Computational approach for the design of potential spike protein binding natural compounds in SARS-CoV2

Anamika Basu
Gurudas College
Anasua Sarkar (anasua.sarkar@jadavpuruniversity.in)
Jadavpur University
Ujjwal Maulik
Jadavpur University

Research Article

Keywords: Spike protein of SARS-CoV2, ACE2, molecular docking, hesperidin, chrysin

Posted Date: June 4th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-33181/v1

Version of Record: A version of this preprint was published on October 19th, 2020. See the published version at https://doi.org/10.1038/s41598-020-74715-4.

Abstract

Angiotensin converting enzyme 2 (ACE2) (EC:3.4.17.23) is a transmembrane protein which is considered as receptor for spike protein binding of novel coronavirus (SARS-CoV2). Since no specific medication is available to treat COVID-19, designing of new drug is important and essential. In this regard, in silico method plays an important role as it is rapid, cost effective, compared to the trial and error methods using experimental studies. Natural products are safe and easily available to treat coronavirus effected patients, in the present alarming situation. In this paper five phytochemicals which belong to flavonoid and anthraquinone subclass, selected as small molecules in molecular docking study of spike protein of SARS-CoV2 with its human receptor ACE2 molecule. From the detail analysis of their molecular binding site on spike protein binding site with its receptor, hesperidin, emodin and chrysin are selected as competent natural products from both Indian and Chinese medicinal plants, to treat COVID-19.

I. Introduction

COVID-19 is caused by novel coronavirus named SARS-CoV-2. Virus particles are spherical in shape having spike proteins around them. These proteins are responsible for virus replication in human host cells. Spike proteins latch onto human cells and undergo a structural change, which results in the fusion of viral membrane with human host cell membrane. Thus, the viral genes enter into the host cell and produces more viruses after coping its genome. SARS-CoV-2 spike proteins bind to the receptor proteins, on the human cell surface, known as angiotensin converting enzyme 2 (ACE2). Atomic level structure of SARS-CoV-2 spike proteins have a Receptor Binding Domain (RBD) for binding to host human cells. Receptor Binding Domain (RBD) of spike glycoprotein (RBD-S) can bind to the ACE2 receptor at the Protease Domain (PD) of the host human cell, causing viral infection.

Considering the preliminary data, it has been suggested that ACE2 is a receptor for the novel coronavirus (SARS-CoV-2), that was identified as the cause of the respiratory disease outbreak in Wuhan in late 2019 [1], [2]. SARS-CoV-2 is a beta coronavirus, having similarity with SARS- CoV virus, in binding with human ACE2 receptor and spike glycoprotein for viral entry [3]. Tai et al, 2020, suggested that RBD fragment (from amino acid residues 331 to 524 of spike protein) in SARS-CoV-2 strongly binds with to human ACE2 (hACE2) and as well as bat ACE2 (bACE2) receptors. Thus, this spike protein fragment can block the entry of SARS-CoV-2 and SARS-CoV into their respective hACE2-expressing cells, resulting in that it may serve as a viral attachment inhibitor against SARS-CoV-2 and SARS-CoV infection.

Every coronavirus contains four structural proteins, for example spike (S), envelope (E), membrane (M), and nucleocapsid (N) proteins. Among them, S protein is the most important protein which controls the biological processes such as viral attachment, fusion and entry into the host cell. As a result, it can be considered as a target for development of antibodies, entry inhibitors and vaccines, similar to SARS-CoV infection [4] [5]. The S protein facilitates viral entry into human host cells by first binding to a host receptor (ACE2) through the receptor- binding domain (RBD) and then fusing with the viral and host membranes. But SARS-CoV-2 spike protein is 10-20 times more likely to bind with ACE2 on human cells,

compared to that of spike protein from the SARS-CoV infection (occurred in 2002). This may enable SARS-CoV-2 to spread more easily than SARS-CoV infection. Despite very much similarity (76.5%) in sequence [6] and structure between the spike proteins of two viruses, three different antibodies against the 2002 SARS virus (SARS-CoV) cannot be successfully administered against SARS- CoV-2, which is popularly known as COVID-19.

ACE2 is a functional receptor for both SARS-CoV and SARS-CoV-2. For SARS-CoV infection, ACE2 is confirmed as receptor in both in vitro and in vivo studies [6]. Similarly, Zhou et al, 2020, has confirmed that SARS-CoV-2 uses ACE2 as a cellular entry receptor in human host [7]. ACE2 enzyme having catalytic activity in maturation of angiotensin, a peptide hormone. ACE2 is a type I membrane protein, expressed in many extrapulmonary tissues including heart, kidney, endothelium, and intestine. ACE2-expressing epithelial cells have high levels of multiple viral replication related genes, [8], signifying that the ACE2-expressing epithelial cells facilitate coronaviral replication in the lung [9]. The presence of ACE2 receptor in other tissues, can explain the cause of kidney damage, heart failure and liver damage in COVID-19 infected patients. Different activities of ACE2 protein and inhibitory role of spike protein, are depicted in Figure 1.

Schematic diagram for structure of transmembrane ACE2 protein is shown in Figure 2. There are three topological domains in ACE2 such as extracellular domain (from 18- 740), a transmembrane helical domain (from 741-761) and cytoplasmic domain (from 762-805) [10].

Several potential therapeutic approaches have been investigated for the treatment of SARS-CoV- 2 infection such as protein-based vaccine design, blocking of ACE2 receptor and effect of phytochemicals on spike protein binding with its ACE2 receptor. Among the various therapeutic strategies that have been proposed for the treatment of SARS-Co V 2 treatment, drug designing with phytochemicals is a well-known method. Several phytochemicals for example, *Ocimum sanctum* extract on main protease protein [11], 5,7,3',4'-tetrahydroxy-2'-(3,3-dimethylallyl) isoflavone from *Psorothamnus arborescens* on 3-chymotrypsin-like protease [12] and curcumin, brazilin, and galangin from *Curcuma sp.*, *Citrus sp.*, *Alpinia galanga*, and *Caesalpinia sappan* on both SARS-CoV-2 protease and RBD (Receptor Binding Domain) of spike glycoprotein (RBD-S) [13] and Belachinal, Macaflavanone E & Vibsanol B on envelop protein [14] are analyzed with the help of molecular docking and molecular dynamics simulation studies. In the last study, hesperidin, one of common flavonoids in Citrus sp., has selected as potent inhibitor with the lowest docking score for protein receptors resulting the highest affinity to bind the receptors.

C Wu et al, 2020 [15] have used homology modeling technique to model 18 viral proteins and 2 human target proteins. They have screened potential small-molecule compounds from a ZINC Drug Database (2924 compounds) and a small in-house database of traditional Chinese medicine and natural products (including reported common anti-viral components from traditional Chinese medicine) and derivatives (1066 compounds) to identify small molecules to treat SARS-CoV-2 infection. Hesperidin molecule, which is known for its anti-inflammatory, anti-oxidant effect, is obtained from *Citrus aurantium*. This is the only compound that could bind the interface between Spike and ACE2. So, they have suggested hesperidin

may disrupt the interaction of ACE2 with RBD. But during molecular docking analysis, they used PDB file SARS_CoV-2_Spike_RBD_homo_Hesperidin considering RBD-S (PDB ID: 6LXT) and PD-ACE2 (PDB ID: 6VWI).

Since, both SARS-CoV-2 spike protein and SARS-CoV spike protein, can bind with human host ACE2 receptor protein, literatures are searched for binding inhibitor for EC 3.4.17.23 - angiotensin-converting enzyme 2 (ACE2) as virus-host interaction in PubMed [16]. Ho et al, 2007 [17] showed that, 1,3,8-trihydroxy-6-methylanthraquinone (emodin) blocks interaction between the SARS corona virus spike protein and its receptor angiotensin-converting enzyme 2, 94.12% inhibition at 0.05 mM. 1,8, dihydroxy-3-carboxyl-9,10-anthraquinone (rhein) and anthraquinone exhibit slight inhibition in spike protein binding. But, 5,7-dihydroxyflavone (chrysin) can act as a weak inhibitor.

To study the effect of Indian phytochemicals on spike protein fragment, molecular docking study is used for spike glycoprotein fragment with human ACE2 receptor. Bound structure of spike glycoprotein with human ACE2 receptor is considered here as target molecule for treatment of COVID-19.

Some phytochemicals, which have been reported earlier as spike protein inhibitor for SARS [16],[17] are considered here as small molecules for protein -ligand molecular docking study. These phytochemicals are present in Indian medicinal plants. Name, source, chemical class and structures of phytochemicals e.g. hesperidin, emodin, anthraquinone, rhein and chrysin are enlisted in Table 1 and Figure 3. This information is collected from IMPPAT: Indian Medicinal Plants, Phytochemistry And Therapeutics a curated database [18].

Table 1 Phytochemicals and their Indian medicinal plant sources

Indian medicinal plant	**Phytochemical identifier**	**Phytochemical name**	**Chemical class of ptyochemicals**
Valeriana Jatamansi	CID:10621	Hesperidin	Flavonoid glycoside
Cassia Angustifolia	CID:6780	Anthraquinone	Anthraquinone
Oroxylum	CID:6780	Anthraquinone	Anthraquinone
Indicum			
Cassia Angustifolia	CID:10168	Rhein	Anthraquinone derivative
Oroxylum Indicum	CID:5281607	Chrysin	Flavone
Rheum Emodi	CID:3220	Emodin	Anthraquinone derivative

Ii. Methodology

1. Protein molecular modeling of spike protein fragment

3D structure of RBD fragment (from amino acid residues 331 to 524 of spike protein) in SARS- CoV-2 is considered in this paper as responsible fragment for strongly binding with to human ACE2 (hACE2) receptor protein. Before molecular docking analysis, following steps are performed with the primary sequence of spike protein fragment.

1. 1 Retrieval of protein sequence for spike protein fragment

The protein sequence of spike glycoprotein from Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV 2) containing 193 amino acid residues from positions 331 to 524 is retrieved from GenBank database (https://www.ncbi.nlm.nih.gov/protein/QHR63250.2) in FASTA format and considered as spike protein fragment in this study.

1. 2 3D structure homology modeling and validation of modeled structure

In modeling 3D structure of the spike protein fragment by using sequence homology approach, first of all sequence alignment method is used. Thus, the best matching PDB structures of other proteins are identified with the help of following steps:

1. 2. 1 Template Search for the spike protein fragment

Template search with Blast [19] and HHBlits [20] has been performed against the SWISS- MODEL template library (SMTL, last update: 2020-04-08, last included PDB release: 2020-04- 03).

The target sequence is searched with BLAST [19] against the primary amino acid sequence contained in the SMTL. A total of 63 templates are found.

An initial HHblits profile has been built using the procedure outlined in [20], followed by 1 iteration of HHblits against NR20. The obtained profile has then been searched against all profiles of the SMTL. A total of 110 templates are found.

1. 2. 2 Template Selection

For each identified template, the template's quality has been predicted from features of the target-template alignment. The templates with the highest quality have then been selected for model building.

1. 2. 3 Model Building

Models are built based on the target-template alignment using ProMod3. Coordinates which are conserved between the target and the template are copied from the template to the model. Insertions and deletions are remodeled using a fragment library. Side chains are then rebuilt. Finally, the geometry of the resulting model is regularized by using a force field. In case loop modelling with ProMod3 fails, an alternative model is built with PROMOD-II [21].

1. 2. 4 Model Quality Estimation

The best model among obtained models by using two types of selection methods are estimated by QMEAN4 scores [22] and Ramachandran plot [23, 24]. The global and per-residue model quality has been assessed using the QMEAN scoring function [22] for both models while the Ramachandran plot for two models are obtained using PROCHECK [23] and MolProbity [24]. Evaluation of backbone conformation of protein molecule is assayed by Ramachandran plot dividing the percentage of amino acid residues of the model in the allowed and disallowed regions [23, 24].

2. Molecular docking between spike protein fragment and human ACE2 receptor

Molecular docking studies between spike protein fragment and human ACE2 receptor are performed using ClusPro [25]. In ClusPro 2.2 web server [25], Cluster scores for lowest binding energy prediction are calculated using the formula-E = 0.40E_{rep} + -0.40E_{att} + 600E_{elec} + 1.00E_{DARS}. Here, repulsive, attractive, electrostatic as well as interactions extracted from the decoys as the reference state, are considered for structure-based pairwise potential calculation in docking [26].

3. Molecular docking study of phytochemicals from Indian medical plants

Docking of bound structure (spike protein fragment and its receptor ACE2) with phytochemicals are carried with SWISSDOCK web server based on EADock DSS [27]. Many binding modes are generated in the vicinity of all target cavities (blind docking). Simultaneously, their CHARMM energies are estimated on a grid with CHARMM force field [28] on external computers from the Swiss Institute of Bioinformatics.

The binding modes with the most favourable energies are evaluated with FACTS [29] and are therefore clustered. Molecular complexes are ranked by the most favourable binding energies. Among those, we select the one structure representing the best binding mode for each phytochemical, based on an energy average value corresponding to the first five ranked structures. The most favourable clusters are visualized by the USCF Chimera software [30].

Iii. Results

1. Protein molecular modeling of spike protein fragment

Gene Bank accession number for SARS-CoV-2 S is QHR63250.2, LOCUS QHR63250, Accession MN996527.1is used for protein molecular modeling of spike protein fragment.

Primary amino acid sequence of spike protein fragment (331 to 524) is as follows

NATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSF
VIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRK
SNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELL HAPATV

2. Primary and Secondary structure analysis

Primary structure analysis shows that this spike protein fragment SARS-CoV-2 have 193 amino acid residues. Secondary structure analysis with PDBsum [31], shows that this protein fragment contains 3 sheets, 1 beta hairpin, 2 beta bulges, 9 strands, 6 helices, 1 helix-helix interaction, 14 beta turns, 4 gamma turns and 2 disulfide bonds.

Table 2 Templates for 3D structure of the spike protein fragment

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Coverage	Description
6lzg.1.B	100.00	monomer	-	HHblits	X-ray	2.50Å	0.62	1.00	SARS-CoV-2 Spike receptor- binding domain
6m0j.1.B	100.00	monomer	-	HHblits	X-ray	2.45Å	0.62	1.00	SARS-CoV-2 receptor- binding domain
6w41.1.C	100.00	monomer	-	HHblits	X-ray	3.08Å	0.62	1.00	Spike glycoprotein receptor binding domain
6m17.1.C	100.00	monomer	-	HHblits	EM	NA	0.62	1.00	SARS-coV-2 Receptor Binding Domain

3. 3D structure modeling and validation

3D structure of the spike protein fragment has been modeled by using SWISSMODEL [32] server. Template 6lzg.1.B is selected for modeling protein with the sequence identity 100% and coverage 100% compared to the other two templates (Table 2) for modeling.

The SWISS-MODEL template library (SMTL version 2020-04-08, PDB release 2020-04-03) is searched with BLAST [19] and HHBlits [20] for evolutionary related structures matching the target sequence in Table 2. Overall, 101 templates are found.

Modelled structure obtained from SWISSMODEL server [32] has -2.87 QMEAN score, shown in Figure 4 (a). QMEAN value is intended as a linear combination of four statistical potential terms and transformed to a Z score relating it to high resolution X-ray structures of similar size. Higher Z score is related to more favorable model. Ramachandran plots are drawn for this model by using two web servers e.g. Molprobity [24] and PDBsum [31], are shown in Figure 4 (b) and 4 (c). For this model the overall average value of G -factors is -0.18 which is not unusual for dihedral angles and main-chain covalent forces. The value of G-factors provides a measure of how unusual or out-of -the ordinary, a property is. From MolProbity version 4.4 [33] it is calculated that, modelled structure has 94.44% residues in favored regions, 0.56% residues in outlier region and 3.18% in rotamer outlier region. Ramachandran plot statistics from PDBsum [34] for modelled structure of spike protein fragment, 136 (86.1%) residues in most favored regions [A, B, L], 21 (13.3%) residues in additional allowed regions [a, b, l, p], 1 (0.6%) residues in generously allowed regions [~a, ~b, ~l, ~p] and 0 (0.0%) residue in disallowed regions [X, X].

4. Molecular docking between spike protein fragment and human ACE2 receptor

Human ACE2 receptor (PDB ID 1R42) [35] is considered as receptor protein for molecular docking study of spike protein fragment with its receptor in human host.

By using ClusPro [25] web server, docking structure of A chain of human ACE2 receptor, binds with spike protein fragment, is obtained. SARS CoV2 spike protein binds with human ACE2 receptor protein with binding energy -779.8 Kcal/mole. A conformational change occurs in ACE2 receptor protein after binding with spike protein fragment (Figure 5).

Amino acids present in distorted site of ACE2 are ASP136, ASN 137, PRO 138, GLN139 and interacting amino acids of spike protein fragment are GLN 403, LYS 451 and ASP 416 (Figure 6).

Bound structure of SARS CoV2 spike protein fragment with ACE2 receptor protein is considered as therapeutic target for SARS-CoV2 treatment.

5. Molecular docking study of phytochemicals from Indian medical plants

5.1 Spike protein binding with ACE2 in presence of hesperidin

In Figure 7, spike protein fragment (331 to 524) is shown in red colour, hesperidin molecule in stick model and human ACE2 is shown in blue colour. Hesperidin binds with spike protein fragment and its receptor ACE2 with binding energy -8.99 Kcal/mole. This docked structure is stabilized by two H binding (shown in Figure with green lines) at PHE 457 of spike protein with O7 atom of hesperidin, with bond length 2.618Å and H atom of small molecule hesperidin with O atom of GLU 455 of spike protein fragment with a distance 2.067 Å. Hesperidin binds at ASN 63, ALA 71, LYS 74 and SER 44 amino acids of ACE2.

5.2 Spike protein binding with ACE2 in presence of emodin

The phytochemical emodin, obtained from *Rheum emodi* or Himalayan rhubarb [36], binds with spike protein fragment and its receptor human ACE2 protein [37], at the same cleft (Figure 8), same to that of hesperidin. But binding energy is less for emodin binding (-6.19 Kcal/mole) compared to that of that of hesperidin (-8.99 Kcal/mole).

5.3 Spike protein binding with ACE2 in presence of anthraquinone

Though anthraquinone can bind with bound structure of spike protein fragment and its receptor ACE2 molecule, with releasing binding energy -6.15 Kcal/mole, but the binding site of this phytochemical is totally different from that of hesperidin and emodin (Figure 9).

5.4 Rhein binding with bound spike protein and ACE2 receptor protein

The phytochemical rhein binds with docked structure of spike fragmented protein and human ACE2 receptor with Δ G value -8.73 Kcal/mole. But the binding site of this chemical totally different from earlier substances (Figure 10). Rhein can bound with only spike protein fragment. It has no interaction with human ACE2 receptor protein molecule.

5.5 Chrysin binding with bound spike protein and ACE2 receptor protein

Chrysin binds with the spike protein fragment and its ACE2 receptor with binding energy -6.87 Kcal/mole (Figure 11). This phytochemical binding site is almost similar with that of spike protein fragment

molecule and its receptor. A conformational change occurs in ACE2 receptor molecule after spike protein fragment binding. Chrysin binding cleft is nearly located to that site as shown in Figure 12.

Energy parameters of bound structure of phytochemicals with spike protein fragment and ACE2 receptor are shown in Table 3.

Table 3 Energy parameters of bound structure of phytochemicals

Name of phytochemicals	Energy/ Simple fitness	FullFitness	ΔGvdw	ΔG (Kcal/mole)
Hesperidin	59.4535	-2147.5469	-52.5659	-8.99
Emodin	19.599	-2301.9927	-23.3637	-6.19
Anthraquinone	17.7976	-2234.7346	-21.5368	-6.15
Rhein	36.5174	-2310.458	-107.401	-8.73
Chrysin	15.8545	-2266.9272	-31.1973	-6.87

Considering the lowest binding energy, the phytochemical hesperidin is considered as most suitable ligand for target molecule, which is formed by binding with spike protein fragment and its human host ACE2 receptor.

In six docking structures interacting amino acids of ACE2 receptor and spike protein fragment are summarized in Table 4.

Table 4 Interacting amino acids in docking structures

Docking structure	Interacting amino acids of ACE2 receptor	Interacting amino acids of spike protein fragment
Spike protein fragment with ACE2	ASP136, ASN 137, PRO 138, GLN 139	GLN 403, LYS 451, ASP 416
Hesperidin binding with spike protein and ACE2	ASN 63, ALA71, LYS 74, SER 44	VAL 472, GLY 474, GLY 471, PHE 475, GLU 473
Emodin binding with spike protein and ACE2	ALA 71, ASP 67, LYS 74	VAL 472, GLY 474, ALA 464, ASN 448
Anthraquinone binding with spike protein and ACE2	SER 105, ASN 103, GLN 102, LEU 100, PHE 28	No interacting amino acids
Rhein binding with spike protein and ACE2	No interacting amino acids	SER 388, VAL 401, THR 333, ASN 332, ASN 353
Chrysin binding with spike protein and ACE2	THR 129, ILE 126, THR 125	ARG 443, SER 448, ASN 449, TYR 410, PHE 486, TYR 484, THR 487, ASN 488, LYS 406

Considering the docking structures and interacting amino acids of both ACE2 receptor and spike protein fragment, chrysin can act as most competent inhibitor for spike protein binding with ACE2 receptor.

Iv. Discussion

With primary sequence from 331 to 524 of Spike protein, a homology modelled structure is built using SWISSMODEL, with template 6lzg.1.B with sequence identity 100.00%, coverage 100%. This modelled structure is validated by Ramachandran plot. This stable spike protein fragment is used for binding with human host ACE2 receptor protein by molecular docking study.

Binding site of spike protein fragment with its ACE2 receptor lying in binding surface with interacting amino acids ASP 136, ASN 137, PRO 138 and GLN 139, forms a beta hairpin motif in between two β strands secondary structure (results from PDBsum). This binding site is present in extracellular domain of ACE2 protein.

Bound structure of SARS CoV2 spike protein fragment with ACE2 receptor protein is considered as therapeutic target for SARS-CoV2 treatment and screened with Indian phytochemicals e.g. hesperidin, emodin, anthraquinone, rhein and chrysin by molecular docking study.

Among them, hesperidin binds with ASN 63, ALA71, LYS 74 of H2 helix and SER 44 of H1 helix of human ACE2 receptor protein. Similarly, emodin binding amino acids i.e. ALA 71, ASP67 and LYS 74 are present on H2 helix of ACE2 molecule. Phytochemical anthraquinone interact with spike protein fragment and rhein has no interacting amino acids with ACE2 receptor. So, both of them are not considered as therapeutic agents in COVID treatment. But the interacting amino acids after chrysin binding with target molecule i.e. THR 129, ILE 126 and THR 125, all are positioned on H5 helix of ACE2 receptor protein. The above mentioned β hair pin motif, which is a supersecondary structure, consists of an antiparallel β sheet formed by sequential segments of polypeptide chain that are connected by a tight reverse turn. Here in ACE2 protein, this antiparallel β sheet is flanked by, in both sides with H5 and H6 helices of that protein. Globular protein ACE2 consists largely of approximately straight runs of secondary structure joined by stretches of polypeptide that abruptly change direction. Such β hair pin motif occurs at protein surface. Here the β hair pin motif contents ASN134, Pro 135, ASP136 and ASN 137 amino acids. Proline is present as second residue, since it can easily achieve the required conformation. This conformation has been changed due to binding of spike protein fragment. Distorted structure of ACE2 contains ASP136, ASN 137, PRO 138, GLN 139 amino acids, which can interact with GLN 403, LYS 451, ASP 416 of spike protein of SARS-CoV 2. FASTA alignment for PDB entry of spike protein fragment with 26 PDB entries, having at least a 30% sequence identity or E values < 0.001, has been executed in PDBsum [31] (results are not shown here). Among three interacting amino acids of spoke protein fragments GLN 403 and ASP 416 are well conserved among all sequences. But LYS 451 is conserved among SARS-CoV2 spike proteins and differed with ARG in SARS-CoV spike proteins. Though arginine is a positively charged, polar amino acid, it can be substituted with the other positively charged amino acid lysine. But a change from arginine to lysine is not always neutral. Arginine contains a complex guanidium group on its positively charged

sidechain and shows a geometry and charge distribution for ideal binding with negatively charged amino acid residues. It can also form multiple hydrogen bonds. But lysine also can interact with negatively charged amino acid residues, but it is more limited in the number of hydrogen bonds it can form [38].

In case of hesperidin, interacting amino acids of spike protein fragment e.g. VAL 472, GLY 474, GLY 471, PHE 475, GLU 473 are well conserved among PDB structures of SARS CoV-2 spike proteins (6m0j:E, 6lzg:B, 6w41:C, 6m17:E and 6vw1:E). But these residues are not present in structures of SARS-CoV spike glycoprotein structures (2dd8:S, 2ghw:A, 1q4z:A, 1t7g:A, 1xjp:A, 5xlr:A, 5x58:A, 6nb6:A, 6nb7:A, 6acc:A, 6acd:A , 6acg:A , 6acj:A, 6ack:A , 2ghv:E, 6waq:D, 5wrg:A, 3bgf:S, 5x5b:A, 6crw:A, 6crx:B, 6crz:A and 6cs0:A).

For emodin phytochemical, other than the interacting amino acids of spike protein fragment, ALA 464 and ASN 448 are also conserved in five SARS CoV-2 spike protein PDB structures and changed in SARS-CoV spike glycoprotein structures.

When chrysin binds with the target molecule, the sequences of interacting amino acids e.g. PHE 486, TYR 484 and THR 487 are same in five SARS CoV-2 spike proteins and changes to SARS- CoV spike glycoprotein structures.

Hesperidin is a major flavonoid compound, present in orange and lemon fruits. Orange juice contains 470-761 mg/l of hesperidin [38]. These phytochemical exhibits various medicinal uses. According to oral toxicity study of hesperidin, it can be concluded that this phytochemical can be safely used in herbal formulations with its LD_{50} value is more than 2000mg/kg [39]. This flavanone glycoside, has a long medicinal history in both Indian and Chinese herbal medications [40]. This phytochemical alone or in combination with chemicals, often be used in various diseases.

Emodin is a polyphenol found in the roots, leaves and bark of several plants including aloe vera, cascara, rhubarb, senna etc. In traditional medicine, emodin has been used for cardiovascular diseases, osteoporosis. It has been suggested earlier that emodin can inhibit Inflenza Avirus replication and influenza viral pneumonia [41] via several cell signaling pathways.

Chrysin a natural flavonoid, is commonly found in propolis and honey and traditionally used in herbal medicine. As reported earlier, chrysin can act as inhibitor during enterovirus 71 (EV71) growth and replication [42]. Similarly, Song et al, 2015 have described antiviral activity of chrysin against coxsackievirus B3 (CVB3) [43].

Considering the results obtained from molecular docking studies, phytochemicals hesperidin, emodin and chrysin can be recommended for the treatment of COVID-19, after in -silico mutagenesis study and experimental verification.

Declarations

Authors' contribution statement

A Basu: Conceptualization, Methodology, Investigation, Writing - Original draft preparation. **A Sarkar:** Data curation, Writing - Reviewing and Editing, **U Maulik:** Supervision, Writing - Reviewing and Editing.

COMPETING INTERESTS STATEMENT

The author(s) declare no competing interests.

References

1. Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C. L., Abiona, O., ... & McLellan, J. S. (2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science, 367(6483), 1260-1263.
2. Walls, A. C., Park, Y. J., Tortorici, M. A., Wall, A., McGuire, A. T., & Veesler, D. (2020). Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein.
3. Tai, W., He, L., Zhang, X. et al. Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine. Cell Mol Immunol (2020). https://doi.org/10.1038/s41423-020-0400-4
4. Du, L., He, Y., Zhou, Y., Liu, S., Zheng, B. J. & Jiang, S. The spike protein of SARS-CoV-a target for vaccine and therapeutic development. Nat. Rev. Microbiol. 7, 226–236 (2009).
5. Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, Somasundaran M, Sullivan JL, Luzuriaga K, Greenough TC, Choe H, Farzan M (2003) Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature 426:450–454
6. Xu X, Chen P, Wang J, Feng J, Zhou H, Li X, Zhong W, Hao P (2020) Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its Spike protein for risk of human transmission. Sci China Life Sci. https://doi.org/10.1007/s11427-020-1637-5
7. Li, , Zhang, C., Sui, J., Kuhn, J. H., Moore, M. J., Luo, S., ... & Murakami, A. (2005). Receptor and viral determinants of SARS- coronavirus adaptation to human ACE2. The EMBO journal, 24(8), 1634-1643.
8. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li B, Huang CL, Chen HD, Chen J, Luo Y, Guo H, Jiang RD, Liu MQ, Chen Y, Shen XR, Wang X, Zheng XS, Zhao K, Chen QJ, Deng F, Liu LL, Yan B, Zhan FX, Wang YY, Xiao GF, Shi ZL (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. https://doi.org/10.1038/s41586-020-2012-7
9. Zhang, H., Penninger, J.M., Li, Y. et al. Angiotensin-converting enzyme 2 (ACE2) as a SARS-CoV-2 receptor: molecular mechanisms and potential therapeutic target. Intensive Care Med 46, 586–590 (2020). https://doi.org/10.1007/s00134-020-05985-9
10. https://www.uniprot.org/uniprot/Q9BYF1

11. Varshney, Krishna Kumar and Varshney, Megha and Nath, Bishamber, Molecular Modeling of Isolated Phytochemicals from Ocimum Sanctum Towards Exploring Potential Inhibitors of SARS Coronavirus Main Protease and Papain-Like Protease to Treat COVID-19 (March 14, 2020). Available at SSRN: https://ssrn.com/abstract=3554371

12. Muhammad Tahir ul Qamar, Safar M. Alqahtani, Mubarak A. Alamri, Ling-Ling Chen, Structural basis of SARS-CoV-2 3CLpro and anti-COVID-19 drug discovery from medicinal plants, Journal of Pharmaceutical Analysis, 2020, ISSN 2095-1779, https://doi.org/10.1016/j.jpha.2020.03.009. (http://www.sciencedirect.com/science/article/pii/S2095177920301271)
13. Utomo, R. Y., & Meiyanto, E. (2020). Revealing the Potency of Citrus and Galangal Constituents to Halt SARS-CoV-2 Infection.
14. Manoj Kumar Gupta, Sarojamma Vemula, Ravindra Donde, Gayatri Gouda, Lambodar Behera & Ramakrishna Vadde (2020) In-silico approaches to detect inhibitors of the human severe acute respiratory syndrome coronavirus envelope protein ion channel, Journal of Biomolecular Structure and Dynamics, DOI: 1080/07391102.2020.1751300
15. Wu, C., Liu, Y., Yang, Y., Zhang, P., Zhong, W., Wang, Y., ... & Zheng, M. (2020). Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. Acta Pharmaceutica Sinica
16. https://ncbi.nlm.nih.gov/
17. Ho, T. Y., Wu, S. L., Chen, J. C., Li, C. C., & Hsiang, C. Y. (2007). Emodin blocks the SARS coronavirus spike protein and angiotensin-converting enzyme 2 interaction. Antiviral research, 74(2), 92-101.
18. Mohanraj, K., Karthikeyan, B. S., Vivek-Ananth, R. P., Chand, R. B., Aparna, S. R., Mangalapandi, P., & Samal, A. (2018). IMPPAT: A curated database of I ndian M edicinal P lants, P hytochemistry A nd T herapeutics. Scientific reports, 8(1), 1-17.
19. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L. BLAST+: architecture and applications. BMC Bioinformatics 10, 421-430 (2009).
20. Remmert, M., Biegert, A., Hauser, A., Söding, J. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat Methods 9, 173-175 (2012)
21. Guex, N., Peitsch, M.C., Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis 30, S162-S173 (2009).
22. Studer, G., Rempfer, C., Waterhouse, A.M., Gumienny, G., Haas, J., Schwede, T. QMEANDisCo - distance constraints applied on model quality estimation. Bioinformatics 36, 1765-1771 (2020).
23. Laskowski R A, MacArthur M W, Thornton J M (2001). PROCHECK: validation of protein structure coordinates, in International Tables of Crystallography, Volume F. Crystallography of Biological Macromolecules, eds. Rossmann M G & Arnold E, Dordrecht, Kluwer Academic Publishers, The Netherlands, pp. 722-725.
24. Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., ... & Richardson, D. C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta

Crystallographica Section D: Biological Crystallography, 66(1), 12-21.

25. Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, Beglov D, Vajda S. The ClusPro web server for protein-protein docking. Nature Protocols. 2017 Feb;12(2):255-278.
26. Kozakov D, Beglov D, Bohnuud T, Mottarella S, Xia B, Hall DR, Vajda, S. How good is automated protein docking? Proteins: Structure, Function, and Bioinformatics. 2013 Dec; 81(12):2159-66.
27. Grosdidier, A., Zoete, V., & Michielin, O. (2011). SwissDock, a protein-small molecule docking web service based on EADock DSS. Nucleic acids research, 39(suppl_2), W270-W277.
28. Grosdidier, A., Zoete, V., & Michielin, O. (2011). Fast docking using the CHARMM force field with EADock DSS. Journal of computational chemistry, 32(10), 2149-2159.
29. Zoete, V., Grosdidier, A., Cuendet, M., & Michielin, O. (2010). Use of the FACTS solvation model for protein–ligand docking calculations. Application to EADock. Journal of Molecular Recognition, 23(5), 457-461.
30. Pettersen, F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T.
31. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. Journal of computational chemistry, 25(13), 1605-1612.
32. Laskowski, R. A. (2009). PDBsum new things. Nucleic acids research, 37(suppl_1), D355-D359.

32. Schwede, T., Kopp, J., Guex, N., & Peitsch, M. C. (2003). SWISS-MODEL: an automated protein homology-modeling server. Nucleic acids research, 31(13), 3381-3385.
33. Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., ... & Jain, S. (2018). MolProbity: More and better reference data for improved all- atom structure validation. Protein Science, 27(1), 293-315.
34. C. Lovell, I.W. Davis, W.B. Arendall III, P.I.W. de Bakker, J.M. Word, M.G. Prisant, J.S. Richardson and D.C. Richardson (2002) Structure validation by Calpha geometry: phi,psi and Cbeta deviation. Proteins: Structure, Function & Genetics. 50: 437-450.
35. Towler, P., Staker, B., Prasad, S. G., Menon, S., Tang, J., Parsons, T., ... & Patane, M. A. (2004). ACE2 X-ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis. Journal of Biological Chemistry, 279(17), 17996-18007.
36. Malik, mushtaq & Bhat, Dr. Showkat & Fatima, Bilquees & Ahmad, Sheikh Bilal & Sidiqui, S. & Shrivastava, Purnima. (2016). Rheum Emodi As Valuable Medicinal Plant. International Journal of General Medicine And Pharmacy. 5. 35-44.
37. Izhaki, I. (2002). Emodin–a secondary metabolite with multiple ecological functions in higher plants. New Phytologist, 155(2), 205-217.
38. Betts, M. J., & Russell, R. B. (2003). Amino acid properties and consequences of substitutions. Bioinformatics for geneticists, 317,
39. Rakesh Sharma, Chapter 59 - Polyphenols in Health and Disease: Practice and Mechanisms of Benefits, Editor(s): Ronald Ross Watson, Victor R. Preedy, Sherma Zibadi, Polyphenols in Human Health and Disease, Academic Press, 2014, Pages 757-778, ISBN 9780123984562,

https://doi.org/10.1016/B978-0-12- 398456-2.00059-1. (http://www.sciencedirect.com/science/article/pii/B9780123984562000591)

40. Anand A. Zanwar, Sachin L. Badole, Pankaj S. Shende, Mahabaleshwar V. Hegde, Subhash L. Bodhankar, Chapter 76 - Cardiovascular Effects of Hesperidin: A Flavanone Glycoside, Editor(s): Ronald Ross Watson, Victor R. Preedy, Sherma Zibadi, Polyphenols in Human Health and Disease, Academic Press, 2014, Pages 989-992, ISBN 9780123984562, https://doi.org/10.1016/B978-0-12-398456-2.00076-1. (http://www.sciencedirect.com/science/article/pii/B9780123984562000761)
41. Dai, J. P., Wang, Q. W., Su, Y., Gu, L. M., Zhao, Y., Chen, X. X., ... & Li, K. S. (2017). Emodin inhibition of influenza A virus replication and influenza viral pneumonia via the Nrf2, TLR4, p38/JNK and NF-kappaB pathways. Molecules, 22(10),
42. Wang, J., Zhang, T., Du, J., Cui, S., Yang, F., & Jin, Q. (2014). Anti-enterovirus 71 effects of chrysin and its phosphate ester. PLoS One, 9(3).
43. Song, J. H., Kwon, B. E., Jang, H., Kang, H., Cho, , Park, K., ... & Kim, H. (2015). Antiviral activity of chrysin derivatives against coxsackievirus B3 in vitro and in vivo. Biomolecules & therapeutics, 23(5), 465.

Figures

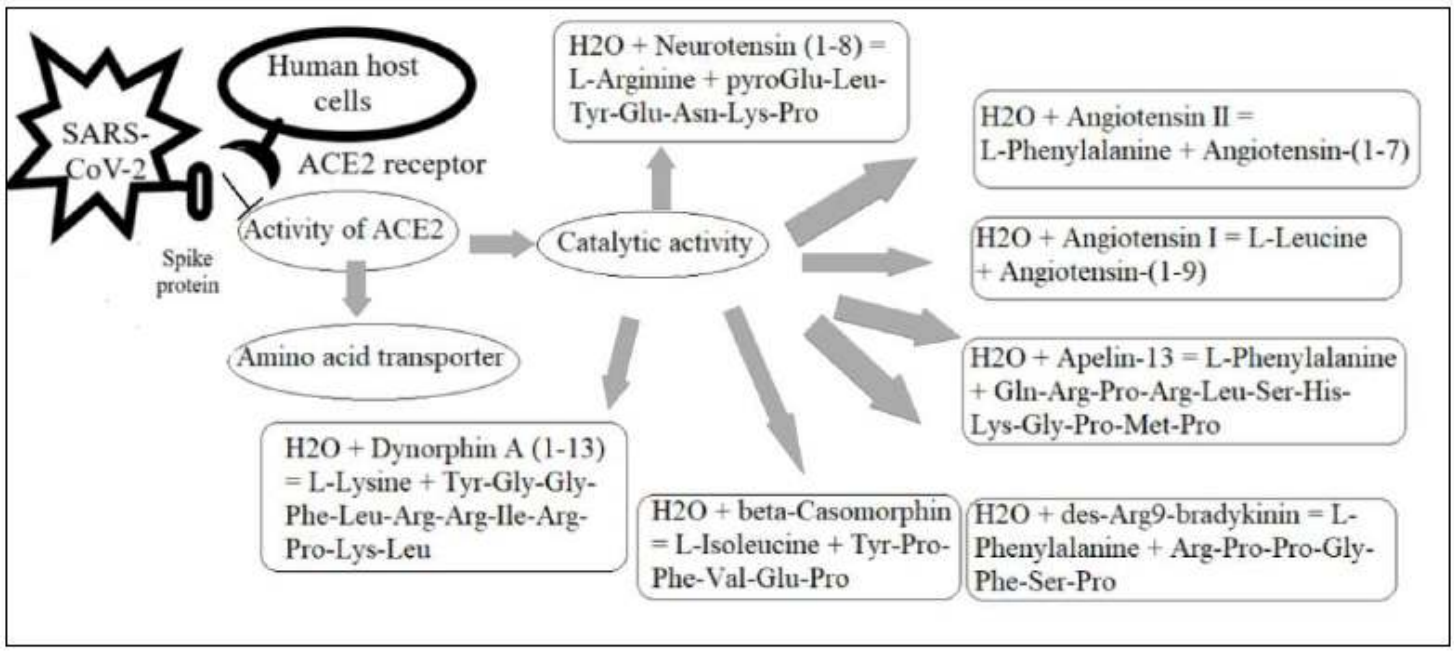

Figure 1

Different activities of ACE2 protein and inhibitory role of spike protein

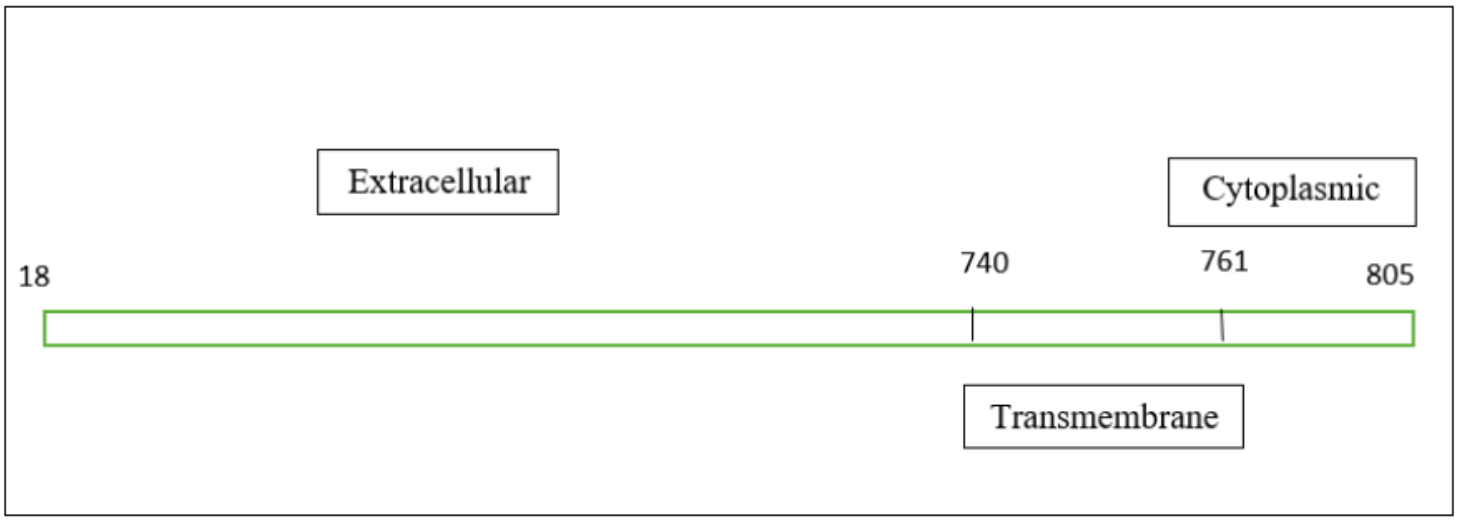

Figure 2

Schematic diagram of ACE2 protein

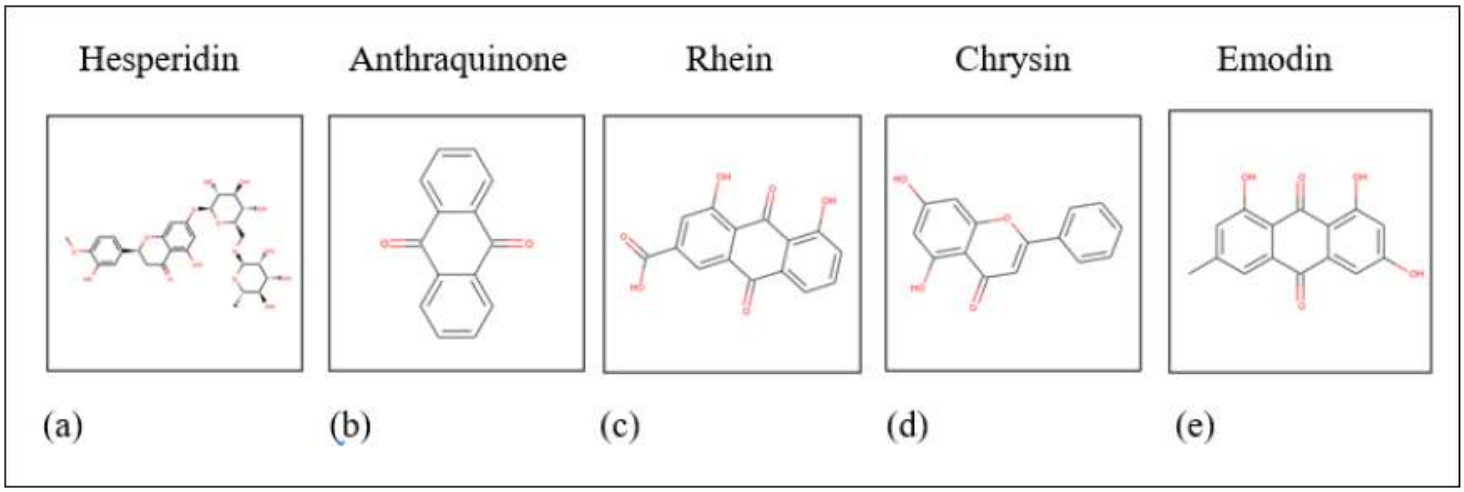

Figure 3

Structures of phytochemicals

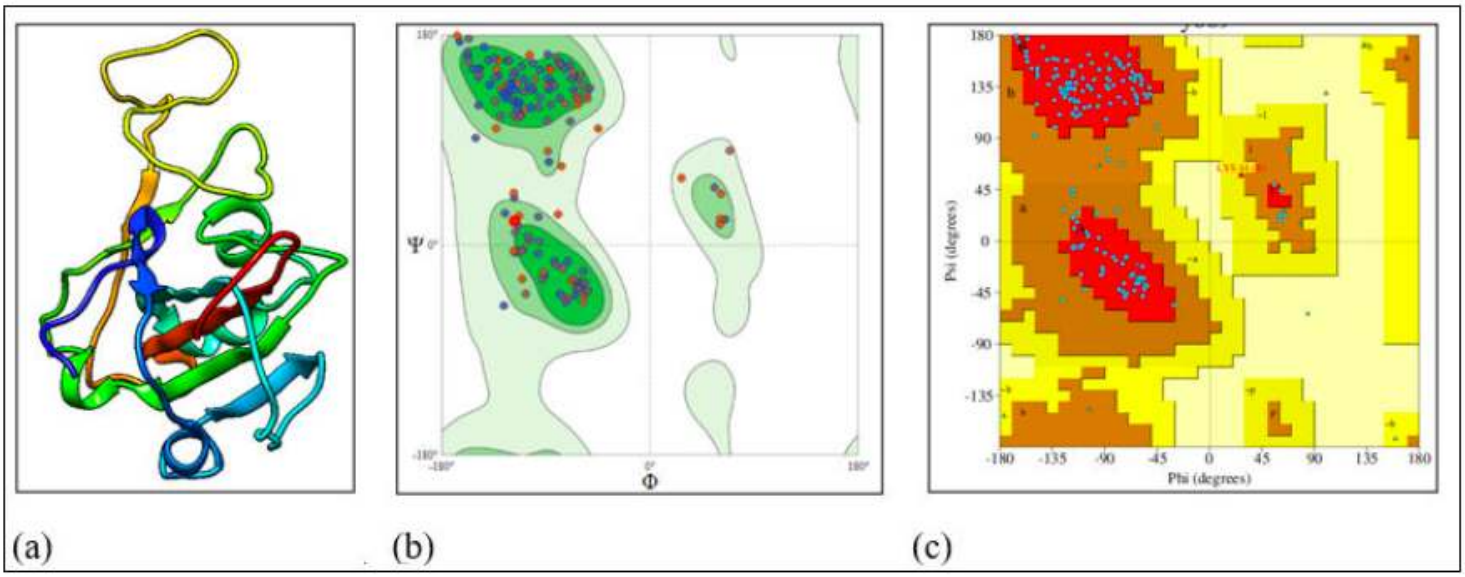

Figure 4

(a) 3D structure for spike protein fragment Rmachandran plot (b) from MolProbity server (c) PROCHECK server

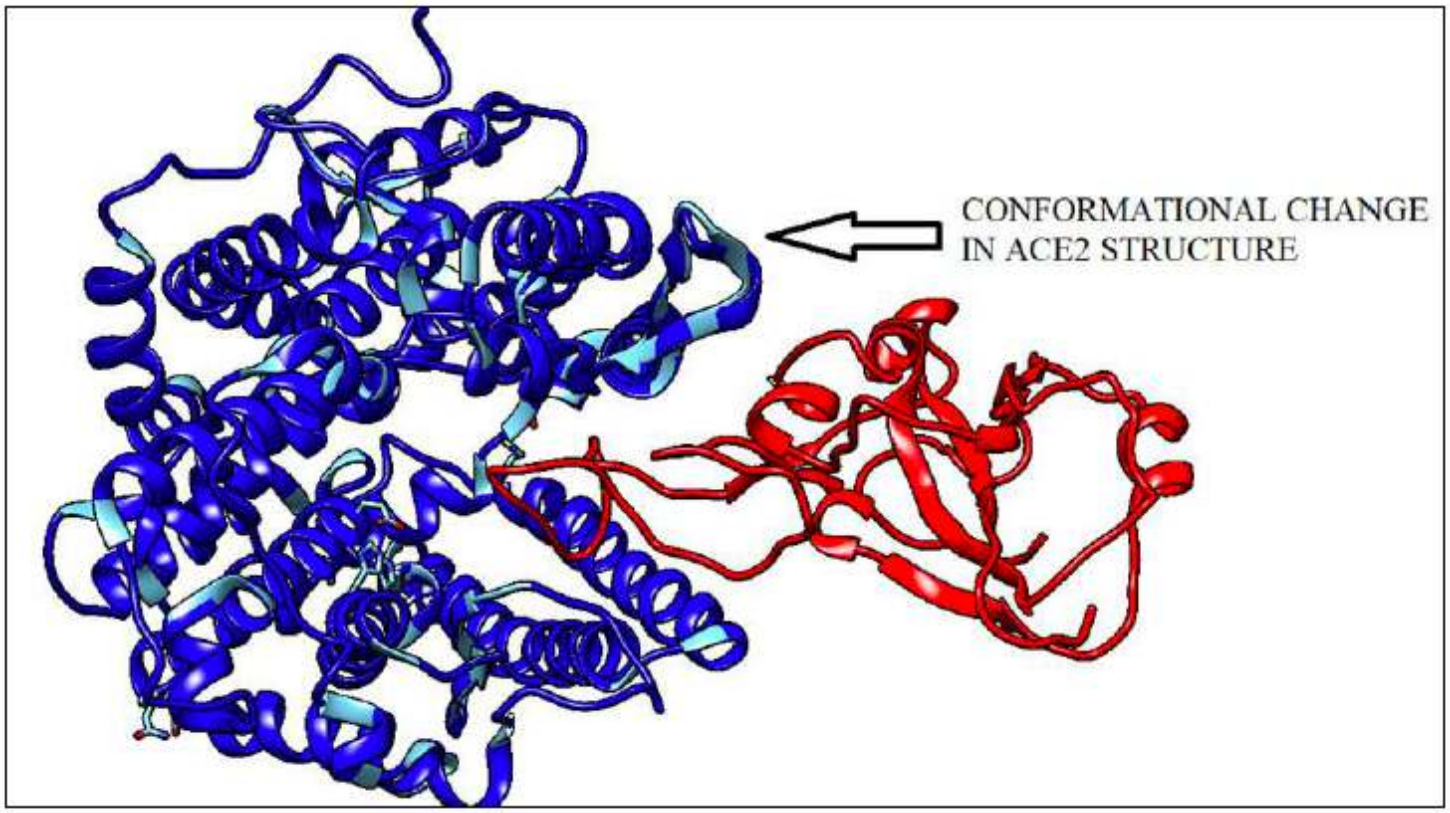

Figure 5

SARS CoV2 spike protein binding with human ACE2 receptor protein

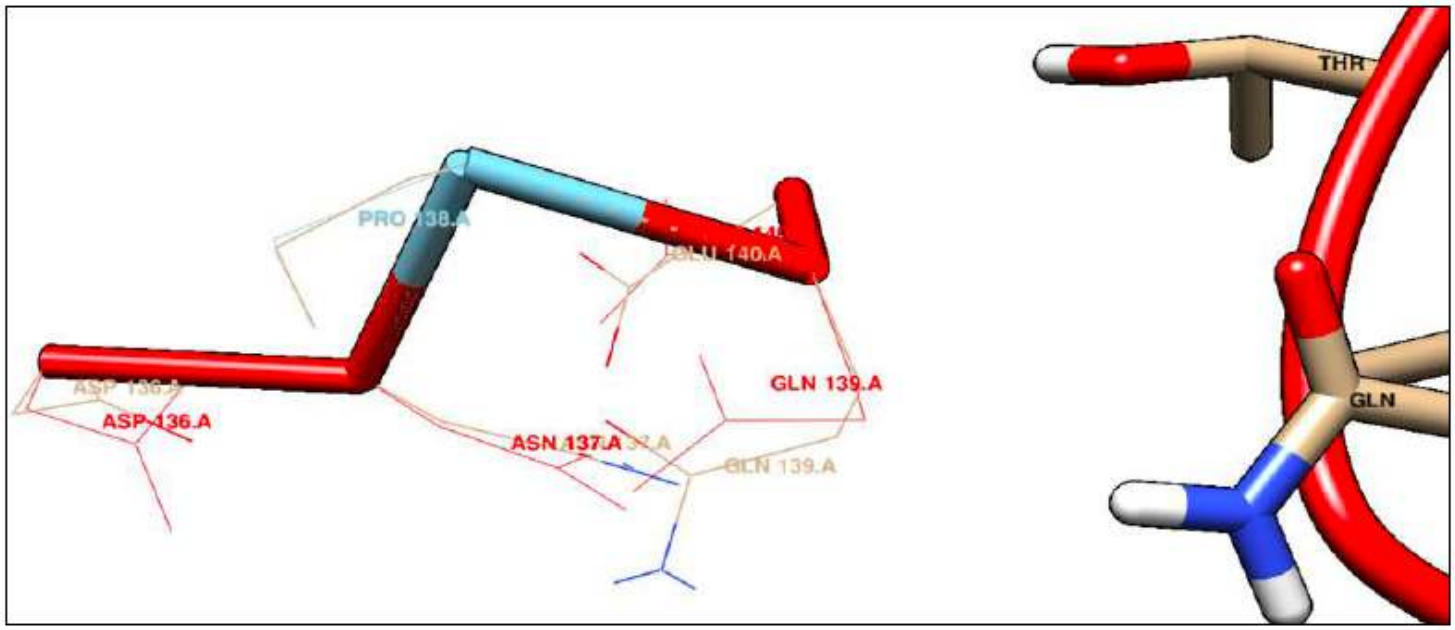

Figure 6

Distorted amino acids after spike protein binding in ACE2 receptor

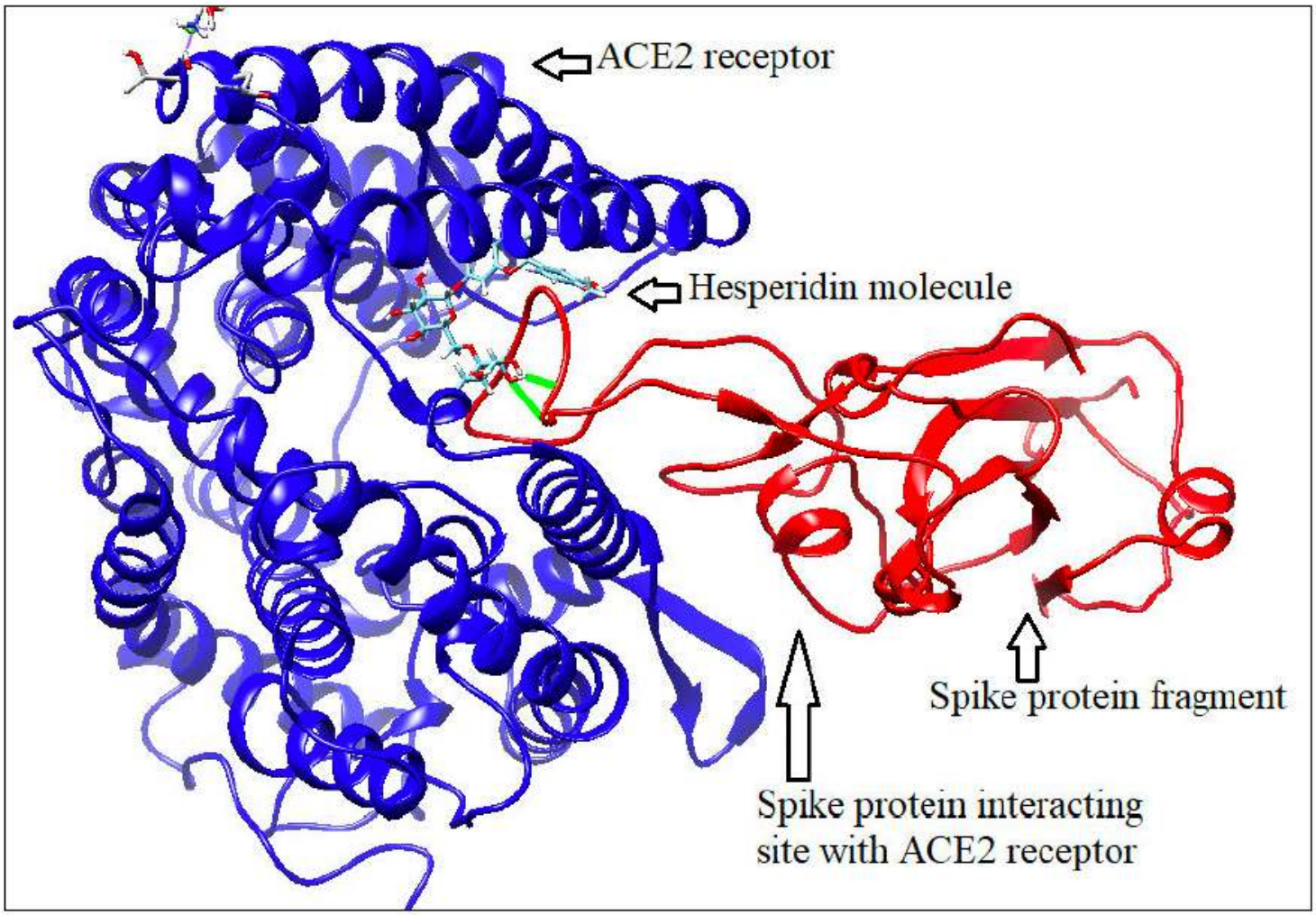

Figure 7

Spike protein binding with ACE2 in presence of hesperidin

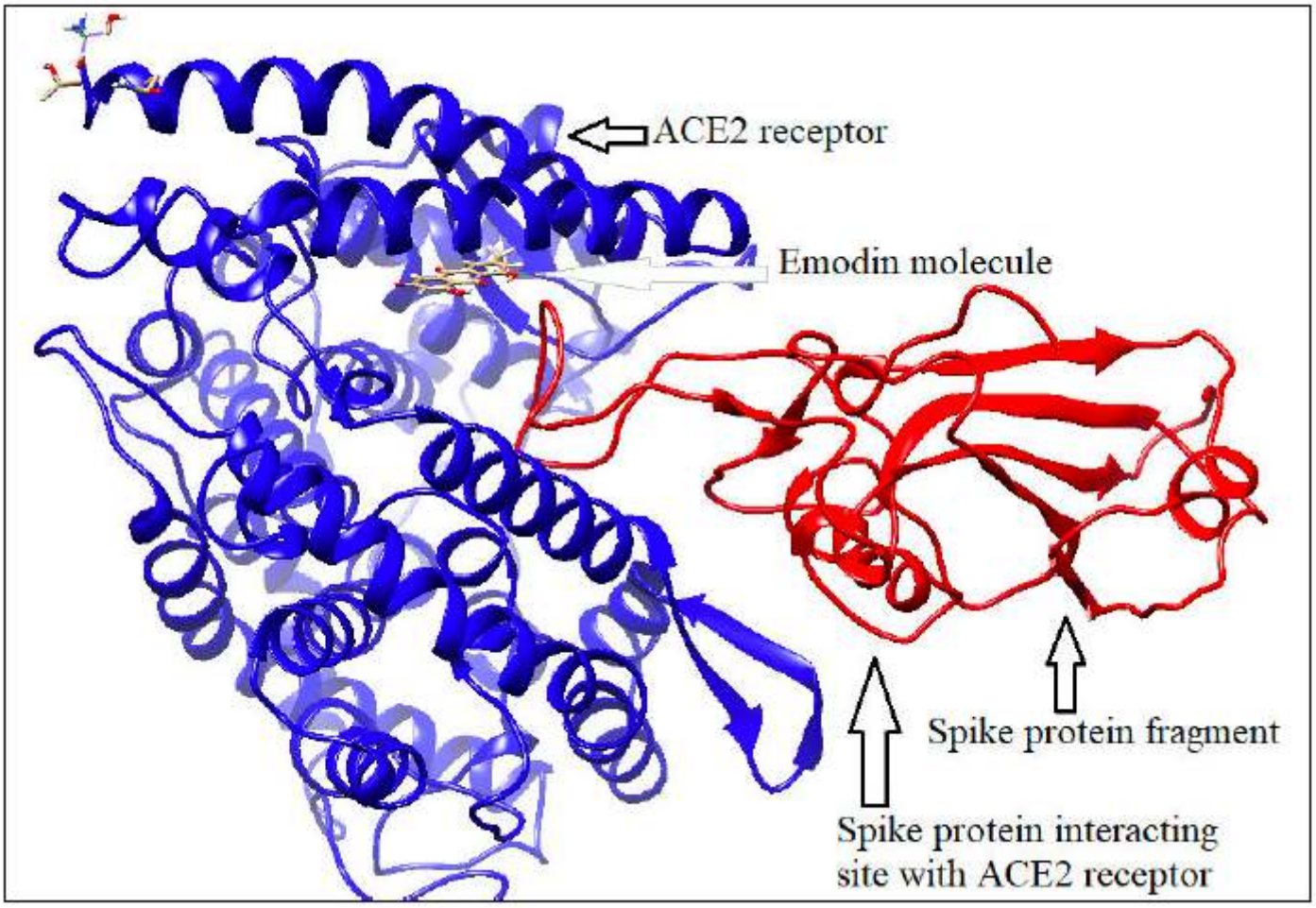

Figure 8

Spike protein binding with ACE2 in presence of emodin

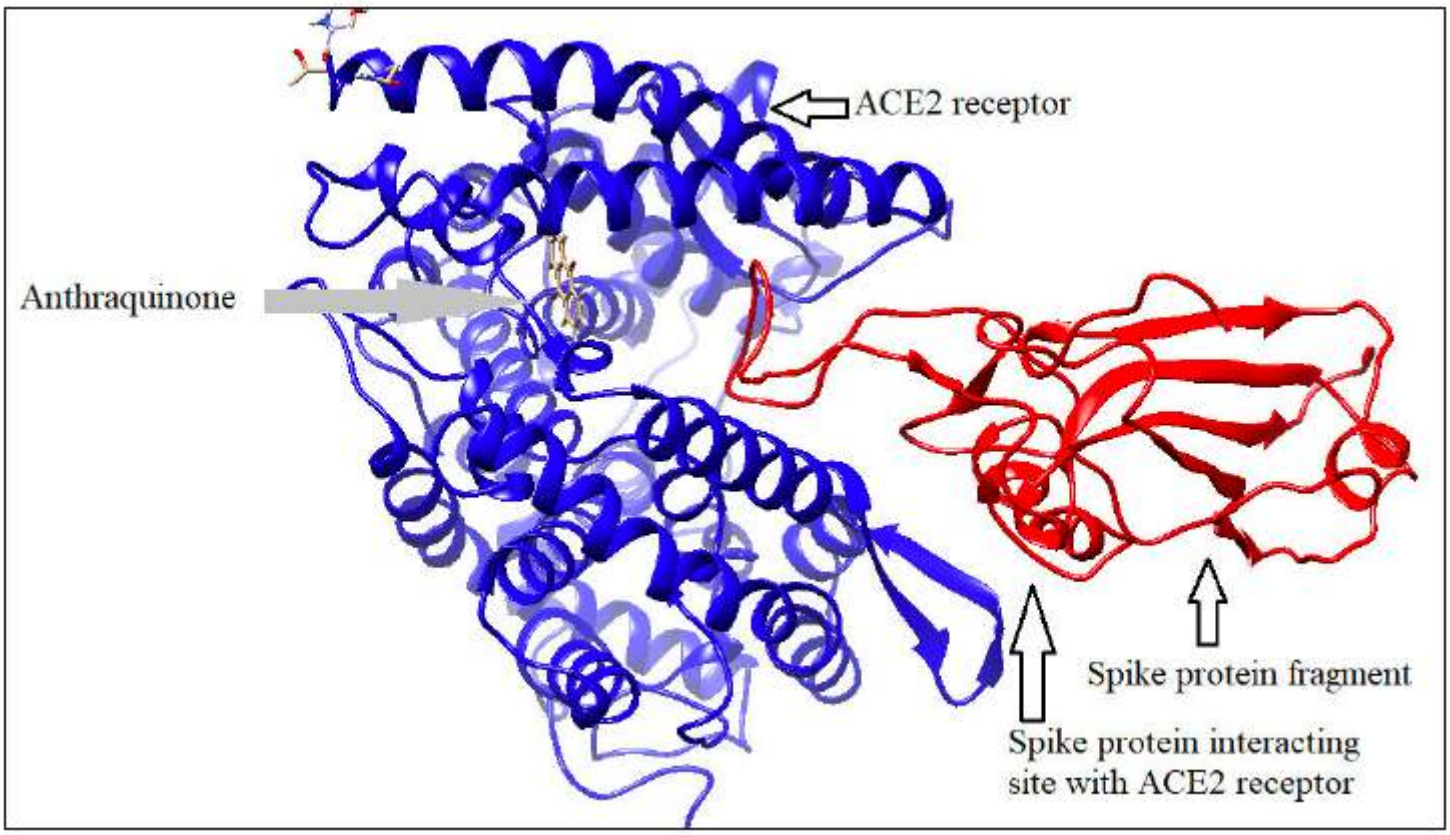

Figure 9

Spike protein binding with ACE2 in presence of anthraquinone

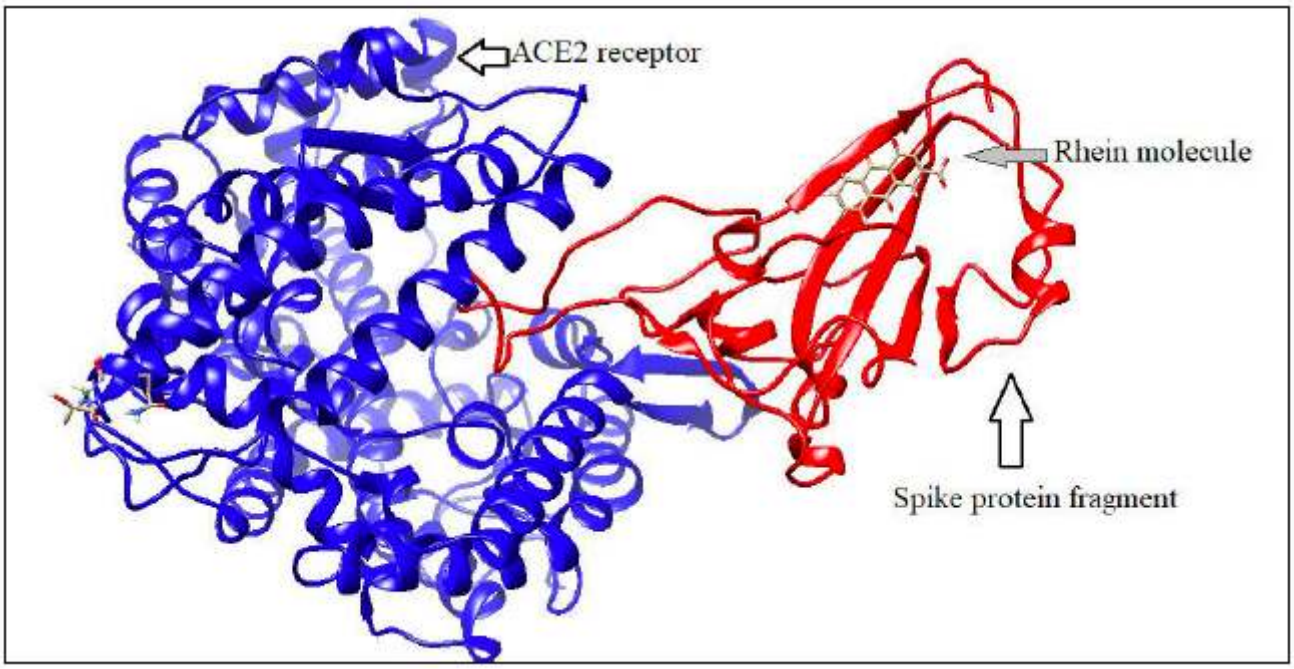

Figure 10

Rhein binding with bound spike protein and ACE2 receptor protein

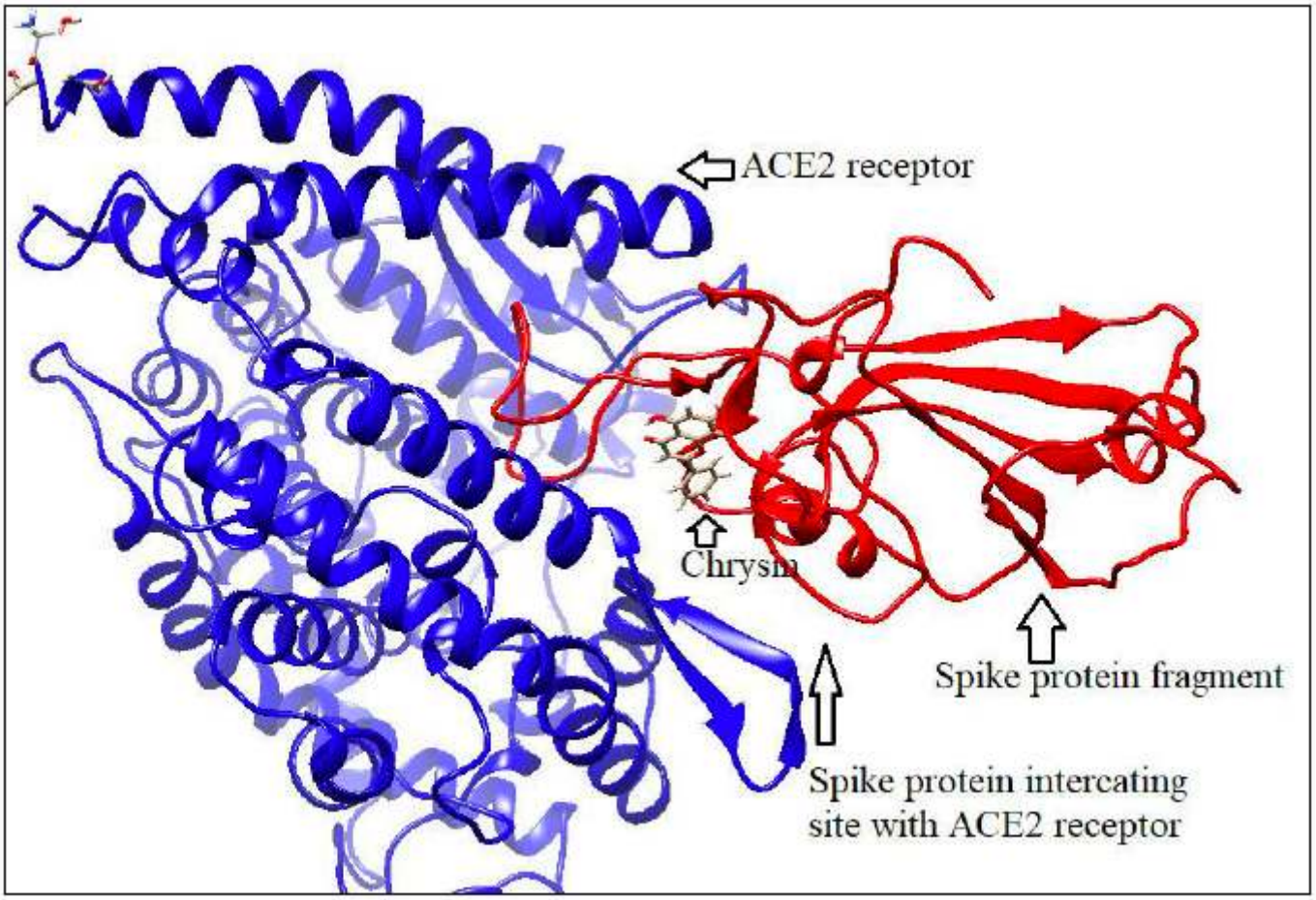

Figure 11

Chrysin binding with bound spike protein and ACE2 receptor protein

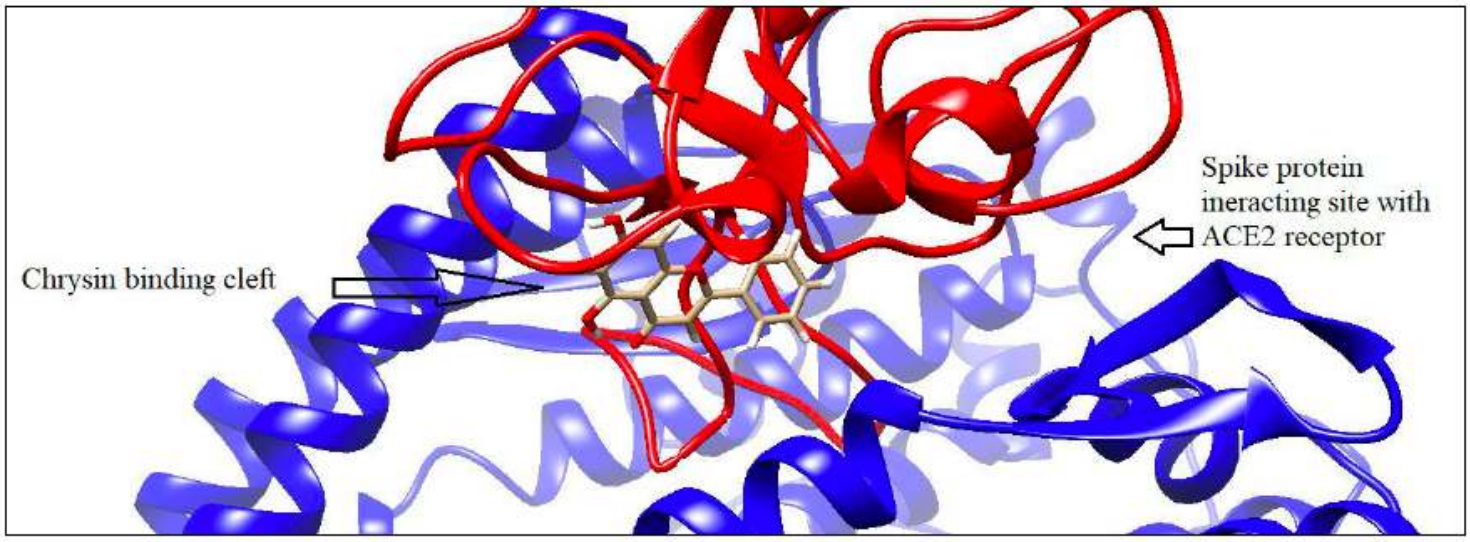

Figure 12

Chrysin binding cleft