



Mutational analysis in international isolates and drug repurposing against SARS-CoV-2 spike protein: molecular docking and simulation approach

Swetha Pulakuntla¹ · Kiran Bharat Lokhande² · Pannuru Padmavathi³ · Meena Pal⁴ · Kakumani Venkateswara Swamy⁵ · Jayashree Sadasivam¹ · Shri Abhiav Singh⁶ · Sree Latha Aramgam^{1,7} · Vaddi Damodara Reddy¹

Received: 31 March 2021 / Accepted: 23 June 2021 / Published online: 15 July 2021
© Indian Virological Society 2021

Abstract The novel SARS-CoV-2 (severe acute respiratory syndrome coronavirus-2) is spreading, as the causative pathogen of coronavirus disease-19 (COVID-19). It has infected more than 1.65 billion people all over the world since it was discovered and reported 3.43 million deaths by mid of May 2021. SARS-CoV-2 enters the host cell by binding to viral surface glycoprotein (S protein) with human ACE2 (angiotensin-converting enzyme2). Spike protein (contains S1 and S2 sub-domains) molecular interaction with the host cells is considered as a major step in the viral entry and disease initiation and progression and this identifies spike protein as a promising therapeutic target against antiviral drugs. Currently, there are no efficient antiviral drugs for the prevention of COVID-19 infection. In this study, we have analyzed global 8719

spike protein sequences from patients infected with SARS-CoV-2. These SARS-CoV-2 genome sequences were downloaded from the GISAID database. By using an open reading frame (ORF) tool we have identified the spike protein sequence. With these, all spike protein amino acid sequences are subjected to multiple sequence alignment (MSA) with Wuhan strain spike protein sequence as a query sequence, and it shows all SARS-CoV strain spike proteins are 99.8% identical. In the mutational analysis, we found 639 mutations in the spike protein sequence of SARS-CoV-2 and identified/highlighted 20 common mutations L5F, T22I, T29I, H49Y, L54F, V90F, S98F, S221L, S254F, V367F, A520S, T572I, D614G, H655Y, P809S, A879S, D936Y, A1020S, A1078S, and H1101Y. Further, we have analyzed the crystal structure of the 2019-nCoV chimeric receptor-binding complex with ACE2 (PDB ID: 6VW1) as a major target protein. The spike receptor binding protein (RBD) used as target region for our studies with FDA-approved drugs for repurposing, and identified few anti-SARS-CoV2 potential drugs (Silmitasertib, AC-55541, Merimepodib, XL413, AZ3451) based on their docking score and binding mode calculations expected to strongly bind to motifs of ACE2 receptor and may show impart relief in COVID-19 patients.

✉ Vaddi Damodara Reddy
damodara.reddyv@reva.edu.in

- ¹ Department of Biochemistry, REVA University, Bangalore, Karnataka 560064, India
- ² Bioinformatics Research Laboratory, Biotechnology and Bioinformatics Institute, Dr. D. Y. Patil Vidyapeeth, Pune, Maharashtra 411033, India
- ³ DR Biosciences, Research & Development Unit, Bettahalasur, Bangalore 562157, India
- ⁴ Department of Molecular Biology, National Institute of Cancer Prevention and Research, Noida, Uttar Pradesh 20130, India
- ⁵ MIT School of Bioengineering Science and Research, MIT Art Design and Technology University, Pune, Maharashtra 412202, India
- ⁶ Department of ISRM, Indian Council of Medical Research, New Delhi 110029, India
- ⁷ Department of Radiation Oncology, Emory University School of Medicine, Atlanta, GA 30329, USA

Keywords COVID-19 · SARS-CoV-2 · Mutation analysis · FDA approved drugs · Molecular docking and simulation

Introduction

COVID-19, an infectious disease caused by coronavirus which is newly discovered in the year 2019. COVID-19 was first begun to emerge in the Wuhan seafood market in the Hubei province of China, it is caused by the SARS-CoV-2 (severe acute respiratory syndrome coronavirus-2). The coronavirus is a single-strand RNA virus (+ ssRNA) that can cause a severe respiratory syndrome in humans [31]. COVID19 is clinically characterized by fever, slight pain in the neck, headache, dry cough, breathing problem, vomit, and it might lead to death through multi-organ failure. COVID-19 has emerged as a severe epidemic declared by WHO and causing more than 3.43 million deaths worldwide between December 2019 to mid-May 2021 in around 200 countries (<https://www.worldometers.info/coronavirus/>) [33, 40].

Two types of coronavirus were reported, α -coronavirus that causes mild infections whereas, β -coronaviruses including SARS-CoV, and MERS-CoV (Middle East respiratory syndrome coronavirus) that caused pandemic in the recent past [18]. There were about 8000 SARS-CoV infected cases reported in the year 2002 and emerged as an epidemic in China [34]. Later, SARS-CoV, and MERS-CoV were reported as responsible for causing infection in humans [8]. The SARS-CoV-2 re-emergence in December 2019, has kept the world under high alert and has made a severe situation demanding fast treatment to prevent the infected patients [20, 32]. Despite expensive efforts globally by the researchers, there are no antiviral drugs or other treatments available to treat COVID-19 infected people. The current prevention methods are directed on quarantine and containment of infected patients for preventing human-to-human transmission along with vaccinations.

The spike protein (S) is a glycoprotein of SARS-CoV-2 and utilizes ACE2 for the host infection [33]. In the case of SARS-CoV, the spike protein on the virion surface mediates receptor identification as a class-1 fusion protein [7, 39]. During viral infection, the trimeric spike protein will cleavage into two subunits, S1 and S2. However, S1 subunits contain longer N-terminal and a receptor binding domain (RBD), while S2 subunit contains C-terminal membrane fusion domain followed by the two heptad regions (HR1 and HR2), the transmembrane (TM) domain and cytosolic tail [2, 27–29]. S2 is highly responsible for membrane fusion and contains the most conserved region. S1 directly binds to the host ACE2 receptor and resulting in

a wide open of cleavage site on S2 by host proteases; this process is critical for the viral infection. [22, 26, 39].

The human ACE2 receptor plays a major role that facilitates the viral entry in to the host and infection. Recent studies on the structure of the spike protein of SARS-CoV-2 revealed that spike protein binds to the Peptide Domain of ACE2 with a dissociation constant (K_d) of ~ 15 Nm [10, 15, 17, 35]. ACE2 is widely expressed in the lungs, heart, intestine, and kidneys [6, 19, 37]. Angiotensin, endogenous ligand of ACE2, is a peptide hormone that regulates vasoconstriction and in turn blood pressure. Many cardiovascular diseases are associated with the decreased expression of ACE2 [5, 23, 41].

The structure of the ACE2-PD alone is complex with the RBD of the spike protein of SARS-CoV. The ACE2 structural information is limited to the protein domain. The human ACE2 structure is combined with a membrane protein that is also called chaperones B⁰AT1 (SLC6A19). This structure of full-length ACE2 with cryo-electron microscopy revealed the presence of ACE2 with neutral amino acid transporter B⁰AT1 with RBD. Moreover, X-ray crystallography studies at a resolution of 2.9 Å of SARS-CoV-2 reported the structure of spike protein [3, 9, 11, 12, 14, 25, 30, 36]. The RBD is identified by the extracellular peptide domain of ACE2 through polar residues. The ACE2 and SARS-CoV-2 surface spike protein assembles into heterodimers with ACE2 homodimers. These findings will give major information insights into the molecular level for coronavirus recognition and infection. In this complex ACE2 is a dimer and further structure forms a basis for the development of novel therapeutic targets to prevent the viral infection.

The details of molecular interactions and characterization of these interactions between S-RBD and the ACE2 receptor of SARS-CoV-2 are crucial to develop the vaccine and therapeutic drug candidates for the prevention of the viral infection. The structural-based virtual screening of small molecules with RBD, spike protein of SARS-CoV-2 and ACE2 (PDB ID: 6VW1) will provide us with the potential drug candidates. Eventually, structure-based molecular docking was used to identify the inhibitor molecules to target SARS-CoV-2 virus. Recently, the crystal structure of the ACE2 complex was resolved along with the whole genome sequence of 2019-nCoV. In this study, we identified the binding interaction of potential drug candidates and validated the drug candidates using computational methods which lead a way for in vitro and in vivo studies.

Material and Methods

Data availability

The 8,719 SARS-CoV-2 isolates' whole-genome sequences were downloaded from the GISAID (<https://www.gisaid.org/>) virus database. The sample data was recorded up to date. The whole-genome data was downloaded in FASTA format and each isolate has a unique number for easy identification.

Open reading frame (ORF) identification

ORF finder (<https://www.ncbi.nlm.nih.gov/orffinder/>) was used to search DNA sequences for potential protein translation. This is the web version; here the query sequence range is 50 kb in length and can identify the spike protein sequence of SARS-CoV-2 for the further alignment process.

Spike protein multiple sequence alignment (MSA) analysis

The amino acid sequences from the different isolates of the SARS-CoV-2 were aligned by the multiple sequence alignment (MSA) method using ClustalW2 program from EMBL-EBI, <https://www.ebi.ac.uk/Tools/msa/clustalw2/server> and analyzed with default parameter to identify the mutations.

In silico structural characterization of SARS-CoV-2 spike protein

The 3-dimensional structure of the 2019-nCoV chimeric receptor-binding domain (RBD) complexed with its receptor human ACE2 was retrieved from the Protein Data Bank (www.rcsb.org) (PDB ID: 6VW1, 2.68 Å resolution). Structural information was examined and analyzed by the PDBsum online server (<https://www.ebi.ac.uk/pdbsum>). It includes the secondary structure, tertiary structure, bonds, and angles, dihedral angles of each amino acid, helix-helix, helix-beta, beta-beta, hydrogen bond, disulfide bonds, and covalent bond interactions. Apart from these studies, with the help of Maestro (**Schrödinger**), further studies like receptor preparation, the binding cavity of the protein analysis, were done.

Molecular modeling and energy minimization of drug library from different database

Total 28 potential anti-viral compounds were retrieved from PubChem (<https://pubchem.ncbi.nlm.nih.gov/>) and

Zinc databases (<http://zinc.docking.org/>) are summarized in Table 1. All retrieved compounds were converted into a 3-dimensional structure and subjected to energy minimized using the OPLS-2005 force field with the help of Maestro, a visualization software. This process of energy minimization was continued until the molecular conformations got stable or completely energy minimized.

Molecular docking studies of antiviral compounds with SARS-CoV-2 RBD protein

Molecular docking is a popular technique in structural bioinformatics to resolve problems related to protein and ligand interaction. In the context of docking, energy evaluations are usually carried out with the help of binding affinity (Gibbs free energy, ΔG), and structure-based drug design (SBDD).

All energetically minimized molecules were subjected to docking with SARS-CoV-2 RBD protein using FlexX software [<https://www.biosolveit.de/FlexX/>]. A grid box was defined based on the inhibitory site (Fig. 1) by considering vital residue coordinates by literature. Grid center was defined for SARS-CoV-2 RBD protein as per the functional and structural information obtained from various databases. FlexX supports SDF (Standard Database Format) file format for ligand and PDB format for the receptor. In FlexX the scoring is done using a modified Böhm scoring function including the loss of entropy upon ligand binding; hydrogen bonding; ionic, electrostatic interactions, aromatic interactions, and lipophilic interactions. Molecular docking interactions were analyzed using Maestro. Based on the high negative binding affinity score, the best confirmation of each ligand was used for further analysis.

Molecular simulation and dynamic studies of antiviral compounds with SARS-CoV-2 RBD protein

The stability of docked complexes and interaction of spike proteins with the antiviral lead compounds was observed by running 100 ns of molecular dynamic (MD) and simulation using the Desmond Molecular Dynamic System (D.E. Shaw Research, New York, NY, 2017). The complex system was solvated using the explicit solvent (TIP3P) system in the size of 10 Å cubic box with periodic boundary condition (PBC) box and the complex system was neutralized by adding counter ions using the System Setup panel of Desmond. The structural changes of the SARS-CoV-2 RBD (C- α backbone atoms) upon binding of lead compounds were compared with the docked structure of the respective complex in terms of RMSD (Root Mean Square Deviation). Also, the amino acids fluctuations

Table 1 Chemical compounds from PubMed and ZINC database

S. No	Compound name	Chemical ID	S. No	Compound name	Chemical ID
1	Silmitasertibl	CID_24748573	15	RVX-208	CID_135564749
2	AC_55541	CID_9589606	16	AZ8838	CID_126961334
3	Merimepodib	CID_153241	17	Daunorubicin	CID_30323
4	XI413	CID_135564632	18	Migalastat	CID_176077
5	AZ3451	CID_126961335	19	PD-144418	CID_9817231
6	E-52862	CID_44247568	20	TMB	CID_46943415
7	Ruxolitinib	CID_25126789	21	Mycophenolic acid	CID_446541
8	Metformin	CID_4091	22	S-verapamil	CID_92305
9	Ponatinib	CID_24826799	23	Valproic Acid	CID_3121
10	Hydroxychloroquine	CID_3652	24	JQ1	CID_46907787
11	GB110	CID_49843508	25	Bafilomycin A1	CID_6436223
12	Apicidin	CID_6918328	26	DCTPP1	ZINC1775962367
13	Haloperidol	CID_3559	27	DCTPP1	ZINC4326719
14	Indomethacin	CID_3715	28	DCTPP1	ZINC4511851

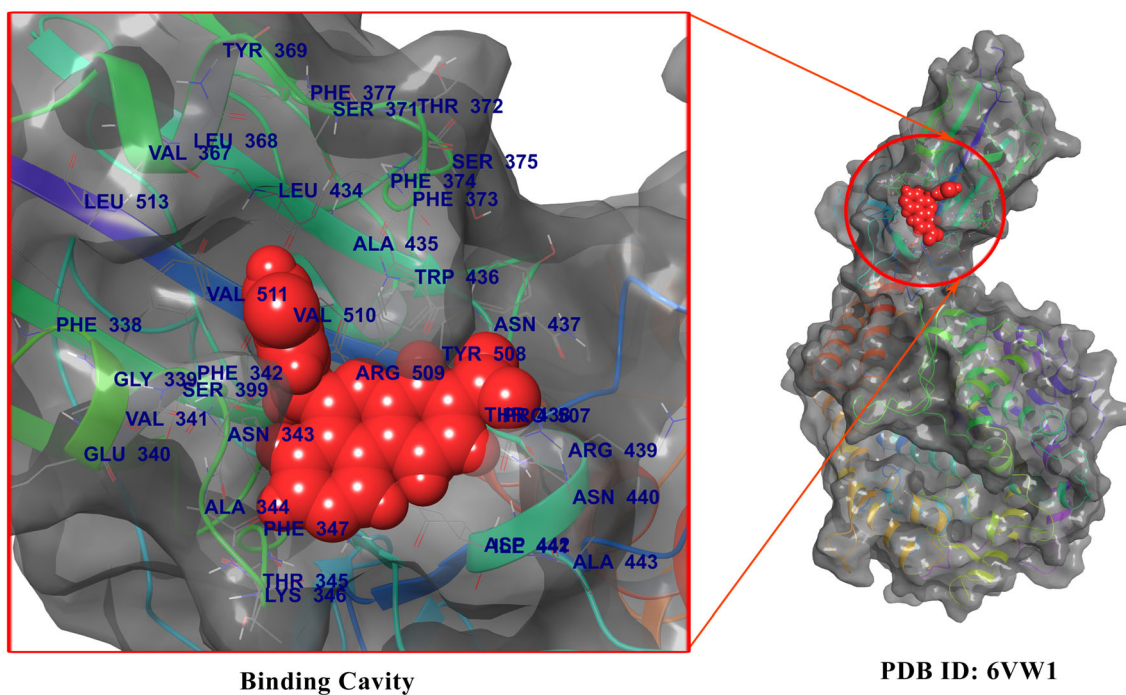


Fig. 1 Binding Pocket of the SARS-CoV-2 Spike Protein RBD region: The protein backbone shown in ribbon form and the cavity of the binding pocket shown in Red color. Amino acid residues that are

contributing to the binding site of spike RBD region are shown blue color (Color figure online)

during 100 ns MD simulation were computed and presented as RMSD (Root Mean Square Fluctuation) Plot.

Results

Sequence alignment

In the present study we have analyzed 8,719 spike protein sequences from all over the globe using the GISAID

database. The data were analyzed and the number of global mutations of spike protein were shown in Table 2. The infection of COVID-19, male-to-female ratio was 3:2 and with a median age range between 7 and 90 years old. We have identified more novel mutations using Multiple Sequence Alignment and mutational analysis. We have presented in Fig. 2 country wise total mutations and frequent mutations occurred in the SARS-CoV-2 Spike Protein RBD region. The total mutations and frequent mutations were listed in Table 3. Our study revealed that a

Table 2 Global spike protein mutations

S. No	Country name	Total no of spike sequences	Mutations
1	USA	2583	L5H (3), L5F (20), L10I (2), N17K (1), C15F (1), L18F (6), T22A (1), T22I (2), A27S (1), A27V (2), T29I (1), S31F (1), H49Y (1), L54F (1), A67V (1), H69Y (1), G75V (1), D80N (1), V83F (1), V90F (1), S94F (1), T95I (1), S98F (3), V127F (1), G142V (1), V143D (1), V143F (1), Y145H (8), N148S (1), M153I (3), M153V (1), V159I (1), G181V (1), R190K (1), R214L (1), D215Y (2), D215H (1), S221L (16), A243S (2), H245Y (1), H245R (1), D253G (1), G261A (1), A263V (1), E298G (1), E309Q (2), T323I (2), A348S (1), S359N (1), V367F (1), P384L (1), P384S (2), V395I (1), A411S (1), A475V (1), N501Y (1), V503F (1), A520S (1), A522S (1), T553I (2), E554D (1), A570T (1), T572I (4), E583D (1), D614G (2,025), T618I (1), V622F (1), A623S (1), P631S (10), S640F (1), S640A (1), D645G (1), A647S (2), A653S (1), H655Y (2), G652A (1), S659L (1), Q675H (1), Q675K (1), Q677H (1), T678I (1), N679K (1), S680F (1), P681L (1), R682W (1), S704P (1), A706V (1), S708F (1), N717T (1), M731I (2), T732I (1), D745G (1), L754F (1), G769V (1), T778I (1), E780V (1), P809S (1), P812L (3), D830G (1), A879S (1), G842S (1), A845S (2), A854S (1), T859I (8), A892S (1), A924V (1), G932D (1), S940F (1), T941I (1), D950H (1), S967R (1), T1027I (1), A1020S (1), M1050I (1), M1050V (1), A1078S (4), G1085R (5), H1101Y (5), D1118Y (1), V1133F (1), P1162L (1), D1163Y (1), V1176F (1), Q1180H (1), K1191N (1), N1194S (1), L1203F (2), K1205N (3), Q1208H (1), G1219C (1), G1219V (2), I1227M (1), V1228L (1), M1237T (5), M1237I (1), C1248F (1), S1252F (1), D1259H (1), P1263L (9)
2	Scotland	1134	M1C (1), F2L (1), L5F (5), L18F (2), R21I (3), P26S (1), A27S (1), T76I (1), V90F (4), T95I (1), E96D (1), V120F (1), G142A (3), H146Y (1), M153I (1), G181R (1), R190K (1), D198Y (1), L216F (1), H245Y (1), S247N (1), S254F (2), W258L (2), P272T (1), F338L (1), A344T (1), N354S (2), P384L (1), N439K (18), S477I (2), A522S (1), P588S (1), Q613H (1), D614G (847), A626S (1), A647S (1), H655Y (1), Y674F (1), A688V (1), I670M (4), Q675H (1), Q677H (1), A706V (1), G769V (3), P809S (1), A846V (1), A879S (7), A879T (5), S929T (2), S943I (2), T961M (1), T1077A (1), A1078V (1), P1079S (1), D1084Y (3), S1097T (1), H1101Y (1), W1102L (1), F1121L (8), G1124V (1), E1150D (1), P1162S (1), D1163Y (1), V1176F (2), K1191N (2), L1200F (6), H1201Y (1), E1207D (1), G1219V (2), V1228L (2), M1233I (1), C1248S (3), P1263L (5)
3	China	734	P9S (1), F32I (1), F32V (1), H49Y (23), S50L (7), L54F (1), A67S (1), H69Y (1), F92L (1), S98F (1), W152R (1), M153T (1), M153I (1), I197V (1), Y200S (1), D215G (1), D215H (1), G261R (2), S247R (2), S254F (1), R273M (1), F275Y (1), L276I (1), E298K (1), K300N (1), T302L (1), T307I (1), N317D (1), N354D (2), S359N (1), D364Y (1), K378R (1), Q409E (1), A411S (1), C432F (1), G485R (1), Q498H (1), H519P (1), T553I (1), S58N (1), A570V (3), P579Q (1), D614G (43), V615L (1), Y636F (1), H655Y (1), S659L (2), R682Q (1), N703S (1), S750R (1), S750I (1), L752F (1), L752R (2), G769A (1), T778N (1), P812T (1), I818S (1), G838S (1), A846V (1), L849F (1), V860Q (1), L861K (1), E868K (1), M900I (1), S943P (3), L948R (1), V952F (1), N969S (1), F970S (1), V1040F (1), V1060F (2), D1084Y (6), G1093A (1), F1103L (1), V1104L (1), V1129L (1), I1130L (3), D1146E (1), T1160K (1), S1161P (1), P1162S (2), P1162L (1), I1216T (1), G1171V (1), V1177A (1), I1216T (1), G1219C (1), S1252F (2), E1262G (1)
4	Japan	625	L5F (3), S13I (2), S50L (1), T95I (1), I119V (1), G181V (1), V213L (5), R357K (1), S477G (1), Q613H (11), D614G (331), P681H (3), S691F (1), R765L (1), I805M (2), A829T (2), A879S (1), D950Y (1), A1020D (1), T1027I (1), P1079T (1), K1086R (1), E1182K (1), E1202A (1), K1205N (1), G1219V (4), V1228L (1), M1229I (7)
5	South Africa	592	M1I (1), L5F (1), L8F (3), V11F (1), L54F (1), H69Y (1), V70F (1), D80Y (1), T95I (1), N121T (1), W152R (1), W152L (1), S155T (1), N188I (1), D198G (1), R214L (2), S221L (2), A222V (1), A344S (1), S373L (1), P384S (2), P385L (1), G413E (1), Y453F (1), I468V (1), Y508H (1), K558R (1), F562Y (1), D614G (579), V622F (1), H625Q (1), A653V (7), H655Y (3), N658D (1), T673P (1), T676P (2), Q677H (1), A688V (24), G769V (6), T778I (1), E780Q (1), P809S (2), P812L (4), A846V (4), N914S (1), L938I (3), G946R (1), A1020S (1), N1074H (2), A1078S (8), E1202D (1), L1203F (1), M1237I (1), C1243F (1), L1244F (1), T1273S (1), D1260Y (3)
6	France	556	L5F (6), L8W (1), L54F (4), A93S (1), L118F (1), V120I (3), S151I (1), W152R (1), M153I (1), V213L (1), T240I (1), L296H (1), V367F (5), C379F (1), V382E (1), Y508H (1), T553N (1), T572I (3), D574Y (1), L611F (1), D614G (517), A623S (1), S640F (1), A653V (1), T676I (3), S686G (1), S689I (1), I720V (1), R765L (1), N801K (1), P812L (1), A845V (1), A845S (9), A924S (1), D936Y (1), S939F (1), S940F (1), A1020V (1), L1063F (1), V1104L (1), I1132V (1), P1143L (2), P1162A (1), P1162L (3), G1219V (3), S1242I (1), D1260N (1), P1263L (1)

Table 2 continued

S. No	Country name	Total no of spike sequences	Mutations
7	India	542	F2L (1), L5F (3), L7I (1), L18F (1), T22I (3), Y28H (2), H49Y (1), L54F (5), K77M (11), R78M (10), P82L (1), E156D (1), S162I (1), M177I (2), E224G (1), A243S (2), G261S (4), Q271R (2), T274I (2), D290Y (1), T299I (1), I402L (1), E471Q (1), A520S (6), S494P (1), K558N (2), T572I (4), D574H (1), E583D (7), T602I (1), D614G (179), R634S (1), Q677H (2), A706S (2), T323I (1), L828P (1), A831S (1), A930T (1), S1021F (1), T1027I (1), H1101Y (1), V1104L (1), G1124V (3), D1163F (1), I1179N (1), K1191N (1), C1243F (1), C1250F (1), G1251V (6), P1263L (1), M731R (1), T761S (1), G769V (1), A771V (1), F797C (1), G857C (1), A879S (1), A892V (1), V1068F (1)
8	Spain	462	V6I (1), T20S (1), P25S (1), F32Y (1), G75D (1), T95I (1), S98F (1), R102I (1), N125H (1), N148K (1), N149H (1), Q183L (1), I197V (5), D198G(1), V213L (1), A222V (1), H245Y (1), S254F (1), V483F (3), S594P (1), D614G (276), A623V (1), Q675R (1), A684V (1), I788M (1), T791K (1), Y837L (1), G838V (1), L841I (1), A871S (1), Q935H (1), Q954K (1), A1020S (1), V1068A (1), V1094F (1), T1120I (1), G1124 (3), G1167V (1) G1219C (1), G1219V (1), M1229I (1), D1259Y (1), L1265P (1)
9	Germany	339	S13I (1), T22I (1), T29I (1), S49L (2), S50L (1), T95I (1), S254F (2), K417N (2), K558R (2), T573I (1), D614G (292), V620I (2), E780Q (1), S789I (1), P809S (1), L938F (1), E990A (1), M1050V (1), D1162G (1), T1117I (1), K1191N (1)
10	Finland	263	L5F (4), A27S (1), G75V (1), S94F (1), D138H (1) Y144F (1), M153I (2), D215H (1), Q239K (1), A243P (1), L244F (1), L249S (1), A435S (1), G446V (1), D614G (253), H625R (5), Q675H (4), A684V (1), A688V (1), I692F (1), V860F (1), A879S (8), A899S (1), D936Y (179), L1063F (1), G1124C (1), V1133F (1), P1162L (1)
11	Russia	240	D614G (211), D261G (1), V193L (1), D198G (2)
12	Brazil	220	D164G (187), V1175F (4), K776T (1), L938F (1)
13	UK	141	L18F (1), R21I (1), V289I (1), D614G (84), A845S (1), A879S (1), I934M (1), D936Y (1), T1117I (1), V1133F (1), P1263L (5)
14	Italy	128	L5F (1), L216J (1), Q239K (1), L300J (1), A570D (2), D614G (120), N960B (1), T630S (1), D839Y (1)
15	Norway	53	L5F (3), V11I (1), A263S (1), G339D (1), Q613H (1), D614G (43), Q675H (1), A845T (1), A892S (1), D936Y (1), A1020V (3)
16	Australia	48	R78K (2), V445A (1), P491R (1), D614G (13), M731I (1), H1058Y (1)
17	Peru	34	L5F (1), I119V (8), D614G (29)
18	Mexico	20	H49Y (5), T573I (1), D614G (9), D1259Y (1)
19	Iran	5	T22I (3)

total of 639 mutations were found globally. The frequent mutations in different isolates are L5F, T22I, T29I, H49Y, L54F, V90F, S98F, S221L, S254F, V367F, A520S, T572I, D614G, H655Y, P809S, A879S, D936Y, A1020S, A1078S, and H1101Y. Moreover, our study found common

mutations and the data was presented in Table 4; L5F mutation was identified in the USA, Scotland, Japan, South Africa, France, India, Italy, Norway, and Peru. T22I mutation was observed in the USA, India, Germany, and Iran. The T29I mutation variant is present in Spain and

Fig. 2 Global numbers of the Spike protein mutations: The country wise total mutations (blue) and frequent mutations (orange) occurred in the SARS-CoV-2 Spike Protein RBD region is shown in the histogram (Color figure online)

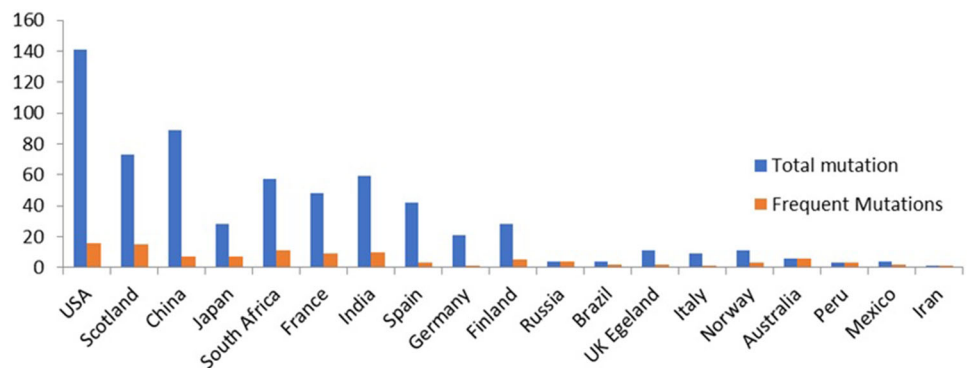


Table 3 Country wide total mutations of spike protein and most frequent mutations in spike protein region of SARS-CoV-2

S. No	Country name	Total mutation	Most frequent mutations (= > 3 times repeated)
1	USA	141	L5H, L18F, Y145H, M153I, D614G, S221L, T572I, T859I, A1078S, G1085R, H1101Y, P1263L, L5F, S98F, P812L
2	Scotland	73	L5F, R211I, V90F, G142A, N439K, D614G, I670M, G769V, A879S, A879T, D1084Y, F1121L, L1200F, C1248S, P1263L
3	China	89	H49Y, S50L, T307I, A570V, D614G, D1084Y, I1130L
4	Japan	28	L5F, V213L, Q613H, D614G, P681H, G1219V, M1229I
5	South Africa	57	L8F, D614G, A653V, H655Y, A688V, G769V, P812L, A846V, L938I, A1078S, D1260Y
6	France	48	L5F, L54F, V120I, V367F, T572I, D614G, T676I, P1162L, G1219V
7	India	59	L5F, T221, L54F, K77M, R78M, D614G, G1251V, A502S, G261S, G1124V
8	Spain	42	I197V, V483F, D614G
9	Germany	21	D614G
10	Finland	28	L5F, D614G, H675H, A879S, D936Y
11	Russia	4	D614G, D198G
12	Brazil	4	D614G, V1175F
13	UK	11	D614G, P1263L
14	Italy	9	D614G
15	Norway	11	D614G, A1020V, L5F
16	Australia	6	R78K, D614G
17	Peru	3	I119V, D614G
18	Mexico	4	H49Y, D614G
19	Iran	1	T22I

Germany, whereas H49Y is the most common mutation in the USA, India, China, and Mexico. The L54F mutation was present in the USA, China, South Africa, France, and India. V90F mutation was observed in the USA and Scotland. The other mutations, S98F mutation (USA, China, and Spain), S221L mutation (USA and South Africa), S254F and V367F mutations (USA and France), A520S mutations (USA and India), T572I mutation (USA, India and along with France) were observed in various countries. However, D614G mutations are high in worldwide isolates. In the USA, the total sequences that we have analyzed are 2,583 and the number of D614G mutations are 2025 (82%). Similarly in Scotland out of 1134 sequences, the number of D614G mutations are 847 (75%). Out of total 734, 625, 592, 556, 542, 462, 339, sequences from China, Japan, South Africa, France, India, Spain and Germany, D614G mutations are 43 (6%), 331 (53%), 579 (98%), 517 (93%), 179 (33%), 276 (60%) and 292 (86%) respectively. Out of total Finland sequences 263, D614G are 253 (96%); Russia sequences 240, D614G are 211 (88%); Brazil sequences 220, D614G are 187 (85%); UK sequences 141, D614G 84 (60%); Italy sequences 128, D614G are 120 (95%); Norway sequences 53, D614G are 43 (81%); Australia sequences 48, D614G are 13 (27%);

Peru sequences 34, D614G are 29 (85%); Mexico sequences 20 of D614G are 9 (45%). H655Y mutation was observed in the USA, Scotland, China, South Africa while P809S mutation was present in the USA, Scotland, South Africa, Germany. A879S mutation was present in the USA, Scotland, Japan, India, Finland, UK. And A1020S and A1078S mutations are present in the USA and South Africa and Spain. H1101Y mutations are present in the USA, Scotland, and India.

Predicted mutations did not occur in the ACE2 binding site and the role of mentioned mutations is yet to be predicted. All these mutations are based on a comparison with the Wuhan genome sequence (NC_045512.1). These spike protein mutations are present at both upstream and downstream of the receptor-binding domain (RBD) and the mutation Q271R has been found to affect the secondary structure of the S1 domain. The other major mutation at D614G at the position 614 with glycine instead of aspartic acid in the spike protein. Generally, all the amino acids can occur in two isomeric forms except glycine. This could be one of the reasons for spread of the disease faster. Hence, these mutations are unique and might impact the spike receptor interaction with changing of its conformations in the protein. Very importantly, positions of these mutations

Table 4 Common mutations of Spike Protein—Global Segregates

S.No	Mutations	Countries
1	L5F	USA, Scotland, Japan, South Africa, France, India, Finland, Italy, Norway, Peru
2	T22I	USA, India, Germany, Iran
3	T29I	Spain, Germany
4	H49Y	USA, India, China, Mexico
5	L54F	USA, China, South Africa, France, India
6	V90F	USA, Scotland
7	S98F	USA, China, Spain
8	S221L	USA, South Africa
9	S254F	USA, France
10	V367F	USA, France
11	A520S	USA, India
12	T572I	USA, France, India
13	D614G	USA, Scotland, China, Japan, South Africa, France, India, Spain, Germany, Finland, Russia, Brazil, UK, Italy, Norway, Australia, Peru, Mexico
14	H655Y	USA, Scotland, China, South Africa
15	P809S	USA, Scotland, South Africa, Germany
16	A879S	USA, Scotland, Japan, India, Finland, UK England
17	D936Y	France, Finland, UK England, Norway
18	A1020S	USA, South Africa, Spain
19	A1078S	South Africa, USA
20	H1101Y	USA, Scotland, India

might be important to find target sites for COVID-19 treatment. Based on the literature review we identified few immune-promoting compounds (Silmitasertib, AC-55541, Merimepodib, XL413, and AZ3451) which could show effectiveness to treat SARS-CoV-2.

Structural analysis

The PROMOTIF (Hutchinson, 1996) program provides information about the secondary structure of a given protein for analysis. As for this analysis, PDB ID: 6VW1 contains chain A and chain B having 596 amino acids with two sheets, two beta hairpins, one beta bulge, six strands, thirty-two helices, fifty six helix-helix interaction, thirty one beta turns, five gamma turns and three disulfides. Other E and F chains contain 193 amino acids with three sheets, one beta bulge, nice strands, seven helices, three helix-helix interaction, thirteen beta turns, three gamma turns and four disulfides.

Docking with SARS-CoV-2 RBD protein

All the designed analogs were docked with SARS-CoV-2 RBD using FlexX software, The docking result shows that the binding score for all the compounds is highly negative

suggesting that their binding affinity is greater towards SARS-CoV-2 RBD protein. The 28 analogs were screened for the interaction with the amino acid of the human ACE2 receptor which are FDA approved drugs for the treatment of different cancers and immune promotion, may work against SARS-CoV-2 Silmitasertib, AC-55541, Merimepodib, XL413, and AZ3451 demonstrated the best docking score with hydrogen bonding and hydrophobic interaction with amino acid residues in active site of RBD protein and selected compounds were analyzed using Maestro. Interacting amino acids, bond types, and their bond distances were noted in Table 5.

We have come to a conclusion based on docking studies that the Silmitasertib showed more affinity towards SARS-CoV-2 RBD protein compared with other compounds. Silmitasertib has a docking score of -26.7713 and also shown a hydrogen bond at a distance 1.85 \AA with Phe342, four aromatic hydrogens (AH) bonds with Phe343, Asn343 (2AH), Ala344 with their respective bond distances 2.73 \AA , 1.81 \AA , 2.33 \AA , and 2.62 \AA . Also, one halogen bond occurs with Asn343 at 3.34 \AA bond distance, two pi-pi interactions with Trp436 and bond distance respectively 4.46 \AA and 4.78 \AA , and one pi-cation with Arg509 at bond distance 5.73 \AA (Fig. 3a).

Table 5 Docking analysis of five lead antiviral compounds against SARS-CoV-2 spike RBD region

Compound name	PubChem CID	Docking score (kcal/mole)	Interacting residue	Type of bond	Bond distance (Å)
Silitasertib	24748573	− 26.7713	Phe342	H-Bond	1.85
			Phe342	Aromatic H-Bond	2.73
			Asn343	Halogen Bond	3.34
			Asn343	2 Aromatic H-Bond	1.81, 2.33
			Trp436	2 Pi–Pi	4.46 and 4.78
			Ala344	Aromatic H-Bond	2.62
			Arg509	Pi-cation	5.73
			AC55541	9589606	− 24.3574
Merimepodib	153241	− 22.2024	Phe373	Pi–Pi	4.91
			Phe373	2 Aromatic H-Bond	2.43, 2.74
			Phe342	2 Aromatic H-Bond	2.23, 2.78
			Phe342	2 H-Bond	2.08, 2.08
			Th372	H-Bond	2.41
			Trp436	Pi–Pi	5.09
			Arg509	Pi-Cation	5.59
			Asn343	Aromatic H-Bond	2.44
XL-413	135564632	− 20.7836	Ala344	Aromatic H-Bond	2.27
			Ala344	2 H-Bond	1.83, 2.00
			Asn343	H-Bond	1.59
			Trp436	H-Bond	1.77
			Trp436	Aromatic H-Bond	3.56
			Phe373	Pi–Pi	5.14
			Asn437	Aromatic H-Bond	2.66
			AZ3451	126961335	− 20.647
			Thr345	Aromatic H-Bond	2.35
			Phe342	Aromatic H-Bond	1.69

The second lead compound was shown in Fig. 3b i.e., AC55541 have a docking score of − 24.35 which forms three hydrogen bonds with Asn343 (bond distances—1.60 Å, 1.85 Å and 2.42 Å), one pi–pi interaction with Phe373 (bond distance 4.91 Å) and four aromatic hydrogen bonds with Phe373 (2AH) and Phe342 (2AH) (bond distance 2.43 Å, 2.74 Å, 2.23 Å, and 2.78 Å respectively).

The third lead compound i.e., Merimepodib was shown in Fig. 3c having docking score of − 22.20 which forms two equal hydrogen bonds with Phe342 (bond distance 2.08 Å), one hydrogen bond with Thr372 (bond distance of 2.41 Å), one pi–pi interaction with Trp436 (bond distance 5.09 Å), one pi-cation interaction with Arg509 (bond distance 5.59 Å) and two aromatic hydrogen bonds with Asn343, Ala344 (bond distances of 2.44 Å and 2.27 Å).

Fourth lead compound was XL-413 having docking score of − 20.7836 which forms four hydrogen bonds with Ala344 (two bonds), Asn343, and Trp436 (bond distance 1.83 Å, 2.00 Å, 1.59 Å and 1.77 Å respectively), two aromatic hydrogen bonds with Trp436 and Asn437 (bond

distance 3.56 Å and 2.66 Å) and one pi–pi interaction with Phe373 (bond length 5.14 Å) (Fig. 3d).

The fifth lead compound AZ3451 shown in Fig. 3e having a docking score of − 20.64 it forms one hydrogen bond with Asn343 (bond distance 1.61 Å) and two aromatic hydrogen bonds with Thr345 and Phe342 (bond distance 2.35 Å and 1.69 Å).

From the literature study, the screened high-affinity lead compounds were reported for the immunomodulatory effects. The screened compounds have shown high docking scores with strong binding energies, and close interaction with conserved catalytic residues. These results indicate that small molecules identified in our study are promising drugs for COVID-19 treatment.

Molecular dynamics and simulation analysis

To explore dynamic perturbation in the conformation of the spike protein and ligand complex molecular dynamic simulation was carried out. The 100 ns simulations have

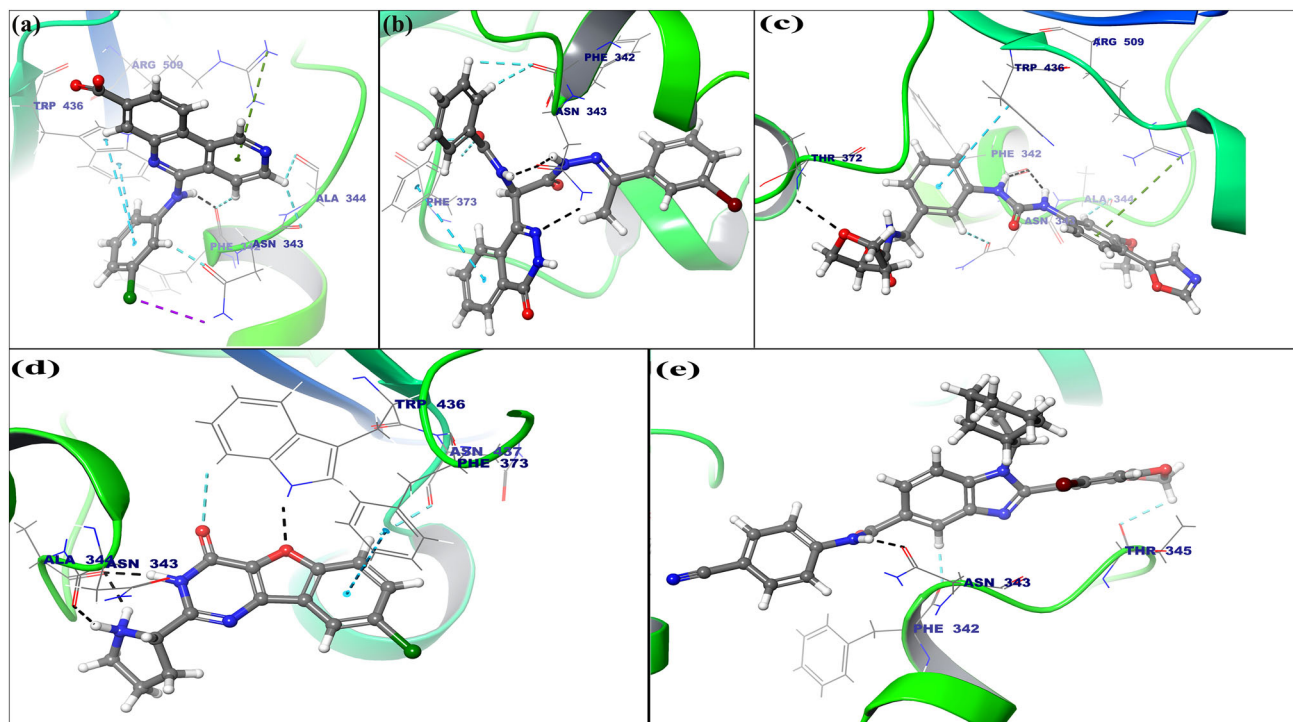


Fig. 3 Intermolecular interactions between docked FDA approved drugs candidates vs SARS-CoV-2 spike RBD region: (a) Binding mode of Siltimasertib with stable SARS-CoV-2 spike RBD, (b) Binding mode of AC55541 with stable SARS-CoV-2 spike RBD (c) The binding mode of Merimepodib within the binding site SARS-CoV-2 spike RBD, (d) Binding mode of XL-413 with SARS-CoV-2 spike RBD protein. (e) Binding mode of AZ3451 within the SARS-CoV-2

spike RBD binding cavity. The hydrogen bonds are shown in dashed Black lines whereas, Pi-Pi stacking, Pi-cation interactions, and aromatic hydrogen bonds were represented in Cyan, Green and Teal color, respectively. The SARS-CoV-2 spike RBD protein is shown in the ribbon form, and the interacting amino acids and the drug candidates are represented in line and ball and stick model (Color figure online)

run for the spike (RBD region) receptor, the potential energy tends to decrease which intimate the stabilization of the system. The conformations obtained during the 100 ns simulation for the RBD-ligand complex were analyzed and RMSD (Root mean square deviation) was calculated during the simulation time. The displacement of selected atoms can be calculated in order of the average change for the particular frame concerning the reference frame.

The RMSD plot of RBD and antiviral drugs is shown in Fig. 4a indicating that all the complex systems get equilibrated from 20 to 100 ns MD simulation. In the initial phase (0–20 ns) of MD simulation, deviations in the complex systems were observed and the complex system gets stabilized from 2.5 to 4.0 Å.

Moreover, to probe the structural changes of Spike RBD upon binding of the antiviral compounds, the RMSF (Root mean square fluctuation) for C-alpha atoms of all the residues were compared using trajectories generated during 100 ns molecular dynamic simulations. It is outstanding that the RMSF trajectories of RBD (Fig. 4b) stability of the system is associated with more rigid and stable conformations. The RMSF graph suggests that Spike RBD shows the minimum fluctuation after binding of antiviral lead

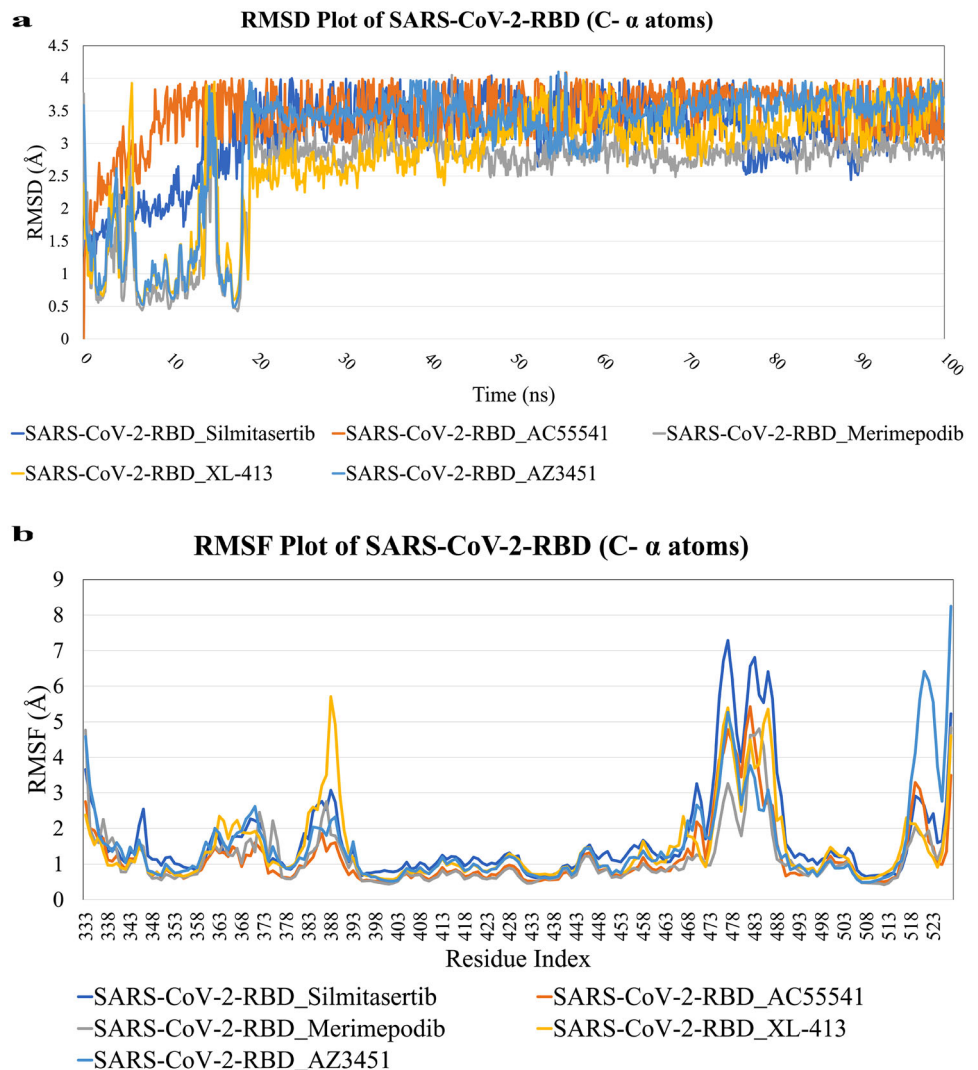
compounds, only the fluctuation in the region of residues (473–490) fluctuated from 3 to 7 Å. From this MD simulation, we can conclude that the lead compounds form the stable complex with Spike RBD and are thermodynamically favorable.

Discussion

When we embarked the SARS-CoV-2 sequence analysis our interest was turned to identify the mutations that could be a potential target for spike protein to find drug compounds. The GISAID database provides data to look at more deep sequence analysis and evolutionary relationships with many SARS-CoV-2 sequences. In this context, our approach revealed various mutations in the spike protein region, the sequence analysis done by comparing with the initial Wuhan isolate spike protein sequence.

We have analyzed clinical data of spike protein sequences thoroughly and found different mutation in spike region. These mutation help virus to spread fast and increase the virus load [1]. Our studies shown that the spike being manifested among with all sites of the frequent

Fig. 4 Molecular dynamic simulation analysis: **a** RMSD plot for the SARS-CoV-2 spike RBD (C- α atoms) molecular simulated with antiviral drug candidates within 100 ns time scale. **b** RMSF plot for the SARS-CoV-2 spike RBD (C- α atoms) docked with antiviral drug candidates within 100 ns time scale in molecular dynamic simulation



mutations viz. L5F, T22I, T29I, H49Y, L54F, V90F, S98F, S221L, S254F, V367F, A520S, T572I, D614G, H655Y, P809S, A879S, D936Y, A1020S, A1078S, and H1101Y. Among them the most appearance mutation is D614G, identified as high frequent mutation worldwide. D614G is an important mutation that causes structural changes which in turn increase the affinity towards ACE2 and leading to functional changes in SARS-CoV-2 spike protein. While the other mutations may have structural changes nearby the receptor-binding domain in the S1 domain of the spike. This mutation could show an impact on the binding activity of the S1 domain, and it can change secondary structure of the protein [1]. The mutations play an important role in to change evolutionary origins of the new isolates. For viruses like SARS-CoV-2 also the natural law of survival of fittest applies. This leads to the creation of new mutations in spike protein that result in new variants with strong binding to the host receptor to increase the virulence capacity and help in escaping the immune system. The hotspot region of

the mutations in the S1 and S2 conserved domain will make a stronger virus to attach the host receptor for viral entry. Hence, exploring this area will help to the best antigen immunogenic region for vaccine development, and other antiviral small molecule drugs. These compounds may be neutralizing the virus potentiality and reduce the infection. For this, we have done a virtual screening of compounds against SARS-CoV-2 binding receptor ACE2.

From the docking analysis, we have selected 5 lead compounds that were having potential inhibitor activity on spike protein RBD binding with ACE2. Among five drugs the recently developed CK2 inhibitor, Silmitasertib, selectively binds and inhibits the enzyme casein kinase II (CK2) in different cancers, also reported anti-fungal activity [21]. Silmitasertib has been tested in phase I/II COVID-19 clinical trials <https://clinicaltrials.gov/ct2/show/NCT04668209>) and shown highest binding affinity, -26.44 kcal/mol among all tested compounds. AC-55541 is a potent and selective protease-activated receptor

2 (PAR-2) agonist [24]. Our docking analysis showed binding affinity score -24.35 kcal/mol. PARs are transmembrane receptors and play a major role in inflammatory responses. They act as a sensor to the proteolytic enzyme formation during viral infection. When tested with human bronchial epithelial cells, the administration of AC-55541 significantly induced the expression of PAR2 and inhibited the Acanthamoeba plasminogen activator.

Merimepodib (MMPD) is a novel selective inhibitor of inosine monophosphate dehydrogenase (IMPDH) and was reported for antiviral effects [38]. Our docking analysis revealed that binding affinity score of MMPD is -22.20 kcal/mol. The inhibition of IMPDH can reduce the available intracellular guanine nucleotide concentration for viral RNA synthesis. Besides, it is an immune modulator and antiviral agent in combination with interferons (INF) [16]. The significant treatment for viral infections and other clinical benefits of MMPD pegylated interferon may be effective for the treatment of COVID-19. XL-413 is a potent and selective ATP competitive CDC7 inhibitor [13]. XL-143 is a novel bezofuroprymidinone inhibitor with a favorable pharmacokinetics property denoted for the significant tumor growth regression and arrest the cell cycle (S phase to the G2 phase) of tumor cells will prevent viral infections at molecular level [13]. Our docking analysis showed XL-143 has strong binding affinity (-20.78 kcal/mol) for spike RBD region. AZ3451 is a PAR2 antagonist and binds to the remote allosteric site outer to the helical bunch. Our results showed AZ3451 has stronger binding affinity (-20.64 kcal/mol) against SARS CoV-2 RBD region. These antagonists prevent structural rearrangements, which required for the receptor activation and signaling cascade mechanisms, and can be effectively used as a therapeutic agent to treat viral infections [4].

In conclusion, we have identified mutations in spike RBD region that lead to structural changes of spike are most important for the interaction with host ACE2 receptor. Our study reveals that Silmitasertib, AC-55541, Merimepodib, XL413 and AZ3451 showed promising results against the novel emerged strain of coronavirus. All these compounds have exhibited excellent binding capacity to SARS-CoV-2 RBD protein. These compounds may be effective to control or stop the viral entry and further infection, as well our study paves a way for further in vivo studies as well clinical trials.

Acknowledgements Authors thank DST-SERB and ICMR for providing financial assistance to Dr. V. Damodara Reddy under Ramanujan Fellowship (SB/S2/RJN-043/2014) and Mrs. Swetha Pulakuntla under ICMR-SRF (ISRM/11(20/2017) scheme. Authors also thank DST-SERB GPU optimized Super computer for calculations at Dr. D. Y. Patil Biotechnology and Bioinformatics Institute, Pune India.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

- Banerjee AK, Begum F, Thagriki D, Tripathi PP, Ray U. Novel mutations in the S1 domain of COVID 19 spike protein of isolate from Gujarat Origin, Western India. Preprints 2020, 2020040450. <https://doi.org/10.20944/preprints202004.0450.v1>.
- Belouzard S, Chu VC, Whittaker GR. Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. *Proc Natl Acad Sci U S A*. 2009;7(106):5871–6. <https://doi.org/10.1073/pnas.0809524106>.
- Bröer A, Klingel K, Kowalczyk S, Rasko JE, Cavanaugh J, Bröer S. Molecular cloning of mouse amino acid transport system B0, a neutral amino acid transporter related to Hartnup disorder. *J Biol Chem*. 2004;279:24467–76. <https://doi.org/10.1074/jbc.M400904200>.
- Cheng RKY, Fiez-Vandal C, Schlenker O, Edman K, Aggeler B, Brown DG, Brown GA, Cooke RM, Dumelin CE, Doré AS, Geschwindner S, Grebner C, Hermansson NO, Jazayeri A, Johansson P, Leong L, Prihandoko R, Rappas M, Soutter H, Snijder A, Sundström L, Tehan B, Thornton P, Troast D, Wiggin G, Zhukov A, Marshall FH, Dekker N. Structural insight into allosteric modulation of protease-activated receptor 2. *Nature*. 2017;545:112–5. <https://doi.org/10.1038/nature22309>.
- Crackower MA, Sarao R, Oudit GY, Yagil C, Kozieradzki I, Scanga SE, et al. Angiotensin-converting enzyme 2 is an essential regulator of heart function. *Nature*. 2002;417:822–8. <https://doi.org/10.1038/nature00786>.
- Donoghue M, Hsieh F, Baronas E, Godbout K, Gosselin M, Stagliano N, Donovan M, Woolf B, Robison K, Jeyaseelan R, Breitbart RE, Acton S. A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1–9. *Circ Res*. 2000;87:E1–9. <https://doi.org/10.1161/01.res.87.5.e1>.
- Gallagher TM, Buchmeier MJ. Coronavirus spike proteins in viral entry and pathogenesis. *Virology*. 2001;279:371–4. <https://doi.org/10.1006/viro.2000.0757>.
- Gaunt ER, Hardie A, Claas EC, Simmonds P, Templeton KE. Epidemiology and clinical presentations of the four human coronaviruses 229E, HKU1, NL63, and OC43 detected over 3 years using a novel multiplex real-time PCR method. *J Clin Microbiol*. 2010;48:2940–7. <https://doi.org/10.1128/JCM.00636-10>.
- Hamming I, Cooper ME, Haagmans BL, Hooper NM, Korstanje R, Osterhaus AD, Timens W, Turner AJ, Navis G, van Goor H. The emerging role of ACE2 in physiology and disease. *J Pathol*. 2007;212:1–11. <https://doi.org/10.1002/path.2162>.
- Hoffmann M, Kleine-Weber H, Schroeder S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*. 2020;181:271–280.e8. <https://doi.org/10.1016/j.cell.2020.02.052>.
- Kirchdoerfer RN, Wang N, Pallesen J, et al. Stabilized coronavirus spikes are resistant to conformational changes induced by receptor recognition or proteolysis. *Sci Rep*. 2018;8:17823. <https://doi.org/10.1038/s41598-018-36918-8>.
- Kleta R, Romeo E, Ristic Z, Ohura T, Stuart C, et al. Mutations in SLC6A19, encoding B0AT1, cause Hartnup disorder. *Nat Genet*. 2004;36:999–1002. <https://doi.org/10.1038/ng1405>.
- Koltun ES, Tshako AL, Brown DS, Aay N, Arcalas A, Chan V, Du H, Engst S, Ferguson K, Franzini M, Galan A, Holst CR, Huang P, Kane B, Kim MH, Li J, Markby D, Mohan M, Noson K,

- Plonowski A, Richards SJ, Robertson S, Shaw K, Stott G, Stout TJ, Young J, Yu P, Zaharia CA, Zhang W, Zhou P, Nuss JM, Xu W, Kearney PC. Discovery of XL413, a potent and selective CDC7 inhibitor. *Bioorg Med Chem Lett*. 2012;22:3727–31. <https://doi.org/10.1016/j.bmcl.2012.04.024>.
14. Kowalczyk S, Bröer A, Tietze N, Vanslambrouck JM, Rasko JE, Bröer S. A protein complex in the brush-border membrane explains a Hartnup disorder allele. *FASEB J*. 2008;22:2880–7. <https://doi.org/10.1096/fj.08-107300>.
 15. Kuba K, Imai Y, Rao S, Gao H, Guo F, et al. A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus-induced lung injury. *Nat Med*. 2005;11:875–9. <https://doi.org/10.1038/nm1267>.
 16. Lauer GM, Walker BD. Hepatitis C virus infection. *N Engl J Med*. 2001;345:41–52. <https://doi.org/10.1056/NEJM200107053450107>.
 17. Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, Somasundaran M, Sullivan JL, Luzuriaga K, Greenough TC, Choe H, Farzan M. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature*. 2003;426:450–4. <https://doi.org/10.1038/nature02145>.
 18. Liu Z, Xiao X, Wei X, Li J, Yang J, Tan H, Zhu J, Zhang Q, Wu J, Liu L. Composition and divergence of coronavirus spike proteins and host ACE2 receptors predict potential intermediate hosts of SARS-CoV-2. *J Med Virol*. 2020;92:595–601. <https://doi.org/10.1002/jmv.25726>.
 19. Ma M, Lin XH, Liu HH, Zhang R, Chen RX. Suppression of DRP1-mediated mitophagy increases the apoptosis of hepatocellular carcinoma cells in the setting of chemotherapy. *Oncol Rep*. 2020;43:1010–8. <https://doi.org/10.3892/or.2020.7476>.
 20. Malik YS, Sircar S, Bhat S, Sharun K, Dhama K, Dadar M, Tiwari R, Chaicumpa W. Emerging novel coronavirus (2019-nCoV)-current scenario, evolutionary perspective based on genome analysis and recent developments. *Vet Q*. 2020;40:68–76. <https://doi.org/10.1080/01652176.2020.1727993>.
 21. Masłyk M, Janeczko M, Martyna A, Kubiński K. CX-4945: the protein kinase CK2 inhibitor and anti-cancer drug shows antifungal activity. *Mol Cell Biochem*. 2017;435:193–6. <https://doi.org/10.1007/s11010-017-3068-z>.
 22. Millet JK, Whittaker GR. Host cell proteases: critical determinants of coronavirus tropism and pathogenesis. *Virus Res*. 2015;202:120–34. <https://doi.org/10.1016/j.virusres.2014.11.021>.
 23. Raizada MK, Ferreira AJ. ACE2: a new target for cardiovascular disease therapeutics. *J Cardiovasc Pharmacol*. 2007;50:112–9. <https://doi.org/10.1097/FJC.0b013e3180986219>.
 24. Seitzberg JG, Knapp AE, Lund BW, Mandrup Bertozzi S, Currier EA, Ma JN, Sherbukhin V, Burstein ES, Olsson R. Discovery of potent and selective small-molecule PAR-2 agonists. *J Med Chem*. 2008;51:5490–3. <https://doi.org/10.1021/jm800754r>.
 25. Seow HF, Bröer S, Bröer A, Bailey CG, Potter SJ, Cavanaugh JA, Rasko JE. Hartnup disorder is caused by mutations in the gene encoding the neutral amino acid transporter SLC6A19. *Nat Genet*. 2004;9:1003–7. <https://doi.org/10.1038/ng1406>.
 26. Simmons G, Gosalia DN, Rennekamp AJ, Reeves JD, Diamond SL, Bates P. Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry. *Proc Natl Acad Sci U S A*. 2005;102:11876–81. <https://doi.org/10.1073/pnas.0505577102>.
 27. Simmons G, Reeves JD, Rennekamp AJ, Amberg SM, Piefer AJ, Bates P. Characterization of severe acute respiratory syndrome-associated coronavirus (SARS-CoV) spike glycoprotein-mediated viral entry. *Proc Natl Acad Sci U S A*. 2004;101:4240–5. <https://doi.org/10.1073/pnas.0306446101>.
 28. Simmons G, Zmora P, Gierer S, Pöhlmann S. Proteolytic activation of the SARS-coronavirus spike protein: cutting enzymes at the cutting edge of antiviral research. *Antivir Res*. 2013;100:605–14. <https://doi.org/10.1016/j.antiviral.2013.09.028>.
 29. Song W, Gui M, Wang X, Xiang Y. Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. *PLoS Pathog*. 2018;14:e1007236. <https://doi.org/10.1371/journal.ppat.1007236>.
 30. Towler P, Staker B, Prasad SG, Menon S, Tang J, Parsons T, et al. ACE2 X-ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis. *J Biol Chem*. 2004;279:17996–8007. <https://doi.org/10.1074/jbc.M311191200>.
 31. Wan Y, Shang J, Graham R, Baric RS, Li F. Receptor recognition by the novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS coronavirus. *J Virol*. 2020;94:e00127–e220. <https://doi.org/10.1128/JVI.00127-20>.
 32. Wilder-Smith A, Freedman DO. Isolation, quarantine, social distancing, and community containment: pivotal role for old-style public health measures in the novel coronavirus (2019-nCoV) outbreak. *J Travel Med*. 2020;27:taaa20. <https://doi.org/10.1093/jtm/taaa020>.
 33. World Health Organization. Public health surveillance for COVID-19: interim guidance. *Who.Int*; 2021. <https://www.who.int/publications/i/item/who-2019-nCoV-surveillanceguidance-2020.8>.
 34. World Health Organization. Summary of probable SARS cases with onset of illness. *Who.int*; 2002. <https://www.who.int/csr/sars/country/en/-2005>.
 35. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, Graham BS, McLellan JS. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*. 2020;367:1260–3. <https://doi.org/10.1126/science.abb2507>.
 36. Zhang H, Wada J, Hida K, Tsuchiyama Y, Hiragushi K, Shikata K, Wang H, Lin S, Kanwar YS, Makino H. Collectrin, a collecting duct-specific transmembrane glycoprotein, is a novel homolog of ACE2 and is developmentally regulated in embryonic kidneys. *J Biol Chem*. 2001;276:17132–9. <https://doi.org/10.1074/jbc.M006723200>.
 37. Zhao Y, Zhao Z, Wang Y, Zhou Y, Ma Y, Zuo W. Single-cell RNA expression profiling of ACE2, the receptor of SARS-CoV-2. *Am J Respir Crit Care Med*. 2020;202:756–9. <https://doi.org/10.1164/rccm.202001-0179LE>.
 38. Zhou S, Liu R, Baroudy BM, Malcolm BA, Reyes GR. The effect of ribavirin and IMPDH inhibitors on hepatitis C virus subgenomic replicon RNA. *Virology*. 2003;310:333–42. [https://doi.org/10.1016/s0042-6822\(03\)00152-1](https://doi.org/10.1016/s0042-6822(03)00152-1).
 39. Zhou P, Yang XL, Wang XG, Hu B, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020;579:270–3. <https://doi.org/10.1038/s41586-020-2012-7>.
 40. Zhu N, Zhang D, Wang W, Li X, Yang B, et al. China novel coronavirus investigating and research team. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med*. 2020;382:727–33. <https://doi.org/10.1056/NEJMoa2001017>.
 41. Zisman LS, Keller RS, Weaver B, Lin Q, Speth R, Bristow MR, Canver CC. Increased angiotensin-(1–7)-forming activity in failing human heart ventricles: evidence for upregulation of the angiotensin-converting enzyme Homologue ACE2. *Circulation*. 2003;108:1707–12. <https://doi.org/10.1161/01.CIR.0000094734.67990.99>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.