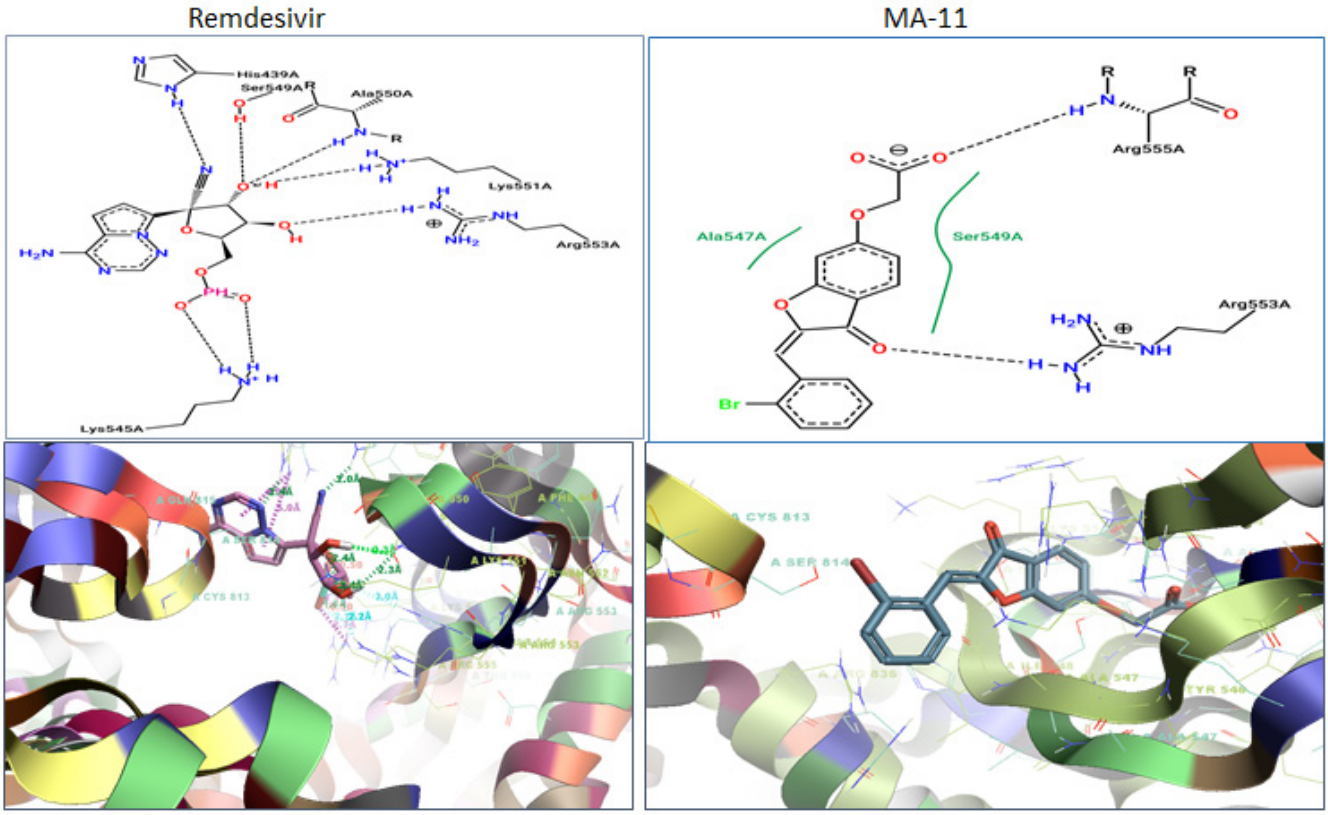

Figure 4. Molecular Docking studies of Remdesivir and MA-11 with mutant type NSP12_97

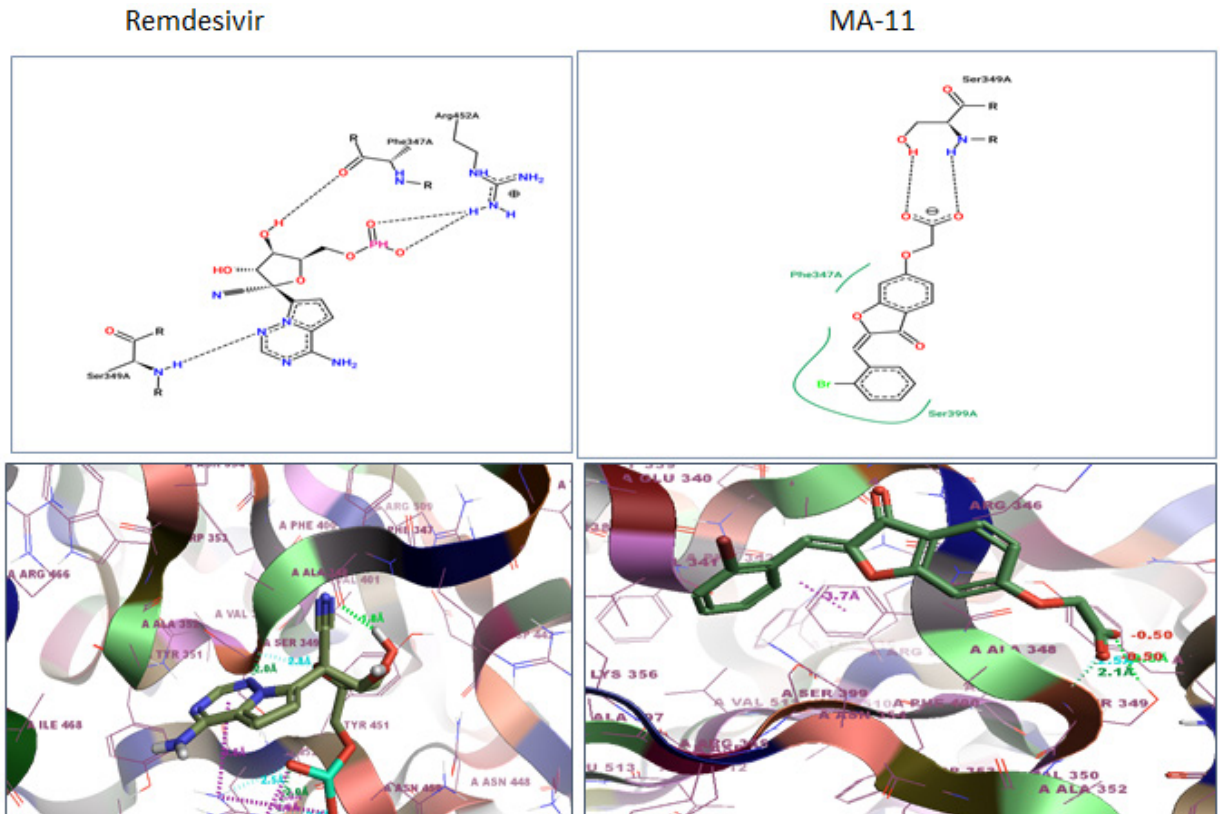

Figure 5. Molecular Docking studies of Remdesivir and MA-11 with mutant type S Spike Protein SARS-CoV-2 S variant B.1.617

Table 1. Similarity Score and Molecular Docking LF Rank Score of Best ligands against SARS-CoV2 Wild and Mutant Targets.

COMPD ID	SIMILARITY SCORE	LF RANK SCORE			
		Wild (7BV2)	NSP12_323	NSP12_97	Spike Protein
MA-01	0.64	-10.28	-9.82	-8.18	-8.18
MA-02	0.636	-10.10	-9.48	-8.37	-9.41
MA-03	0.625	-9.82	-9.12	-7.29	-9.34
MA-04	0.625	-9.35	-8.48	-6.42	-8.70
MA-05	0.623	-11.12	-10.21	-9.58	-10.14
MA-06	0.623	-9.48	-7.93	-6.73	-9.70
MA-07	0.623	-8.89	-8.69	-8.49	-9.08
MA-08	0.621	-9.12	-9.29	-9.47	-8.62
MA-09	0.62	-9.08	-9.73	-9.48	-8.75
MA-10	0.62	-10.35	-9.46	-8.18	-8.58
MA-11	0.613	-10.98	-10.35	-9.87	-10.25
RTP	--	-10.43	-10.93	-10.05	-9.50

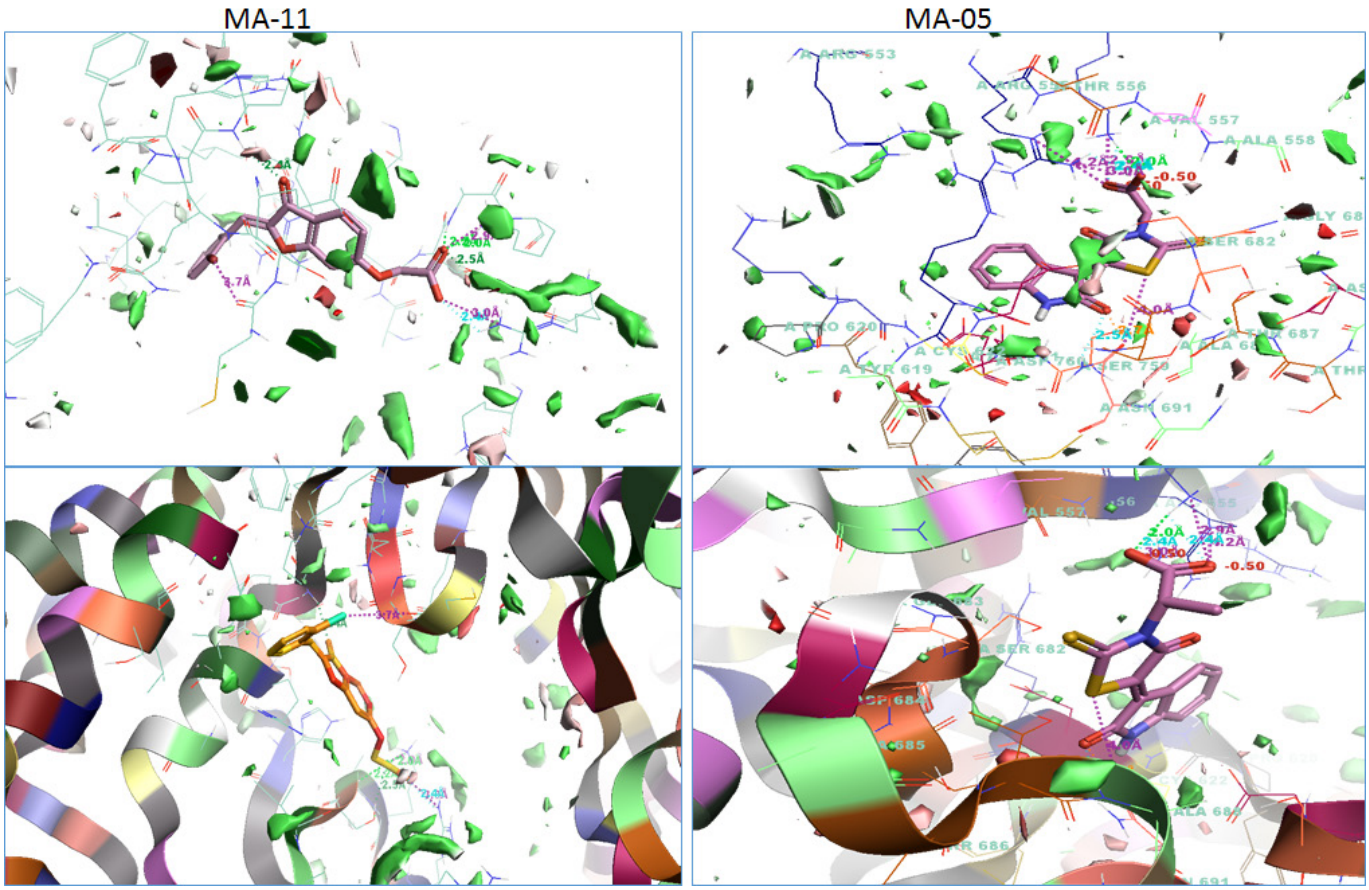

Figure 6. 3D RISM solvation studies for MA-11 and MA-05 in complex with mutant NSP12_323 Protein

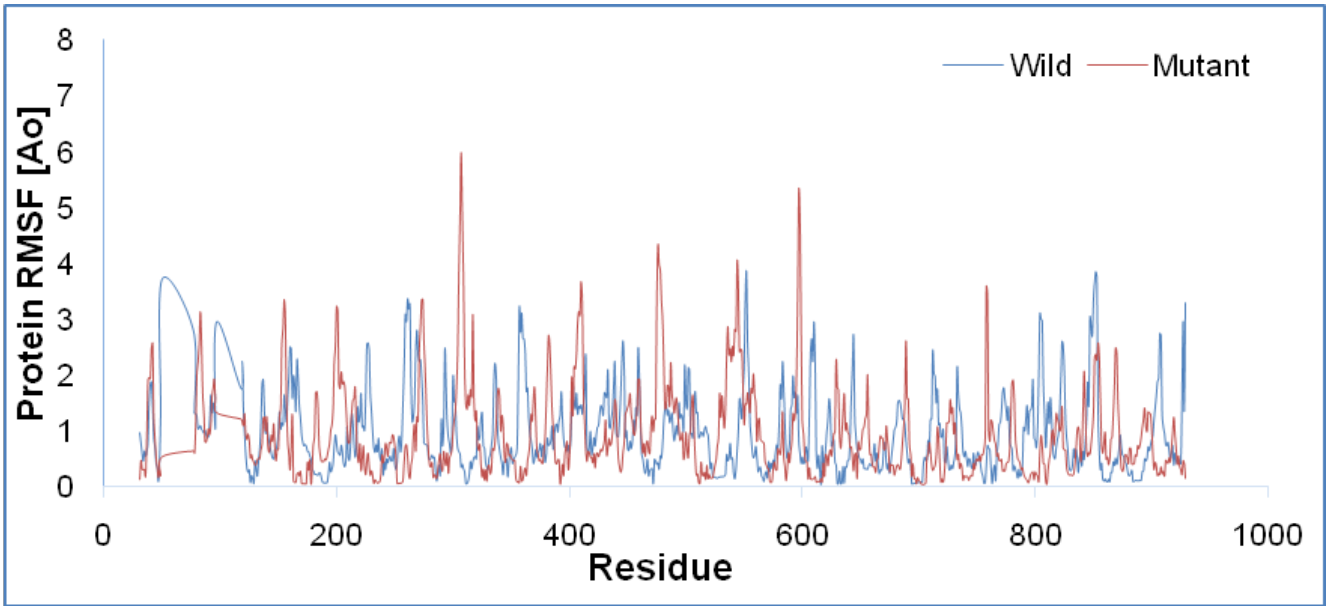

Figure 7. RMSF plots of Wild protein, MA-11-NSP12 complex (blue), MA-11-NSP12_323 mutant complex (pink).

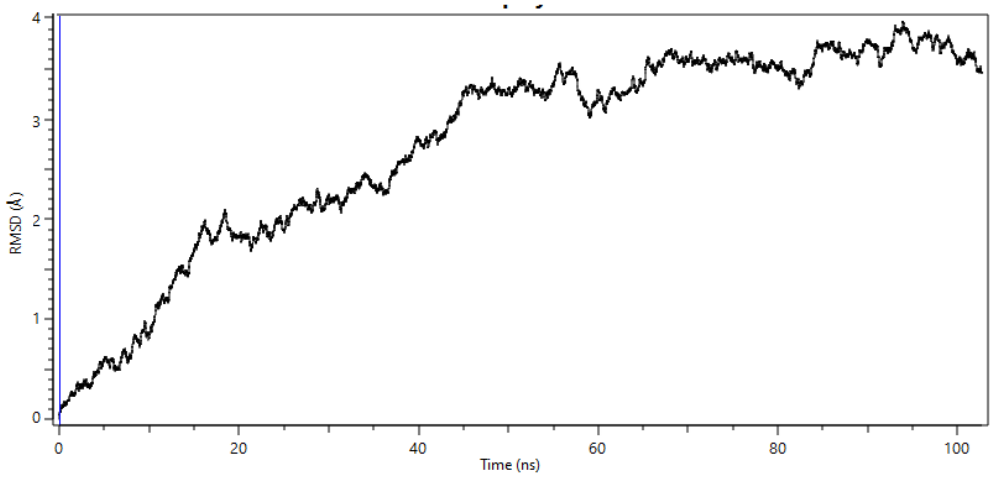

Figure 8. MD simulation of Protein-ligand complex of Compound MA-11 and NSP12_323 Mutant Protein

Molecular Dynamic Simulation

In this report, MD simulations were carried out to predict the protein-ligand binding stabilities between compound MA-11 and NSP12_323, S Spike mutant proteins. The flare module in the Cresset Software Suite was used for the MD simulations. After completing the MD simulation, the trajectory file was opened from Open Trajectory, and RMSD plots were observed. The RMSD plot for protein-ligand complexes was stable during the simulations. The results were analyzed using the Open Trajectory wizard, and observations showed that the RMSD plots were stable during the MD simulations. The simulation was carried out for MA-11 and NSP12_323 and S Spike protein complex for 100ns.

The RMSF and RMSD plots of the protein-ligand complex were monitored during the whole MD simulations. The RMSD plot measures the protein conformational changes and identifies when the simulation reaches equilibrium. The RMSF plot identifies the local changes/movements in each amino acid residue along the protein chain. The RMSF plots of the wild and mutant proteins from the simulation studies were overlaid and shown in **Figure 7** and **9**. The fluctuations at most amino acid residues were <10Å, and several huge fluctuations at some residues indicate the protein loop with high flexibility. A plot of RMSD was computed between the initial frame of the protein and chosen frame. The **Figure 8** and **Figure 10** illustrate the RMSD plot of the protein-ligand complex of MD simulation. The molecular docked complex of protein-ligand was taken as the reference frame, and the movement in this original alignment during MD simulation was gauged by aligning all the protein frames over corresponding times. The RMSD value

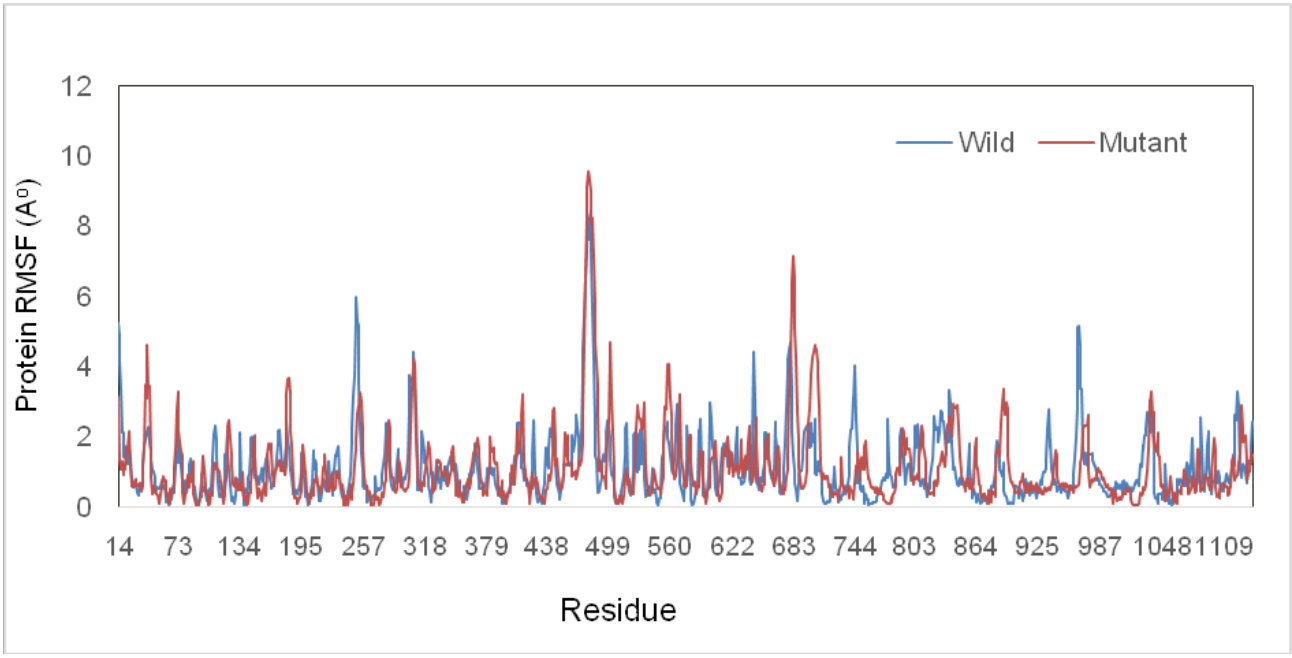

Figure 9. RMSF plots of Wild protein, MA-11- S Spike Protein (blue), MA-11- S Spike Protein mutant complex (pink).

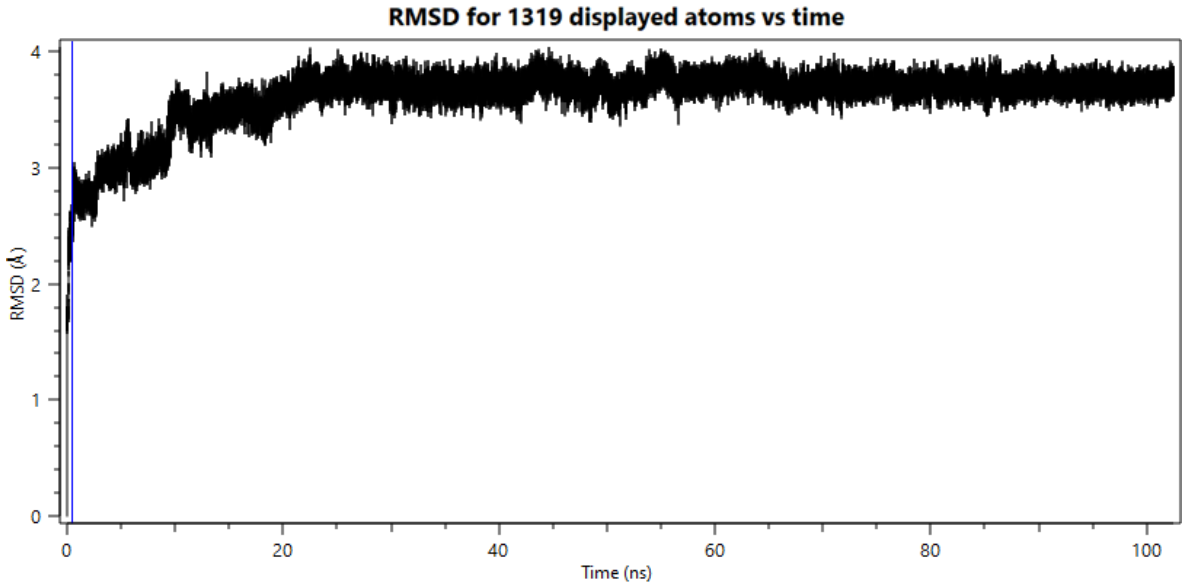

Figure 10. MD simulation of Protein-ligand complex of Compound MA-11 and S Spike Protein Mutant Protein.

of 1-4Å is considered satisfactory for globular proteins (24). For our MD simulation of the MA-11-NSP12_323 complex, the value of RMSD did not exceed 4.0Å showing the stability of the protein conformation. The H-bond analysis of the MA-11-NSP12 complex shows its inter action with the amino acid residues Phe442, Gln444, Arg555, and Arg553. Whereas, for MD simulation of compound MA-11- S Spike Protein mutant complex, the value of RMSD did not exceed 4.0 Å showing the stability of the protein conformation. **Figure 7** depicts the RMSD paths associated with Compound MA-11 and NSP12_323 Mutant Protein. It is evident from the figure that deviation in RMSD between 1 Å and 3 Å was during the 100 ns simulation time. For instance, RMSD fluctuation was significantly higher in the time range between 20 ns and 40 ns. After 40 ns, minimal fluctuation in RMSD was observed, and the system attained stability with 3.5 Å towards the end of simulation time (100 ns). The H-bond analysis of theMA-11-NSP12 complex shows the interactions with the amino acid residues Phe442, Gln444, Arg555, and Arg553.

The RMSD trajectories of compound MA-11- S Spike Protein mutant complex is shown in Figure 10. It is demonstrated in

the figure that deviation in RMSD is between 1Å and 3.5 Å was observed during the 100 ns simulation time. No significant RMSD fluctuation were observed, and the system reached stable conformation at the end of simulation. The RMSD value of compound MA-11- S Spike Protein mutant complex did not exceed 4.0 Å showing the stability of the protein conformation. The H-bond analysis of the MA-11- S Spike Protein mutant complex shows the interactions with following amino acid residues Lys356, Ser349, Ala348, Phe347, and Arg346.

Discussion

The recent increase in COVID-19 infections and deaths in India is due to the rapid spread of the novel variant B.1.617 SARS-CoV-2. This variant has mutations in the RBD and other domains of the S protein, which could affect the virus's biological features, such as its ability to enter target cells efficiently and its susceptibility to antibodies that affect the entry process(25). Sarah Cherian *et al.,* 2021(preprint) have reported the whole genome sequence of SARS-CoV-2 from different regions of Maharashtra state. The structural analysis has shown that the RBD mutations at L452 and E484Q residues of S spike protein will stabilize the binding with ACE2 and increase the viral infectivity (26). Markus Hoffmann *et al.,* 2021 analyzed the B.1.617 susceptibility in entering the cells or evading antibody response and being resistant against the Bamlanivimab drug (27).According to WHO, the various global variants of concern (VOCs)that emerged in different regions in the world are B.1.525 (U.K), B.1.149 (Denmark), and A.23.1 (Uganda) (28). These variants are associated with increased transmission. VOC 202012/01 and 501Y.V2 are also involved in disease severity. 501Y.V2 and P.1 are involved in immune escape and neutralization. Also, these could cause reinfection or an increased risk of vaccine failure (29). Nupur Biswas and the team have investigated the structural information of mutant SARS-CoV-2 NSPs, and all the 3D protein structures were deposited in a database called DbNSPInC (20). The platform provides updated patient information on NSPs specific to the Indian variant of SARS-CoV-2. 3D coordinates of 5 wild and 36 mutant strains of NSPs are available in the DbNSPInC database. The present study involves the *In silico* screening of a dataset containing potent antiviral compounds obtained from Life chemicals, which is the supplier of HTS compounds for SARS-CoV-2. Firstly, we performed the Field-based molecular alignment using the Align Molecule Wizard in the Forge module of Cresset Software using Remdesivir as a reference ligand bound to the template-primer RNA Protein (7BV2). As illustrated in **Figure 1**, shape and field templates (positive electrostatic, negative electrostatic, van der Waals, and hydrophobicity) were generated for all ligands, and we identified the highest alignment score of compounds. Most of the dataset ligands have aligned with the reference molecule (Remdesivir) and showed similarity scores in the range from 0.640 to 0.491. a dataset of 50 compounds having the highest similarity scores was selected for further *In silico* computational studies.

We performed molecular docking studies for 50 compounds with a similarity score >0.6 against NSP12 wild and mutant proteins and with B.1.617 S Spike protein. Among the docked ligands, MA-05 and MA-11 have shown good binding affinity against the NSP12 (7BV) with LF Rank Score -11.2 and -10.98, respectively. Similarly, when performing docking with NSP12_323 mutant protein, Remdesivir has showed LF Rank Score -10.93 and LF dG -5.55 kcal/mol, MA-11 with LF Rank Score -10.35 and LF dG -6.93 kcal/mol. With NSP12_97 mutant structure Remdesivir shown LF Rank Score -10.05 and LF dG -6.19 kcal/mol and MA-11 with LF Rank Score -9.87 and LF dG -5.97 kcal/mol, involves H-bond interactions with Arg553, and Arg555 amino acid residues and hydrophobic interactions with Cys813, Ser549 Ala547, and Arg836 amino acid residues. The docking studies have identified a potent analog, MA-11, which exhibits the highest binding affinity for both wild and mutant structures.

The docking studies were also performed with S Spike Protein SARS-CoV-2 S variant B.1. 617.The RBD of the S protein harbors two mutations at L452R and E484Q. Interestingly, the MA-11 compound also had shown the highest binding affinity with LF Rank Score -10.25 and binding energy LF dG -5.57 kcal/mol. Further, we also performed solvation studies to determine the stability of biomolecules in bulk water at the active site using the 3D-RISM protocol. This study helped in gaining molecular insights into the active sites of mutant NSP12. MD simulations of protein-ligand complex MA-11-NSP12_323 have shown the most stable complex as seen in the generated RMSD and RMSF plots. From these studies, we identified the potential ligands against SARS-CoV-2:MA-05 and MA-11. MA-11 similar series of analogs were reported by Cole *et al.* group which synthesized substituted chalcones, coumaranones and flavonoids as an anti-HIV agent. These analogs have shown the most promising anti-HIV activity (30).

Conclusion

In the present work, *In silico* computational techniques like field template, molecular docking, molecular dynamics, and 3D-RISM studies were utilized towards a dataset containing antiviral compounds to target wild and mutant proteins of NSP12 and Spike variant proteins. Most of the dataset ligands have aligned with the reference molecule (Remdesivir) and showed similarity scores from 0.640 to 0.491. The ligands MA-05 and MA-11 have shown good binding affinity against NSP12 wild and mutant proteins and with B.1.617 S Spike protein. 3D-RISM Solvation studies demonstrated that the protein-ligand complexes formed were stable in the presence of water. The hits identified in this study have potential against the mutant strains of SARS-CoV-2 and can be of clinical use against COVID-19.

Competing interests

The authors do not declare any conflict of interest about this research.

References:

1. Rajnik M, Cascella M, Cuomo A, Dulebohn SC, Di Napoli R. Features, evaluation, and treatment of coronavirus (COVID-19). Uniformed Services University Of The Health Sciences; 2021 Mar 1.
2. Sheng J, Shao A. The epidemiological and clinical features of COVID-19 and lessons from this global infectious public health event; 2020.
3. Singhal T. A Review of Coronavirus Disease-2019 (COVID-19). *Indian J Pediatr* 2020; *87*(4): 281-286.
4. Wang W, Tang J, Wei F. Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCoV) in Wuhan, China. *J Med Virol* 2020; *92*(4): 441-447.
5. Lai CC, Shih TP, Ko WC, Tang HJ, Hsueh PR. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. International journal of antimicrobial agents 2020; 55(3):105924.
6. Prompetchara E, Ketloy C, Palaga T. Immune responses in COVID-19 and potential vaccines: Lessons learned from SARS and MERS epidemic. Asian Pacific journal of allergy and immunology 2020;38(1):1-9.
7. Fan Y, Zhao K, Shi Z. L, Zhou P. Bat coronaviruses in China. Viruses 2019;11(3):210.
8. Lefkowitz EJ, Dempsey DM, Hendrickson RC, Orton RJ, Siddell SG, Smith DB. Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). Nucleic acids research 2018; 46(D1):D708-717.
9. Sexton NR, Smith EC, Blanc H, Vignuzzi M, Peersen OB, Denison MR. Homology-based identification of a mutation in the coronavirus RNA-dependent RNA polymerase that confers resistance to multiple mutagens. Journal of virology 2016; 90(16):7415-7428.
10. Simmons G, Zmora P, Gierer S, Heurich A, Pöhlmann S. Proteolytic activation of the SARS-coronavirus spike protein: cutting enzymes at the cutting edge of antiviral research. Antiviral research 2013;100(3):605-614.
11. Fehr AR, Perlman S, Maier HJ, Bickerton E, Britton P. An overview of their replication and pathogenesis; section 2 genomic organization. Methods in Molecular Biology 2015;1282:1-23.
12. Yadav R, Chaudhary JK, Jain N, Chaudhary PK, Khanra S, Dhamija P, Sharma A, Kumar A, Handu S. Role of structural and non-structural proteins and therapeutic targets of SARS-CoV-2 for COVID-19. Cells 2021;10(4):821.
13. Lei J, Kusov Y, Hilgenfeld R. Nsp3 of coronaviruses: Structures and functions of a large multi-domain protein. Antiviral research 2018;149:58-74.
14. Walensky RP, Walke HT, Fauci AS. SARS-CoV-2 variants of concern in the United States—Challenges and opportunities. *Jama* 2021;*325*(11):1037-1038.
15. Shastri J, Parikh S, Agrawal S, Chatterjee N, Pathak M, Chaudhary S, Sharma C, Kanakan A, Srinivasa Vasudevan J, Maurya R, Fatihi S. Clinical, serological, whole genome sequence analyses to confirm SARS-CoV-2 reinfection in patients from Mumbai, India. Frontiers in medicine 2021; 215.
16. Yadav PD, Sapkal GN, Abraham P, Ella R, Deshpande G, Patil DY, Nyayanit DA, Gupta N, Sahay RR, Shete AM, Panda S. Neutralization of variant under investigation B. 1.617. 1 with sera of BBV152 vaccinees. Clinical Infectious Diseases 2022; 74(2):366-368.
17. Cherian S, Potdar V, Jadhav S, Yadav P, Gupta N, Das M, Rakshit P, Singh S, Abraham P, Panda S, Team NI. SARS-CoV-2 spike mutations, L452R, T478K, E484Q and P681R, in the second wave of COVID-19 in Maharashtra, India. Microorganisms 2021;9(7):1542.
18. https://lifechemicals.com/screening-libraries/hts-compound-collection
19. Cheeseright TJ, Mackey MD, Melville JL, Vinter JG. FieldScreen: virtual screening using molecular fields. Application to the DUD data set. Journal of chemical information and modeling 2008;48(11):2108-2117.
20. Bordoli L, Schwede T. Automated protein structure modeling with SWISS-MODEL Workspace and the Protein Model Portal. Homology Modeling: Methods and Protocols 2012;107-36.
21. Biswas N, Kumar K, Mallick P, Das S, Kamal IM, Bose S, Choudhury A, Chakrabarti S. Structural and drug screening analysis of the non-structural proteins of severe acute respiratory syndrome coronavirus 2 virus extracted from Indian coronavirus disease 2019 patients. Frontiers in genetics 2021;12:626642. (DbNSP, http://www.hpppi.iicb.res.in/covid19/index.php)
22. Bauer MR, Mackey MD. Electrostatic complementarity as a fast and effective tool to optimize binding and selectivity of protein–ligand complexes. Journal of medicinal chemistry 2019;62(6):3036-3050.
23. Kalhotra P, Chittepu VC, Osorio-Revilla G, Gallardo-Velazquez T. Field-template, QSAR, ensemble molecular docking, and 3D-RISM solvation studies expose potential of FDA-approved marine drugs as SARS-CoVID-2 main protease inhibitors. Molecules 2021; 26(4):936.
24. Rao P, Patel R, Shukla A, Parmar P, Rawal RM, Saraf M, Goswami D. Identifying structural–functional analogue of GRL0617, the only well-established inhibitor for papain-like protease (PLpro) of SARS-CoV2 from the pool of fungal metabolites using docking and molecular dynamics simulation. Molecular diversity 2021; 6:1-21.
25. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, Ludden C, Reeve R, Rambaut A, COVID-19 Genomics UK (COG-UK) Consortium, Peacock SJ. SARS-CoV-2 variants, spike mutations and immune escape. Nature Reviews Microbiology 2021;19(7):409-424.
26. Gan HH, Twaddle A, Marchand B, Gunsalus KC. Structural modeling of the SARS-CoV-2 spike/human ACE2 complex interface can identify high-affinity variants associated with increased transmissibility. Journal of Molecular Biology 2021;433(15):167051.

27. Hoffmann M, Hofmann-Winkler H, Krüger N, Kempf A, Nehlmeier I, Graichen L, Arora P, Sidarovich A, Moldenhauer AS, Winkler MS, Schulz S. SARS-CoV-2 variant B. 1.617 is resistant to bamlanivimab and evades antibodies induced by infection and vaccination. Cell reports 2021; 36(3):109415.
28. Jaspe RC, Sulbaran Y, Loureiro CL, D'Angelo P, Rodríguez L, Garzaro DJ, Rangel HR, Pujol FH. Importance of E484K and N501Y mutations in SARS-CoV-2 for genomic surveillance: rapid detection by restriction enzyme analysis. medRxiv 2021;:2021-05.
29. Focosi D, Maggi F. Neutralising antibody escape of SARS-CoV-2 spike protein: risk assessment for antibody-based Covid-19 therapeutics and vaccines. Reviews in medical virology 2021; 31(6):e2231.
30. Cole AL, Hossain S, Cole AM, Phanstiel IV O. Synthesis and bioevaluation of substituted chalcones, coumaranones and other flavonoids as anti-HIV agents. Bioorganic & medicinal chemistry 2016; 24(12):2768-2776.