Structure-based bioactive phytochemical design from Ayurvedic formulations towards Spike and M^{pro} druggable targets of SARS-CoV-2

Anju Choorakottayil Pushkaran[1], Prajeesh Nath EN[2], Shantikumar V Nair[1],

Rammanohar Puthiyedath[2], C Gopi Mohan[1*]

[1]*Computational Biology and Bioinformatics Lab,*

Center for Nanosciences and Molecular Medicine,

Amrita Vishwa Vidyapeetham, Kochi 682041, Kerala, India.

[2]*Amrita School of Ayurveda,*

Amrita Vishwa Vidyapeetham, Kollam 690525, Kerala, India.

*Corresponding author Dr. C Gopi Mohan
E-mail id: cgmohan@aims.amrita.edu;
cgopimohan@yahoo.com

Abstract

The present COVID-19 global crisis invoked different disciplines of the biomedical research community to address the contagious human to human viral transmission and infection severity. Traditional or de novo drug discovery approach is a very time consuming process and will conflict with the urgency to discover new anti-viral drugs for combating the present global pandemic. Modern anti-viral drugs are not very effective and show resistance with serious adverse effects. Thus, identifying bioactive natural ingredients (phytochemical) from different medicinal plants or Ayurvedic formulations provides an effective alternative therapy for SARS-CoV-2 viral infections. We performed structure-based phytochemical design studies involving bioactive phytochemicals from medicinal plants towards two key druggable targets, spike glycoprotein and main protease (M^{pro}) of SARS-CoV-2. Phyllaemblicin class of phytocompounds showed better binding affinity towards both these SARS-CoV-2 targets and its binding mode revealed interactions with critical amino acid residues at its active sites. Also, we have successfully shown that the SARS-CoV-2 spike glycoprotein interaction towards human ACE2 receptor as its port of human cellular entry was blocked due to conformational variations in ACE2 receptor recognition by the binding of the phytocompound, Phyllaemblicin C at the ACE2 binding domain of spike protein. Our study shows that the Phyllaemblicin class of phytochemicals can be a potential dual inhibitor of spike and M^{pro} proteins of SARS-CoV-2 and could be promising for the treatment of COVID-19.

Keywords: COVID-19; SARS-CoV-2 spike protein; Main protease; phytochemical; molecular docking; Ayurveda

Abbreviations: Main protease (M^{pro}); three dimensional (3D); protein data bank (PDB); angiotensin converting enzyme 2 (ACE2); receptor binding domain (RBD); root mean square deviations (RMSD); Absorption, Distribution, Metabolism, Excretion and Toxicity (ADMET).

1. Introduction

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection commonly referred as novel coronavirus disease 2019 (COVID-19) has declared pandemic by the World Health Organization. This contagious viral infection around the world reminds us the vulnerability of the human race towards this invisible enemy. The emergence of different zoonotic viral diseases earlier, including SARS, Ebola, and Middle East Respiratory Syndrome were highly infectious and resulted in millions of death [1]. However, these viruses are successfully contained without causing a global pandemic. In contrast, the present COVID-19 is highly contagious and more than 21 million people were infected globally with a mortality rate of ~4%. Different neurological symptoms and respiratory manifestations associated with the COVID-19 patients were revealed. The most prevalent symptom is the severe acute respiratory complications leading to ventilator support; other neurological signs include: headache, fever, nausea, pneumonia, loss of smell, unconsciousness, ataxia, epilepsy, neuralgia, cerebrovascular and musculoskeletal disease [2, 3]. In some patients, this viral infection causes acute encephalopathy and acute hemorrhagic necrotizing encephalopathy, which might lead to blood brain barrier breakdown [4].

The bioactive phytochemical compounds are well known in pharmaceutical industries for developing treatment for various health conditions like inflammation, cancer, infectious diseases *etc*. Phytochemicals isolated from various natural sources has been of great interest in the development of anti-viral treatment for Chikungunya, HIV, Influenza, Dengue, and SARS [5]. Since the outbreak of SARS-CoV-1 in 2003, several phytochemicals from flora and fauna were tested for its anti-viral activity, which include alkaloids, flavones, flavonols, fatty acids, tannins and terpenes. Jo *et al.* [6] reported key pharmacophoric chemical features such as hydrophobic groups, electron donor (hydroxyl) and carbohydrate moieties for anti-SARS-CoV-1 infection apart from lipophilic and

hydrophilic groups seen in other anti-viral drugs. Further, different computational and *in vitro* studies were reported recently demonstrating the anti-viral properties of bioactive phytochemicals from medicinally important plant sources like green tea, turmeric, gooseberry, basil *etc*. against SARS-CoV-2 [7-9]. It is to be noted at this juncture that the development of bioactive natural product is more feasible than specific vaccine design for this new virus despite heavy experimental measurements involving extraction, chemical complexity and diversity of natural compounds. Thus, computational strategy can be employed to identify bioactive compounds from natural sources by virtual screening technique in order to shorten the time period required for anti-viral drug discovery [6, 10, 11].

Structure-based drug design is playing an important role to understand the molecular mechanism of interaction at the atomic level of SARS-CoV-2 druggable proteins with the host cell receptor for the discovery of clinically efficacious drugs. This is possible due to the powerful atomic scale view of the three dimensional (3D) structure of different druggable proteins solved for SARS-CoV-2 using X-ray crystallography, NMR and cryo-EM techniques. At present, ~330 structures accounts for SARS-CoV-2 druggable targets in protein data bank (PDB) which includes non-structural proteins: Main protease (M^{pro}) [12], Papain-like protease (PLpro) [13], Helicase (Nsp13) [14], RNA-dependent RNA polymerase (RdRp) [15], Nsp14 (N-terminal exoribonuclease and C-terminal guanine-N7 methyl transferase, nsp10/14), Nsp15 (Uridylate-specific endoribonuclease), Nsp16 (2'-O-methyltransferase, nsp10/16), NSP10, and structural proteins: Spike protein bind with the cellular receptor angiotensin converting enzyme 2 (ACE2) [16], N (Nucleocapsid protein) [17] and is freely available to the scientific community in PDB.

Several studies have been reported recently involving computer-aided designing of inhibitors against SARS-CoV-2 [18-22]. Khan *et al.* performed molecular dynamics simulation studies using

marine natural products against main protease of SARS-CoV-2 [18]. The drug repurposing approach was adopted by Shah *et al.* for identifying SARS-CoV-2 potent inhibitors by molecular docking techniques [19]. Other research groups were also reported different computational approaches using phytochemicals to identify novel SARS-CoV-2 inhibitors [20-22]. The present work constitutes the structure-based phytochemical design studies on two key pharmacological SARS-CoV-2 targets: (i) the receptor binding domain (RBD) of spike glycoprotein and (ii) Main protease (M^{pro}). The spike glycoprotein is present on the surface of the SARS-CoV-2 which interacts directly with the peptidase domain of human ACE-2 receptor found on the surface of the human epithelial cells and forms the fundamental key mechanism of the virus gain entry into the host living cell [16]. M^{pro} is an enzyme required for SARS-CoV-2 viral replication where it is involved in the proteolytic processing of viral polyproteins [12].

The main objective of the present work is to identify the best phytochemical from five different medicinal herbs for inhibiting the two key SARS-CoV-2 druggable targets and further to understand its molecular mechanism of inhibition. These include (i) *Phyllanthus emblica* (Amalaki), (ii) *Cinnamomum zeylanica* (Tvak), (iii) *Embelia ribes* (Vidanga), (iv) *Curcuma longa* (Haridra) and (v) *Justicia adhatoda* (Vasa) herbs which are well documented in classical Ayurvedic texts. Out of these five herbs, *P. emblica* is ascribed with immunomodulatory (rasayana) and anti-ageing (vayasthapana) properties [23], *C. zeylanica* is indicated for the management of oropharyngeal (kantharuk) and respiratory afflictions (kasa) [24]. *E. ribes* [25] and *C. longa* [26] are attributed with microbicidal activity (krimighna). *J. adhatoda* is specifically indicated in the management of respiratory illnesses (kasa and svasa) [27]. In addition, the present computational modeling revealed for the first time the influence of key phytochemical in blocking the human ACE-2 receptor recognition site towards the binding of SARS-CoV-2 spike protein by docking and electrostatic recognition studies and which would further aid in the computer-aided phytochemical design.

2. Materials and Methods:

2.1 COVID-19 Protein target selection

The two important molecular targets of novel SARS-CoV-2 contributing to its virulence are the spike protein and M^{pro}. The X-ray crystallographic coordinate of the RBD domain of spike protein in complex with human cell receptor ACE2 with a resolution of 2.45 Å was retrieved from PDB having an accession code '6M0J' [16]. Initially, the crystal structure was optimized using the Protein Preparation Wizard module of the Schrödinger software. The ACE2 protein and the water molecules were removed from this complex structure. Further, the RBD of SARS-CoV-2 spike protein structure was pre-processed by adding hydrogen atoms, removing the alternate conformations of the amino acids and also added the missing atoms. The geometry of the pre-processed structure is optimized followed by energy minimization by employing OPLS2005 force field. Similarly, the X-ray crystal structure of M^{pro} of SARS-CoV-2 (PDB code - 6LU7) [12] was optimized with the same protocol as mentioned above. Finally, these two key druggable targets of SARS-CoV-2 were used for molecular docking studies to understand the molecular mechanism of inhibition using these Ayurvedic active ingredients (phytochemical) to suppress the severity of virus infections.

2.2 Ligand preparation:

The bioactive Ayurvedic ingredients (phytochemical): Procyanidin A2, Procyanidine B1, Cinnamtannin B1 from Twak (*Cinnamomum zeylanica*); Phyllaemblicin B, Phyllaemblicin C from Amalaki (*Phyllanthus emblicalinn*); Germacrone from Haridra (*Curcuma longa*); Embelin from Vidanga (*Embelica ribes*) and Vasicine from Vasa (*Justicia adhatoda*) were used to study their anti-viral activities targeting SARS-CoV-2 viral proteins. The chemical structure of these bioactive

phytochemical compounds was retrieved from PubChem database. These ligand structures were inspected and corrected for bond lengths, angles and missing hydrogen atoms were added. The geometry of the bioactive phytochemical structures was optimized by employing the MMFF94 force field using the LigPrep module of Schrödinger software. These optimized conformations of the phytochemical structures were used further for our molecular docking studies.

2.3 Molecular docking of the phytochemicals towards SARS-CoV-2 targets

Initially, the molecular docking grid box was generated using the Glide grid generation module of Schrödinger software by defining the reported active sites of both spike protein and M^{pro} crystal structure. The amino acid residues, N487, K417, Q493, Y505, Y449, T500, N501, G446, Y449, Y489, N487 and G502 present inthe RBD of spike protein are reported to be involved in direct hydrogen bonding interactions with the host receptor ACE2 [16]. Therefore, the docking grid box for the spike protein was generated by keeping the centroid of these residues as the center of the grid box. The grid box generated for spike protein was enclosing all these interface amino acid residues important for binding to the host cell receptor. Jin *et al*. has identified a small molecule inhibitor against SARS-CoV-2 M^{pro} protein and the key residues involved in the ligand binding areY54, H41, S46, M49, D187, Q189, M49, F185, Q192, T190, A191, P168, H164, C145, M165, H163, H172, G143, and L167[12]. These M^{pro} residues were set as the center of the docking grid box. Further, using the GLIDE module of Schrödinger software [28], docking of the phytochemicals were performed at the active sites of both spike and M^{pro} SARS-CoV-2 protein targets.

Glide molecular docking was performed in two sequential steps: (i) standard precision (SP) docking, ligands which are able to bind with protein are screened [29]; and (ii) extra precision (XP) docking step involve a strict scoring function and which will eliminate the false positive hits [30].

The docking protocol using the GLIDE module has been successfully implemented in our previous studies [31-34]. The bioactive phytochemicals from different medicinal plants were ranked based on their docking score (or binding affinity) against the SARS-CoV-2 protein targets. The best phytochemicals were further analyzed using PyMol and Discovery Studio visualizer to understand its molecular mechanism of action for SARS-CoV-2 infections.

2.4 Prediction of the pharmacokinetic properties of Bioactive Phytochemicals

The pharmacokinetic properties of eight phytochemicals selected in this study were predicted using the QikProp module of Schrödinger software. QikProp predicts the pharmacokinetic parameters which include: absorption, distribution, metabolism, excretion and toxicity of the compounds with high accuracy.

2.5 SARS-CoV-2 Spike-ACE2 complex binding studies in the presence of Phytochemical

The binding mode of spike protein with ACE2 protein was analyzed in the presence of highest binding affinity phytochemical obtained in our molecular docking studies. The crystal structure of the spike-ACE2 protein complexes solved by protein crystallography technique (PDB id: 6M0J) was used as reference for the present study. To evaluate the accuracy of the docking predictions, we first performed the molecular docking of Spike protein with the ACE2 receptor using ZDOCK software. ZDOCK is an interactive web-based server for performing the protein–protein docking simulations [35]. Different output poses generated by the ZDOCK server was compared with the binding mode of spike protein and ACE2 in the crystal structure complex. The best mode of binding of ACE2 with the SARS-CoV-2 spike protein was structurally compared by superimposing both structures (6M0J crystal structure with the docked structure) and evaluating its root mean square deviations (RMSD).

Further, the complex of SARS-CoV-2 spike protein and the best phytochemical was allowed to dock with the human ACE2 protein and the mode of binding of ACE2 with the ligand (phytochemical) free and ligand bound spike proteins was compared. Further, the binding affinity of these docked protein complexes was calculated using PRODIGY web server [36, 37].

3. Results and discussion

3.1 Molecular docking of bioactive phytochemicals towards SARS-CoV-2 Spike protein

The molecular mechanism of the SARS-CoV-2 entry into the human host epithelial cell via the interaction between the viral Spike protein and the ACE2 host cell receptor was studied extensively [16, 38, 39]. It was reported that the spike protein of SARS-CoV-2 is ten times potent than SARS-CoV [40]. This makes spike protein a prime druggable target in the present scenario for preventing the SARS-CoV-2 viral entry into the host cell. Many studies are undergoing in order to identify suitable inhibitors targeting the spike protein of SARS-CoV-2. The spike protein is composed of two subunits namely S1 and S2, among them the S1 subunit possess the receptor binding domain (RBD). The RBD domain residues directly recognize the ACE2 host cell receptor and the molecular interaction between them allows the human cell attachment and subsequent viral entry into the host cell. In the present study, we used X-ray crystal structure of the spike protein RBD domain of SARS-CoV-2 (6M0J) for the identification of bioactive phytochemical compounds as an entry inhibitor by preventing its direct interaction towards its natural ACE2 host cell receptor binding interface [16].

The selected bioactive phytochemical compounds (active ingredients) from different medicinal plants showing anti-viral activity include (i) Procyanidin A2, Procyanidine B1 and

Cinnamtannin B1 from Twak (*C. zeylanica*); (ii) Phyllaemblicin B and Phyllaemblicin C from Amalaki (*P.emblica*) ; (iii) Germacrone from Haridra (*C. longa*); Embelin from Vidanga (*E. ribes*) and Vasicine from Vasa (*J. adhatoda*) respectively (**Figure 1**). All these eight phytochemical compounds were docked into the active site of the RBD of SARS-CoV-2 spike glycoprotein. The

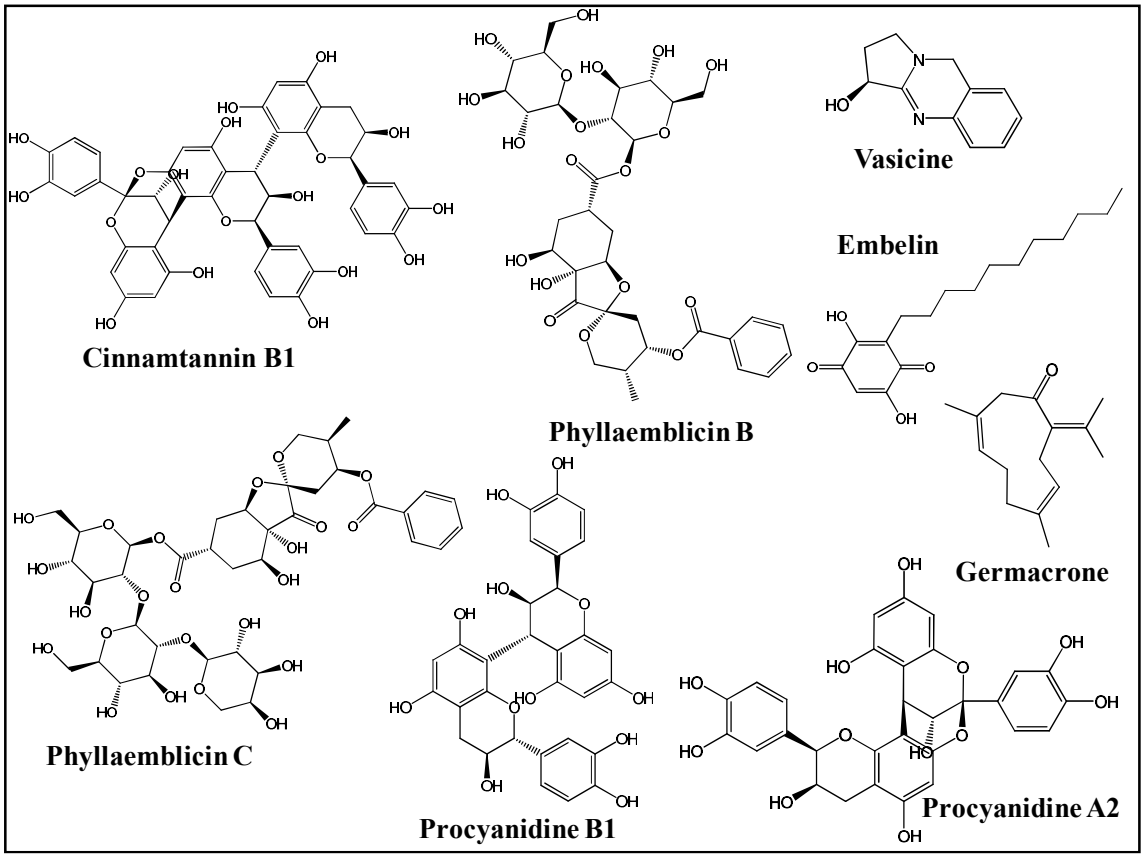

Figure 1 Chemical structure of bioactive phytochemicals from different medicinal plants.

molecular docking using Glide software revealed that the key ingredients present in Amalaki and Twak showed highest binding affinity towards the active sites of the RBD domain of SARS-CoV-2 spike protein by making strong interactions with the key residues involved in ACE2 binding. Phyllaemblicin C, a bioactive phytochemical present in Amalaki showed the highest binding affinity towards active site of RBD of spike protein with a docking score of -9.131 kcal/mol (**Table 1**). Cinnamtannin B1 and Phyllaemblicins B bioactive phytochemicals were obtained as the second and third best compounds with a docking score of -9.008 kcal/mol and -7.381 kcal/mol respectively, as

shown in **Table 1**. Further, the key ingredients Embelin, Vasicine and Germacrone present in other medicinal plants showed only weak binding affinity towards SARS-CoV-2 spike protein in the range of -3.0 to -4.0 kcal/mol.

Table 1: The binding affinity of phytochemicals (active ingredients) against spike protein and M^{pro} of SARS-CoV-2.

SARS-CoV-2 target	Phytochemicals	Binding affinity (kcal/mol)	Interacting amino acid residues with phytochemical
Spike protein (RBD)	Phyllaemblicin C	-9.131	**Y453, Q496, Q498, N501**, **Y449,** S494, **Q493, G496**, **T500, Y505**, F497, R403, Y495, **L455, Q493, K417**
	Cinnamtannin B1	-9.008	R403, **Y453, G502, G496**, **Q498**, S494, Q406**, Q493**, **Y505, N501**, Y495, Y449, **K417**, F497
	Phyllaemblicin B	-7.381	R403, Q409, **K417, Y453**, **Y505,** D405, Q406, Y495
	Procyanidine B1	-6.275	R403, Q409, **G496**, **N501**, E406, **Y505**, **K417**
	Procyanidine A2	-5.023	**Y453**, **G496**, **Q498**, E406, **Y449**, Y495, R403, **Y505**, **K417**
Main protease (M^{pro})	Phyllaemblicin C	-9.723	N142, **Q189**, E166, **H164**, **H163**, **P168**, **H41**, **L167**, **Q192**, **M165**, **C145**, **Y54**, **M49**, **Q189**
	Phyllaemblicin B	-9.151	N142, **Q189**, E166, **L167**, **T190**, **P168**, **C145**, **H41**, **M49**, **M165**
	Procyanidine B1	-9.128	E166, N142, **T190**, **H164**, F140, **Q189**, **H41**, **P168**, **M165**
	Cinnamtannin B1	-8.385	E166, N142, **T190**, **P168**, F140, **Q189**
	Procyanidine A2	-7.105	**H41**, **M165**, E166, T26, E166, **C145**

*The underlined residues are the reported key amino acids of Spike protein of SARS-CoV-2 for the direct binding with human ACE2 receptor [16, 41] and for M^{pro} protein these residues corresponds to its inhibitor binding [12].

The X-ray crystal structure of SARS-CoV-2 spike glycoprotein in complex with human ACE2 receptor has already been solved experimentally [16]. The contact residues of the RBD domain of spike protein interacting with ACE2 are K417, G446, Y449, Y453, L455, F456, A475, F486, N487, Y489, Q493, G496, Q498, T500, N501, G502, and Y505. Among these K417, G446, Y449, N487, Y489, Q493, T500, N501, G502 and Y505 residues made direct hydrogen bonding interactions with the human ACE2 residues [16]. Yi *et al.* demonstrated that the single amino acid substitutions at positions R439, K452, E484, T470, Q498, and N501 residues of SARS-CoV-2 spike protein resulted in a reduced binding affinity with the human ACE2 receptor protein. Also the substitutions at P499, Q493, F486, A475 and L455 residues has enhanced the binding of spike protein towards the human ACE2 receptor [41]. Hence, these residues play a critical role in binding with the host cell receptor and subsequent viral entry into the host cell.

The detailed molecular interaction analyses of these phytochemical compounds at the active site of SARS-CoV-2 spike protein revealed that most of the interactions are within the experimentally reported active site amino acid residues. Phyllaemblicin C, a bisabolane-type sesquiterpenoids isolated from Gooseberry was shown to bind efficiently with the RBD of Spike protein in comparison to other studied phytochemical compounds. This compound has several medicinal properties including anti-proliferative and anti-viral activities [42]. Phyllaemblicin C also made strong hydrogen bonding interactions with the Y453, G496, Q498, N501, Y449, S494, Q493, and G498 RBD residues. Apart from these Y505, G496, F497, R403, Y495, Y453, L455 and K417 RBD residues also made direct interactions with the Phyllaemblicin C, as shown in **Figure 2a and Table 1**. Most of these residues are reported to be important for binding of SARS-CoV-2 spike protein with the ACE2 host cell receptor. Therefore, Phyllaemblicin C would be able to block the binding of the spike protein with the host cell receptor ACE2 and thereby preventing the virus entry into the host cell.

The second best binding compound against the RBD domain of spike protein was Cinnamtannin B1. Cinnamtannin B1 is a type of proanthocyanidin which is commonly found in the cortex of cinnamon [43] and possesses anti-microbial, anti-platelet and anti-oxidant properties [43, 44]. The molecular interactions of Cinnamtannin B1 with the spike protein- RBD consisted of hydrogen bonding interactions with Y453, G502, G496, Q498, S494, Q406, G493, R403 and Y505 residues, electrostatic interactions with R403 and hydrophobic interactions with Y453 and Y505

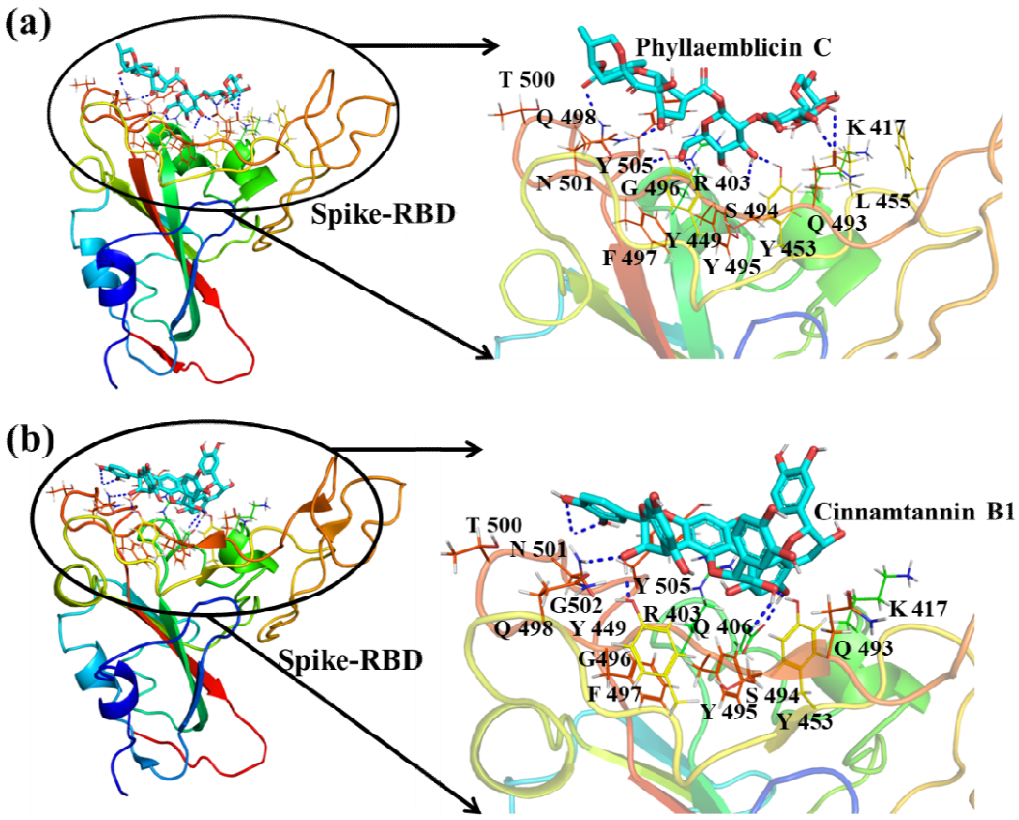

Figure 2: Spike protein RBD domain of SARS-CoV-2 in complex with the phytochemical obtained by molecular docking. (a) The molecular interactions of Phyllaemblicin C compound with the RBD domain of spike protein with the interacting amino acid residues (b) The molecular interaction of Cinnamtannin B1 compound with the RBD of spike protein alongwith the interacting residues. The interacting residues of spike protein are shown as lines, the ligands (Phyllaemblicin C and Cinnamtannin B1) is represented as cyan sticks and the hydrogen bonding interactions between the molecular complexes are represented as blue dotted lines.

residues as shown in **Figure 2b and Table 1**. Many of these Cinnamtannin B1 interactions are with the key amino acid residues of SARS-CoV-2 Spike protein-RBD contributing its direct interactions with the ACE2 host receptor for the viral entry.

3.2 Binding effect of Phyllaemblicin C at the interface of SARS-CoV-2 Spike protein – RBD and the host ACE2 receptor

The viral entry into the host cell is made possible by the atomic level direct interactions of the SARS-CoV-2 spike protein with the host cell receptor ACE2 [38, 40]. We have investigated how the phytochemical binding at the RBD domain of spike protein has affected its ACE2 binding efficiency by adopting protein-protein docking technique using ZDOCK software. Initially, the accuracy of the docking program and our computational strategy was validated by performing the docking of the RBD domain of Spike protein with the human ACE2 receptor. Further, the mode of binding was compared with its X-ray crystal structure complex (PDB id: 6M0J). Interestingly, our ZDOCK binding study predicted a highly similar complex association of SARS-CoV-2 Spike protein - RBD with the ACE2 host cell receptor and which was highly similar with the 3D folding observed in the corresponding X-ray crystal structure (**Figure 3a**). The RMSD between the predicted and corresponding X-ray complex structure showed a perfect 3D alignment of the backbone atoms, as shown in supplementary **Figure S1.**

The binding affinity between the spike protein-RBD and ACE2 receptor was computed by PRODIGY web server and is found to be -13.8 kcal/mol. Further, the amino acid residues involved in molecular interactions between RBD and ACE2 complex were identified, and is shown in **Figure 3a** and supplementary **Table S2.** The RBD of spike protein exhibits a concave surface with a ridge on one side and this molecular surface is where the extracellular part of the host ACE2 receptor will

make molecular interactions. It was reported that most of the key residues of SARS-CoV-2 RBD spike protein interact with the N-terminal helix of the human ACE2 protein which include: Q24, D30, E35, E37, D38, Y41, Q42, Y83, Q325, E329, N330, K353 and R393 residues [16]. Interestingly, in our molecular docking simulations using ZDOCK software, we also obtained similar binding residue pairs at its complex binding interface between SARS-CoV-2 Spike- RBD and ACE2 receptor, as shown in **Figure 3a** and supplementary **Table S2**.

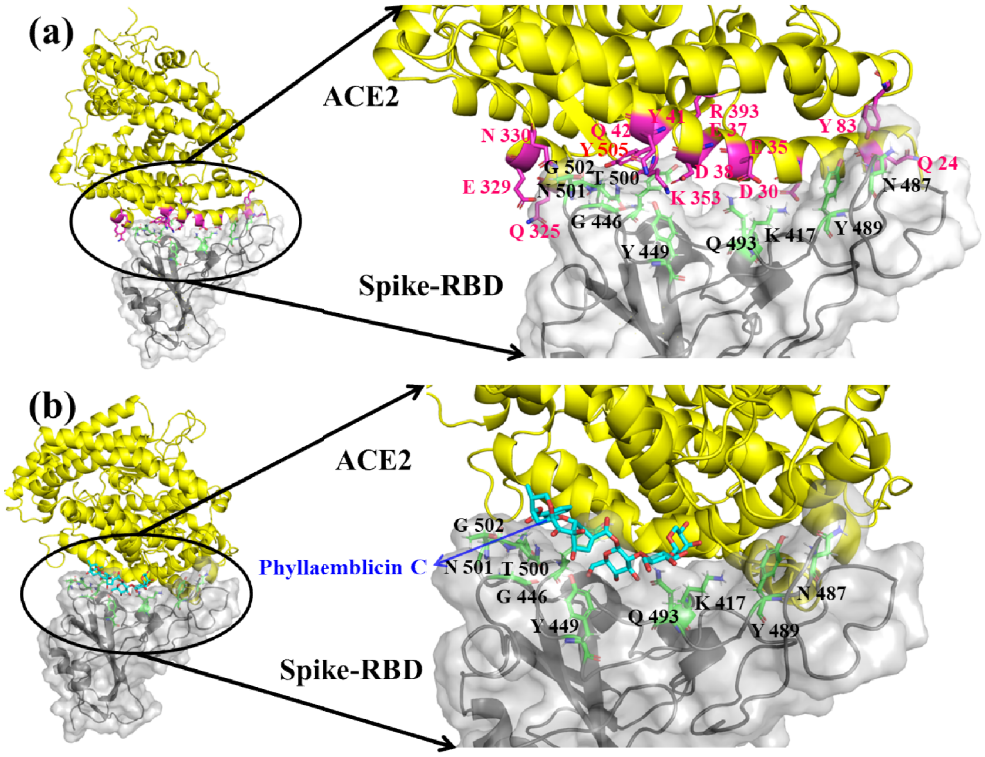

Figure 3: Molecular docking of human ACE2 receptor in complex with RBD spike protein without phytochemical (a) and with phytochemical Phyllaemblicin C (b). a) The docked pose of spike protein RBD interacting residues (black color)with ACE2 interacting residues (magenta color) is shown in magnified view (right zoom out panel) and **b)** the docked pose of spike protein RBD interacting residues (black color) with Phyllaemblicin C (Cyan sticks) at the binding interface of RBD-spike-ACE2 complexes. The ACE2 protein is shown as yellow cartoon with the interacting amino acid residues in magenta sticks. The RBD of spike protein is represented by grey cartoon and its key interacting residues are shown as green sticks (right zoom out panel).

There are different experimental studies reported involving the elucidation of molecular recognition between SARS-CoV-2 spike protein and ACE2 host receptor [16, 41]. These bio-molecular observations are in good agreement with our predicted complex model system obtained using ZDOCK and Prodigy server. In order to understand the influence of small molecule binding at this interface, we used the bound structure of spike-RBD and Phyllaemblicin C, and docked this complex against the human ACE2 protein using ZDOCK software. The mode and mechanism of binding of the spike protein -Phyllaemblicin C complex with the ACE2 receptor was analyzed. Their binding pose revealed that the binding of spike protein with the ACE2 receptor was highly distorted due to the presence of Phyllaemblicin C at its binding interface. In the absence of a small molecule at the interface of spike-RBD and human ACE2 receptor, most of the interactions of RBD residues are directed towards the N-terminal alpha helix of human ACE2 protein. Here, we observed that the presence of a small molecule (Phyllaemblicin C) at its binding interface has resulted in the binding of spike protein to the C-terminal region of the ACE2 instead of the N-terminal helix, as shown in **Figure 3b**. The interacting amino acid residues of spike-Phyllaemblicin C complex and the ACE2 receptor are presented in supplementary **Table S1.** This clearly indicates that the binding of Phyllaemblicin C at the active site of the Spike protein-RBD caused large structural variations resulting in the steric hindrance with the host receptor ACE2. Even though, the binding affinity between Spike protein RBD - Phyllaemblicin C complex with the ACE2 receptor was -12.7 kcal/mol, and was found to be far away from the active sites of ACE2 receptor depicting that this binding affinity has no biological (or pharmacological) significance. These docking simulations further signifies that the phytochemical, Phyllaemblicin C block the direct binding of Spike protein RBD with the ACE2 receptor active sites and thereby preventing the viral entry into host cells. So, this important molecular mechanism was revealed in the present modeling work in which blocking the direct interaction between spike protein and ACE2 receptor should have key influence on the SARS-CoV-2 viral entry and thereby controlling the severity of the infection, spread and morbidity.

3.3 Molecular docking of phytochemical compounds towards M^{pro} of SARS-CoV-2

M^{pro} protein of SARS-CoV-2 represents another important druggable target for the anti-viral drug discovery. Many studies were reported recently regarding the discovery and evaluation of small molecule inhibitors targeting M^{pro} protein. Using X-ray crystallography technique, the key active site residues of M^{pro} for the ligand binding are reported consisting of Y54, H41, S46, M49, D187, Q189, M49, F185, Q192, T190, A191, P168, H164, C145, M165, H163, H172, G143, and L167[12] residues. The function of M^{pro} enzyme in SARS-CoV-2 is to cleave the polypeptide into its functional proteins and which is carried out by an acylation and deacylation steps [12]. M^{pro} is a cysteine protease enzyme having a cysteine-histidine (C145-H41) catalytic dyad [12, 45].

In the present study, different phytochemicals were docked into the active site of M^{pro} protein of SARS-CoV-2 and ranked based on the docking score (binding affinity). Phyllaemblicin C was obtained as best hit with a docking score of (-9.723 kcal/mol) towards the active site of M^{pro} protein of SARS-CoV-2. Phyllaemblicin B and Procyanidin B1 showed second and third best binding affinity towards M^{pro} with a docking score of -9.151 kcal/mol and -9.128 kcal/mol respectively, as shown in **Table 1**. Both Phyllaemblicin C and Phyllaemblicin B are found in Indian gooseberry and Procyanidin B1 is found in cinnamon medicinal plant. Procyanidin B1 is a polyphenolic flavonoid having anti-inflammatory and immune-modulatory activities [46]. The anti-viral activity of Procyanidin B1 from Cinnamomi Cortex was already reported. Procyanidin B1 inhibited the progression of the infection of vesicular stomatitis virus and Hepatitis C Virus pseudotype by inhibiting the viral replication [47].

The molecular interaction of Phyllaemblicin C towards M^{pro} protein of SARS-CoV-2 was analyzed and it formed strong hydrogen bonding interactions with N142, Q189, E166, H164, H163, P168 and H41 residues. Moreover, L167, Q192, M165, C145, Y54, and M49 residues were involved in other non-bonded interactions as shown in **Figure 4a and Table 1**.

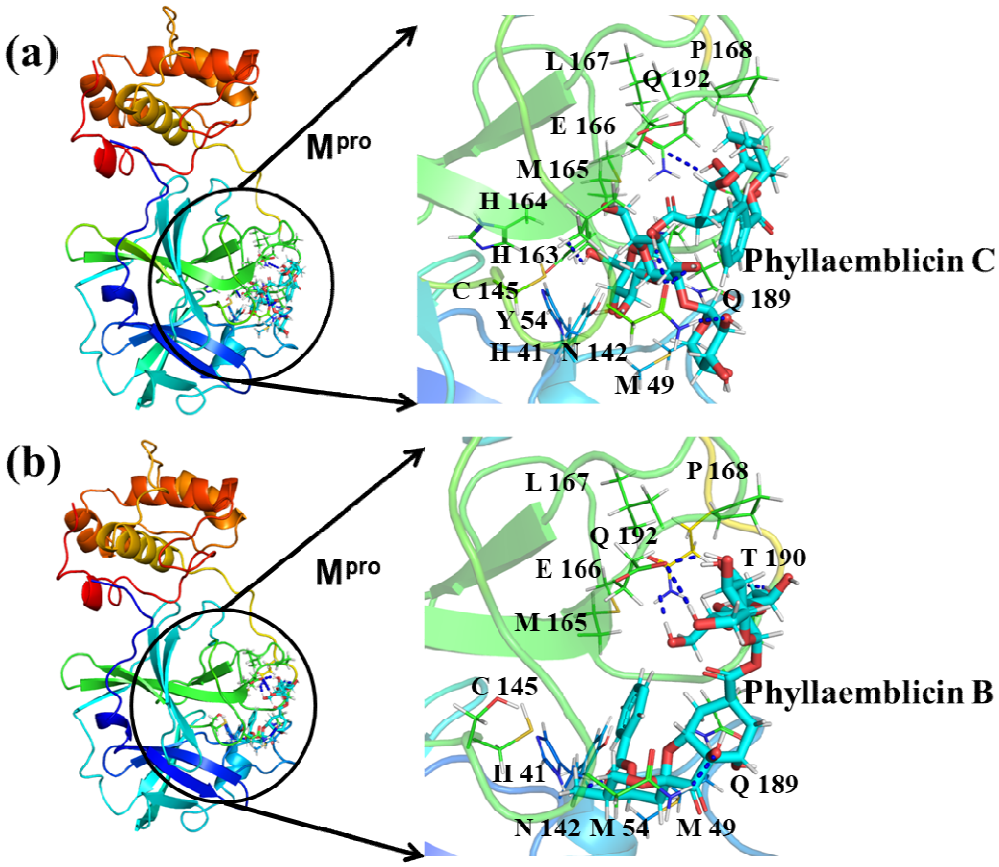

Figure 4. SARS-CoV-2 M^{pro} in complex with phytochemicals: The molecular interactions of Phyllaemblicin C with the M^{pro} protein (a) and Phyllaemblicin B with the M^{pro} protein (b). The protein is represented as cartoon, the interacting amino acid residues are shown as lines and the hydrogen bonding interactions are represented using blue dotted lines. The ligand (phytochemical) is represented as cyan sticks.

Further, the interaction of the next best phytochemical compound, Phyllaemblicin B with M^{pro} was analyzed. This compound made strong hydrogen bonding interactions as well as hydrophobic interactions with the M^{pro} active site residues. The M^{pro} residues H41, N142, Q189, E166, L167,

T190 and P168 are involved in the direct hydrogen bonding with the phytochemicals. The hydrophobic interactions are with C145, H41, M49 and M165 residues of M^{pro}, as shown in **Figure 4b** and **Table 1**. In the present case, both Phyllaemblicin C and Phyllaemblicin B compounds are interacting with the M^{pro} catalytic dyad residues. Also, many of these interacting residues are reported as key amino acids involved in the ligand binding at the active site of M^{pro}.

3.4 Pharmacokinetic properties of the Phytochemicals

The pharmacokinetic properties of the eight phytochemical compounds used in this study were predicted and is presented in **Table 2**. The Absorption, Distribution, Metabolism, Excretion and Toxicity (ADMET) properties of most of these phytochemical compounds were in the satisfactory range. Generally, phytochemical compounds have large

Table 2: ADMET properties of the studied phytochemicals predicted using QikProp module of Schrödinger software.

Phytochemical compounds	MW	QPlogPo/w	QPlogS	QPlog HERG	QPPCaco	QPlog Khsa	Human oral absorption
Phyllaemblicin C	876.8	-3.91	-0.54	-4.69	0.72	-2.01	1
Cinnamtannin B1	864.8	0.37	-5.15	-6.81	0.06	-0.14	1
Phyllaemblicin B	744.7	-2.44	-1.91	-5.02	1.69	-1.52	1
Procyanidin B1	578.5	0.24	-3.60	-5.21	1.28	-0.23	1
Procyanidin A2	576.5	0.36	-4.08	-5.60	1.64	-0.19	1
Embelin	294.4	2.15	-3.76	-4.78	201.35	-0.18	3
Germacrone	218.3	3.38	-3.75	-2.81	4641.3	0.43	3
Vasicine	188.2	1.90	-2.47	-4.10	2420.9	-0.19	3

*Molecular weight (MW) 130 – 725; Predicted octanol/water partition coefficient (QPlogPo/w) -2.0 to 6.5; Predicted aqueous solubility (QPlogS) -6.5 – 0.5; Predicted IC_{50} value for blockage of HERG potassium channel (QPlogHERG) – concern below -5; Predicted apparent Caco-2 cell permeability in nm/sec (QPPCaco) <25 poor, >500 great; Prediction of binding to human serum albumin (QPlogKhsa) -1.5 to 1.5; Human oral absorption – 1 low, 2 medium, 3 high.

molecular weight with many stereo centers and do not follow the basic Lipinski's rule of five. Some of the present active phytochemicals studied also follow in this category.

The phytochemical compounds have been highly promising since ancient times for treating a wide range of infectious diseases [5-9]. In the present study, eight phytochemicals from five different medicinal plants were virtually screened against two important SARS-CoV-2 druggable targets, Spike protein and M^{pro}. Among these eight compounds, Phyllaemblicin C showed highest binding affinity towards both spike and M^{pro} protein targets. It was reported earlier that Procyanidin A2, Procyanidine B1, Cinnamtannin B1 of Twak has been found to have activity against coronavirus (wt SARS-CoV) and SARS-CoV S pseudovirus infections [48]. Phyllaemblicin B and Phyllaemblicin C of Amalaki has been found to have activity against Influenza A virus strain H3N2 causing respiratory infections [42]. Germacrone of Haridra has been found to have activity against H1N1 and H3N2 virus causing respiratory illness [49]. Embelin from Vidanga has been found to have activity against influenza virus H1N1, H5N2 and H3N2 causing respiratory illness [50]. Vasicine from Vasa also has been found to have anti-viral activity against Influenza virus causing respiratory illness [51].

Zhuang *et al*. [48] reported the anti-viral activities of the active ingredients of Cinnamomi Cortex against other family of coronavirus (SARS-CoV). Three compounds Procyanidin A2, Procyanidin B1 and Cinnamtannin B1 exhibited good *in vitro* anti-viral activities towards SARS-CoV (**Table 3**). Among them, Cinnamtannin B1 showed the most potent inhibitory activity in micromolar concentrations [48]. Our structure-based virtual screening study revealed that these

phytochemicals from cinnamon can also bind with the spike and M^{pro} proteins of SARS-CoV-2 (**Table 3**). Further, weak binding affinity in the range of -2 to -3 kcal/mol with the M^{pro} protein was

Table 3. Comparison of the binding affinity of phytochemicals against the SARS-CoV-2 molecular targets with the experimentally reported *in vitro* anti-viral activity of the phytochemicals.

Bioactive Phytochemicals	SARS-CoV-2		Anti-viral activity	
	Binding affinity (kcal/mol)		Cytotoxic concentration CC_{50} (μM)	Inhibitory concentration IC_{50} (μM)
	Spike (RBD)	M^{pro}		
Phyllaemblicin C	-9.131	-9.723	67.7[a]	11.0[a]
Cinnamtannin B1	-9.008	-8.385	242.3[b]	32.9 [b]
Procyanidin B1	-6.275	-9.151	656.2[b]	161.1[b]
Procyanidin A2	-5.023	-7.105	796.6[b]	120.7[b]
Phyllaemblicin B	-7.381	-9.151	50.2[a]	7.8[a]

[a]*in vitro* activity towards Coxsackie Virus [52]
[b]*in vitro* activity towards SARS-CoV[48]

observed for key ingredients Embelin, Vasicine and Germacrone present in other medicinal plants. These docking results follow a similar trend on the basis of binding affinity towards two key SARS-CoV-2 targets. Hence, the phytochemicals identified in the present study could be able to inhibit both Spike and M^{pro} viral proteins and can serve as an alternative pre-clinical finding for the treatment of COVID-19 infections.

4. Conclusions:

The present work forms the basis for identifying best phytochemical compound towards inhibiting two key SARS-CoV-2 functional druggable targets by structure-based phytochemical design. Molecular docking of eight key ingredients from different medicinal plants revealed the strong

binding affinity of Phyllaemblicin C, Phyllaemblicin B, Procyanidin B1 and Cinnamtannin B1 compounds towards the active sites of Spike and M^{pro} viral proteins. Binding affinity obtained from our docking studies on these phytochemicals are in good agreement with the corresponding *in vitro* anti-viral activity studies performed in other close homolog of the coronavirus family. Another interesting conclusion emerged from the present modeling study was the mode of binding of SARS-CoV-2 spike protein with the human ACE2 receptor in the presence of Phyllaemblicin C showed significant structural variations and the receptor recognition of its binding interface was completely obstructed in the presence of phytochemical. Thus, the present computational study forms the basis of further pre-clinical evaluation of the phytochemicals for controlling the SARS-CoV-2 infections.

5. Acknowledgement:

The authors acknowledge Computational biology and Bioinformatics Lab, Centre for Nanosciences and Molecular Medicine, Amrita Vishwa Vidyapeetham, Kochi for the computing infrastructure support. CGM gratefully acknowledges the Department of Biotechnology (DBT), India, for providing financial support (Grant No. BT/PR21018/BID/7/776/2016) for procuring Schrödinger molecular modeling software license and which was used in some part of the present computational study.

6. Declaration of competing interest:

The authors declare no competing interests.

7. References:

[1] Wang L, Crameri G. Emerging zoonotic viral diseases. Rev Sci Tech. 2014;33:569-81.

[2] Lai C-C, Ko W-C, Lee P-I, Jean S-S, Hsueh P-R. Extra-respiratory manifestations of COVID-19. Int J Antimicrob Agents. 2020; 56:106024.

[3] Wang H-Y, Li X-L, Yan Z-R, Sun X-P, Han J, Zhang B-W. Potential neurological symptoms of COVID-19. Ther Adv Neurol Disord. 2020;13:1-2.

[4] Franceschi A, Ahmed O, Giliberto L, Castillo M. Hemorrhagic Posterior Reversible Encephalopathy Syndrome as a Manifestation of COVID-19 Infection. Am J Neuroradiol. 2020;41(7): 1173-76.

[5] Ilkay EO, F.Sezer SD, Natural Products as Potential Leads Against Coronaviruses: Could They be Encouraging Structural Models Against SARS‑CoV‑2? Natural Products and Bioprospecting 2020; 10:171–86.

[6] Jo S, Kim S, Shin DH, Kim M-S. Inhibition of SARS-CoV 3CL protease by flavonoids. J Enzyme Inhib Med Chem. 2020;35:145-51.

[7] Chikhale RV, Sinha SK, Patil RB, Prasad SK, Shakya A, Gurav N, et al. *In-silico* investigation of phytochemicals from Asparagus racemosus as plausible antiviral agent in COVID-19. J Biomol Struct Dyn. 2020:1-15, https://doi.org/10.1080/07391102.2020.1784289.

[8] ul Qamar MT, Alqahtani SM, Alamri MA, Chen L-L. Structural basis of SARS-CoV-2 3CLpro and anti-COVID-19 drug discovery from medicinal plants. J Pharma Anal. 2020;10:313-319.

[9] Dhivya S, Prabitha P, Prashantha KB, Suresh B. Curcumin to inhibit binding of spike glycoprotein to ACE2 receptors: computational modelling, simulations, and ADMET studies to explore curcuminoids against novel SARS-CoV-2 targets, RSC Advances, 2020;10(52):31385-99.

[10] Structural Bioinformatics- Applications in Preclinical Drug Discovery Process, Editor: C Gopi Mohan, Springer Nature Publisher, 2019. ISBN 978-3-030-05282-9.

[11] da Silva Antonio A, Wiedemann LSM, Veiga-Junior VF. Natural products' role against COVID-19. RSC Advances. 2020;10:23379-93.

[12] Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, et al. Structure of M^{pro} from SARS-CoV-2 and discovery of its inhibitors. Nature. 2020:582;289-293.

[13] Rut W, Lv Z, Zmudzinski M, Patchett S, Nayak D, Snipas SJ, et al. Activity profiling and structures of inhibitor-bound SARS-CoV-2-PL^{pro} protease provides a framework for anti-COVID-19 drug design. bioRxiv. 2020.

[14] Chen J, Malone B, Llewellyn E, Grasso M, Shelton PM, Olinares PDB, et al. Structural basis for helicase-polymerase coupling in the SARS-CoV-2 replication-transcription complex. Cell. 2020;183 ; 1560-1573.

[15] Gao Y, Yan L, Huang Y, Liu F, Zhao Y, Cao L, et al. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. Science. 2020;368:779-82.

[16] Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature. 2020;581:215-20.

[17] Ye Q, West AM, Silletti S, Corbett KD. Architecture and self-assembly of the SARS-CoV-2 nucleocapsid protein. bioRxiv. 2020.

[18] Khan MT, Ali A, Wang Q, Irfan M, Khan A, Zeb MT, et al. Marine natural compounds as potents inhibitors against the main protease of SARS-CoV-2. A molecular dynamic study. J Biomol Struct Dyn. 2020:1-11, 10.1080/07391102.2020.1769733.

[19] Shah B, Modi P, Sagar SR. In silico studies on therapeutic agents for COVID-19: Drug repurposing approach. Life Sci. 2020;252:117652.

[20] Yu R, Chen L, Lan R, Shen R, Li P. Computational screening of antagonist against the SARS-CoV-2 (COVID-19) coronavirus by molecular docking. Int J Antimicrob Agents. 2020; 56:106012.

[21] Amit KS, Sanjeev KG, Umesh K. Computational studies towards identification of lead herbal compounds of medicinal importance for development of nutraceutical against COVID-19, ChemRxiv, (2020).

[22] Canrong Wu, Yang Liu, Yueying Y, Peng Z, Wu Z, et al., Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods, Acta Pharmaceutica Sinica B, 2020, 10:766-88.

[23] Murthy S, editor. AshtangaSangraha of Vagbhata, SūtraSthana; Dvividhaushadhavijnaniya. Chowkhambha Orientalia. 2005; Chapter 12, (Students edition),Verse 39-40:268.

[24] Vinayaka Gansha Apatte, editor. Rajanighantu Sahita Dhanwantari Nighantu, Shatapusphadivarga Anandasrama Samskrita Grandhavali. 1925 (Second edition) Verse 51.:79.

[25] Apatte VG, editor. Rajanighantu Sahita Dhanwantari Nighantu, Shatapusphadivarga. Anandasrama Samskrita Grandhavali. 1925 (Second edition) Verse 12.

[26] Apatte VG, editor. Rajanighantu Sahita Dhanwantari Nighantu, Guduchyadivarga. Anandasrama Samskrita Grandhavali. 1925 (Second edition) Verse 55-56.

[27] Murthy S, editor. AshtangaSangraha of Vagbhata, SūtraSthana; Annasvarupavijnaniya. Chowkhambha Orientalia. 2005; Chapter 7, (Students edition) Verse 119.

[28] Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, et al. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. J Med Chem. 2004;47:1750-9.

[29] Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem. 2004;47:1739-49.

[30] Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, et al. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein−ligand complexes. J Med Chem. 2006;49:6177-96.

[31] Pushkaran AC, Nataraj N, Nair N, Götz F, Biswas R, Mohan CG. Understanding the structure–function relationship of lysozyme resistance in *Staphylococcus aureus* by peptidoglycan O-acetylation using molecular docking, dynamics, and lysis assay. J Chem Inf Model. 2015;55:760-70.

[32] Pushkaran AC, Vinod V, Vanuopadath M, Nair SS, Nair SV, Vasudevan AK, et al. Combination of repurposed drug diosmin with amoxicillin-clavulanic acid causes synergistic inhibition of mycobacterial growth. Sci Rep. 2019;9:1-14.

[33] Vijayrajratnam S, Pushkaran AC, Balakrishnan A, Vasudevan AK, Biswas R, Mohan CG. Bacterial peptidoglycan with amidated meso-diaminopimelic acid evades NOD1 recognition: an insight into NOD1 structure–recognition. Biochem J. 2016;473:4573-92.

[34] Panicker PS, Melge AR, Biswas L, Pavithran K, Mohan CG, Epidermal growth factor receptor (EGFR) structure-based bioactive pharmacophore models for identifying next-generation inhibitors against clinically relevant EGFR mutations. Chem Biol & Drug Design, 2017;90:629-36.

[35] Pierce BG, Wiehe K, Hwang H, Kim B-H, Vreven T, Weng Z. ZDOCK server: interactive docking prediction of protein–protein complexes and symmetric multimers. Bioinformatics. 2014;30:1771-3.

[36] Vangone A, Bonvin A. PRODIGY: a contact-based predictor of binding affinity in protein-protein complexes. Bio-protocol. 2017;7:e2124.

[37] Xue LC, Rodrigues JP, Kastritis PL, Bonvin AM, Vangone A. PRODIGY: a web server for predicting the binding affinity of protein–protein complexes. Bioinformatics. 2016;32:3676-8.

[38] Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell. 2020;181:271-280.

[39] Sungnak W, Huang N, Bécavin C, Berg M, Queen R, Litvinukova M, et al. SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. Nat Med. 2020;26:681-7.

[40] Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh C-L, Abiona O, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science. 2020;367:1260-3.

[41] Yi C, Sun X, Ye J, Ding L, Liu M, Yang Z, et al. Key residues of the receptor binding motif in the spike protein of SARS-CoV-2 that interact with ACE2 and neutralizing antibodies. Cell Mol Immunol. 2020;17:621-630.

[42] Lv J-J, Yu S, Xin Y, Cheng R-R, Zhu H-T, Wang D, et al. Anti-viral and cytotoxic norbisabolane sesquiterpenoid glycosides from *Phyllanthus emblica* and their absolute configurations. Phytochemistry. 2015;117:123-34.

[43] Anderson RA, Broadhurst CL, Polansky MM, Schmidt WF, Khan A, Flanagan VP, et al. Isolation and characterization of polyphenol type-A polymers from cinnamon with insulin-like biological activity. J Agric Food Chem. 2004;52:65-70.

[44] López JJ, Jardín I, Salido GM, Rosado JA. Cinnamtannin B-1 as an antioxidant and platelet aggregation inhibitor. Life Sci. 2008;82:977-82.

[45] Dai W, Zhang B, Jiang X-M, Su H, Li J, Zhao Y, et al. Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. Science. 2020;368:1331-5.

[46] Xing J, Li R, Li N, Zhang J, Li Y, Gong P, et al. Anti-inflammatory effect of procyanidin B1 on LPS-treated THP1 cells via interaction with the TLR4–MD-2 heterodimer and p38 MAPK and NF-κB signaling. Mol Cell Biochem. 2015;407:89-95.

[47] Li S, Kodama EN, Inoue Y, Tani H, Matsuura Y, Zhang J, et al. Procyanidin B1 purified from Cinnamomi cortex suppresses hepatitis C virus replication. Antivir Chem Chemother. 2010;20:239-48.

[48] Zhuang M, Jiang H, Suzuki Y, Li X, Xiao P, Tanaka T, et al. Procyanidins and butanol extract of Cinnamomi Cortex inhibit SARS-CoV infection. Antiviral Res. 2009;82:73-81.

[49] Liao Q, Qian Z, Liu R, An L, Chen X. Germacrone inhibits early stages of influenza virus infection. Antiviral Res. 2013;100:578-88.

[50] Hossan MS, Fatima A, Rahmatullah M, Khoo TJ, Nissapatorn V, Galochkina AV, et al. Antiviral activity of *Embelia ribes* Burm. f. against influenza virus in vitro. Arch Virol. 2018;163:2121-31.

[51] Amber R, Adnan M, Tariq A, Mussarat S. A review on antiviral activity of the Himalayan medicinal plants traditionally used to treat bronchitis and related symptoms. J Pharm Pharmacol. 2017;69:109-22.

[52] Liu Q, Wang Y-F, Chen R-J, Zhang M-Y, Wang Y-F, Yang C-R, et al. Anti-coxsackie virus B3 norsesquiterpenoids from the roots of *Phyllanthus emblica*. J Nat Prod. 2009;72:969-72.