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

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

Computational drug re-purposing targeting the spike glycoprotein of SARS-CoV-2 as an effective strategy to neutralize COVID-19

Himanshu G. Toor\textsuperscript{a,b}, Devjani I. Banerjee\textsuperscript{c,*}, Soumya Lipsa Rath\textsuperscript{d}, Siddhi A. Darji\textsuperscript{c}

\textsuperscript{a} P.G. Department of Genetics, Ashok and Rita Patel Institute of Integrated Study and Research in Biotechnology and Allied Sciences (ARIBAS), Charutar Vidyamand Mandd University, P.O. Box No. 61, New Vallabh Vidyanagar, Vithal Udyognagar, 388121, Anand, Gujarat, India
\textsuperscript{b} Sardar Patel University, Gujarat, India
\textsuperscript{c} Dr. Vikram Sarabhai Institute of Cell and Molecular Biology, Faculty of Science, The Maharaja Sayajirao University of Baroda, Pratapnagri, Vadodara, 390002, Gujarat, India
\textsuperscript{d} National Institute of Technology, Warangal, Telangana, 506004, India

ARTICLE INFO

Keywords:
Coronavirus
COVID-19
Drug repurposing
Molecular docking
SARS-CoV
SARS-CoV-2

ABSTRACT

COVID-19 has intensified into a global pandemic with over a million deaths worldwide. Experimental research analyses have been implemented and executed with the sole rationale to counteract SARS-CoV-2, which has initiated potent therapeutic strategy development in coherence with computational biology validation focusing on the characterized viral drug targets signified by proteomic and genomic data. Spike glycoprotein is one of such potential drug target that promotes viral attachment to the host cellular membrane by binding to its receptor ACE-2 via its Receptor-Binding Domain (RBD). Multiple Sequence alignment and relative phylogenetic analysis revealed significant sequential disparities of SARS-CoV-2 as compared to previously encountered SARS-CoV and MERS-CoV strains. We implemented a drug re-purposing approach wherein the inhibitory efficacy of a cluster of thirty known drug candidates comprising of antivirals, antibiotics and phytochemicals (selection contingent on their present developmental status in underway clinical trials) was elucidated by subjecting them to molecular docking analyses against the spike protein RBD model (developed using homology modelling and validated using SAVES server 5.0) and the composite trimeric structures of spike glycoprotein of SARS-CoV-2. Our results indicated that Camostat, Favipiravir, Tenofovir, Raltegravir and Stavudine showed significant interactions with spike RBD of SARS-CoV-2. Proficient bioavailability coupled with no predicted in silico toxicity rendered them as prospective alternatives for designing and development of novel combinatorial therapy formulations for improving existing treatment regimes to combat COVID-19.

1. Introduction

SARS-CoV-2 (Severe Acute Respiratory Syndrome Coronavirus-2), a member of the beta-coronavirus genus, which includes SARS-CoV (Severe Acute Respiratory Syndrome Coronavirus), MERS-CoV (Middle East Respiratory Syndrome Coronavirus) and other coronaviruses, is a novel discovered virus that causes COVID-19 (A Chronicle on the SARS Epidemic, 2003). The occurrence of SARS-CoV-2 was first epidemiologically linked in December 2019 to patients suffering from pneumonia of an unknown cause who were linked to a seafood market in Wuhan, Hubei province, China. On January 30, 2020, WHO (World Health Organisation) declared SARS-CoV-2 as a “Public Health Emergency of International Concern” and officially named it as Coronavirus Disease 2019(COVID-19/2019-nCoV) (Wu et al., 2020; Li et al., 2020). As of July 27, 2020, Global statistics reported by WHO regarding this intensifying pandemic indicate 16,114,449 confirmed COVID-19 cases, including 6,46,641 deaths collectively in the 216 affected countries (https://www.who.int/emergencies/diseases/novel-coronavirus-2019).

Pathological diagnoses of SARS-CoV-2 infections designate severe acute respiratory distress syndrome along with secondary ailments like sore throat, high fever, muscle weakness, encephalitis and diarrhoea in infected patients (Zhu et al., 2020). It primarily targets the lower respiratory system by invading the pulmonary epithelial cells, releasing the nucleocapsid component and arrogating the host’s cellular machinery for cytoplasmic replication (Shah et al., 2020).

\* Corresponding author.
E-mail addresses: himanshutoor@aribas.edu.in (H.G. Toor), devjani.chak-cmb@msubaroda.ac.in, devjani.chakraborty@yahoo.com (D.I. Banerjee), slrath@nitw.ac.in (S. Lipsa Rath), siddhiadarji@gmail.com (S.A. Darji).

https://doi.org/10.1016/j.ejphar.2020.173720

Received 28 July 2020; Received in revised form 30 October 2020; Accepted 30 October 2020
Available online 6 November 2020
0014-2999/© 2020 Elsevier B.V. All rights reserved.
Comprehensive structural genomics analyses exemplify that the SARS-CoV-2 genome codes for a variety of structural and non-structural proteins which constitute potential drug targets for scientific experimental investigations to impede the COVID-19 epidemic (Prajapati et al., 2020). The spike glycoprotein (S protein) is a prospective drug target as it initiates the first crucial step in SARS-CoV-2 attachment to the host cell. It is a homotrimeric protein (each chain containing 1273 amino acids) which recognizes and binds to its receptor Angiotensin Converting Enzyme-2 (ACE-2) with a greater affinity attributed to the mutations at L455, F486, Q493, S494, N501 and Y505 residues (Yan et al., 2020). Microbiological digital resources data instigate a comprehensive schematic of the spike glycoprotein (Figs. 1 and 2).

1.1. Alternative strategies targeting SARS-CoV-2 spike glycoprotein

Several potential therapeutic strategies are being explored to combat SARS-CoV-2 targeting the spike glycoprotein as depicted in Fig. 3 (Wu et al., 2020). Of these strategies, ‘Drugs Re-purposing’ has emerged as a strong pharmaceutical rationale due to its reduced perils of drug development costs and shortening of the time gap between identification of prospective drug candidates and designing of treatment regimes for patients. This is due to the availability of extensive data regarding its pharmacokinetics, pharmacodynamics, toxicity, clinical trial results, etc (Zhou et al., 2020; Cilibrerto, 2020).

Drug-spike protein molecular docking studies were executed to elucidate their inhibitory competence to counteract SARS-CoV-2 infections targeting spike glycoprotein receptor-binding domain. Computational pharmacokinetic analyses coupled with in silico toxicity prediction studies were used to evaluate the therapeutic proficiency of the selected drug candidates for futuristic COVID-19 treatment regimes.

2. Materials and methods

2.1. Workstation

The entire study was executed on a Dell Laptop having an i5-6200U Intel core (2.30 GHz) processor, 8 GB RAM, 1 TB hard disk and a 64-bit operating system. All softwares utilized for the present study were open-source tools and freely downloadable.

2.2. Sequence retrieval

The protein sequences, structures and other relevant data regarding the surface spike glycoprotein from SARS-CoV-2 and the whole genome sequence of SARS-CoV-2 Wuhan-Hu-1 isolate were downloaded from digital resources such as NCBI (National Centre for Biotechnology Information) (https://www.ncbi.nlm.nih.gov/), Universal Protein Knowledgebase (UniProtKB) and RCSB-PDB (Research Collaboratory for Structural Bioinformatics-Protein Data Bank) databases (Sayers et al., 2018; Apweiler et al., 2004; Berman et al., 2000). For comparative analysis, SARS-CoV (first identified in Foshan, Guangdong, China in 2002 from bats) data was also collected and analyzed (Chinese Law and Government, 2003; Li et al., 2005) (Table 1).

2.3. Crystallizability and protein- disorderness

The crystallization propensity of the SARS-CoV-2 spike glycoprotein sequence (UniProtKB ID: P0DDTC2) was predicted by the PPrCpred, CRYSTALP2 and XTalPred servers respectively (Mizianty and Kurgan, 2011; Kurgan et al., 2009; Slabinski et al., 2007). XTalPred also disclosed diverse physicochemical information such as Molecular Weight (MW), hypothetical isoelectric point (pi), amino-acid length, instability index and Grand Average of Hydropathicity (GRAVY) index, etc. The prediction of the conformational diversity in the ordered globular domains and intrinsically disordered regions of the spike glycoprotein sequence was derived using IUPred2A, MobiDB web tool, Meta disorder web server and PONDOR servers (Meszaros et al., 2018; Piovesan et al., 2018; Kozlowski and Bujnicki, 2012; Romero et al., 1997). The unstructured regions of proteins are stipulated to play a crucial role in diverse cellular activities.

2.4. Physicochemical characterisation and post-Translational modifications (PTMs)

The computation of the physicochemical properties and amino acid scale representation to determine the significant amino acid enrichment and/or depletion patterns were resolved by ProtParam and ProtScale tools of ExPasy (a SIB Bioinformatics resource portal) and Protein Calculator v 3.4 servers. (Gasteiger et al., 2005; Anthis and Clore, 2013).

SignalP 4.1 server was used to identify the secretory signal peptides and its cleavage sites (Nielson, 2017). N-Terminal acetylation sites were predicted by NetAcet 1.0 server (Kiemer et al., 2005). High specificity Coronavirus 3C-like protease cleavage sites were predicted by the NetCorona-1.0 server (Kiemer et al., 2004).

N-linked and O-linked glycosylation sites were predicted using the NetNGlyc-1.0 and NetOGlyc-4.0 servers respectively (Gupta and Brunak, 2002; Steenopt et al., 2013). NetPhos 3.1 server was used to predict the serine, threonine or tyrosine phosphorylation sites (Blom et al., 1999). DIANA 1.1 webserver was used to predict the disulfide bond connectivity (Ferre and Clore, 2005). The surface accessibility, secondary structure, disorder, and phi/psi dihedral angles of amino acids were predicted by NetSurfP-2.0 server (Klausen et al., 2019). DeepLoc 1.0 server was used to predict the sub-cellular localization of the spike glycoprotein by deep neural network algorithms (Armenteros et al., 2017).

2.5. Secondary structure prediction, tertiary structure prediction, refinement and model evaluation

PSIPRED 4.0 and JPred v.4 servers were utilized to elucidate the secondary structure characteristics of the spike glycoprotein (McGuffin et al., 2000; Drozdetskiy et al., 2015). The UniProtKB spike glycoprotein sequence (P0DDTC2) was submitted in FASTA format to Modeller v9.24 which generated its 3D structure employing a restraint-based approach

### Table 1

| Sd. No. | Digital resource | ID | Description |
|---------|------------------|----|-------------|
| 1       | UniProtKB        | P0DDTC2 | SARS-CoV-2’ spike glycoprotein |
| 2       | NCBI             | YP_009724390.1 | SARS-CoV-2 surface glycoprotein |
| 3       | NCBI             | MN908947.3 | Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome |
| 4       | RCSB-PDB         | 6YYB | SARS-CoV-2 spike ectodomain structure (open state) |
| 5       | RCSB-PDB         | 6VSB | Prefusion 2019 nCoV spike glycoprotein with a single receptor-binding domain up |
| 6       | RCSB-PDB         | 6VWI | Structure of SARS-CoV-2 chimeric receptor-binding domain complexed with its receptor human ACE2 |
| 7       | RCSB-PDB         | 6LZG | Structure of novel coronavirus spike receptor-binding domain complexed with its receptor ACE2 |
| 8       | RCSB-PDB         | 6LYA | Crystal structure of the SARS-CoV-2 receptor binding domain in complex with CR3022 Fab |
| 9       | RCSB-PDB         | 2AF  | Structure of SARS coronavirus spike receptor-binding domain complexed with its receptor |

| a | Identification. |
| b | Severe Acute Respiratory Syndrome – Coronavirus-2. |
| c | Angiotensin Converting Enzyme-2. |
| d | antigen-binding Fragment. |
using a model-single module (Webb and Sali, 2017; Sali and Blundell, 1993). The best model was selected based on low discrete optimized protein energy (DOPE) score. It was then submitted to RAMPAGE server to determine the geometrical structural validity (Lovell et al., 2003). Loop refinement of the outlier residues was carried out by ModLoop server (Fiser et al., 2000).

The above steps were repeated until the outlier residues were within the designated values. The final model geometry was subjected to energy minimization by Chiron server wherein the atomic steric clashes between the side-chain and backbone were resolved (Ramachandran et al., 2011). The overall quality assessment of the final model in terms of packing density (total void volume), unsatisfied hydrogen bonds, steric clashes and the accessible surface area scaling was exemplified by Gaia server (Kota et al., 2011). The stereochemical quality evaluation of the pre- and post-refinement 3D models of spike glycoprotein was done via PROCHECK, ERRAT (v 2.0) and Verify3D environment profile analysis methods by submitting to SAVES v5.0 (Structural Analysis and Verification Server) (Pradeepkiran et al., 2015).

The Root Mean Square Deviation (RMSD) statistics between the target model and template were calculated by structural superimposition in SuperPose v1.0 program which utilizes a modified quaternion eigenvalue approach under the default parameters (Mafti et al., 2004). This final model was selected as the receptor for further studies. Composition profiler was utilized to investigate the amino acid composition bias in the modeled spike glycoprotein and the template protein (Vacic et al., 2007).

2.6. Multiple Sequence Alignment (MSA) and comparative phylogeny

MSA of the sequence of the generated 3-D model with the existing PDB structures was carried out using Jalview v2.11.1.0 (Waterhouse et al., 2009). Evolutionary analyses of the selected sequences was carried out by phylogenetic tree construction using Mega X v10.1.8 maintaining the default values of the selected parameters (Kumar et al., 2018).

2.7. Binding sites prediction

Binding site prediction to determine the active site residues of the 3-D spike glycoprotein model was carried out by COACH server (Yang et al., 2013; Yang et al., 2013). These were also determined by the receptor cavity method using Biovia Discovery studio v17.2.0.16349 (Dassault Systemes, 2017). The functional domain prediction was carried out by MOTIF search, a subserver of GenomeNet, using the Pfam and PROSITE data (Al-Khayyat and Al-Dabbagh, 2016). The cut-off score (E-value) was set at 1.0.

2.8. Ligand selection and molecular docking

The computational molecular docking of ACE-2 (PDB ID: 1R42) (receptor) with the spike glycoprotein model (ligand) was performed using the HawkDock server. It is an integrated web server that utilizes ATTRACT for protein-protein docking, HawkRank scoring function for re-ranking docked complexes based on desolvation potentials and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) free energy decomposition methodology for the identification of key residues in Protein-Protein Interactions (PPIs) (Weng et al., 2019; Sarkar et al., 2020).

The drug candidates for virtual screening against the spike glycoprotein model were selected from an extensive literature survey of potential therapeutic agents undergoing trials against SARS-CoV-2 and/or those which were previously developed as antiviral/anti-malarial therapeutic inhibitors as shown in Table 1. The sdf files of all molecules were retrieved from the PubChem database and were converted to pdb files and saved for further docking protocols (Kim et al., 2015).

Molecular docking of the selected ligands with the spike glycoprotein 3-D model (receptor) was carried out by iGEMDOCK v2.1 (Hsu et al., 2011; Kashyap, 2019). It is an integrated virtual screening graphical automatic drug design system dealing with preparations through post-screening analysis depicting crucial pharmacological interactions based on protein-compound interaction profiles. It employs a generic evolutionary method algorithm. The docking accuracy settings (GA parameters) for all docking operations were set as follows: Population size: 300; Generations: 80; Number of solutions: 10. The mode of docking was stable (slow docking).

The scoring function was set to GEMDOCK and the hydrophobic and electrostatic preferences for the ligands were set to 1.0. The molecular interactions of the ligand-receptor docked complexes were analyzed using Discovery studio v20.1.0.19295, PyMol v1.7.4.5 and UCSF Chimera v11.1.2 respectively (Seelig and de Groot, 2010; Sadat et al., 2019; Peetersen et al., 2004). The best docked complex of each ligand-receptor combination was selected based on the least binding energy values for further analysis.

2.9. Galaxy refine complex

Refinement of protein-protein complex structure was done by GalaxyWEB. This tool allows us to understand the flexibility, at the protein interface (modeled RBD with RBD-ligand complex) and gives us an idea regarding overall conformational changes that occurs upon binding (Ko et al., 2012). GalaxyRefine refines the docked structure by a hybrid technique that involves the triaxial loop closure (TLC) algorithm during global optimization by conformational space annealing (CSA) technique (Park and Seok, 2012). The molecular interactions of the ligand-receptor docked complexes were analyzed using PyMol v1.7.4 (Seelig and de Groot, 2010). Further, Favipiravir, Hydroxychloroquine and Nafamostat were taken to study the possible structural changes occurring to spike protein when drug binds.

2.10. Argus Lab analysis

Molecular docking interactions of the selected ligands with the SARS-CoV-2 whole trimeric spike glycoprotein (PDB ID’s 6VSB and 6VYB respectively) were analyzed and visualized using Argus Lab 4.0.1 (Thompson, 2004). Argus Lab platform has distinct advantages over other softwares because it performs meticulous docking analyses of huge peptides/proteins with a precise chain or domain selection within minutes. It follows the Lamarckian genetic algorithm along with precise grid resolution calculation in coherence with the ligand binding sites.

It was performed under the following parametric considerations: scoring function selected as ‘AScore’, docking algorithm as Argus dock (exhaustive search), a grid resolution of 0.4 A³ and the binding site box dimensions measuring 60 A³ x 60 A³ x 60 A³ respectively. The docking precision was set to “Regular precision” and “Flexible” ligand docking mode was utilized for all docking runs. The individual docked pose stability was evaluated by ArgusLab energy calculations and the number of hydrogen bonds formed (Hafeez et al., 2013).

The best docking model was selected according to the lowest AScore computed and the most pertinent binding conformation was selected on the basis of ligand-receptor hydrogen bond interactions in the vicinity of the substrate binding site. The lowest energy poses implied the highest binding affinity because high energy values represent unstable conformations.

2.11. In silico toxicity prediction

The analysis of the pharmacokinetic properties of the selected ligands such as absorption, distribution metabolism and excretion (ADME), drug-likeness and medicinal chemistry features was carried out by SWISS ADME (Daina et al., 2014, 2017; Daina and Zoete, 2016). OSIRIS Property Explorer was used to evaluate the risks of side effects, such as mutagenicity, tumorigenicity, irritant and reproductive
consequences, as well as drug-relevant properties including cLogP, LogS (solubility), molecular weight, Topological Polar Surface Area (TPSA), drug-likeness and overall drug-score (Ayati et al., 2012).

3. Results

3.1. Crystallization propensity and protein- disorderness

PPCPred and CRYSTALP2 servers both predicted the sequence to be non-crystallizable. PPCPred predicted the crystallization propensity to be 0.084 which was significantly lower than the probability threshold value of 0.43. The individual step values of the crystallization process were as follows: probability that production of protein material fails: 0.831, probability that purification fails: 0.806, probability that crystallization fails: 0.01 and probability that target will yield diffraction-quality crystals: 0.355. CRYSTALP2 predicted the sequence as non-crystallizable with a confidence of 0.28. Xtalpred results depicted that the protein was non-crystallizable (EP crystallization class 5: very difficult; RF protein crystallization class 4). It also enumerated an assortment of molecular characterization parameters (Table 2). Overall, the spike glycoprotein RBD model was deemed non-crystallizable in its native form.

The MobiDB web tool indicated 0% disorderness in the spike glycoprotein sequence (UniProtKB: P0DTC2) while IUPred2A and Meta-disorder web server (global disorder tendency = 0.3474) indicated disordered residues. PONDR server depicted 11 disordered regions, comprising a total of 98 disordered amino acid residues (7.70%) with the longest disorder region comprising of 38 amino residues (Average prediction score = 0.1374). DEPP (Disorder Enhanced Phosphorylation Predictor) in the PONDR server predicted the phosphorylation potentiality of serine (6 out of 99; 6.06%), threonine (2 out of 97; 2.06%) and tyrosine (3 out of 54; 5.56%) residues of the spike glycoprotein. (Table S1; Fig. S1 in supplementary data).

3.2. Physicochemical characterization and post-translational modifications (PTMs)

SARS-CoV-2 encompasses a variety of structural proteins such as the spike glycoprotein, small envelope protein (E), Matrix protein (M) and the nucleocapsid protein (N) while the ORF1a/b region (initial ~20 kb) is translated into two polyproteins (PP1A and PP1B) which encode the majority of the Non-Structural Proteins (NSPs) (Khodadadi et al., 2020, Wu et al., 2020). ProtParam and ProtScale servers explicated several physicochemical parameters of SARS-CoV-2 spike glycoprotein viz. Molecular weight: 141.178Kda; theoretical pl: 6.24; Instability index: 33.01 (stable); aliphatic index: 84.67; GRAVY: −0.079. The estimated half-life (t1/2) was predicted to be 30 h in mammalian reticulocytes (in vitro); >20 h in yeast (in vivo) and >10 h in E. coli (in vivo) (Kar et al., 2020). ProtScale also depicted an anthology of hydrophobicity plots of the spike glycoprotein. Protein calculator v3.4 further corroborated the above results. SignalP 4.1 server predicted residues 1–13 as the signal peptide (cleavage site between the 13th and 14th residues) with a score (D value) of 0.837 which was higher than the threshold cut-off value (D = 0.450).

NetAcet 1.0 server predicted no N-terminal acetylation as alanine, glycine, serine or threonine residues were not present at positions 1–3. One specific coronavirus 3C-like protease cleavage site was found by NetCorona 1.0 server at position 1002 with a score of 0.685 (TGRLG’SLQTY). N-linked glycosylation sites predicted by NetNGlyc 1.0 server were in accordance with the data retrieved from UniProtKB for the spike glycoprotein while O-linked glycosylation sites were predicted at positions 673, 678 and 686 by NetOglyc server 4.0 which was in accordance with the previous research analysis (Andersen et al., 2020).

NetPhos 3.1 server predicted phosphorylation of serine, tyrosine and threonine residues at 136 positions in the spike glycoprotein. DiANNA 1.1 web server revealed 20 disulfide bond connectivity predictions. NetSurfP 2.0 results validated the disorder predictions and also depicted surface accessibility and phi/psi angles for each amino acid. DeepLoc 1.0 server predicted that the spike glycoprotein was a soluble (score = 0.0577), membrane type protein (score = 0.9423).

3.3. Secondary structure prediction, tertiary structure prediction, refinement and model evaluation

PSIPRED 4.0 and Jpred v.4 servers deduced the secondary structure of the spike glycoprotein which comprised of α-helices, β-sheets and coils. It also predicted a putative domain boundary at position 164 (Asparagine) (Fig. S2 in supplementary data).

The spike glycoprotein sequence from UniProtKB was submitted to MODELLER v9.24. It generated two models whose characteristics are listed in Table 3. GA341 (range: 0–1, value near 1 being preferred), z-

| Table 2 | Molecular characterization of P0DTC2 sequence predicted by XtalPred server. |
|---------|--------------------------------------------------------------------------------|
| Sr. No. | Server | Molecular feature (s) | Prediction |
| 1       | XtalPred | Molecular weight | 141.196 KDa |
|         |         | Amino acid length | 1273 |
|         |         | Theoretical pl | 6.24 |
|         |         | Instability index | 33.02 |
|         |         | GRAVY | −0.08 |
|         |         | Molecular surface feature (EP and RF crystallization): |  |
|         |         | a) Surface entropy | −1.06 |
|         |         | b) Surface hydrophobicity | −0.99 |
|         |         | c) Surface ruggedness | 4.04 |
| 2       | COILS | Coiled-coils regions | 28 |
| 3       | DISOPRED | Longest disorder region | 16 (1%) |
| 4       | TMHMM | Transmembrane helices | 1 |
| 5       | SignalP | Signal peptides | No |
| 6       | SEG | Longest-low complexity region | 23 |
| 7       | PSIPRED | Secondary structure prediction: |  |
|         |         | Coil residues % | 38 |
|         |         | Helix residues % | 22 |
|         |         | Strand residues % | 40 |
| 8       | – | Residue content |  |
|         |         | Cysteine residues | 40 |
|         |         | Methionine residues | 14 |
|         |         | Tyroptphan residues | 12 |
|         |         | Tyrosine residues | 54 |
|         |         | Phenylalanine residues | 77 |

pl: Isoelectric point; GRAVY: Grand Average of Hydropathicity; EP: Expert Pool; RF: Random Forest.

| Table 3 | Model characteristics generated by MODELLER v.9.24. |
|---------|--------------------------------------------------------------------------------|
| Sr.No.  | Features | Model 1 | Model 2 |
| 1       | Target region | 14-302 | 330-529 |
| 2       | Protein length | 1273 | 1273 |
| 3       | PDB template code | 5 × 4S A | 6LYA E |
| 4       | Template region | 18-289 | 330-529 |
| 5       | Sequence identity | 54% | 100% |
| 6       | E-value | 0 | 0 |
| 7       | GA341 | 1 | 1 |
| 8       | MPQS | 0.786723 | 1.37941 |
| 9       | z-DOPC | −0.3 | −1.32 |
| 10      | TSVMod method | MSALL | MSALL |
| 11      | TSVMod RMSD | 3.711 | 0.464 |
| 12      | TSVMod NO35 | 0.923 | 0.972 |

E-value: no: of expected hits of similar quality; MPQS: ModPipe Quality Score; z-DOPC: Discrete Optimized Protein Energy Score of Salllab Model evaluation server; TSVMod RMSD: Root Mean Square Deviations; TSVMod NO35: Native Overlap prediction of fraction of Cz atoms within 3.5 Å of the native structure.
DOPE (lowest score preferred), MPQS (ModPipe Quality Score; highest value preferred) and sequence identity (highest value preferred) scores were employed for model quality assessment.

The results indicated that model 2 was the most reliable one as compared to model 1. The MPQS score of model 1 was also predicted to be unreliable as compared to that of model 2. Model 2 represented a segment of the S1 subunit of the SARS-CoV-2 spike glycoprotein as deduced from the UniProtKB data while model 1 corresponded to the N-Terminal Domain (NTD) region. It depicted 2 N-Acetyl glycosylation sites at asparagine residues (positions 331 and 343 respectively). The model (residues 1–200) formed the majority portion of the Receptor Binding Domain (RBD: residues 319–541) containing the Receptor Binding Motif (RBM; residues 437–508).

The selected model (Fig. 4a) was analyzed by Ramachandran plot generation by submission to the RAMPAGE server which elucidated 186 favored residues (93.9%), 10 allowed residues (5.1%) and 2 outlier residues (TRP353 and PHE497; 1%) respectively. The model was subjected to further structural refinement by ModLoop server which then depicted 186 favored residues (93.9%), 12 allowed residues (6.1%) and zero outlier residues respectively (Fig. 4b). The refined model was then subjected to energy minimization by Chiron server with side chain constraints.

After energy minimization, the model was subjected to structural validation by Gaia server (with side chain constraints) which elucidated the following data: Clash score = 0.0381; 69 unsatisfied HBS (Hydrogen Bonds in Shell); 4 HBC (Hydrogen Bonds in Core); solvent accessible (10225.4 Å²) and solvent excluded (8623.33 Å²) surface areas and a void volume score of 0.53. All bond lengths were within the acceptable range. The side-chain integrity discrepancies were refined and the model was then submitted to the SAVES V5.0 server for structure validation.

SAVES server validated the refined model yielding an overall quality factor of 83.333. VERIFY3D substantiated that about 80% amino acids scored >0.2 in the 3D/1D profile. PROVE results depicted about 25 buried outlier protein atoms (3.6%). Overall results validated the refined model (spike RBD model). Discovery studio construed characteristics of the spike RBD model viz. 1947 atoms; Molecular formula: C₁₀₁₆H₁₅₂₂N₃₅₀O₂₉₆S₉; molecular weight: 22.411Kda and a net formal charge of +5.

The spike RBD model and the template (6YLA chain E) were superimposed via Superpose 1.0 server which depicted a resulting local and global RMSD value of 0.37 over 200 Ca atoms and a value of 0.43 over 800 atoms in the protein backbone (Fig. 5a).

Composition profiler revealed that no significant enrichment or depletion patterns in the amino acid composition had occurred. The relative entropy had a value of 0.00092 with a p-value of 1. The spike model was also superposed with the SARS-CoV-2 spike open-state ectodomain structure (PDB ID: 6VYB) and the Prefusion 2019-nCoV spike glycoprotein with a single receptor-binding domain up structure (PDB ID: 6VS6). The spike RBD model-6VYB superimposition illustrated a local and global RMSD value of 1.183 over the Ca atoms and atoms in the protein backbone whereas the spike RBD model-6VS6 superimposition elucidated a local and global RMSD values of 1.655 and 1.648 over the Ca atoms and atoms in the protein backbone respectively (Fig. 5b and c). The MOTIF server detected diverse proteomic segments in the whole spike glycoprotein and the spike RBD model (Table 4).

Sequential analysis of the spike RBD model and the template (6YLA chain E) revealed that our RBD model lacked the first three residues (GLU327, THR328 and GLY329) incorporated by 6YLA chain E. In addition, the residues 1, 2, 3, 4 ... in the spike RBD model corresponded to residues 330, 331, 332, 333 ... and so on. Hence, the numbering of residues in the spike RBD model could be correlated to the residue position in 6YLA chain E as follows:

\[ \text{R.N in 6YLA chain E = R.N in spike RBD model + 329} \]

Where R.N = Residue Number.

3.4. Multiple Sequence Alignment (MSA), comparative phylogeny and binding sites prediction

SARS-CoV-2 is a more potent virus as it depicts important genomic dissimilarities when compared to the SARS-CoV (79%) and MERS-CoV (50%) genomes (Walls et al., 2020). Previous studies elucidated that amino acid mutations in the Spike glycoprotein influenced its credited ACE-2 (receptor) binding capability in humans (Walls et al., 2020; Coutard et al., 2020).

MSA of the selected sequences listed in Table 1 including the spike RBD model was carried out using Jalview v2.11.1.0 (Fig. 6). It revealed that large number of significant amino acid substitutions/mutations had occurred in the SARS-CoV-2 as compared to the SARS-CoV spike protein sequences reported earlier viz. R403K, E406D, K417V, L441I, S443A, and so on. Hence, the numbering of residues in the spike RBD model could be correlated to the residue position in 6YLA chain E as follows:

\[ \text{R.N in 6YLA chain E = R.N in spike RBD model + 329} \]

Where R.N = Residue Number.

3.5. ACE-2-spike RBD model molecular docking and analysis

The spike RBD model (ligand) and ACE-2 (receptor) were docked using the HawkDock server which generated 10 ligand-receptor docked complexes. The best one was selected based on the MM/GBSA score.
(binding free energy in Kcal/mol) having a value of $-52.33$ (Fig. 8). A total of 20 molecular interactions were observed upon analysis using the aforementioned molecular analysis software, of which 15 interactions were considered significant (bond-length of $\leq 4.0$ Å) (Table 5).

Earlier studies expounded that the spike RBD comprised of significant loops (a-f) which were then re-categorized into loop 1, 2 and 3 sub-segments. It was observed that the spike protein interactions with ACE-2 receptor comprised of loops 3a, 3b and 3f respectively (Robson, 2020). It also contained nine cysteine residues of which eight are involved in disulfide bond formation viz. CYS336-CYS361, CYS379-CYS432, CYS391-CYS525 and CYS480-CYS488 respectively. Of these, the first three disulphide bonds facilitate structural stabilization while the last one acts as a loop connector in the distal RBM region (Lan et al., 2020). From the results obtained, it was observed that the loop regions of the
spike RBD model (2 g, 3 b and 3f respectively) displayed maximum significant interactions.

A total of thirty ligands were selected from an extensive literature survey, WHO (www.who.int/covid-19/information) and FDA (www.fda.gov/drugs/emergency-preparedness-drugs/) organizations’ list of approved drugs designated for the treatment of SARS-CoV-2 to evaluate their therapeutic inhibitory potential against the spike Glycoprotein. These ligands were docked against Spike RBD model (receptor) utilizing iGEMDOCK v2.1 and their molecular interactions were analyzed (Table 6).

The criteria for selection of the potential drug candidates as potential inhibitors of SARS-CoV-2 spike glycoprotein RBD model hinged on various factors viz. number and strength of the hydrogen bonds, number and strength of other interactions (electrostatic/hydrophobic/van der Waals), number of loop interactions (2 g, 3 b and 3f respectively), number of mutated residues involved, and number of non-significant interactions involving loop residues, etc. Based on the above factors, we selected ten drugs out of the thirty studied for further analysis whose detailed significant molecular interactions are entailed (Table 7). All selected drug candidates are illustrated in Fig. 9.

Most significant interactions were observed in case of Stavudine, Doxycycline and Favipiravir (Fig. 10a, b and 10c). Our results also indicated that Tenofovir, Eugenol, Allicin, Raltegravir, Zalcitabine, Camostat and Ivermectin depicted significant molecular interactions with the ACE-2 residues involved in spike glycoprotein recognition elucidating their inhibitory potentiality. Moreover, Raltegravir also interacted with loop 1e (LYS441) but via a halogen (fluorine) interaction. (Fig. S3a/b - S9a/b in supplementary data).

Remdesivir, Liquiritin, Nafamostat, Oseltamivir, Zanamivir, Lopinavir, EmblicaninA and Ivermectin depicted strong interactions with CYS336 and CYS379 residues of the Spike RBD model, thus hypothesizing their role in structural destabilization of the Spike protein (Table S2; Fig. 6 supplementary data).

In Fig. 11 we can see the residues at the interface of RBD of the spike protein and the ACE-2 receptor as listed earlier (Table 5). To understand whether the binding of inhibitor causes any structural changes in the RBD-ACE2 complex, we superimposed the structure of the RBD with and without the inhibitors. Favipiravir was found to interact directly with residues Y505, R403 and L441 of the RBD, where R403 is considered to be a possible mutation (Fig. 12a, Fig. 6, Table 7). Surprisingly, upon superimposition of RBD and the RBD-Favipiravir complex we found deviation in side chain conformation of R403 (Fig. 12a). Apart from that...
few changes were seen in V401 and R509. This may envisage that Favipiravir alters the spike binding efficiency with ACE-2 receptor. Further, we also did a similar study with Hydroxychloroquine and Nafamostat, both of which did not directly bind near the interface region of the Spike and ACE2 proteins. As it can be seen from Fig. 1b, amino acids L517, T430, F429, F464, E516 and D428 were directly interacting with the hydroxychloroquine. The large number of non-polar residues promotes primarily hydrophobic interactions with the ligand. After superimposition we could observe the switch in the polar side chain of D428 with the inhibitor, similar to favipiravir. Nafamostat showed interactions with residues and not much difference was seen in the absence of the ligand.

Validation of the drug-Spike molecular interactions were achieved by docking studies of the selected drugs with the Spike glycoprotein trimeric structures available (PDB ID’s 6VSB and 6VYB respectively) (Tables S4–S5 in supplementary data). We observed that Eugenol, Allicin, Doxycycline and Ivermectin did not depict any significant interactions with 6VSB. Furthermore, Tenofovir, Zalcitabine and Favipiravir depicted four significant interactions each with the Extra-cellular Topological Domain (ETD) of the S1 subunit while Camostat exhibited five significant interactions with the S2 subunit ETD. Likewise, Allicin, Stavudine, Raltegravir and Ivermexitin elucidated no interactions with 6VYB. However, Camostat and Zalcitabine exhibited 1 and 4 interactions respectively with the Coiled-Coil Region 1 (CCR-1) of the S2 subunit. Eugenol and Doxycycline exhibited 1 interaction each with the ETD of S2’ domain and S1 subunit respectively (Figs. S10–S11 in supplementary data).

### 3.6. In silico toxicity prediction

SWISS-ADME determined the in silico physico-chemistry and pharmacokinetic analysis of the selected drugs. It was perceived that Ivermectin had the highest $P_{o/w}$ coefficient ($iLoG P_{o/w} = 5.86$) while Favipiravir had the lowest value ($iLoG P_{o/w} = 0.39$) signifying its higher solubility. Furthermore, Allicin exhibited the lowest skin permeation rate ($k_p = -9.89 \text{ cm/s}$) whereas Eugenol depicted the highest ($k_p = -5.69 \text{ cm/s}$). All drugs, except Ivermectin, complied with the Lipinski rules accompanied by minor violations. The GI absorption rate was high for all drugs except Tenofovir, Raltegravir, Doxycycline and Ivermectin. Out of all drugs, Eugenol and Allicin were predicted as Blood-Brain Barrier (BBB) permeants. Ultimately, all drugs had an acceptable bioavailability score, except for Doxycycline and Ivermectin (Tables S6–S7 in supplementary data).

OSIRIS Property Explorer was utilized to compute the systemic toxicity of the selected drugs. Our analysis showed that Raltegravir displayed the highest drug-likeness capacity (score = 0.999) while Tenofovir had the least possible value (0.0). Moreover, Eugenol and exemplified a high risk of mutagenicity, tumorigenicity and irritant capacity (scores of 0.6 in each) while Zalcitabine and Doxycycline exhibited a high risk of reproductive toxicity (score of 0.6). Favipiravir had the best drug score (0.933) amongst all potential drug candidates (Tables S8–S9 in supplementary data).

### 4. Discussion

Coronaviruses (CoV) are positive single-stranded RNA viruses classified as pathogenic agents reported since the 1960’s. These viruses are responsible for causing acute and chronic respiratory diseases as well as enteric, hepatic and neurologic infections as it has a broad host range (avian, murine, porcine, bovine and other domestic mammalian species including humans) (Weiss and Martin, 2005; Di Gennaro et al., 2020).

SARS-CoV-2 has materialized as a potent etiological agent with a current infection rate of ~20,000 individuals/day. Till date, no confirmed drug/vaccine candidates have been reported to counteract COVID-19 epidemic rendering healthcare facilities to face comprehensive obstacles to restore the healthiness of COVID-19 affected individuals. This is predominantly accredited to the genomic discrepancy of SARS-CoV-2 as compared to previously established SARS-CoV and MERS-CoV strains, making SARS-CoV-2 more virulent than its predecessors. These anomalies have commanded stringent amendments in conventional treatment regimes in terms of efficacy for eradication of SARS-CoV-2 (Fig. 6).

Diverse experimental research analyses have yielded abundant information regarding SARS-CoV-2 life cycle, epidemiology, etc (Wu et al., 2020; Li et al., 2020). Molecular experimentation in coherence with computational biology has resulted in the fabrication of credible therapeutic strategies for combating COVID-19 viz. identification of proteomic and genomic drug targets, traditional epidemiological drug re-purposing, peptidomimetics studies for antibody/vaccine development against structural and non-structural proteins of SARS-CoV-2 targeting host attachment and viral replication processes, etc. (Alanagreh et al., 2020) (Figs. 1 and 2).

The present study focused on the drug re-purposing based inhibition of SARS-CoV-2 attachment wherein the spike glycoprotein of SARS-CoV-2 undergoes conformational reorganization to distinguish its receptor ACE-2 and initiates a cascade of molecular progressions resulting in the integration with the cell membrane and liberation of its genome inside the host cell for replication.

We developed our validated model of the spike glycoprotein RBD based on the UniProtKB sequence P0DTC2 utilizing 6YLA chain E as the template (Fig. 5a). Multiple sequence alignment studies of the constructed spike RBD model with previously constructed 3-D structures demonstrated that SARS-CoV-2 exhibited significant mutations as compared to SARS-CoV and MERS-CoV strains resulting in an amplified
affinity for ACE-2 (Fig. 6). The protein was predicted to be non-crystallizable and depicted minute disorderness. The spike glycoprotein RBD harboured many such mutations and exhibited additional sequential diversity which designates it as a latent therapeutic drug target (Tables 2–4; Fig. 10 a/10b/10c).

We implemented a computational biology–based approach for drug re-purposing to screen drug candidates as potential therapeutic inhibitors via molecular docking analyses impacting the spike glycoprotein mechanism of action (Tables 6–7). Molecular docking analysis involving the spike glycoprotein RBD model (ligand) with ACE-2 (receptor) depicted twenty interactions of which fifteen were deemed significant and exhibited additional sequential diversity which designates it as a latent therapeutic drug target (Tables 2–4; Fig. 10 a/10b/10c).

We implemented a computational biology–based approach for drug re-purposing to screen drug candidates as potential therapeutic inhibitors via molecular docking analyses impacting the spike glycoprotein mechanism of action (Tables 6–7). Molecular docking analysis involving the spike glycoprotein RBD model (ligand) with ACE-2 (receptor) depicted twenty interactions of which fifteen were deemed significant and exhibited additional sequential diversity which designates it as a latent therapeutic drug target (Tables 2–4; Fig. 10 a/10b/10c).

We implemented a computational biology–based approach for drug re-purposing to screen drug candidates as potential therapeutic inhibitors via molecular docking analyses impacting the spike glycoprotein mechanism of action (Tables 6–7). Molecular docking analysis involving the spike glycoprotein RBD model (ligand) with ACE-2 (receptor) depicted twenty interactions of which fifteen were deemed significant and exhibited additional sequential diversity which designates it as a latent therapeutic drug target (Tables 2–4; Fig. 10 a/10b/10c).

We implemented a computational biology–based approach for drug re-purposing to screen drug candidates as potential therapeutic inhibitors via molecular docking analyses impacting the spike glycoprotein mechanism of action (Tables 6–7). Molecular docking analysis involving the spike glycoprotein RBD model (ligand) with ACE-2 (receptor) depicted twenty interactions of which fifteen were deemed significant and exhibited additional sequential diversity which designates it as a latent therapeutic drug target (Tables 2–4; Fig. 10 a/10b/10c).

We implemented a computational biology–based approach for drug re-purposing to screen drug candidates as potential therapeutic inhibitors via molecular docking analyses impacting the spike glycoprotein mechanism of action (Tables 6–7). Molecular docking analysis involving the spike glycoprotein RBD model (ligand) with ACE-2 (receptor) depicted twenty interactions of which fifteen were deemed significant and exhibited additional sequential diversity which designates it as a latent therapeutic drug target (Tables 2–4; Fig. 10 a/10b/10c).

We implemented a computational biology–based approach for drug re-purposing to screen drug candidates as potential therapeutic inhibitors via molecular docking analyses impacting the spike glycoprotein mechanism of action (Tables 6–7). Molecular docking analysis involving the spike glycoprotein RBD model (ligand) with ACE-2 (receptor) depicted twenty interactions of which fifteen were deemed significant and exhibited additional sequential diversity which designates it as a latent therapeutic drug target (Tables 2–4; Fig. 10 a/10b/10c).

We implemented a computational biology–based approach for drug re-purposing to screen drug candidates as potential therapeutic inhibitors via molecular docking analyses impacting the spike glycoprotein mechanism of action (Tables 6–7). Molecular docking analysis involving the spike glycoprotein RBD model (ligand) with ACE-2 (receptor) depicted twenty interactions of which fifteen were deemed significant and exhibited additional sequential diversity which designates it as a latent therapeutic drug target (Tables 2–4; Fig. 10 a/10b/10c).

We implemented a computational biology–based approach for drug re-purposing to screen drug candidates as potential therapeutic inhibitors via molecular docking analyses impacting the spike glycoprotein mechanism of action (Tables 6–7). Molecular docking analysis involving the spike glycoprotein RBD model (ligand) with ACE-2 (receptor) depicted twenty interactions of which fifteen were deemed significant and exhibited additional sequential diversity which designates it as a latent therapeutic drug target (Tables 2–4; Fig. 10 a/10b/10c).

We implemented a computational biology–based approach for drug re-purposing to screen drug candidates as potential therapeutic inhibitors via molecular docking analyses impacting the spike glycoprotein mechanism of action (Tables 6–7). Molecular docking analysis involving the spike glycoprotein RBD model (ligand) with ACE-2 (receptor) depicted twenty interactions of which fifteen were deemed significant and exhibited additional sequential diversity which designates it as a latent therapeutic drug target (Tables 2–4; Fig. 10 a/10b/10c).

We implemented a computational biology–based approach for drug re-purposing to screen drug candidates as potential therapeutic inhibitors via molecular docking analyses impacting the spike glycoprotein mechanism of action (Tables 6–7). Molecular docking analysis involving the spike glycoprotein RBD model (ligand) with ACE-2 (receptor) depicted twenty interactions of which fifteen were deemed significant and exhibited additional sequential diversity which designates it as a latent therapeutic drug target (Tables 2–4; Fig. 10 a/10b/10c).
Zalcitabine, another thymidine analogue, acts in the same way as Stavudine and is more potent (Leandro et al., 2013). It depicted two interactions with ARG403 (loop 1e) of the spike protein with a high risk of hazardous reproductive consequences. Tenofovir, an acyclic adenosine nucleotide analogue, is used in combinatorial therapy with other anti-retroviral drugs to treat HIV and singularly to treat Hepatitis B infections, acts as a reverse transcriptase inhibitor and depicted no predicted systemic toxicity upon administration (Delaney et al., 2006) (Fig. S7; Table S7).

Raltegravir is an anti-retroviral drug belonging to the class of integrase inhibitors, ultimately affecting viral DNA insertion and its subsequent integration into the host cell’s DNA. It is primarily used to treat HIV-1 infections (Boesecke and Gelgor, 2009). It exhibited seven significant interactions of which only one targeted the loop residue of spike protein (LEU441) (Fig. S7). Similarly, Camostat (a carbonyl compound) is a synthetic serine protease inhibitor targeting the host cell TMRPS2 (viral entry mediator) thereby inhibiting viral infection and replication (Uno, 2020). It also exhibits anti-inflammatory and anti-fibriotic capacity. Like Raltegravir, it interacted with LEU441 but overall had only two significant bonds. Both drugs lacked predicted toxicity upon administration (Fig. S8).

Doxycycline (a semi-synthetic derivative of oxytetracycline) is a protein synthesis inhibitor and exhibits broad spectrum anti-bacterial potency (Holmes and Charles, 2009). It also exhibits antiviral competence by inhibition of matrix metalloproteases in case of Dengue and other retroviruses or transcriptional up-regulation of intracellular zinc finger antiviral protein (ZAP) (Malek et al., 2020). Although it illustrated three significant spike RBD loop interactions, it was predicted to have a low bioavailability and posed a high risk of reproductive toxicity (Fig. 10b; Table S9). Ivermectin (an avermectin derivative), a broad-spectrum anti-parasitic agent is a potent endectocide and has been reported to inhibit the HIV-1 integrase and host cell Importin (responsible for nuclear import of viral proteins) subsequently hindering viral replication and also exhibits antiviral potency against other retroviruses (Caly et al., 2020). Ivermectin did not interact with any ACE-2 binding regions of Spike protein. However, it did depict a significant interaction with CYS379 which is involved in disulfide bond formation (CYS379-CYS432) and depicted no systemic toxicity whatsoever. Docking of the selected drug candidates with the spike glycoprotein trimeric structures elucidated distinctive subsidiary spike protein target regions in addition to the RBD, contriving the segmented drug targets predicted systemic toxicity upon administration (Delaney et al., 2006) (Fig. S7; Table S7).

Raltegravir is an anti-retroviral drug belonging to the class of integrase inhibitors, ultimately affecting viral DNA insertion and its subsequent integration into the host cell’s DNA. It is primarily used to treat HIV-1 infections (Boesecke and Gelgor, 2009). It exhibited seven significant interactions of which only one targeted the loop residue of spike protein (LEU441) (Fig. S7). Similarly, Camostat (a carbonyl compound) is a synthetic serine protease inhibitor targeting the host cell TMRPS2 (viral entry mediator) thereby inhibiting viral infection and replication (Uno, 2020). It also exhibits anti-inflammatory and anti-fibriotic capacity. Like Raltegravir, it interacted with LEU441 but overall had only two significant bonds. Both drugs lacked predicted toxicity upon administration (Fig. S8).

Doxycycline (a semi-synthetic derivative of oxytetracycline) is a protein synthesis inhibitor and exhibits broad spectrum anti-bacterial potency (Holmes and Charles, 2009). It also exhibits antiviral competence by inhibition of matrix metalloproteases in case of Dengue and other retroviruses or transcriptional up-regulation of intracellular zinc finger antiviral protein (ZAP) (Malek et al., 2020). Although it illustrated three significant spike RBD loop interactions, it was predicted to have a low bioavailability and posed a high risk of reproductive toxicity (Fig. 10b; Table S9). Ivermectin (an avermectin derivative), a broad-spectrum anti-parasitic agent is a potent endectocide and has been reported to inhibit the HIV-1 integrase and host cell Importin (responsible for nuclear import of viral proteins) subsequently hindering viral replication and also exhibits antiviral potency against other retroviruses (Caly et al., 2020). Ivermectin did not interact with any ACE-2 binding regions of Spike protein. However, it did depict a significant interaction with CYS379 which is involved in disulfide bond formation (CYS379-CYS432) and depicted no systemic toxicity whatsoever. Docking of the selected drug candidates with the spike glycoprotein trimeric structures elucidated distinctive subsidiary spike protein target regions in addition to the RBD, contriving the segmented drug targets predicted systemic toxicity upon administration (Delaney et al., 2006) (Fig. S7; Table S7).

Raltegravir is an anti-retroviral drug belonging to the class of integrase inhibitors, ultimately affecting viral DNA insertion and its subsequent integration into the host cell’s DNA. It is primarily used to treat HIV-1 infections (Boesecke and Gelgor, 2009). It exhibited seven significant interactions of which only one targeted the loop residue of spike protein (LEU441) (Fig. S7). Similarly, Camostat (a carbonyl compound) is a synthetic serine protease inhibitor targeting the host cell TMRPS2 (viral entry mediator) thereby inhibiting viral infection and replication (Uno, 2020). It also exhibits anti-inflammatory and anti-fibriotic capacity. Like Raltegravir, it interacted with LEU441 but overall had only two significant bonds. Both drugs lacked predicted toxicity upon administration (Fig. S8).

Doxycycline (a semi-synthetic derivative of oxytetracycline) is a protein synthesis inhibitor and exhibits broad spectrum anti-bacterial potency (Holmes and Charles, 2009). It also exhibits antiviral competence by inhibition of matrix metalloproteases in case of Dengue and other retroviruses or transcriptional up-regulation of intracellular zinc finger antiviral protein (ZAP) (Malek et al., 2020). Although it illustrated three significant spike RBD loop interactions, it was predicted to have a low bioavailability and posed a high risk of reproductive toxicity (Fig. 10b; Table S9). Ivermectin (an avermectin derivative), a broad-spectrum anti-parasitic agent is a potent endectocide and has been reported to inhibit the HIV-1 integrase and host cell Importin (responsible for nuclear import of viral proteins) subsequently hindering viral replication and also exhibits antiviral potency against other retroviruses (Caly et al., 2020). Ivermectin did not interact with any ACE-2 binding regions of Spike protein. However, it did depict a significant interaction with CYS379 which is involved in disulfide bond formation (CYS379-CYS432) and depicted no systemic toxicity whatsoever. Docking of the selected drug candidates with the spike glycoprotein trimeric structures elucidated distinctive subsidiary spike protein target regions in addition to the RBD, contriving the segmented drug targets predicted systemic toxicity upon administration (Delaney et al., 2006) (Fig. S7; Table S7).
within the spike protein for future drug development experiments (Fig. S9).

Several studies corroborate the usage of the drugs selected in the present study to counteract COVID-19. Remdesivir has reached the phase III level of clinical trials for the treatment of mild and moderate SARS-CoV-2 infections (Senanayake, 2020). Favipiravir, an antiviral/anti-influenza drug has been shortlisted in ten hospitals across India for a phase III trial with mild and moderately infected COVID-19 patients.
Table 7
Detailed analysis of selected drugs with spike glycoprotein RBD model involving significant interactions.

| Sr. No. | Drug    | Significant spike RBD model interacting residues | Bond Type | Bond Length (in Å) | Interacting residues of Actual spike Glycoprotein | Precise Loop Interactions | Contact residues of the SARS-CoV-2 RBD-ACE2 | Cysteine residue interactions |
|---------|---------|-----------------------------------------------|-----------|-------------------|-----------------------------------------------|--------------------------|----------------------------------------|------------------------------|
|         |         |                                               | RBD       |       |       |       |       |       |       |
|         |         |                                               | RBM (319-541) |       |       |       |       |       |       |
|         |         |                                               | RBM (437-508) |       |       |       |       |       |       |
| 1       | Tenofovir | W436                                         | Conventional Hydrogen bond | 2.62 | Yes | No | No | No | No |
|         |         | F342                                         | Carbon Hydrogen bond | 3.30 | Yes | No | No | No | No |
|         |         | W436                                         | Electrostatic (pi-anion) | 3.48 | Yes | No | No | No | No |
|         |         | L441*                                        | Hydrophobic (pi-sigma) | 3.58 | Yes | Yes | 2 g | No | No |
|         |         | W436                                         | Hydrophobic (pi-sigma) | 3.80 | Yes | No | No | No | No |
| 2       | Eugenol  | E471                                         | Carbon Hydrogen bond | 3.30 | Yes | Yes | 3a | No | No |
|         |         | E471                                         | Carbon Hydrogen bond | 3.29 | Yes | Yes | 3a | No | No |
|         |         | D467                                         | Electrostatic (pi-anion) | 3.74 | Yes | Yes | No | No | No |
|         |         | K458                                         | Hydrophobic (pi-sigma) | 3.81 | Yes | Yes | 2 g | No | No |
| 3       | Allicin  | E471                                         | Electrostatic interaction | 3.78 | Yes | Yes | No | No | No |
|         |         | K458                                         | Hydrogen bond | 2.70 | Yes | Yes | 2 g | No | No |
|         |         | S469                                         | Hydrogen bond | 3.08 | Yes | Yes | No | No | No |
|         |         | P491                                         | Hydrophobic interaction | 3.98 | Yes | Yes | 3 b | No | No |
| 4       | Stavudine | Y453                                         | Conventional Hydrogen bond | 2.35 | Yes | Yes | 2 g | Yes | No |
|         |         | N501*                                        | Conventional Hydrogen bond | 3.05 | Yes | Yes | 3f | Yes | No |
|         |         | R403*                                        | Carbon Hydrogen bond | 3.30 | Yes | Yes | 1e | No | No |
|         |         | Y495                                         | Carbon Hydrogen bond | 3.23 | Yes | Yes | 3f | No | No |
|         |         | S494                                         | Carbon Hydrogen bond | 3.30 | Yes | Yes | 3 b | No | No |
| 5       | Raltegravir | S349                                        | Conventional Hydrogen bond | 2.59 | Yes | No | No | No | No |
|         |         | S399                                         | Conventional Hydrogen bond | 2.98 | Yes | No | No | No | No |
|         |         | S349                                         | Conventional Hydrogen bond | 2.96 | Yes | No | No | No | No |
|         |         | F347                                         | Conventional Hydrogen bond | 2.90 | Yes | No | No | No | No |
|         |         | L441*                                        | Conventional Hydrogen bond | 3.65 | Yes | Yes | 2 g | Yes | No |
|         |         | N354                                         | Pi-Donor Hydrogen bond (alkyl) | 2.44 | Yes | No | No | No | No |
|         |         | A348                                         | Hydrophobic (alkyl) | 3.54 | Yes | No | No | No | No |
| 6       | Zalcitabine | R403*                                       | Carbon Hydrogen bond | 3.30 | Yes | No | 1e | No | No |
| 7       | Favipiravir | R403*                                        | Conventional Hydrogen bond | 3.05 | Yes | No | 1e | No | No |
|         |         | Y505                                         | Conventional Hydrogen bond | 3.30 | Yes | Yes | 3f | No | No |
|         |         | R403*                                        | Carbon Hydrogen bond | 3.40 | Yes | No | 1e | No | No |
| 8       | Camostat | L441*                                        | Conventional Hydrogen bond | 2.91 | Yes | Yes | 2 g | No | No |
|         |         | N354                                         | Carbon Hydrogen bond | 3.30 | Yes | No | No | No | No |
| 9       | Doxycycline | Y449                                        | Conventional Hydrogen bond | 1.87 | Yes | Yes | 2 g | Yes | No |
|         |         | N501*                                        | Conventional Hydrogen bond | 2.35 | Yes | Yes | 3f | Yes | No |
|         |         | S494                                         | Conventional Hydrogen bond | 3.22 | Yes | Yes | 3 b | No | No |
|         |         | Y505                                         | Hydrophobic (pi-sigma) | 3.63 | Yes | Yes | 3f | Yes | No |
| 10      | Ivermectin | G381                                         | Conventional Hydrogen bond | 2.32 | Yes | No | No | No | No |
|         |         | F377                                         | Conventional Hydrogen bond | 2.99 | Yes | No | No | No | No |

(continued on next page)
patients while another combination of Doxycycline and Ivermectin has been recommended to treat acute symptoms of COVID-19 in Bangladesh which has reached to the experimental stage in India (Hafeez et al., 2020).

Another example of combinatorial therapy combines Emtricitabine/Tenofovir-Alafenamide and Lopinavir/Ritonavir to treat COVID-19 patients (Duan et al., 2020). Camostat and Nefamostat and their derivatives are certified protease inhibitors which are currently being investigated in Japan as therapeutic alternatives to combat SARS-Co-V-2 (Hoffman et al., 2020). No significant interaction was observed between RBD and hydroxychloroquine thus concluding that the mode of action of hydroxychloroquine is different than that of direct interaction with spike protein (Table 6; Fig. 12b). Chloroquine inhibits the viral attachment to its receptor along with the subsequent molecular processes involved in the final viral particle extrusion (Narkhede et al., 2020; Devaux et al., 2020). Nafamostat, primarily used for anticoagulant therapy, exhibited spike-mediated membrane fusion inhibitory potentiality (Li and Clercq, 2020). These experimental studies advocate drug-repurposing as a viable and resourceful technique to eradicate the infectivity of SARS-CoV-2 because of its cost-effectiveness and reduction of administration complexity in treatment regimes with high compatibility efficacy.

5. Conclusion

The present work comprised of selection and molecular docking analysis of a collection of 30 plausible drug candidates having diverse in silico mechanisms with the spike glycoprotein of SARS-CoV-2 aimed at identifying novel inhibitory competency of the same. Our results suggested that ten drugs out of the thirty selected could be utilized as promising drugs as they interacted with the experimentally validated ACE-2 binding residues and also depicted additional interactions with the ETD regions. Camostat, Favipiravir, Tenofovir, Raltegravir and Stavudine illustrated maximum interactions with the spike glycoprotein RBD model and the inherent trimeric structure displaying optimal bioavailability (score = 0.55) with absence of predicted systemic toxicity. As these drugs can be a good candidate for further in vitro or in vivo studies so dosage appropriate amalgamation of these drugs/drug derivatives in conjunction with refined experimental validation can serve as the platform for combinatorial drug therapy design and development to counter COVID-19 for futuristic applications.

Support

The authors have credited the sources of support received in the Acknowledgement page.
Permissions

The authors have duly credited all the sources which were utilized during the manuscript preparation.

Sources of funding

The present work has received no specific grant from any funding agency in the public, commercial or not-for-profit sectors in India or elsewhere.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

CRediT authorship contribution statement

Himanshu G. Toor: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing - original draft, Writing - review & editing, Final manuscript approval. Devjani I. Banerjee: Conceptualization, Data curation, Formal analysis, Investigation, Writing - review & editing.
Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing, Final manuscript approval. **Soumya Lipsa Rath:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing - review & editing, Final manuscript approval. **Siddhi A. Darji:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing - original draft, Writing - review & editing, Final manuscript approval.

**Declaration of competing interest**

The authors declare that they have no conflict(s) of interest.

**Acknowledgements**

The authors would like to express their sincere gratitude to the Dr. Vikram Sarabhai Institute of Cell and Molecular Biology, Faculty of Science, The Maharaja Sayajirao University of Baroda, Gujarat, India for providing the platform to conduct the present study. We are also thankful to Dr. Rina Soni (DST-Woman Scientist), Department of Chemistry, The Maharaja Sayajirao University of Baroda for her valuable suggestions. We are also grateful to Ashok and Rita Patel Institute of Integrated Study and Research in Biotechnology and Allied Sciences (ARIBAS), Gujarat; Charutar Vidya Mandal (CVM) and Sardar Patel University Gujarat, India for providing unconditional support for the duration of the present study.

**Appendix A. Supplementary data**

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejphar.2020.173720.

**References**

Al-Khayyat, M.Z.S., Al-Dabbagh, A.G.A., 2016. *In silico* prediction and docking of tertiary structure of LuxI, an inducer synthase of Vibrio fischeri. Rep. Biochem. Mol. Biol. 4 (2), 66–75.
