Emerging Strains Of SARS-Cov2 And Their Inhibition by The Use of Phytochemicals: An *In-Silico* Analysis

Mohammad Faheem Khan (✉ drfaheemkhan@erauniversity.in)

Era's Lucknow Medical College   https://orcid.org/0000-0002-1943-7160

Tanveer Ahamad

Era's Lucknow Medical College   https://orcid.org/0000-0002-1062-084X

Waseem Ahmad Ansari

Era's Lucknow Medical College

Mohsin Ali Khan

Era's Lucknow Medical College

Zaw Ali Khan

Era's Lucknow Medical College

Mohd Aamish Khan

Era's Lucknow Medical College

Abstract

Recently identified coronavirus namely severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a single-stranded positive-sense RNA virus with a genome of 29.9 kb in size encoding 14 open reading frames (ORFs) and 27 different structural and non-structural proteins. Among the structural proteins, trimeric-shaped spike glycoprotein is responsible for the entry of the SARS-CoV2 genome into host cells by interacting with human angiotensin-converting enzyme 2 (ACE2) receptors that are present on the cell surface with high affinity. Notably, inhibition of spike protein is considered a prime target for the development of drugs against COVID-19. Viruses can mutate, and SARS-CoV-2 is no exception. Since the first whole genome of SARS-CoV2 was published in February 2020, at least 4400 amino acid substitutions and several thousand mutations have been identified to date. As of today, more than 3500 new variants of SARS-CoV2 have been sequenced with a high spreading and infectivity rate which makes the virus more contagious. These new variants have been spread to several countries including United States (US), United Kingdom (UK), Brazil, South Africa, India, and other countries, etc. Therefore, herein, we analysed the new SARS-CoV2 strains, constructed the 3D homology models of Brazil P.1 and Indian B.1.617 variants, and screened them against 100 phytochemicals having previously identified anti-viral activity. Our study revealed that the top three phytochemicals for each of the new strains might serve as potential anti- SARS-CoV-2 agents for further drug discovery and development process to tackle COVID-19.

Introduction

Recently identified coronavirus namely severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a single-stranded positive-sense RNA virus that shows similar genomic organization (99.8% sequence identity) with genes of previously identified SARS-CoV (1). Recent studies showed that the genome of SARS-CoV2 is approximately 29.9 kB in size encoding 14 open reading frames (ORFs) and 27 different structural and non-structural proteins. The structural proteins including spike glycoprotein (S), an envelope protein (E), nucleocapsid protein (N), and matrix protein (M) are noted to have a key role in binding to host receptors. The non-structural proteins such as papain-like protease (nsp3), main protease (nsp5), RNA-dependent RNA polymerase (nsp12), helicase (nsp13), and 2'-ZO-methyltransferase (nsp16) are involved from transcription to replication to pathogenesis processes during the life cycle of SARS-CoV2 (2). Among them, spike protein (S-protein) which is trimeric in shape is responsible for the entry of the SARS-CoV2 genome into host cells by interacting with human angiotensin-converting enzyme 2 (ACE2) receptors that are present on the cell surface with high affinity. The S-protein has two subunits as the N-terminal S1 domain and membrane-proximal S2 domain. The S1 domain recognizes the cell surface ACE2 receptor through the receptor-binding domain (RBD) while the S2 domain intervenes in the fusion of viral particles with the host cell membrane. The RBD and other binding domains of S-protein are known to neutralize antibodies that are formed against SARS-CoV2 (3). As a result, inhibition of S-protein and ACE2 human receptor can be considered as prime targets for the development of drugs against COVID-19. Notably, most of the vaccines which are in clinical trials or are being used to neutralize SARS-CoV2 infection are generally based on S-protein sequences. Among the anti-SARS-CoV2 vaccines, several candidates use the RBD site as the sole antigen

Viruses can mutate, and SARS-CoV-2 is no exception. Since the first whole genome of SARS-CoV2 was published in February 2020, at least 4400 amino acid substitutions and several thousand mutations have been identified for the same virus so far (4). Among them, a total of 725 mutations in S-protein with 89 on the RBD segment has occurred, suggesting that RBD is more dispose site to mutation (5). The S-protein mutation such as D614G has already appeared prominently in many places around the world during early pandemic situations through single amino acid substitution at 614 positions from an aspartic acid (D) to a glycine (G) residue. This mutation has been considered potentially more transmissible and infective (6). Recent data provided some shreds of evidence that D614G mutation is associated with the higher load of viral nucleic acid in the upper respiratory tract. Though this mutation is potentially more transmissible and infective, still, the disease severity is low (7). As a consequence of numerous mutations, more than 3500 new variants of SARS-CoV2 have been sequenced with a high spreading and infectivity rate which makes the virus more contagious. As of year-end 2020, these new variants had spread to several countries including United States (US), United Kingdom (UK), Brazil, South Africa,

etc. A highly mutated variant (20I/501Y.V1 or lineage B.1.1.7) had been identified in early September of 2020 in Eastern England, hence also called as UK variant. Early pieces of evidence showed that this variant may be associated with increased transmissibility and higher death risk compared to wild type variants (8). In Nelson Mandela Bay, South Africa, another variant (20H/501Y.V2 or B.1.351) was also identified with multiple mutations. Preliminary studies suggested that it has a low impact on disease severity, but it may affect the efficacy of vaccines because of the presence of E484K mutation in spike protein (9). A Brazil variant (known as lineage P.1), first detected in four travellers from Brazil at Japan airport, is receiving more attention because it may allow immune scape as evidenced in a report. In October 2020, a new Indian B.1.617 variant has been detected in Maharashtra state but in April 2021, the number of new infections has rapidly jumped to rise in an unprecedented manner because of existing of double mutations in spike glycoprotein which has predicted by preliminary study. Although this variant has spread to other countries like Singapore, Australia, Germany, Belgium, United Kingdom as well as the United States, still, some experts have a deep concern that it may turn into a super spreader mutant that will be continued to spread in other countries of the world. Health workers are uncertain that whether the Indian B.1.617 variant is more infectious or may inhibit vaccine efficacy. In addition, some other strains including B. 1. 429 California, B.1.427 California, B.1.526 New York, B.1.526.1 New York, B.1.526.2 New York have also been detected with mild to moderate severity

Current COVID-19 vaccines are generally based on the spike protein of SARS-CoV-2, but for the emerging strains having mutations in spike proteins, it is a question of concern that how much these vaccines are efficient than the wild-type SARS-CoV-2 because new strains seem to be more transmissible and deadlier. To date, at least 308 vaccine candidates are in various clinical stages of development including 24 in Phase I trials, 33 in Phase II trials, and 16 in Phase III development. Most of them showed efficacy as high as 95% in preventing symptomatic COVID-19 infections. As of April 2021, regulatory authorities from different countries approved 13 vaccines for public use. Depending on the mode of action, all these vaccines are divided into four different groups including two RNA vaccines, five conventional inactivated vaccines, four viral vector vaccines, and two protein subunit vaccines (11). Medicinal plants and herbs employed in traditional medicines have attracted significant attention because of having bioactive molecules that may act as therapeutic agents for the prevention and treatment of several diseases with no or minimal side effects (12). In the last few months, several studies, as well as our studies, have shown the promising role of natural products or phytochemicals to inhibit the SARS-CoV2 targeting their structural and non-structural proteins including spike protein by the use of computational tools (13–16). Therefore, the aim of the present study was to unfold the insights into the new strains of SARS-CoV2 and their inhibition by phytochemicals from different medicinal plants and herbs for the prevention of COVID-19 by using computational tools.

Material And Methods

Data Collection

All the data related to mutations that gave rise to new strains of SARS-CoV-2 were searched and collected from online literature databases including Wikipedia, google scholar, PubMed, SciFinder, Scopus, NCBI, Web of Science as well as various scientific journals like Nature, NatureSpringer, ScienceDirect, Wiley, ACS, RSB, RSC etc. along with other relevant websites

Homology modelling

The experimental 3D structure of SARS-CoV-2 spike glycoprotein of Brazil Strain P.1 and Indian variant B.1.617 were unavailable on the RCSB protein data bank (https://www.rcsb.org/). Therefore, to get 3D structure, they were modeled with the help of the Swiss Model server (https://swissmodel.expasy.org/). The target sequence of both the strains with GenBank QRX39425.1 (https://www.ncbi.nlm.nih.gov/protein/QRX39425) and QUA70603.1 (https://www.ncbi.nlm.nih.gov/protein/QUA70603.1) was retrieved in FASTA format from NCBI (https://www.ncbi.nlm.nih.gov/) website. Later on, the amino acid sequences of the targeted strain were pasted in the Swiss Model server to build their 3D model (17).

Template selection

BLAST and HHblits methods were applied for template structure against the SWISS-MODEL template library. The template selection of the targeted sequence of Brazil Strain P.1 and Indian variant B.1.617 was performed on the basis of coverage of template sequence with the targets sequence, identity (similarity between them), and global model quality estimation (GMQE). The two PDB IDs viz. 6ZWV and 7KRS were used as templates against Brazil Strain P.1 Spike glycoprotein and Indian B.1.617 spike glycoprotein sequences, respectively. The range of GMQE relies on a range from 0 to 1. Herein, the higher value of GMQE represented the maximum accuracy (18,19).

Protein structure validation

The quality of predicted models of Brazil Strain P.1 and Indian variant B.1.617 was analysed with the help of the PROCHECK Ramachandran plot. Ramachandran statistics were used to figure out the range of amino acids available in the favourable as well as allowed and disallowed regions. To find out the overall quality score of generated models, the ERRAT server was used which was available in SAVES v6.0. The root-mean-square deviation (RMSD) values between template structure and modelled structure were obtained by aligning them in PyMol software

Protein preparation

The UK strain B.1.1.7 and South Africa strain B.1.351 were available and downloaded from the protein data bank website. All spike proteins (for Brazil Strain P.1, Indian variant B.1.617, UK strain B.1.1.7, and South Africa strain B.1.351) were subjected to pre-process and refinement by using protein preparation wizard in the Schrödinger suite. The selected receptor was pre-processed by adding missing hydrogen atoms. All hydrogen bonds were optimized using sample water orientation whereas energy minimization of the receptor was performed by using default RMSD value 0.30 Å and OPLS3e force field methods

Phytochemical hits screening, preparation and optimization as ligands

A literature review of phytochemical with antiviral activities was performed and a dataset comprising 100 phytochemicals with antiviral activities was prepared. The chemical structures of all the selected phytochemicals were retrieved from PubChem (https://pubchem.ncbi.nlm .nih.gov/) website along with their PubChem IDs. Their spatial data file in SDF format was utilized for prediction purposes. The ligand structures were corrected for bond length and bond angles. Later on, missing hydrogen atoms were also added. The geometry of all ligands was optimized by employing force field OPLS3e using the LigPrep module of Schrodinger software. Finally, the optimized structures were further used for docking studies.

Molecular docking

Glide v8.8 (Schrodinger, LLC, New York) software was used for docking studies to identify the binding affinities of ligands within the binding pockets of target proteins. For each ligand, binding affinities were expressed as Glide scores. The Discovery studio was used to visualize the docked poses. Among them, the best poses were ranked on the basis of an energy function that combines the empirical and force-field algorithm simultaneously.

Results And Discussion

Analysis of mutation patterns

The UK strain (20I/501Y.V1 or lineage B.1.1.7) is originated by the replacement of amino acid asparagine with tyrosine at position 501 of the RBD subunit of the spike protein along. Several other mutations including 69/70 deletion as well as P681H near the S1/S2 furin cleavage sites were also seen. The South Africa strain (20H/501Y.V2 or B.1.351) was detected to have numerous mutations in the spike protein, including K417N, E484K, and N501Y. The first Brazil P.1 strain of SARS-CoV2 contains 10 spike protein mutations N501Y, E484K, and K417T along with 17 other unique mutations. In contrast, the

second Brazilian P.2 lineage was also identified as having three mutations in spike proteins namely E484K, N501Y, and K417T. As per evidence, the E484K mutation has been noticed in the South African variant but not in the UK variant (10). The Indian B.1.617 variant consists of two mutations namely E484Q and L452R in the spike glycoprotein which were already in circulation globally. Because of the current surge or worst condition, scientists are worried that B.1.617 variant is not a double mutant but it may consist of more mutations including E154K, P681R, and Q1071H. The three mutations E484Q, L452R, and P681R have been found in other variants of concern from the UK, South Africa, and Brazil. The mutation E484Q is similar to the E484K mutation that was previously seen in the Brazilian and South African variants. On the other hand, L452R is similar to the California variant, an immune escape strain, thus, may affect the vaccine efficacy. Further, the P681R mutation showed similarities to that of mutations seen in the United Kingdom variant. Furthermore, other strains that are present in many countries exhibited different mutations in their spike glycoproteins as depicted in table 1.

Table 1: Analysis of presence of mutations in spike proteins of various emerging strains of SARS-CoV2

Mutations in spike protein	B.1.1.7 UK strain	P.1 Brazil	B.1.429 California	B.1.351 South Africa	B.1.427 California	B.1.526 New York	B.1.526.1 New York	B.1.526.2 New York	P.2 Brazil	B.1.617 India
DEL69/70.0	+	-	-	-	-	-	-	-	-	-
DEL144/144.0	+	-	-	-	-	-	+	-	-	-
N501Y	+	+	-	+	-	-	-	-	-	-
A570D	+	-	-	-	-	-	-	-	-	-
D614G	+	+	+	+	+	+	+	+	+	+
P681H	+	-	-	-	-	-	-	-	-	-
T716I	+	-	-	-	-	-	-	-	-	-
S982A	+	-	-	-	-	-	-	-	-	-
D1118H	+	-	-	-	-	-	-	-	-	-
L18F	-	+	-	-	-	-	-	-	-	-
T20N	-	+	-	-	-	-	-	-	-	-
P26S	-	+	-	-	-	-	-	-	-	-
D138Y	-	+	-	-	-	-	-	-	-	-
R190S	-	+	-	-	-	-	-	-	-	-
K417T	-	+	-	-	-	-	-	-	-	-
E484K	-	+	-	+	-	-	-	-	-	-
H655Y	-	+	-	-	-	-	-	-	-	-
T1027I	-	+	-	-	-	-	-	-	-	-
S13I	-	-	+	-	-	-	-	-	-	-
W152C	-	-	+	-	-	-	-	-	-	-
L452R	-	-	+	-	+	-	+	-	-	+
K417N	-	-	-	+	-	-	-	-	-	-
A701V	-	-	-	+	-	-	-	-	-	-
D253G	-	-	-	-	-	-	-	+	+	-
T95I	-	-	-	-	-	+	-	+	+	-
F157S	-	-	-	-	-	-	+	-	-	-
T859N	-	-	-	-	-	-	+	-	-	-
D950H	-	-	-	-	-	-	+	-	-	-
D80G	-	-	-	-	-	-	+	-	-	-

D253G	-	-	-	-	-	+	-	-	-	-
S477N	-	-	-	-	-	-	-	+	+	-
L5F	-	-	-	-	-	-	-	+	+	-
Q957R	-	-	-	-	-	-	-	+	+	-
P681R	-	-	-	-	-	-	-	-	-	+
E484Q	-	-	-	-	-	-	-	-	-	+
E154K	-	-	-	-	-	-	-	-	-	+

Generation of 3D Protein Structures: Validation and Analysis

Homology modelling is a key procedure to predict the protein structures relating to their functions. The SWISS-MODEL server is used to design the protein structure by the use of the homology modelling method. Based on query coverage, identity and GMQE scores, the BLAST and HHblits search revealed that PDB IDs 6ZWV and 7KRS may act as the best possible template match in the homology modelling for spike glycoprotein of Brazil Strain P.1 and Indian B.1.617 strains with 100% query coverage respectively. The sequence identity (or sequence similarity) was found to 99.29 (or 0.62) and 99.14 (or 0.62) for Brazil P.1 and Indian B.1.167 strains respectively. Furthermore, the quaternary structure quality estimate (QSQE) values were 0.93 for Brazil P.1 strain and 0.78 for Indian B.1.617 strain. The values are mentioned in table 2

Table 2: The properties and their values of generated 3D structures of spike glycoproteins of new strains

Properties	Brazil Strain P.1	Indian Strain B.1.617
Template	7krs.1.A	6zwv.1.A
Oligostate	Homo trimer	Homo trimer
QSQE	0.93	0.78
Found by	HHblits	HHblits
Sequence identity	99.29	99.14
Sequence similarity	0.62	0.62
Range	14-1160	2-1151
Coverage	1.00	1.00
GMQE	0.72	0.66
QMEAN	-1.51	-2.25

The overall molecular structures of Brazil P.1 and Indian B.1.617 strains were inconsistent with the crystal structure of spike glycoprotein of wild type of SARS-CoV2 (PDB ID: 3M3V).

The analyses of the Ramachandran plot also revealed that 99.9% of residues were in the allowed region while only 0.1% of residues were in the outlier region (Figure 1)

Evaluation and Analysis of docking studies

The four new strains such as Indian B.1.617 variants, UK B.1.1.17 variant, South Africa B.1.351 variant and Brazil P.1 variant have been emerged by mutations in spike glycoprotein of SARS-CoV2 which interact with the transmembrane protein (ACE2) of human cell receptor. The interactions between S-protein and ACE2 is the key point to enter the virus into host cells, thereafter facilitate the replication of the viral genome. Therefore, inhibition of Spike and ACE2 proteins are the therapeutic targets for the anti-COVID-19 inhibitors. Now it is essential to search for novel compounds as potential SARS-CoV2 inhibitors targeting the inhibition of spike glycoprotein. In continuation of our research to find the phytochemicals as SARS-CoV2 inhibitors, we prepared a library that consists of 100 phytochemicals possessing potential biological activities against different types of viruses and screened them against the spike glycoprotein of the aforementioned new strains of wild SARS-CoV-2. Recently, several repurposing studies suggested that some already FDA-approved drugs (remdesivir, favipiravir, nafamostat etc.) could be used to treat COVID-19. Thus, herein, we selected nafamostat amongst them as a control to compare our results because it is a spike glycoprotein inhibitor. We performed preliminary docking screening of all 100 phytochemicals and identified some of them as potential inhibitors that are predicted to bind within the receptor binding pockets of four new strains (Table 3; Fig. 2) by performing molecular docking using Glide v8.8 (Schrodinger, LLC, New York) software, a computational tool. As discussed above, a threshold value ($\geq$ -7.00) of docking score was selected for further analysis. Thus, to this, the top three phytochemicals against each of the new strain (Indian B.1.617, UK B.1.1.17, South Africa, and Brazil P.1 Brazil Strains) were found to pass the threshold value as summarized in table 3.

Examining the docking study of phytochemicals with spike protein of B.1.617 Indian variant, the docking score revealed the binding order as rutin > EGCG > hesperidin. Rutin docked with a docking score of -8.160 by forming seven hydrogen bonds with Ser50, Cys299, Thr300, Asp735, Thr737, Gly755 and Thr759 amino acid residues. Two hydrophobic bonds were observed with Leu51 and Leu52 amino acid residues. EGCG docked with docking score -7.993 while interacting with Thr272, Thr300, Thr313, Asn315, Thr737, Thr759, Asn762, Leu751 and Leu752 amino acid residues by forming seven hydrogen and two hydrophobic bonds respectively. Moreover, hesperidin docked with docking score -7.873, showing hydrogen bond as well as hydrophobic bond interactions at active sites with Gln52, Ser314, Gly755, Thr759, Leu301, Leu752 and Phe757 amino acid residues. Docking analysis of phytochemicals with spike protein of B.1.1.17 UK strain revealed that hesperidin, withanolide G and rosmarinic acid docked with docking score of -8.993, -8.766 and -8.761 respectively. Hesperidin formed seven hydrogen bonds with Thr547, Thr573, Asp745, Asn856, Leu977, Asn978 and Arg1000 amino acids of target protein whereas ten hydrophobic bonds were also found through interactions with Val320, Pro322, Phe541, Leu546, Ile587, Pro589, Cys590, Met740, Val976, and Leu977 amino acids. Moreover, withanolide G exhibited molecular interactions with Thr549, Asn856, Asn978, Pro322, Val320, Phe541, Leu546, Ile587, Pro589, Cys590, Phe592, Met740 and Val976 amino acid residues whereas Thr547, Thr549, Thr573, Tyr741, Asn856, Asn578, Arg1000, Phe541, Leu546, Ile587, Pro589, Cys590, Met740, Ile742, Cys743, Leu966, Val976 and Leu977 amino acid residues were found to interact with the hydroxyl and phenyl functional moieties present in rosmarinic acid.

Table 3: Detail account of phytochemical, their docking score and molecular interactions with the amino acid residues at binding sites

Strains (Spike glycoprotein)	Phytochemicals	Docking score	Molecular interactions with the amino acids	
			Hydrogen bonds	Hydrophobic bonds
B.1.617 Indian variant	Rutin	-8.160	Ser50, Cyx299, Thr300, Asp735 Thr737, Gly755, Thr759	Leu51, Leu52
	EGCG	-7.993	Thr272, Thr300, Thr313, Asn315, Thr737, Thr759, Asn762	Leu751, Leu752
	Hesperidin	-7.873	Gln52, Ser314, Gly755, Thr759	Leu301, Leu752, Phe757
	Nafamostat	-5.665	Leu301, Asn315, Asp735, Asn762	-
B.1.1.7 UK strain	Hesperidin	-8.993	Thr547, Thr573 Asp 745, Asn856, Leu977, Asn978, Arg1000	Val320, Pro322, phe541, leu546, Ile587, pro589, Cys590, Met 740, Val976, Leu977
	Withanolide G	-8.766	Thr549, Asn856, Asn978	Pro322, Val320, Phe541, Leu546, Ile587, Pro589, Cys590, Phe592, Met740, Val976
	Rosmarinic acid	-8.761	Thr547, Thr549, Thr573, Tyr741, Asn856, Asn578, Arg1000	Phe541, Leu546, Ile587, Pro589, Cys590, Met740, Ile742, Cys743, Leu966, Val976, Leu977
	Nafamostat	-5.340	Val320, Asp571, Leu966	Pro322, Pro589, Cys590, Met740, Leu966, Val976, Leu977
B.1.351 South Africa strain	EGCG	-8.369	Thr549, Thr573, Asp574, Met740, Gly744, Asn856	Phe541, Ile569, Ala570, Ile587, Pro589, Tyr741, Phe855, Val976
	Diosmetin	-8.200	Thr549, Thr573, Ile587, Met740 Asn856	Phe541, Ala570, Pro589, Tyr741, Cys743, Leu966, Val976, Leu977
	Myricetin	-8.102	Thr549, Thr573, Met740, Gly744	Phe541, Ile587, Pro589, Tyr741, Cys743, Leu966, Val976, Leu977
	Nafamostat	-4.260	Asp574, Asp745, Asn978, Phe855	Ile587, Pro589, Leu977, Leu981
P.1 Brazil Strain	Rosmarinic acid	-9.235	Tyr369, Phe374, Phe377, Lys378, Asp405, Glu406, Gln409, Thr415, Thr417	Leu368, Ala372, Tyr495
	Epicatechin	-8.248	Tyr369, Phe374, Arg408, Gln409, Thr415	Ala372, Phe377, Pro384
	Quercetin	-7.925	Tyr369, Ser371, Phe374, Phe377, Lys378, Arg408, Gln409, Thr417	Leu368, Ala372
	Nafamostat	-5.325	Tyr369, Ser371, Phe374, Lys378, Glu988	Ala372, Pro384, Phe377, Val987

In the case of B.1.351 South Africa strain, the phytochemicals EGCG, diosmetin and myricetin showed strong binding affinities. EGCG docked with docking score -8.369 which was attributed to the presence of six hydrogen bonds as well as eight hydrophobic bonds with Thr549, Thr573, Asp574, Met740, Gly744, Asn856, Phe541, Ile569, Ala570, Ile587, Pro589, Tyr741, Phe855 and Val976 amino acids respectively. Additionally, diosmetin docked with docking score -8.200 via

interacting with Thr549, Thr573, Ile587, Met740, Asn856, Phe541, Ala570, pro589, Tyr741, cys743, leu966, Val976, and Leu977 amino acid residues whereas myricetin also displayed a significant docking score (-8.102) through molecular interactions with Thr549, Thr573, Met740, Gly744, Phe541, Ile587, Pro589, Tyr741, Cys743, Leu966, Val976 and Leu977 amino acid residues. Furthermore, another study of protein-ligand interactions revealed that rosmarinic acid binds with a higher dock score (-9.235) at the active sites of the spike protein of P.1 Brazil strain followed by the epicatechin (-8.248) and quercetin (-7.925). Rosmarinic acid exhibited the nine-strong hydrogen bond interactions with Thr369, Lys378, Asp405, Glu406, Gln409, Thr415, Phe374, Phe377, Thr417, amino acid residues as well as three hydrophobic bonds with Leu368, Ala372 and Tyr495 amino acid residues. On the other hand, epicatechin and quercetin exhibited molecular interactions with Tyr369, Ser371, Phe374, Phe377, Lys378, Arg408, Gln409, Thr417, Leu368 and Ala372 amino acid residues respectively. On comparing the results with nafamostat, a reference drug, it was quite significant that the top three phytochemicals among the selected ones had been found to have promising binding affinities with better binding scores -5.665, -5.340, -4.260, -5.325 against all new strains B.1.617 Indian variant, B.1.17 UK strain, B.1351 South Africa and P.1 Brazil strain respectively. The binding pose of the best phytochemical within the binding pocket of the spike protein of each of the new strains is collectively shown in Fig. 3. Thus, the present work revealed that rutin, hesperidin, EGCG, and rosmarinic acid displayed the potential inhibition against the spike protein of B.1.617 Indian variant, B.1.17 UK strain, B.1351 South Africa and P.1 Brazil strain respectively and they may serve as promising leads for further optimisation and drug development process to manage the COVID-19.

Conclusion

In concluding remarks, the present study revealed that because of mutations in spike glycoprotein, many thousands of new variants of SARS-CoV2 have been sequenced with a high spreading and infectivity rate which makes the virus more contagious. As of May 2020, these new variants have been spread to several countries including United States (US), United Kingdom (UK), Brazil, South Africa, India, and other countries. The unavailable experimental 3D structure of SARS-CoV-2 spike glycoprotein of Brazil Strain P.1 and Indian variant B.1.617 on the RCSB protein data bank is modeled by our research group with the help of the Swiss Model server. After structure validation, both the strains have shown > 99.0 sequence identity, and thus, these modeled structures may help researchers working on findings of therapeutic agents against COVID-19. Phytochemicals have already been shown to have a promising inhibitory effect for the successful treatment against several viral diseases. Herein, we screened 100 phytochemicals having anti-viral effects and selected the top three hits as potential inhibitors of spike glycoprotein of each of the new strains. The order of inhibitory effect of top three phytochemicals against Indian B.1.617 variant is rutin > EGCG > hesperidin; against UK B.1.1.17 variant is hesperidin > withanolide G > rosmarinic acid; against South Africa B.1.351 variant is EGCG > diosmetin > myricetin; against Brazil P.1 variant is rosmarinic acid > epicatechin > quercetin. Thus, our in-silico analysis explicitly demonstrated that rutin, EGCG, hesperidin, withanolide G, rosmarinic acid, diosmetin, myricetin, epicatechin, and quercetin may be considered as promising candidates as inhibitors of the spike protein of SARS-CoV2 new strains. Further, in-vitro and in-vivo experiments are urgently needed to prove them as clinical drugs to combat COVID-19.

Declarations

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgement

References

1. Kumar A, Ansari W, Ahamad T, Saquib M, Khan M. Safe Use of Sodium Dodecyl Sulfate (SDS) to Deactivate SARS-CoV-2: An Evidence-Based Systematic Review. *Coronaviruses*. 2021;01. doi:10.2174/2666796701666210105114804

2. Tahir ul Qamar M, Alqahtani S, Alamri M, Chen L. Structural basis of SARS-CoV-2 3CLpro and anti-COVID-19 drug discovery from medicinal plants. *J Pharm Anal*. 2020;10(4):313–319. doi:10.1016/j.jpha.2020.03.009

3. Jaimes J, Millet J, Whittaker G. Proteolytic Cleavage of the SARS-CoV-2 Spike Protein and the Role of the Novel S1/S2 Site. *iScience*. 2020;23(6):101212. doi:10.1016/j.isci.2020.101212

4. Shariare M, Parvez M, Karikas G, Kazi M. The growing complexity of COVID-19 drug and vaccine candidates: challenges and critical transitions. *J Infect Public Health*. 2021;14(2):214–220. doi:10.1016/j.jiph.2020.12.009

5. Chen J, Wang R, Wang M, Wei G. Mutations Strengthened SARS-CoV-2 Infectivity. *J Mol Biol*. 2020;432(19):5212–5226. doi:10.1016/j.jmb.2020.07.009

6. Groves D, Rowland-Jones S, Angyal A. The D614G mutations in the SARS-CoV-2 spike protein: Implications for viral infectivity, disease severity and vaccine design. *Biochem Biophys Res Commun*. 2021;538:104–107. doi:10.1016/j.bbrc.2020.10.109

7. Korber B, Fischer W, Gnanakaran S et al. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell*. 2020;182(4):812–827.e19. doi:10.1016/j.cell.2020.06.043

8. Kirby T. New variant of SARS-CoV-2 in UK causes surge of COVID-19. *The Lancet Respiratory Medicine*. 2021;9(2):e20-e21. doi:10.1016/s2213-2600(21)00005-9

9. Weisblum Y, Schmidt F, Zhang F et al. Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. *Elife*. 2020;9. doi:10.7554/elife.61312

10. Toovey O, Harvey K, Bird P, Tang J. Introduction of Brazilian SARS-CoV-2 484K.V2 related variants into the UK. *Journal of Infection*. 2021;82(5):e23-e24. doi:10.1016/j.jinf.2021.01.025

11. Chakraborty S, Mallajosyula V, Tato C, Tan G, Wang T. SARS-CoV-2 vaccines in advanced clinical trials: Where do we stand?. *Adv Drug Deliv Rev*. 2021;172:314–338. doi:10.1016/j.addr.2021.01.014

12. Yuan H, Ma Q, Ye L, Piao G. The Traditional Medicine and Modern Medicine from Natural Products. *Molecules*. 2016;21(5):559. doi:10.3390/molecules21050559

13. Khan M, Khan M, Khan Z, Ahamad T, Ansari W. In-Silico Study to Identify Dietary Molecules as Potential SARS-CoV-2 Agents. *Lett Drug Des Discov*. 2020;17. doi:10.2174/1570180817999201209204153

14. Ansari W, Ahamad T, Khan M, Khan Z, Khan M. Luteolin: A Dietary Molecule as Potential Anti-COVID-19 Agent. 2020. doi:10.21203/rs.3.rs-35368/v1

15. Khan M, Khan M, Khan Z, Ahamad T, Ansari W. Identification of Dietary Molecules as Therapeutic Agents to Combat COVID-19 Using Molecular Docking Studies. 2020. doi:10.21203/rs.3.rs-19560/v1

16. Huang J, Tao G, Liu J, Cai J, Huang Z, Chen J. Current Prevention of COVID-19: Natural Products and Herbal Medicine. *Front Pharmacol*. 2020;11. doi:10.3389/fphar.2020.588508

17. Oduselu, G. O., Ajani, O. O., Ajamma, Y. U., Brors, B., & Adebiyi, E. (2019). Homology modelling and molecular docking studies of selected substituted Benzo [d] imidazol-1-yl) methyl) benzimidamide scaffolds on Plasmodium falciparum adenylosuccinate lyase receptor. *Bioinformatics and biology insights*, *13*, 1177932219865533.

18. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., … Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic acids research*, *46*(W1), W296-W303.

19. Messaoudi, A., Belguith, H., & Hamida, J. B. (2013). Homology modeling and virtual screening approaches to identify potent inhibitors of VEB-1 β-lactamase. *Theoretical Biology and Medical Modelling*, *10*(1), 1–10.
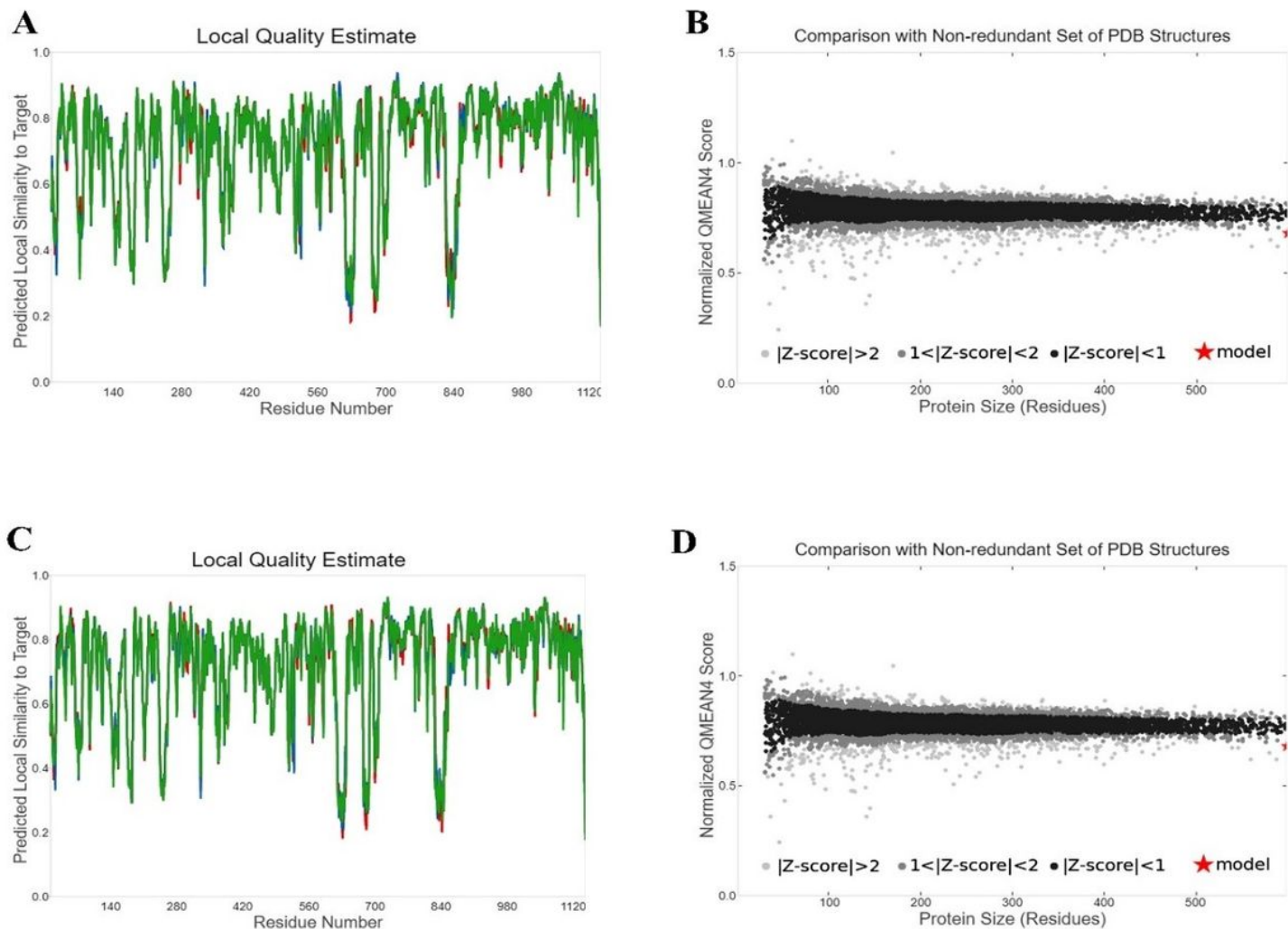
Figures

Figure 1

A: Local quality estimate of modelled spike glycoprotein of UK B.1.1.7 strain, B: Comparison of the modelled spike glycoprotein of UK B.1.1.7 strain with a non-redundant set of PDB structure, C: Local quality estimate of modelled spike glycoprotein of South Africa B.1.351 strain, D: Comparison of the modelled spike glycoprotein of South Africa B.1.351 strain with a non-redundant set of a PDB structure.
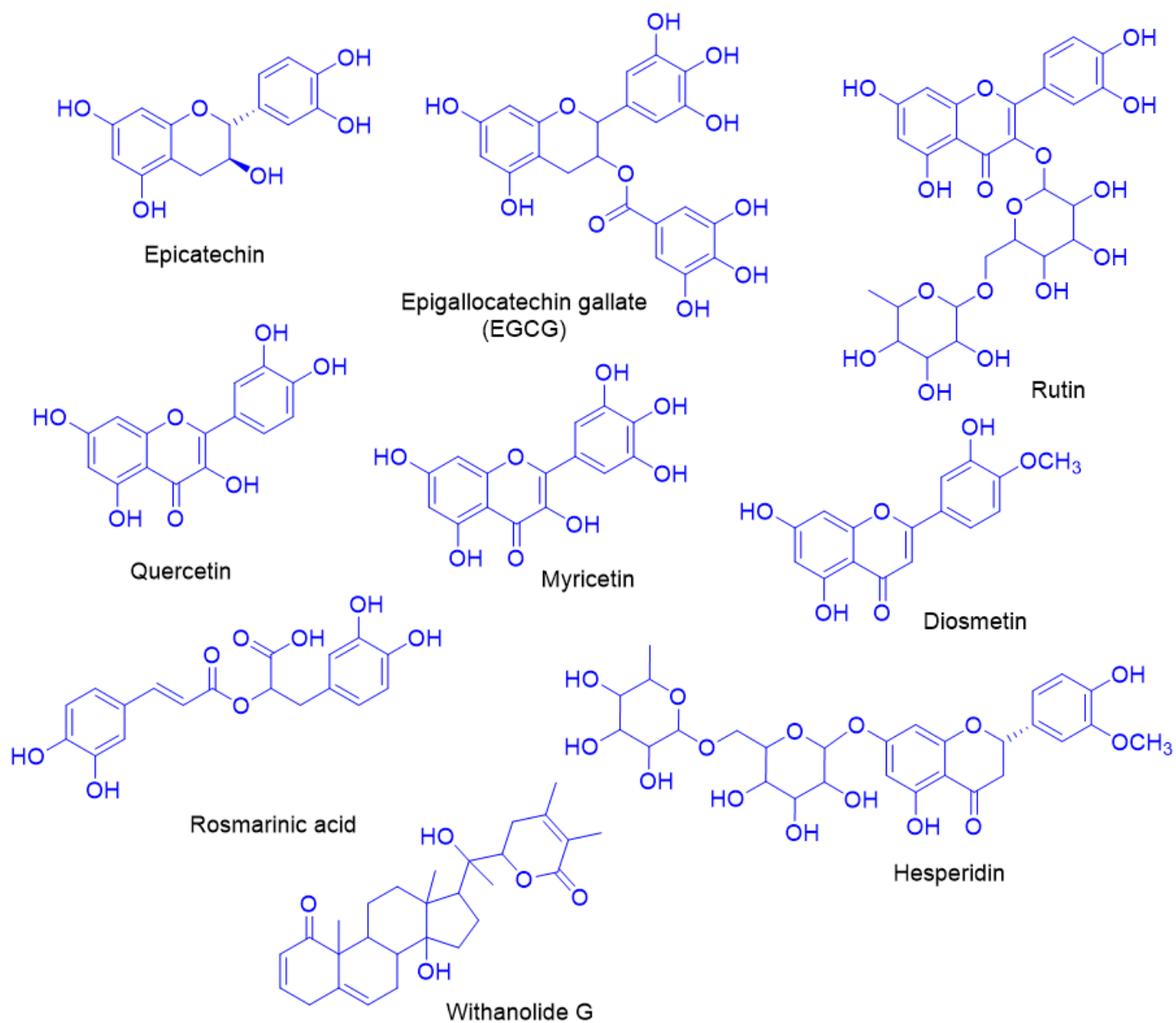
Figure 2

The chemical structures of top three phytochemicals as SARS-CoV2 inhibitors for each of the new strains
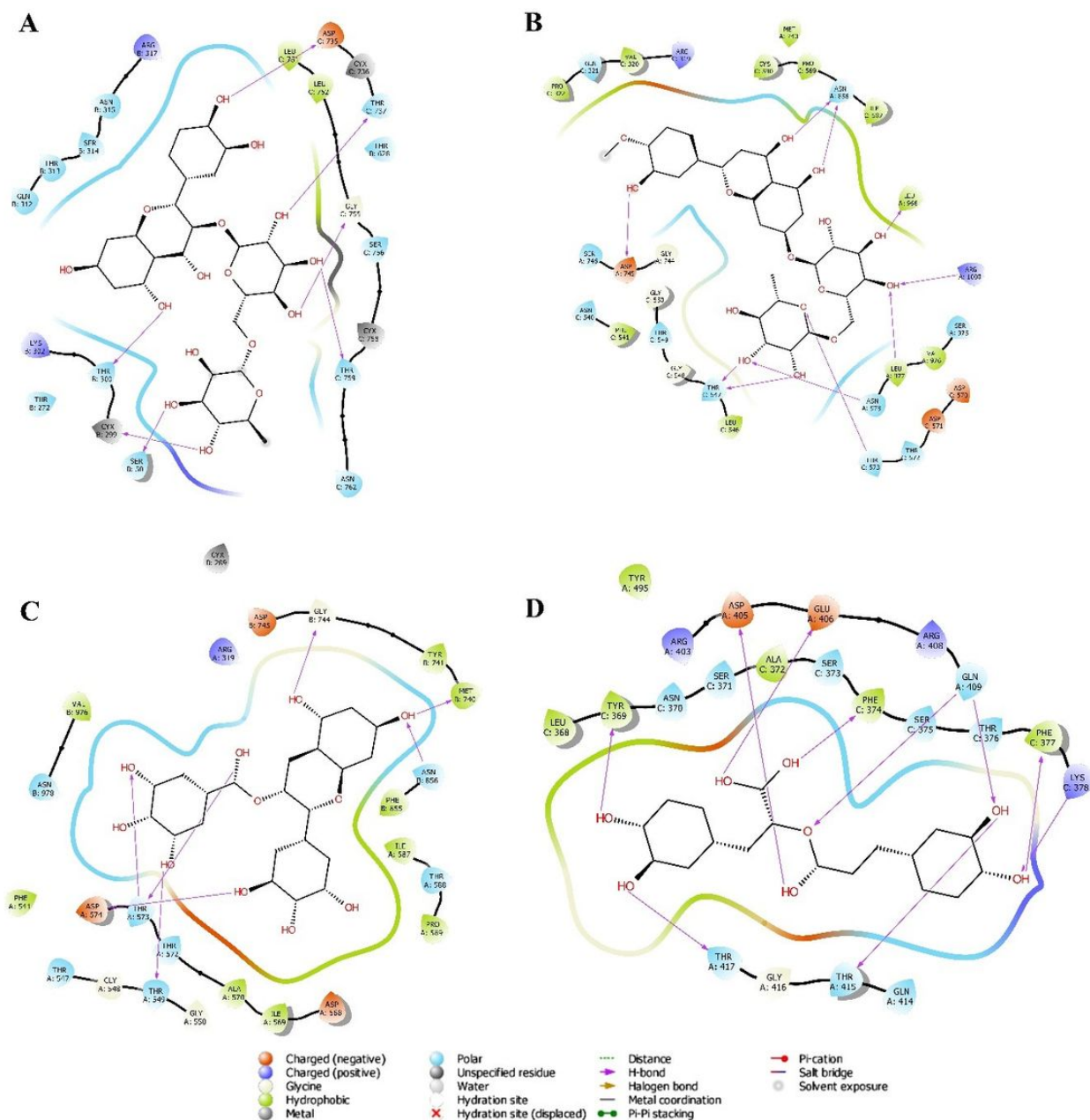
Figure 3

Two-dimensional view of molecular interactions in between (A) Rutin and spike protein of B.1.617 Indian variant, (B) Hesperidin and spike protein of B.1.1.17 UK strain, (C) EGCG and spike protein of B.1.351 South Africa strain, (D) Rosmarinic acid and spike protein of P.1 Brazil Strain