

Exploring structurally diverse plant secondary metabolites as a potential source of drug targeting different molecular mechanisms of Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) pathogenesis: An in silico approach

Hariprasad. P. (✉ phari@iitd.ac.in)

Indian Institute of Technology Delhi <https://orcid.org/0000-0002-6598-5301>

Gowtham, H. G.

Indian Institute of Technology Delhi

Monu, D. O.

Indian Institute of Technology Delhi

Ajay, Y.

Indian Institute of Technology Delhi

Gourav, C.

Indian Institute of Technology Delhi

Vasantharaja, R.

Indian Institute of Technology Delhi

Bhani, K.

Indian Institute of Technology Delhi

Koushalya, S.

Indian Institute of Technology Delhi

Shazia, S.

Indian Institute of Technology Delhi

Priyanka, G.

Indian Institute of Technology Delhi

Faiz. A.

Indian Institute of Technology Delhi

Leena, C.

Indian Institute of Technology Delhi

Research Article

Keywords: Plant Secondary Metabolites, Molecular Docking, SARS-CoV-2, Spike Protein, Bioavailability

Posted Date: May 11th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-27313/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Plants are endowed with a large pool of structurally diverse small molecules known as secondary metabolites. Present study aims to virtually screen these plant secondary metabolites (PSM) for their possible anti-SARS-CoV-2 properties targeting four protein/enzymes which determines viral pathogenesis. Results of molecular docking and data analysis revealed a unique pattern of structurally similar PSM interacting with the target protein. Among the top-ranked PSM with lower binding energy, >50% were triterpenoids against viral spike protein, >32% were flavonoids and their glycoside against Human transmembrane serine protease, >16% were flavonol glycosides and >16% were Anthocyanidine against viral main protease and >13% were flavonol glycoside against viral RNA dependent RNA polymerase. The primary concern about these PSM is their bioavailability. However, several PSM recorded higher bioavailability score and found fulfilling drug-likeness characters as per Lipinski's rule. Natural occurrence, biotransformation, bioavailability of selected PSM and their interaction with the target site of selected proteins were discussed in detail. Further, we hypothesized the use of selected PSM to cure the COVID-19 by inhibiting the process of viral host cell recognition and replication in host cell. However, these PSM needs thorough *in vitro* and *in vivo* evaluation before taking them to clinical trials.

Introduction

Coronaviruses (CoV) are spherical or pleomorphic enveloped, non-segmented particles contains positive-sense single-stranded RNA (Zhu et al., 2020). There are several types of low pathogenic CoV which cause mild respiratory symptoms. In general, it is classified under four genera such as α , β , γ , and δ CoV. α and β -CoV are reported to cause fatal respiratory tract infections in mammals. Severe Acute Respiratory Syndrome-CoV (SARS-CoV) is grouped under β -CoV. Whereas, γ , and δ CoV infects birds (Yin and Wunderink, 2018). Based on genomic sequence evidence, bat CoV RaTG13 shares 96.2 % similarity with SARS-CoV-2. Hence, bats are considered as a primary source of SARS-CoV-2 and infected humans through several intermediate hosts (Zhou et al., 2020).

Coronavirus disease 19 (COVID-19) was first spotted in a seafood market of Wuhan city, Hubei Province of China and now it had reached all continents across the globe (Hui et al., 2020). The SARS-CoV-2 virus causes SARS by the onset of pneumonia, acute respiratory distress syndrome (ARDS) and multi-organ dysfunction. Moreover, the SARS-CoV-2 enters the human body through the mouth and other openings and spreads primarily through droplets, saliva, or discharges from the nose/mouth of an infected person after sneezing or coughing (Pant et al., 2020). On 11th March 2020, the World Health Organization (WHO) declared COVID-19 as a pandemic. As on 4th May 2019 the global COVID-19 data shows, 34,42,234 confirmed cases and 2,39,740 confirmed deaths (www.who.int).

At present, there are no pharmaceutical products approved by the Food and Drug Administration (FDA) as safe and effective for the treatment of COVID-19. Preliminary *in vitro* studies and clinical trials carried

out by scientists on COVID–19 patients disclosed the effectiveness of Hydroxychloroquine, an anti-malarial drug in combination with azithromycin, a broad spectrum anti-bacterial drug in reducing the disease severity (Gautret et al., 2020). Similarly, several off-label use of medicines (such as lopinavir-ritonavir, favipiravir, remdesivir, ribavirin, anti-Interleukin–6 inhibitors, etc.) has been suggested to treat COVID–19 as potential investigational drugs. However, several antiviral drugs are in clinical trials/advanced testing, and major antiviral drugs have been developing based on different molecular targets including compounds that block the spike protein-human Angiotensin-converting enzyme 2 (hACE2) mediated viral entry, main protease (M^{pro}), papain-like protease 2 (PLP2), RNA dependent RNA polymerase (RdRp) and SARS-CoV helicase (Wu et al., 2020). Additionally, these drugs require scientific evaluation in terms of their antiviral potencies, efficacy, adverse effect and safety before available for the public use. Further, as a precautionary measures, surveillance borders, personal hygiene, social distancing, lockdown, supportive care and personal protective equipment are practiced across the world to prevent disease spreading among the communities.

Plants are known to inherently contain a large number of structurally diverse secondary metabolites which are developed during the course of evolution primarily as defence mechanisms against herbivores and predators, mediating pollination and for protection against abiotic stress (Pichersky and Gang, 2000). The variability of metabolites can be seen across different plant families, genus, species and different parts of the same plant species (Holeski et al., 2012). The concentrations of PSM varies according to the growth stages and in response to biotic and abiotic stress to which plant exposed. The development of drugs from phytopharmaceuticals is a trending approach to look for eco-friendly therapeutic molecules with no or minimal side-effects. This time-bound situation requires an efficient and effective method to develop therapeutics which disables the virus molecular machinery. Considering the safety of the users, any conventional drug discovery plan is time-consuming process that sometimes takes decades to complete. Thus, repurposing the already available FDA approved drugs, use of plant based herbal medicines, or edible plant parts rich in anti-viral PSMs are other strategies appears to be promising under current situations.

In-Silico or computational approaches are algorithm-based virtual screening methods developed for screening large number of molecules in shorter time and identification of probable potent drug candidate. In a recent research, 1903 approved drugs were virtually screened through molecular docking and binding free energy calculations suggested nelfinavir as a potential inhibitor against SARS CoV–2 (Xu et al., 2020). Similarly, phytochemicals such as kaempferol, quercetin, luteolin–7- glucoside, demethoxycurcumin, naringenin, apigenin–7-glucoside, oleuropein, curcumin, catechin, and epicatechin-gallate have been reported potential viral M^{pro} inhibition (Khaerunnisa et al., 2020). Elfiky (2019) reported Ribavirin, Remdesivir, Sofosbuvir, Galidesivir, and Tenofovir as potent drugs candidates against RdRp of SARS-CoV–2 through molecular docking studies.

Most of the earlier PSM based virtual screening were either limited by number of PSM or structural diversity of test ligands and number of target proteins. Hence, the present study is aimed to (i) Creating a structurally diverse PSM library, (ii) Finding potent PSM which binds to the target site of selected

protein/enzymes with low binding energy (BE), (iii) analyzing structural and functional relation of top scored PSMs, and (iv) Analyzing the bioavailability of selected PSM using Swiss ADME.

Materials And Methods

Preparation of plant secondary metabolites library

To prepare PSM library, an extensive literature survey was conducted on selected plants and the general and species-specific PSMs including Alkaloids, Phenolics and Terpenoids were listed. The 3D and 2D structures (SDF Files), and canonical smiles of the selected PSMs were retrieved from online databases such as PubChem (<https://pubchem.ncbi.nlm.nih.gov>) and ChemSpider (<http://www.chemspider.com>). The 2D structures were converted into 3D coordinates and geometries were optimized by using the free offline tool Marvin Sketch (<http://www.chemaxon.com/products/marvin/marvinsketch>). As several PSM are present in multiple plant species, a 6% duplication was allowed in the main PSM library. Additionally, PMS isomers were considered as a separate ligand. All the files were coded and used for further studies.

Target Proteins

In the present study we selected four target protein, one from human (human transmembrane serine protease 2, TMPRSS2) and three from SARS-CoV-2 (spike protein, M^{Pro} and RdRp). The crystal structure of SARS-CoV-2 spike receptor-binding domain bound with ACE2 (PDB ID: 6M0J) (2.45Å) was retrieved from Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB) (RCSB, www.rcsb.org) (Lan et al., 2020) and ACE2 was removed, processed and used for docking studies. The 14 amino acid residues (AAR) (THR415, ASN439, TYR449, TYR453, LEU455, PHE486, ASN487, TYR489, GLN493, GLN498, THR500, ASN501, GLY502 and TYR505) of spike protein that are key for binding hACE2 (Walls et al., 2020) were considered as an active site for molecular docking process. The TMPRSS2 sequence (NP_001128571.1) was retrieved from the National Center for Biotechnology Information (NCBI) protein database. The structure of TMPRSS2 was generated by using the SWISS-MODEL online server (Biasini et al., 2014). The structures were marked, superimposed and visualized by using Chimera (Pettersen et al., 2004). Model 1 of our work and O15393 present in uniprot were found 100% similar. Three amino acid residues (HIS296, ASP345 and SER441) of catalytic site (Meng et al., 2020) were considered as a key residue of TMPRSS2 in the molecular docking process. The crystal structure of SARS-CoV-2 Mpro (PDB ID: 6LU7; 2.16 Å) (Jin et al., 2020) was retrieved and used for docking studies after processing. Mpro has three domains, and the active site is located between domain I and II. Here, the CYS145-HIS41/SER144-HIS163 can act as nucleophilic agent and GLY143 and GLU166 can form hydrogen bonds with the "CO-NH-Cα-CO-NH-Cα" structure of the backbone of the substrate protein. These 6 residues were considered as the key residue of Mpro in the molecular docking process (Li et al., 2020). The SARS-CoV-2 RdRp protein sequence (YP_009725307.1) was retrieved from the National Center for Biotechnology Information (NCBI) protein database. The structure of RdRp was generated by using the SWISS-MODEL online server (Biasini et al., 2014). The structures were marked, superimposed

and visualized by using Chimera (Pettersen et al., 2004). The amino acid residues VAL557 in motif F (Gao et al., 2020) was considered as a key residue of RdRp in the molecular docking process.

Removal of water molecules, metal ions, cofactors, and addition of charges and hydrogen atoms was done by UCSF Chimera tool. (Pettersen et al., 2004) Computing energy minimization and reconstruction of missing atoms and to perform stereo-chemical quality checks to come up at the best possible three-dimensional structures were done through discovery studio software (Dassault Systemes BIOVIA, Discovery Studio, Version 3, San Diego: Dassault Systèmes: 2019).

Molecular docking

The ligands were energy minimized by conjugate gradients optimization algorithm with total numbers of 200 steps performed as a default universal force field (UFF) parameters (Rappe et al., 1992). The capability of ligands to interact with the target site of selected protein was studied following computational ligand-target docking approach. Molecular docking was carried out using PyRx, AutoDock Vina option based on scoring functions (Trott et al., 2010; Dallakyan et al., 2015). The least binding energy (BE, Kcal/mol) conformation was considered as the most favourable docking pose. The interaction between ligand and protein were analyzed using pymol (De Lano., 2020) and Discovery Studio 3.5 (Accelrys Software Inc. Accelrys, San Diego, CA, US).

Structural activity relationship Analysis

According to their BE, all the ligands were assigned ranking and aligned in ascending order. For the ease of the study top 250 molecules (268 for Mpro) were subjected to ligand structure similarity analysis (based on canonical SMILES) and BE using Data Warrior software (version 5.2.1).

Physicochemical properties and Bioavailability of PSM

The drug-likeness and the physicochemical properties were studied using SwissADME (www.swissadme.ch). The canonical SMILES of the selected PSMs were subjected to SwissADME analysis. The PSMs were analyzed for their drug-likeness properties following Lipinski's rule of five (Lipinski et al., 2001). The bioavailability radar chart obtained were analyzed for their drugs like properties, i.e., lipophilicity (XLOGP3 between -0.7 and +5.0), Molecular weight (150 and 500 g/mol), Topological polar surface area (between 20 and 130 Å²), Solubility (log S not higher than 6), Flexibility (no more than 9 rotatable bonds) and Saturation (fraction of carbons in the sp³ hybridization not less than 0.25) (Daina et al., 2017). The bioavailability score (BAS) for selected PSM was also recorded (Martin, 2005)

Results And Discussion

The PSM library contains 4704 molecules collected from 203 plant species belongs to diverse plant families (Sup File 1). Over 22000 docking reactions were run (which includes replications, to confirmation the activity of several molecules with BE) using 4750 ligands against four selected target proteins/enzymes which involved in host cell recognition, entry and replication of SARS-CoV-2. Upon molecular docking, a wide range of BE was obtained for all four target proteins. The obtained results were arranged in ascending order of BE. For the ease of the study, we selected the top 268 molecules against Mpro and 250 molecules for all other target protein for further analysis (Sup. File 2) viz, structural similarity and activity relationship and Physicochemical characterization to evaluate its drug-likeness.

SARS-CoV-2 Spike Protein

Spike protein is a class I fusion protein present within the envelope as a homotrimer and consists of three S1- S2 heterodimers. The receptor-binding domain (RBD) is located on the head of S1 (Gui et al., 2017) and binds with the cellular receptor hACE2 (Kong et al., 2020). Any PSM interacts with these selected AARs of spike protein by forming multiple H bonds and other interactions with lower BE may interfere with the spike protein - hACE2 interactions, thereby preventing the initial recognition of host cell by SARS-CoV-2.

Interestingly, the large pool of PSM found interacting with the exposed surface of spike protein were found belonging to class triterpenoids. More than 50% of the PSM among top 250 molecules belongs to Triterpenoids and their derivatives, largely includes triterpenes and sterols. Here sterol lactones alone represent >14% of total PSM (Fig. 1a). Few biflavonoids, flavonoid glycoside and hydrolysable tannins were also showed promising results.

Triterpenoids are a large pool of diverse phyto-molecules grouped under terpenoids, which includes triterpenes, steroids, limonoids, quassinoids, triterpenoidal and steroidal saponins (Mahato et al., 1992; Dzubak et al., 2005) found widely distributed across the plant kingdom. Yin et al. (2012) reported the bioavailability of eight triterpenes (oleanolic acid, ursolic acid, arjunolic acid, asiatic acid, boswellic acid, corosolic acid, madecassic acid, and maslinic acid) from different vegetables and fruits, in their intact form in the different organs of Mice. A similar kind of study proved the bioavailability of betulinic acid exerting its anti-tumor properties (Godugu et al., 2014). The data obtained from our study indicated that possibilities of developing triterpenoid based drug molecules targeting SARS-CoV-2 Spike protein. Coagulins from *Withania coagulans* recorded lower BE with target AAR of spike protein. Similar observations were made with structurally similar triterpene and steroid, i.e. steroidal lactones, steroidal saponins, steroidal glycoalkaloids, triterpene glycosides, triterpene saponins, and triterpene sterols. Coagulin N recorded a lest of -9.1 BE followed by Coagulin K (BE -8.9). Coagulin N forms H bond with ARG403, TYR449 and GLY496 of spike protein, whereas Coagulin K forms H bond with ARG403, GLY496, GLN498 and TYR505, which may interfere with viral host cell recognition process. Also, both the Coagulin K recorded 0.55 BAS, and passed all the Lipinski's rule of drug-likeness except MW.

Withanolides are naturally occurring C₂₈-Steroidal lactone triterpenoids build on an intact or rearranged estrogen framework (Christen, 1986; Glotter, 1991; Mirjalili et al., 2009). Withanolides reported in 15 Solanaceous genera and commonly found in *Withania*, *Withania somnifera* and *W. coagulans*. Withanolide E recorded the least BE of -8.4 followed by Withanone (-8.3), Ashwagandhanolide (-8.3) and Withastramonolide (-8.2) (Sup. File 2). Withanolide E recorded BAS 0.55 and also, found following all drug-likeness rule of Lipinski's. It recorded H bond formation with GLN493 and TYR453 AAR of spike protein (Fig. 1b; Sup. File 2). Withanolides are smaller in molecular weight with high GI absorption and follow all the characters of the drug-like molecules as per Lipinski's rule (Sup file 3).

Saponins are naturally occurring plant glycosides found in a wide range of plant spices. They are high molecular weight amphiphilic compounds having triterpenoids and steroid aglycon as lipophilic moiety and sugars as hydrophilic moiety (Vicken et al., 2007). Another class of basic steroidal saponins contain nitrogen analogues of steroid sapogenins as aglycones. Almost all saponins have unfavourable physicochemical properties for drug likeliness rule. Glycosylation leading to the increased number of sugars makes the molecule more complex and poorly bioavailable. The biotransformation of saponins mainly occurs at intestine with the aid of gut microbes leading to the generation of the rare low molecular weight saponins containing no or lower number of sugar moiety (He et al., 2019). These hydrolyzed products are higher in bioavailability and bioactivity compared to their parental compounds (Gao et al., 2012; Ramasamy et al., 2015; Del Hierro et al., 2018). Graecunin are a group of spirostanol saponins isolated from the leaves of *Trigonella foenum-graceum* (Fenugreek) (Varshney et al., 1994). In our study, Graecunin E recorded a least of -8.6 BE and found interacting with THR415 and GLN493 AAR's of spike protein (Fig. 1b; Sup. File 2). Other Graecunin related compounds, Trigofenoside E1 (BE -8.2), Uttronin B (BE -8.3), Stigmasteryl glucoside (BE -7.9) and Yuccagenin (BE -7.6) also showed promising results. It was noticed that BE of above-mentioned molecules was related to their number of the sugar moiety. As the number of sugar moiety reduces, BE also found reducing. However, with the loss of sugar moiety, their bioavailability is increasing as observed between Graecunin E and Yuccagenin (aglycon form) (Fig. 1b; Sup. File 2 and 3). Also, Yuccagenin was found forming H bonds with GLY496 and ASN501 AARs of spike protein which are crucial to interact with AARs of Human ACE2 and found fulfilling all drug-likeness characters as per Lipinski's rule (Fig. 1b).

Alkaloids are low molecular weight and heterocyclic N containing basic compounds widely distributed across the plant kingdom with >20000 compounds identified (Yang and Stockigt, 2010). Based on their biosynthetic origin, they are broadly classified as indole, Pyrrolizidine and Quinolizidine alkaloids (Seigler, 1998). They are preferably used at a lower dosage for pharmacological applications, at the higher dose they may adversely affect the consumers. Additionally, several unfavourable pharmacokinetics behaviours of alkaloids, such as low oral absorption/bioavailability, short half-life, and rapid clearance, have been reported to significantly decline in their bioavailability and bioactivity (Li et al., 2004; Zaho et al., 2011; Pirillo and Catapano, 2015). Bismahanine, a carbazole alkaloid isolated from leaves of *Murraya koenigii* (Husson, 1985; Chakraborty and Roy, 1991) and some other Rutaceae members (Chakraborty, 1993). Bismahanin and related compounds are reported for their wide biological activities such as anti-

oxidant, anti-diabetic, anti-inflammatory, anti-microbial, anti-cancerous, anti-viral, etc. (Knolker and Kethiri, 2008; Nalli et al., 2016; Yu et al., 2020). In our studies, physicochemical properties of Bismahanin revealed it as a low polar, non-soluble, high MLOGP value and high molecular weight PSM according to Lipinski's rule of drug-likeness, which reduces its bioavailability (BAS 0.17). The molecule interacted with residue GLU406 through H bonds and showed Van der Waals interaction with other residues (TYR495, LEU455, ASN501 and GLY502) of spike protein (Fig. 1b; Sup. File 2). Pseudojervine, a steroidal alkaloid was first isolated from the rhizome of *Veratrum album* by Wright and Luff (1879) which is regularly used in Chinese traditional medicine. Alkaloids of Veratrum rhizome have been reported for their hypotensive, anti-thrombotic and anti-tumor activity (Xu et al., 2001; Berman et al., 2002; Thayer et al., 2003). In the present study, Pseudojervin recorded -8.7 BE and found interacting with AAR's GLN493, GLY496 and SER494 of spike protein through H bonds. Also, it passed all the Lipinski's rule except molecular weight (587.74 g/mol) and recorded 0.55 BAS (Fig. 1b; Sup. File 2 and 3).

Tannic acid is a polyphenolic compound which is well known for protein precipitation properties. It is a specific tannin that contains 10 galloyl (3,4,5-trihydroxyphenyl) unit surrounding a glucose centre. It is mainly used for industrial applications and as raw materials for gallic acid and pyrogallol production through hydrolysis using enzyme tannase (Lekha and Lonsane, 1997). As it is high molecular weight compound and highly soluble in water (low lipophilicity), its GI absorption is almost nil in its original form. But it is hypothesized that bio-transformed products may enter into plasma and exert biological activities which still needs a thorough study. In our study, Tannic acid recorded lower BE of -8.9 , and some of the structurally related hydrolyzable tannins such as Strictinin (BE -8.6), Tercatin (BE -8.3), Punicalgin (BE -8.5), Terchabulin (BE -8.4), Terflavin (BE -8.1) also showed promising results (Sup. File 2). Details of biotransformation and bioavailability of hydrolysable tannins are discussed somewhere else in this paper.

Arecatannins are a class of condensed tannins (Procyanidins) which contains catechin or epicatechin as oligomers. Arecatannin A3 contains epicatechin-epicatechin-epicatechin-epicatechin-catechin as their basic structure and in its original form may be poorly available because of their unfavourable physicochemical properties (Sup. File 3). But, its bio-transformed products, especially monomers, dimers or trimers, may show a varied degree of bioavailability and bioactivity. Their monomers catechin and epicatechin are reported as bioavailable better than their parental molecules (Baba, 2001). However, their dimer and trimers are poorly bio-available (Serra et al., 2010). In our studies, none of the catechin or epicatechin, and their derivatives were appeared in top 250 PSM rank indicating its inability to bind an open surface of the spike protein. Some Arecatannin related molecules such as Arecatannin B1 (BE -8.1), Proanthocyanidine A-6 (BE -7.9) and Proanthocyanidine A1 (BE -7.6) also found interacting with spike protein (Sup. File 2).

Chalcones are a group of plant-derived aromatic ketone precursors of flavonoid biosynthesis in plants and forms central core for a variety of biological compounds. They are reported for their anti-oxidant, anti-bacterial, anthelmintic, anti-ulcer, anti-viral and anti-cancerous activity (Valavanidis et al., 2013; Kulkarni et al., 2014). Kamalachalcone C is a dimeric chalcone first isolated from *Mallotus philippensis*

by Furusawa et al. (2005). It recorded -8.8 BE and was found interacting with AAR's GLN493, GLN492, ARG403 and GLU406 of spike protein. Bioavailability score 0.55 was recorded by Kamalachalcone C, indicating it as a potent drug candidate (Fig. 1b; Sup. File 2). Amentoflavone, a biflavone also recorded a lower BE of -8.7 and showed H bond formation with AAR's GLN493, SER494 and GLY496 of SARS-CoV-2 Spike protein (Fig. 1b).

Human Transmembrane serine protease (TMPRSS2)

Structural similarity and BE correlation analysis indicated that a large number of flavonoid glucoside ($> 32\%$) were found interacting with the target site of TMPRSS2, followed by ellagitannins and triterpenoids. Interestingly, two triterpenoid saponins (Liquorice and Glycyrrhizic acid) and a stilbenoid (Cis-Miyabenol C) recorded lowest BE (Fig. 2a). The data obtained from this study indicated that possibilities of developing Flavonoid glycoside or triterpenoid based drug molecules targeting human TMPRSS2 which process SARS-CoV-2 Spike protein and facilitate the entry of the virus into the host cell.

Flavonoids are a ubiquitous group of naturally occurring PSM. Basically, they are polyphenolic compound characterized by a flavan nucleus. They exist in every part of the plant body in various forms. Vegetables, grains, leaves, bark, flowers, roots, stems, flowers, fruits, tea and wines are a rich source of dietary flavonoids. Their acceptance as health beneficial PSM by the scientific community is because of their broader spectrum biological activities viz. anti-oxidant, anti-inflammatory, anti-mutagenic, anti-carcinogenic, anti-choline esterase, immunomodulator, anti-thrombogenic, etc. as reviewed by Velayutham et al. (2009) and Panche et al. (2016). Largely flavonoids are classified into seven classes viz., Flavanol (Catechin, Epicatechin, Epigallocatechin, etc.), Flavanonol (Taxifolin, Aromadedin, etc.), Flavone (Chrysin, Apigenin, Rutin, Luteolin, etc.), Flavonol (Kempferol, Quercetin, Myricetin, Tamarixetin, etc.), Flavanone (Naringin, Naringenin, Taxifolin, Hesperidin, etc.), Isoflavone (Gensitin, Daidzin, etc.), Anthocyanidin (Apigenidin, Cyanidin, etc.), Chalcone (Derricin, Lonchocarpin, etc.) (Kumar and Pandey, 2013; Panche et al., 2016).

In plants, most flavonoids occurs in their β -glycosidic form bound to one or more sugar molecules. The exceptions are catechin and proanthocyanidins (Willimson, 2004). In our study, a large number of flavonoid glycosides were found efficiently binding to the target site of protein with lower BE. Similarly, in most of the earlier *in vitro* studies, flavonoid glycosides are proved to be a better candidate for enzyme inhibition activity related to health benefits. However, a major concern about flavonoid glycosides is their bioavailability. Upon consumption, at the small intestine, they absorbed mostly in their aglycon form at the cellular level. Absorption of glycosidic form mainly depends on their number of the sugar moiety. Through many research studies, it was confirmed that flavonoid glycosides are converted to aglycones by gut microbes and decomposed to yield two different phenolic products. Whereas, catechin and procyanidins are biotransformed into 5-(hydroxyphenyl)- γ -valerolactone found in blood plasma and urine as sulfate and glucuronide metabolites (Monagas et al., 2010). However, several earlier studies supported the hypothesis that glycosidic forms are more bioavailable and resistant to microbial degradation. Also,

they can deliver aglycone moiety better when administrated as glycosidic form. Participation of active uptake of flavonoid glycosides by enterocytes was researched in detail. During this process, flavonoid glycosides are converted into aglycones by membrane-bound beta-glucosidase (Reviewed by Kumar and Pandey, 2013). Here, Agatisflavone, a bioflavonoid recorded a BE of -8.9 and found interacting with catalytic site AAR's SER441 of TMPRSS2 forming hydrogen bonds. Nirurin, a prenylated flavonone glycoside found in *Phyllanthus niruri* recorded a BE of -8.9 and showed H bond formation with HIS296 and SER441 of TMPRSS2 catalytic site and with VAL290, SER436, GLY439 and CYS465 in close vicinity of the catalytic site, presenting itself as a strong irreversible inhibitor of TMPRSS2 (Fig. 2b; Sup. File 2). Following to this, Naringin, a flavanone-7-O-glycoside recorded BE of -8.3 and showed similar interaction with AARs of TMPRSS2 (Sup. File 2). All these molecules are low in their BAS and their physicochemical characters are not favorable as per Lipinski's rule (Sup. File 3).

Though the flavonoids are dominating the group in numbers, the lowest BE was recorded by two triterpenoid saponins, i.e., Liquorice (-9.7) and Glycyrrhizic acid (-9.5) followed by a stilbenoid Cis-Miyabenol (-9.4). Here, Liquorice was found forming H bond with SER441 of TMPRSS2 catalytic site AAR through glucose moiety. Similarly, Glycyrrhizic acid also showed H bond formation with HIS296 of TMPRSS2 through Oxygen involved in a glycosidic bond. Liquorice and Glycyrrhizic acid are isomers, triterpenoid glycosides obtained from roots of *Glycyrrhiza glabra* (Liquorice). Roots of this plant are traditionally used to alleviate jaundice, gastritis and bronchitis. Gancao, a Chinese herbal decoction of dried plant roots and stem, contains 3.63 – 13.06% Glycyrrhizin and well known for their therapeutic properties, including antiviral (Pompei et al., 1979). Glycyrrhizic acid was reported as an active component from the roots of *Glycyrrhiza glabra* which inhibited the growth and cytopathology of both DNA and RNA virus viz., Herpes simplex type I, Newcastle disease virus, Vesicular stomatitis virus and Polio type I virus without affecting host cell activity and replication. Research regarding the bioavailability of Glycyrrhizic acid revealed that after oral administration, only bio-transformed Glycyrrhetic acid was detected in plasma at higher concentration (Stormer et al., 1993; Takeda et al., 1996). It is because of the complete biotransformation of Glycyrrhizic acid to Glycyrrhetic acid by the activity of gut microbes. Hence to know the capability of aglycon form, i.e. Glycyrrhetic acid (Enoxolone, not present in main PSM library) was docked and found it also form H bond with SER441 of TMPRSS2, but the BE was reduced to -7.7 (Fig. 2b).

Cis-Miyabenol C, a stilbenoid (resveratrol timer) found in fennel (*Foeniculum vulgare*).. Stilbenoids are characterized by two phenyl group linked by a transethane bond and reported to exhibit a wide range of biological activities and pharmacological properties (Akinwaumi et al., 2018). In the present study, Cis-Miyabenol C recorded a lower BE of -9.4 and found interacting with catalytic site AAR of TMPRSS2, i.e., ASP345 through H bond (Fig 2b; Sup. File 2). This compound was low in GI absorption, highly lipophilic, insoluble in water and found violating Lipinski's rule of drug-likeness. The low bioavailability of stilbenoids is largely due to their rapid, extensive metabolism in the intestine and liver during and after adsorption giving rise to a lower level of the free parent compound (Walle et al., 2011). Pharmacokinetics studies showed a higher level of sulphate and glucuronides (metabolites of stilbene) in plasma.

Therefore it was hypothesized that these metabolites might act as a reservoir for the stilbenoids either by direct action or via enterohepatic recycling (Marier et al., 2002; Walle et al., 2004; Calamini et al., 2010).

Chrysophenols (Anthraquinones) are anthracene derivatives, and structurally they are tricyclic aromatic quinones with two ketone group attached to the central benzene ring. Anthraquinones are reported for their wide spectra of biological activity, including anti-oxidant, anti-ulcer, anti-inflammatory, anti-cancer, neuroprotective, anti-ageing, hepatoprotective, etc. (Diaz-Munoz et al., 2018). Here, two anthraquinone glycosides, Chrysophanein and Rhein-8-glycoside recorded lower BE of -8.9 and -8.9, respectively which was lower than their structurally related flavonoid glycosides, Baicalein 6-glucoside (BE -8.7) and Luteolin 3'-xyloside (BE -8.5). Chrysophaenein showed H-bond with SER441, an AAR of the catalytic site of TMPRSS2 and VAL280 and GLY439 in the catalytic pocket. Chrysophaenein passes all the parameters of Lipinski's rule with 0.55 BAS (Fig. 2b; Sup. File 2 and 3) proving itself as a better drug candidate to inhibit the activity of TMPRSS2.

Plumbagin is a naphthoquinone derivative from the roots of *Plumbago zeylanica* and well studied for their anti-cancerous properties (Hazar et al., 2002; Sugie et al., 1998). In our study, 3,3'-Biplumbagin recorded BE of -8.9 and found interacting with HIS296 and SER441 of the catalytic site of TMPRSS2 and with VAL280 and GLY439 AAR's in the close vicinity, showing its strong affinity towards the catalytic domain of TMPRSS2. Also, it recorded higher BAS of 0.55 and found fulfilling all necessary physicochemical character for a drug-like molecule as per Lipinski's rule (Fig 2b; Sup. File 2 and 3).

Quinic acid is a carboxylated cyclohexanepolyol that is found in several plants like coffee, tomato, carrot, etc. and exist either in free form or as esters (de Maria et al., 1999). Quinic acid is a starting compound used to synthesize "Tamiflu", an antiviral drug used to treat Influenza A and B virus (Federspiel et al., 1999). Derivatives of Quinic acid were reported to be antiviral in nature against Human immunodeficiency virus (HIV), Hepatitis B virus (HBV), Herpes simplex virus 1 and Dengue virus (Mc Dougall et al., 1998; Ikeda et al., 2011; Kamng'ona et al., 2011; Zanello et al., 2015). 3-Caffeoyl-5-Feruloylquinic acid, in our study, recorded -9.0 BE and found interacting with SER441 of TMPRSS2 through H bonding. Following this, similar structure 3-O-Caffeoyloleanolic acid also recorded BE of -8.7. Both the molecules recorded low BAS of 0.11 and violated Lipinski's rule of drug likeness (Fig. 2b; Sup. File 2). Studies of Gonthier et al. (2003) showed that chlorogenic acids, the esters of caffeic and quinic acid are poorly bioavailable. However, their microbial bio-transformed metabolites such as m-coumaric acid and derivatives of phenylproponic, benzoic and hippuric acid were found in higher concentration after oral administration of chlorogenic acids. The anti-viral activity of biotransformed molecules are yet to be studied.

Ellagitannins are the largest group of tannins containing over 500 molecules (Khanbabaee and Van Ree, 2001; Landete et al., 2011). They are bioactive polyphenols present in several fruits and nuts such as Pomegranates, Grapes, Cloudberry, Raspberries, Strawberries, Green tea, Seabuckthorn, Walnuts, Almonds, etc. (Amakura et al., 2000; Clifford and Scalbert, 2000; Landete et al., 2011). Structurally, they are complex class of polyphenols characterized by one or more hexahydroxydiphenoyl (HHDP) moieties esterified to a polyol (Khanbabaee and Van Ree, 2001; Niemetz and Gross, 2005). Ellagitannins are the

most potential anti-oxidant agents because of their high polyphenolic nature. In the present study, group of bulky Ellagitannins such as Terflavin A, Terflavin B, Punicalin, Strictinin, Pedunculagin, Punicafofin, Tellimagradin I, Tercatain, Punicalin, Emblicanin A, Phyllanemblinin B, etc. recorded lower binding efficiency indicating their capability to inhibit the activity of TMPRSS2 enzyme. Similarly, Granatin B an ellagitannin commonly found in the pericarp of *Punica granatum* recorded lower BE of -9.1 followed by Granatin A (-8.9). Granatin B recorded H bond formation with HIS296 of TMPRSS2. Bioavailability score of Granatin B was 0.11 and found violating Lipinski's rule (Fig. 2b; Sup. File 2 and 3). However, they are all high molecular weight, and according to previous studies, they are poorly/not bioavailable. Further, they hydrolyze into Ellagic acid and their complex derivatives in the intestine and can cross the gastrointestinal barrier into the bloodstream (Gil et al., 2000). Whereas gut microflora is reported to convert Ellagitannins to Urolithins an anti-cancerous compound (Seeram et al., 2004; Heber, 2011). Bioavailability of Ellagitannins metabolites was successfully proved in humans in the form of Ellagic acid in blood plasma (Seeram et al., 2008).

Ellagic acid consists of a hydrophilic domain made up of four phenolic groups and two lactones, and a lipophilic domain made up of four rings. These domains, particularly the hydrophilic, which can hydrogen bond and accept an electron, thus determining a structural activity relationship. In our experiments, ET derivatives 3'-O-Methyl ellagic acid 4-xyloside recorded BE of -8.1 against TMPRSS2 (Sup. File 2), indicating the possibilities of Ellagitannin metabolized products as TMPRSS2 inhibitors.

SARS-CoV-2 Main Protease (M^{Pro})

The key enzyme in proteolytic processing of SARS-CoV-2 replication is M^{Pro}. It is initially released by the auto-cleavage of pp1a and pp1ab. Then M^{Pro}, in turn, cleaves pp1a and pp1ab to release functional proteins necessary for viral replication (Krichel et al., 2020). Any PSM binding to the AAR's of the catalytic site or pocket with H bonds and other interactions may interfere with the viral replication in host cell, thereby reducing the severity of the COVID-19.

When BE and canonical smiles structural similarity of top 250 PSM were analyzed, it was observed that molecules from two major groups, i.e., Flavonol glycosides and Anthocyanidine were dominating with $>16\%$ and $>16\%$ PSM, respectively. Other flavonoids and triterpenes also recorded promising results (Fig. 3a). But the least BE was recorded by Hypericine - a naphthodianthrone and Amentoflavone - a biflavonoid.

Hypericin, a naturally occurring chromophore "Naphthodianthrone" compound, derivative of anthraquinone found in common St. John's Wort (*Hypericum* species) and in some fungi. *Hypericum perforatum* a source of Hypericin has been used as folk medicine. Also, Hypericin is reported as anti-depressive, anti-tumor, anti-viral, antineoplastic, etc. (Kubin et al., 2005). Here, Hypericin recorded least BE of -10.4 . Further, it was found forming H bond with GLU166 residues in catalytic sits of M^{Pro} (Fig. 3b; Sup. File 2). Physiochemical parameters of Hypericin showed that it is poorly soluble in water and

violates Lipinski's rule of drug-likeness (Fig. 3b; Sup. File 2) with 0.17 BAS. It is a naturally occurring photosensitizer reported to accumulate in tumour cells and upon illuminating release ROS killing the cancerous cells (Watson et al., 2014). Recent in silico studies revealed its potential to bind SARS-CoV-2 spike protein (Smith and Smith, 2020). Additionally, its antiviral properties against Bronchitis virus (Chen et al., 2019), Hepatitis C virus (Shih et al., 2018) and Human Coronavirus Oc43 (Kim et al., 2019) represents it as a potential antiviral drug material.

Amentoflavone, a biflavonoid (dimer of two apigenin) occurs naturally in more than 120 plants belong to families; Selaginellaceae, Cupressaceae, Euphorbiaceae, Podocarpaceae and Calophyllaceae (Yu et al., 2017). They are reported for wide spectrum of biological activities. Sotetsuflavone, a biflavonoids from *Dacrydium balance* (Gymnosperm) was found to be most potent inhibitor of Dengue virus NS5 RdRp with IC₅₀ of 0.16 µM, among 23 bioflavonids screened. Further, their enzyme inhibitory activity was related to the number and position of methyl groups on the biflavonoid moiety and the degree of oxygenation on flavonoid monomers (Couleric et al., 2013). Gastrointestinal absorption of Amentoflavone was poor and their physicochemical parameters violated Lipinski's rule (Sup file 3). However, its higher bioavailability as intravenous (>70%) administration was reported in rats (Liao et al., 2015). In our studies, Amentoflavone recorded least of -9.7 BE against M^{Pro} and found interacting with target AAR by forming H bonds with GLU166 and other residues in the vicinity of catalytic site (Fig 3b; Sup. File 2). Similarly, other biflavonoids, Agatisflavone also recorded lower BE of -9.3, followed by Ginkgetin and Isoginkgetin, derivatives of Amentoflavone (from *Ginkgo biloba*) were recorded BE -9.5 and -9.5, respectively.

Flavonol glycosides are the most predominant group of PSM found interacting with the target site of M^{Pro}. Flavonols are the class of flavonoids that has 3-hydroxyflavone backbone. Flavonols are present in a large group of plants and their parts. Lower BE of -9.6 was recorded by Quercetin 3,5-diglucoside followed by Myricetin 3-rutinoside (BE -9.4), Rutin (BE -9.3), Kempferol (BE -9.2), Myricetin 3-rhamnoside (BE -9.2), Robinetin 3-rutinoside (BE -9.1), etc. (Fig. 3b; Sup. File 2). It was observed that flavonol with two glucose moiety recorded lower BE when compared to flavonol with one or three glucose moiety. Bioavailability of flavonoids is discussed somewhere else in this paper. Graefe et al. (2001) reported a peak of 2.3 and 2.1 µg/ml of quercetin glucuronides in plasma of human volunteers after oral administration of onion or quercetin- 4'-O Glucoside, respectively. Similarly, a peak of 4.3 and 4.0 µg/ml was recorded for buckwheat tea and rutin oral administration. The authors hypothesized that quercetin-4'-O glucoside and quercetin-3-O glucoside absorption may happen at a different site and depends on the number of sugar moiety. Further, plant matrix may have an additional effect on GI absorption of these flavonol glycosides.

Proanthocyanidins are a class of oligomeric flavonoids (condensed tannins) found in a variety of plants. They are formed by the condensation of catechine and epicatechin units. They are well known for their anti-oxidant, anti-cancerous, anti-diabetic, immunomodulatory, cardioprotective, and neuroprotective activities. Found in flowers, fruits, bark, nuts and seeds of various plants (Rauf et al., 2019; Bao et al., 2015). Studies on Proanthocyanidins reveals it is poorly bioavailable, its monomers viz., (+)-Catechin and (-)-epicatechin have been previously reported to cross GI track in its monomeric form or as conjugated

metabolites (Baba et al., 2001). Further, its dimers are also reported to present in plasma and urine after consuming food and beverages rich in Proanthocyanidine (Holt et al., 2002; Urpi-Sarda et al., 2009). In our studies, Proanthocanidin A2 and Proanthocanidin B2 recorded least of -9.4 BE followed by Procyanidine B1 (BE -9.3). Further, their monomers/dimers/conjugates viz., (-)-Gallocatechin gallate (BE -9.1), Epicatechin-3-gallate (BE -9.0), Gallocatechin - (4 α \rightarrow 8) epigallocatechin (BE -8.6) were found effective (Sup. File 2).

Glycosidic anthocyanidin (anthocyanins), water-soluble pigments found in many fruits like grapes, bilberry, raspberry, blackberry, cherry, blueberry, etc. Here, Cyanidine-3,5, diglucoside recorded -9.4 BE followed by Penodine 3,5 diglucoside (BE -9.3). Further H bond formation between Cyanidine-3,5, diglucoside and AAR GLY143 and SER144 of the catalytic site of M^{pro} was observed. It recorded a BAS of 0.17 and did not fulfil the criteria of a drug molecule as per Lipinski's rule (Fig 3b; Sup. File 2). Bioavailability of Cyanidine-3-glucoside was studied in human trials by Czank et al. (2013) following the isotopically labelled compound. They found radioactivity in plasma, urine and breathe, indicating the absorption of the labelled compound. However, the study did not reveal the existence of bioactive metabolites; it may be in degraded forms of the parental compound.

Terflavin B, an ellagitannin (found in *Terminalia chebula* and *Terminalia catapa*) recorded BE of -9.7 . Also, it was found forming H bonds with AARs HIS41, GLY143, SER144, CYS145 and GLU166 present in the catalytic site of SARS-CoV-2 M^{pro} . It is water-soluble, low in GI absorption and violates Lipinski's rule (Fig 3b; Sup. File 2). Aqueous extracts of *Terminalia chebula* was reported to be inhibitory to Hepatitis B virus infection in Hep G 2.2.15 cells (Kim et al., 2001). Not much research studies on their bioavailability and antiviral properties have been done with purified Terflavin B. Vescalagin, and Castalagin are ellagitannins found in oak and chestnut wood (Peng et al., 1991; Tanalka et al., 1996). They are water-soluble, high oxidizable and astringent (Vivas et al., 2004). Here, Viscalagin recorded -9.6 BE and found interacted with GLU166 residues of M^{pro} catalytic site through H bond, and Castalagin showed -8.8 BE (Fig. 3b; Sup. File 2). Viscalagin is high molecular weight, and high polar molecule recorded lower BAS 0.17. Vilhelmova et al. (2011) demonstrated the antiviral activity of Vescalagin and Castalagin against Herpes simplex virus type I and II. Also, these PMS synergistically inhibited the multiplication of test virus along with antiviral compound Acyclovir.

Mudanpioside J is a monoterpene glucoside. This compound is one among those identified in *Paeonia delavayi* (chinese medicinal plant), which was found inhibitory to influenza virus neuraminidase (Li et al., 2016). In our studies, Mudanpioside J recorded BE of -9.6 and showed strong interaction with GLY143, SER144 and CYS145 residues of M^{pro} catalytic site through H bonds (Fig 3b; Sup. File 2). Studies on Pharmacokinetics properties of monoterpene glycosides are limited. Paeoniflorin and albiflorin, Mudanpioside J related compounds were reported as low in oral bioavailability due to its poor membrane permeability and gut microbes-induced metabolism (Takeda et al., 1997; Liu et al., 2006). Swiss ADME analysis of Mudanpioside J also indicated its P-glycoprotein substrate nature and violation of Lipinski's rule with 0.17 BAS (Sup file 3).

Arylnaphthalene lignan glycosides are PSM widely distributed in family Acanthaceae. Lignans were reported for their wide biological activities, including antiviral against Hepatitis B, Hepatitis C, Herpes simplex virus type 1 and 2, Epstein-Barr virus and Cytomegalovirus (Ye et al., 2005; Charlton, 1988). Justalakonin, an Arylnaphthalene lignan glycoside isolated from *Justicia purpurea*. In the present study, it recorded BE of -9.4 and showed interaction with AAR GLY143 and SER144 of the catalytic site of M^{Pro} through H bonding (Fig. 3b; Sup. File 2). Physicochemical characterization revealed it was found violating Lipinski's rule with 0.17 BAS. Regarding its metabolism and bioavailability, not much research studies have been done. Similarly, structurally related PSM's, Elenoside (BE -8.9), Patentiflorin B (BE -8.9) and Patentiflorin A (BE -8.6) also recorded lower BE against M^{Pro} (Sup. File 2).

SARS-CoV-2 RNA Dependent RNA Polymerase (RdRp)

The central component of coronaviral replication/transcription machinery is RNA-dependent RNA polymerase (RdRp, also named nsp12) that constructs copies of its RNA genome playing the key role in replication and transcription of SARS-CoV-2 in the host cell (Subissi et al., 2014; Gao et al., 2020). Because of its high sequence and structural conservation, it remains the target of choice for the prophylactic or curative treatment of several viral diseases. Studies on structural activity relationship and BE of PSM revealed that a large number of flavonol glycoside ($>13\%$) followed by hydrolysable tannins, anthocyanins and triterpenes were found interacting with target site of RdRp with lower BE (Fig. 4a). Though the number of flavonol glycoside was high, the lowest BE was recorded by Erodictyol-7-O-glycoside and Narirutin belongs to group flavanone glycosides. Interestingly none of the PSM analyzed was found interacting with VAL557 of RdRp, but they found forming H bond with AAR's in the catalytic pocket probably sterically hinders the substrate interaction with the catalytic site.

Erodictyol-7-O-glycoside is a flavanone glycoside commonly found in lemons (Miyake et al., 1970), also called as lemon or citrus flavonoid. It recorded -9.9 BE and found interacting with ARG553, ALA554, THR556, ASP618, TRY619 and ASP623 in the active site of RdRp. Both glycone and aglycone moiety of this molecule were found involved in H bond formation with the target site of RdRp. Similarly, Narirutin another flavanone glycoside which naturally present in sweet oranges recorded lower BE of -9.7 and recorded a similar pattern of H bond formation with active site of RdRp. Both molecules recorded BAS of 0.17 , indicating their low bioavailability and found violating Lipinski's rule (Fig. 4b; Sup. File 2). Structurally similar compounds, Nirurin from *Phyllanthus niruri* and Naringin from grapefruits also showed promising results with BE -9.0 and -8.9 , respectively.

Myricetin 3-rutinoside is flavonol glycoside isolated from *Chrysobalanus icaco* (Barbosa et al., 2006) and from other related plant species recorded -9.5 BE. This molecule was found forming a large number of H bond with AAR's ASP452, TYR456, ARG553, ALA554, ARG555, CYS622, ASP623 and SER682 showing its capability of strongly interacting with the target protein. Similarly, Kempferol (flavonol glycoside) also recorded -9.5 BE. For both the compound BAS was found to 0.11 (Fig. 4b; Sup. File 2 and 3). Rhoifolin and Isorhoifolin are flavone glycoside isolated from *Rhus succedanea* (Hattori et al., 1952) and several

citrus plants and onions. It is reported as anti-oxidant, anti-inflammatory, anti-microbial, hepatoprotective and anti-cancerous (reviewed by Refaat et al., 2015). Though they recorded lower BE, Rhoifolin and Isorhoifolin were not interacting with AAR's close to the active site of RdRp. Isoginkgetin, a biflavonoid from *Ginkgo biloba* recorded lower BE of -9.5 followed by Agatisflavone (BE -9.4) and ginkgetin (BE -9.2). Isoginkgetin was found interacting with ARG553, THR556, THR687 and ALA685 of RdRp catalytic pocket. Compared to flavonoids glycosides, aglycons were frequently reported to be bioavailable. Hence, Myricetin, a monomeric aglyconic flavonoid, was analyzed, and record BE -8.4. Further, Myricetin showed H bond with ARG553, ARG555, ARG624 and THR680 of RdRp active site, proving itself as a better candidate for developing drug targeting viral RdRp. Similarly, Quercetin glycosides were found promising, but quercetin was not able to appear in the list of top ranked PSM.

Rotundioside B is a triterpenoid glycoside (Saponin) found in *Bupleurum rotundifolium* (Akai et al., 1985) and reported as anti-inflammatory, anti-proliferative (Navarro et al., 2001; Fujioka et al., 2003). Here, Rotundioside B recorded -9.5 BE and formed H bond with ASP452, SER501, ALA554, ASP623, ARG624, ASN691, SER759 and ASP760 of RdRp catalytic pocket proving itself as a strong contender of antiviral drug. However, this molecule recorded poor BAS of 0.11 and violated Lipinski's rule of drug-likeness (Fig. 4b; Sup. File 2). Similar to this, Ginsenoside Ro, a triterpene saponin from roots of *Panax ginseng* and related plants recorded -9.4 BE. Trigofoenoside G, a steroidal saponin from plants and seeds of *Trigonella foenum-graecum* was found interacting with RdRp with -9.3 BE and forms H bonds with ASP452, LYS500, ALA554, ALA558, LYS621 and ALA685 of catalytic pocket (Fig 4b; Sup. File 2). Following to this, Trigofoenoside F and Trigofoenoside A recorded -8.9 and -8.4 BE, respectively (Sup file 2).

Hippomanin A, an ellagitannin found in *Hippomane mancinella* plant is known for its toxic properties. It causes oropharyngeal and gastrointestinal tract lesions, hypotension and bradycardia (Boucaud-Maiter et al., 2019; Rao, 1974). Here, Hippomannin A recorded -9.6 BE and forms H bond with ARG553, TYR619, ARG624, THR680, SER682 and ASP760 with target site of RdRp. A structurally similar molecule, Tellimagrandin I found widely in fruits, nuts and vegetables. Tellimagrandin was reported for their wide spectra of biological activities (reviewed by Zheng et al., 2012). Here, it recorded -9.5 BE followed by Punigluconin (BE -9.3). Other hydrolyzable tannins viz., Emblicanin A and Phyllanemblinin C found in fruits of Indian Gooseberry (*Emblica officinalis*) were recorded lower BE of -9.4 and -9.3, respectively. As the above discussed hydrolyzable tannins are highly water-soluble and larger in molecular weight, their bioavailability in original form is a major concern. Some of the molecules, structure similar to their biotransformed metabolites with higher bioavailability, like Ellagic acid (BE -8.3), 3'-O-methyl ellagic acid-4-xyloside (BE -8.4), and 3,3'-Di-O-methyl ellagic acid (BE -8.2) also recorded promising results. Indicating the possible involvement of ellagitannins-biotransformed products to inhibit RdRp activity thereby reducing the COVID-19 severity.

Conclusion

From the obtained results, it could be concluded that virtual screening through molecular docking is a promising preliminary step towards developing an effective drug against a desired target protein/enzyme. Here, more than the BE of a PSM, its bioavailability play a crucial role in determining its biological activity under actual conditions. Triterpenoid based PSM structures (Coagulins, Withanolides, Pseudojervine, Kamalachalcone, etc.) are hypothesized as potent drug molecules which can be used to block surface AARs of spike protein which interacts with hACE2, thereby preventing host cell recognition by SARS-CoV2. In the case of TMPRSS2, Mpro and RdRp, molecules belongs to group flavonoid glycosides, biflavonoids, ellagitannins, anthocyanidins, triterpens, etc. can be explored. Though the large number of PSMs were found violating Lipinski's rule and recorded lower BAS, these PSMs cannot be ignored. Because, several biotransformed structure of these PSMs are highly bioavailable and they may retain structural moiety of the parental compound. These biotransformed molecules may further interact with target site of protein and exert a similar results as observed in molecular docking studies. Considering the safety, though these PSM are of natural origin, they should be tested thoroughly under *in vitro* and *in vivo* conditions for their biological activity, biotransformation, bioavailability and toxicity before taking them into the clinical trials.

Declarations

Acknowledgment

The authors would like to thank all the research scholars of Environmental Biotechnology lab, Centre for Rural Development and Technology, Indian Institute of Technology Delhi, India for their constant support in conducting this work during the period of lockdown. Our special thanks to Head, Centre for Rural Development and Technology and Director, Indian Institute of Technology Delhi, India, for providing all facilities needed to conduct this research work.

Competing interests

The authors declare no competing interests.

References

- Amakura, Y., Okada, M., Tsuji, S. & Tonogai, Y. High-performance liquid chromatographic determination with photodiode array detection of ellagic acid in fresh and processed fruits. *J. Chromatogr. A* 896, 87–93, [https://doi.org/10.1016/s0021-9673\(00\)00414-3](https://doi.org/10.1016/s0021-9673(00)00414-3) (2000).
- Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J. R. & Hilgenfeld, R. Coronavirus main proteinase (3CL^{pro}) structure: basis for design of anti-SARS drugs. *Science* 300, 1763–1767, <https://doi.org/10.1126/science.1085658> (2003).
- Baba, S. *et al.* *In vivo* comparison of the bioavailability of (+)-Catechin, (–)-Epicatechin and their mixture in orally administered rats. *J. Nutr.* 131, 2885–2891, <https://doi.org/10.1093/jn/131.11.2885> (2001).

- Barbosa, W. L. R., Peres, A., Gallori, S. & Vincieri, F. F. Determination of myricetin derivatives in *Chrysobalanus icaco* L. (Chrysobalanaceae). *Rev. Bras. Farmacogn.* 16, 333–337, <https://doi.org/10.1590/S0102-695X2006000300009> (2006).
- Belouzard, S., Chu, V. C. & Whittaker, G. R. Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. *Proc. Natl. Acad. Sci. U.S. A.* 106, 5871–5876, <https://doi.org/10.1073/pnas.0809524106> (2009).
- Biasini, M. *et al.* SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* 42, W252–W258, <https://doi.org/10.1093/nar/gku340> (2014).
- Calamini, B. *et al.* Pleiotropic mechanisms facilitated by resveratrol and its metabolites. *Biochem. J.* 429, 273–282, <https://doi.org/10.1042/BJ20091857> (2010).
- Calligari, P., Bobone, S., Ricci, G. & Bocedi, A. Molecular investigation of SARS-CoV-2 proteins and their interactions with antiviral drugs. *Viruses* 12, 445. <https://doi.org/10.3390/v12040445> (2020).
- Chen, H. *et al.* Antiviral activity against infectious bronchitis virus and bioactive components of *Hypericum perforatum* L. *Front. Pharmacol.* 10, 1272, <https://doi.org/10.3389/fphar.2019.01272> (2019).
- Clifford, M. N. & Scalbert, A. Ellagitannins—nature, occurrence and dietary burden. *J. Sci. Food Agric.* 80, 1118–1125, [https://doi.org/10.1002/\(SICI\)1097-0010\(20000515\)80:7<1118::AID-JSFA570>3.0.CO;2-9](https://doi.org/10.1002/(SICI)1097-0010(20000515)80:7<1118::AID-JSFA570>3.0.CO;2-9) (2000).
- Coulerie, P. *et al.* Structure-activity relationship study of biflavonoids on the Dengue virus polymerase DENV-NS5 RdRp. *Planta Med.* 79, 1313–1318, <https://doi.org/10.1055/s-0033-1350672> (2013).
- Daina, A., Michielin, O. & Zoete, V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* 7, 42717, <https://doi.org/10.1038/srep42717> (2017).
- Dallakyan, S. & Olson, A. J. Small-molecule library screening by docking with PyRx. *Methods Mol. Biol.* 1263, 243–250, https://doi.org/10.1007/978-1-4939-2269-7_19 (2015).
- Del Hierro, J. N., Herrera, T., Fornari, T., Reglero, G. & Martin, D. The gastrointestinal behavior of saponins and its significance for their bioavailability and bioactivities. *J. Funct. Foods* 40, 484–497, <https://doi.org/10.1016/j.jff.2017.11.032> (2018).
- Doubl  , S. & Ellenberger, T. The mechanism of action of T7 DNA polymerase. *Curr. Opin. Struct. Biol.* 8, 704–712, [https://doi.org/10.1016/S0959-440X\(98\)80089-4](https://doi.org/10.1016/S0959-440X(98)80089-4) (1998).
- Duquerroy, S., Vigouroux, A., Rottier, P. J. M., Rey, F. A. & Bosch, B. J. Central ions and lateral asparagine/glutamine zippers stabilize the post-fusion hairpin conformation of the SARS coronavirus spike glycoprotein. *Virology* 335, 276–285, <https://doi.org/10.1016/j.virol.2005.02.022> (2005).

- Dzubak, P. *et al.* Pharmacological activities of natural triterpenoids and their therapeutic implications. *Nat. Prod. Rep.* 22, 487, <https://doi.org/10.1039/B515312N> (2005).
- Elbe, S. & Buckland-Merrett, G. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob. Chall.* 1, 33–46, <https://doi.org/10.1002/gch2.1018> (2017).
- Elfiky, A. A. Anti-HCV, nucleotide inhibitors, repurposing against COVID–19. *Life Sci.* 248, 117477, <https://doi.org/10.1016/j.lfs.2020.117477> (2020).
- Elfiky, A. A. & Ismail, A.M. Molecular docking revealed the binding of nucleotide/side inhibitors to Zika viral polymerase solved structures. *SAR QSAR Environ. Res.* 29, 409–418, <https://doi.org/10.1080/1062936X.2018.1454981> (2018).
- Fujioka, T. *et al.* Antiproliferative constituents from Umbelliferae plants VI. New ursane-type saikosaponin analogs from the fruits of *Bupleurum rotundifolium*. *Chem. Pharm. Bull. (Tokyo)* 51, 365–372, <https://doi.org/10.1248/cpb.51.365> (2003).
- Gao, Y. *et al.* Structure of the RNA-dependent RNA polymerase from COVID–19 virus. *Science* eabb7498, <https://doi.org/10.1126/science.abb7498> (2020).
- Gautret, P. *et al.* Hydroxychloroquine and azithromycin as a treatment of COVID–19: results of an open-label non-randomized clinical trial. *Int. J. Antimicrob. Agents* 105949, <https://doi.org/10.1016/j.ijantimicag.2020.105949> (2020).
- Gil, M. I., Tomás-Barberán, F. A., Hess-Pierce, B., Holcroft, D. M. & Kader, A. A. Antioxidant activity of pomegranate juice and its relationship with phenolic composition and processing. *J. Agric. Food Chem.* 48, 4581–4589, <https://doi.org/10.1021/jf000404a> (2000).
- Glötter, E. Withanolides and related ergostane-type steroids. *Nat. Prod. Rep.* 8, 415–440, <https://doi.org/10.1039/np9910800415> (1991).
- Godugu, C., Patel, A. R., Doddapaneni, R., Somagoni, J. & Singh, M. Approaches to improve the oral bioavailability and effects of novel anticancer drugs berberine and betulinic acid. *PLoS One.* 9, e89919, <https://doi.org/10.1371/journal.pone.0089919> (2014).
- Gordon, C. J., Tchesnokov, E. P., Feng, J. Y., Porter, D. P. & Götte, M. The antiviral compound remdesivir potently inhibits RNA-dependent RNA polymerase from Middle East respiratory syndrome coronavirus. *J. Biol. Chem.* 295, 4773–4779, <https://doi.org/10.1074/jbc.AC120.013056> (2020).
- Graefe, E. U. *et al.* Pharmacokinetics and bioavailability of quercetin glycosides in humans. *J. Clin. Pharmacol.* 41, 492–499, <https://doi.org/10.1177/00912700122010366> (2001).
- Gui, M. *et al.* Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding. *Cell Res.* 27, 119–129, <https://doi.org/10.1038/cr.2016.152>

(2017).

Hattori, S. & Matsuda, H. Rhoifolin, a new flavone glycoside, isolated from the leaves of *Rhus succedanea*. *Arch. Biochem. Biophys.* 37, 85–89, [https://doi.org/10.1016/0003-9861\(52\)90164-1](https://doi.org/10.1016/0003-9861(52)90164-1) (1952).

He, Y. *et al.* Recent advances in biotransformation of saponins. *Molecules*. 24, 2365, <https://doi.org/10.3390/molecules24132365> (2019).

Hilgenfeld, R. From SARS to MERS: crystallographic studies on coronaviral proteases enable antiviral drug design. *FEBS J.* 281, 4085–4096, <https://doi.org/10.1111/febs.12936> (2014).

Holeski, L. M., Hillstrom, M. L., Whitham, T. G. & Lindroth, R. L. Relative importance of genetic, ontogenetic, induction, and seasonal variation in producing a multivariate defense phenotype in a foundation tree species. *Oecologia* 170, 695–707, <https://doi.org/10.1007/s00442-012-2344-6> (2012).

Hui, D. S. *et al.* The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health—The latest 2019 novel coronavirus outbreak in Wuhan, China. *Int. J. Infect. Dis.* 91, 264–266, <https://doi.org/10.1016/j.ijid.2020.01.009> (2020).

Hulswit, R. J. G. *et al.* Human coronaviruses OC43 and HKU1 bind to 9-O-acetylated sialic acids via a conserved receptor-binding site in spike protein domain A. *Proc. Natl. Acad. Sci. U.S. A.* 116, 2681–2690, <https://doi.org/10.1073/pnas.1809667116> (2019).

Jacobson, J. M. *et al.* Pharmacokinetics, safety, and antiviral effects of Hypericin, a derivative of St. John's Wort plant, in patients with chronic hepatitis C virus infection. *Antimicrob. Agents Chemother.* 45, 517–524, <https://doi.org/10.1128/AAC.45.2.517-524.2001> (2001).

Jin, Z. *et al.* Structure of M^{Pro} from COVID–19 virus and discovery of its inhibitors. *Nature* <https://doi.org/10.1038/s41586-020-2223-y> (2020).

Khaerunnisa, S., Kurniawan, H., Awaluddin, R., Suhartati, S. & Soetjipto, S. Potential inhibitor of COVID–19 main protease (M^{Pro}) from several medicinal plant compounds by molecular docking study. *Preprints* 2020030226, <https://doi.org/10.20944/preprints202003.0226.v1> (2020).

Khanbabaee, K. & van Ree, T. Tannins: classification and definition. *Nat. Prod. Rep.* 18, 641–649, <https://doi.org/10.1039/b101061I> (2001).

Kim, D. E. *et al.* Natural bis-benzylisoquinoline alkaloids-tetrandrine, fangchinoline, and cepharanthine, inhibit human coronavirus OC43 infection of MRC–5 human lung cells. *Biomolecules* 9, 696, <https://doi.org/10.3390/biom9110696> (2019).

Kim, T. G. *et al.* Antiviral activities of extracts isolated from *Terminalis chebula* Retz., *Sanguisorba officinalis* L., *Rubus coreanus* Miq. and *Rheum palmatum* L. against hepatitis B virus. *Phytother. Res.* 15,

718–720, <https://doi.org/10.1002/ptr.832> (2001).

Knölker, H. J. & Reddy, K. R. Biological and pharmacological activities of carbazole alkaloids. *Alkaloids Chem. Biol.* 65, 181–193, [https://doi.org/10.1016/S1099-4831\(07\)00004-1](https://doi.org/10.1016/S1099-4831(07)00004-1) (2008).

Kong, R. *et al.* COVID–19 Docking Server: An interactive server for docking small molecules, peptides and antibodies against potential targets of COVID–19. *arXiv* <https://arxiv.org/abs/2003.00163> (2020).

Kulkarni, R. R. *et al.* Antifungal dimeric chalcone derivative kamalachalcone E from *Mallotus philippinensis*. *Nat. Prod. Res.* 28, <https://doi.org/10.1080/14786419.2013.843178> (2014).

Kumar, S. & Pandey, A. K. Chemistry and biological activities of flavonoids: an overview. *ScientificWorldJournal* 2013, 162750, <https://doi.org/10.1155/2013/162750> (2013).

Lan, J. *et al.* Structure of the SARS-CoV–2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* <https://doi.org/10.1038/s41586-020-2180-5> (2020).

Lehmann, K. C. *et al.* Discovery of an essential nucleotidylating activity associated with a newly delineated conserved domain in the RNA polymerase-containing protein of all nidoviruses. *Nucleic Acids Res.* 43, 8416–8434. <https://doi.org/10.1093/nar/gkv838> (2015).

Li, Z. *et al.* FEP-based screening prompts drug repositioning against COVID–19. *bioRxiv* <https://doi.org/10.1101/2020.03.23.004580> (2020).

Li, Z. W., Ma, X. Y., Wang, M. N., Liang, W. W. & Wang, Y. H. Investigation of pharmacokinetics of oxymatrine soft capsule in healthy volunteers. *Northwest Pharm. J.* 19, 149–151, (2004).

Liao, S. *et al.* Liquid chromatography–tandem mass spectrometry determination and pharmacokinetic analysis of Amentoflavone and its conjugated metabolites in rats. *J. Agric. Food Chem.* 63, 1957–1966, <https://doi.org/10.1021/jf5019615> (2015).

Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 46, 3–26, [https://doi.org/10.1016/S0169-409X\(00\)00129-0](https://doi.org/10.1016/S0169-409X(00)00129-0) (2001).

Liu, Z. Q., Jiang, Z. H., Liu, L. & Hu, M. Mechanisms responsible for poor oral bioavailability of paeoniflorin: Role of intestinal disposition and interactions with sinomenine. *Pharm. Res.* 23, 2768–2780, <https://doi.org/10.1007/s11095-006-9100-8> (2006).

Lung, J. *et al.* The potential chemical structure of anti-SARS-CoV-2 RNA-dependent RNA polymerase. *J. Med. Virol.* 92, 693–697, <https://doi.org/10.1002/jmv.25761> (2020).

Madu, I. G., Roth, S. L., Belouzard, S., & Whittaker G. R. Characterization of a highly conserved domain within the severe acute respiratory syndrome coronavirus spike protein s2 domain with characteristics of

- a viral fusion peptide. *J. Virol.* 83, 7411–7421, <https://doi.org/10.1128/JVI.00079-09> (2009).
- Martin, Y. C. A bioavailability score. *J. Med. Chem.* 48, 3164–3170, <https://doi.org/10.1021/jm0492002> (2005).
- McDougall, B. *et al.* Dicafeoylquinic and dicafeoyltartaric acids are selective inhibitors of human immunodeficiency virus type 1 integrase. *Antimicrob. Agents Chemother.* 42, 140–146, (1998).
- Meng, T. *et al.* The transmembrane serine protease inhibitors are potential antiviral drugs for 2019-nCoV targeting the insertion sequence-induced viral infectivity. *bioRxiv* <https://doi.org/10.1101/2020.02.08.926006> (2020).
- Millet, J. K. & Whittaker, G. R. Host cell proteases: Critical determinants of coronavirus tropism and pathogenesis. *Virus Res.* 202, 120–134, <https://doi.org/10.1016/j.virusres.2014.11.021> (2015).
- Miyake, Y., Yamamoto, K. & Osawa, T. Isolation of eriocitrin (eriodictyol 7-rutinoside) from lemon fruit (*Citrus limon* BURM. f.) and its antioxidative activity. *Food Sci. Technol. Int. Tokyo* 3, 84–89, <https://doi.org/10.3136/fsti9596t9798.3.84> (1997).
- Nalli, Y. *et al.* Four new carbazole alkaloids from *Murraya koenigii* that display anti-inflammatory and antimicrobial activities. *Org. Biomol. Chem.* 14, 3322–3332, <https://doi.org/10.1039/C6OB00267F> (2016).
- Niemetz, R. & Gross, G. G. Enzymology of gallotannin and ellagitannin biosynthesis. *Phytochemistry* 66, 2001–2011, <https://doi.org/10.1016/j.phytochem.2005.01.009> (2005).
- Pachetti, M. *et al.* Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant. *J. Transl. Med.* 18, 179, <https://doi.org/10.1186/s12967-020-02344-6> (2020).
- Panche, A. N., Diwan, A.D. & Chandra, S. R. Flavonoids: an overview. *J. Nutr. Sci.* 5, e47, <https://doi.org/10.1017/jns.2016.41> (2016).
- Pant, S., Singh, M., Ravichandiran, V., Murty, U.S. N. & Srivastava, H. K. Peptide-like and small-molecule inhibitors against Covid-19. *J. Biomol. Struct. Dyn.* <https://doi.org/10.1080/07391102.2020.1757510> (2020).
- Pettersen, E. F. *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612, <https://doi.org/10.1002/jcc.20084> (2004).
- Pichersky, E. & Gang, D. R. Genetics and biochemistry of secondary metabolites in plants: an evolutionary perspective. *Trends Plant Sci.* 5, 439–445, [https://doi.org/10.1016/S1360-1385\(00\)01741-6](https://doi.org/10.1016/S1360-1385(00)01741-6) (2000).
- Pirillo, A. & Catapano, A. L. Berberine, a plant alkaloid with lipid- and glucose-lowering properties: From *in vitro* evidence to clinical studies. *Atherosclerosis* 243, 449–461, <https://doi.org/10.1016/j.atherosclerosis.2015.09.032> (2015).

- Pompei, R., Flore, O., Marccialis, M. A., Pani, A. & Loddo, B. Glycyrrhizic acid inhibits virus growth and inactivates virus particles. *Nature* 281, 689–690, <https://doi.org/10.1038/281689a0> (1979).
- Rao, K. V. Toxic principles of *Hippomane mancinella*. *Planta Med.* 25, 166–171, <https://doi.org/10.1055/s-0028-1097927> (1974).
- Rappe, A. K., Casewit, C. J., Colwell, K. S., Goddard III, W. A. & Skiff, W. M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.* 114, 10024–10035, <https://doi.org/10.1021/ja00051a040> (1992).
- Refaat, J., Desoukey, S. Y., Ramadan, M. A. & Kamel, M. S. Rhoifolin: A review of sources and biological activities. *Int. J. Pharmacogn.* 2, 102–109, [https://doi.org/10.13040/IJPSR.0975-8232.IJP.2\(3\).102-09](https://doi.org/10.13040/IJPSR.0975-8232.IJP.2(3).102-09) (2015).
- Rossen, J. W. A. *et al.* The viral spike protein is not involved in the polarized sorting of coronaviruses in epithelial cells. *J. Virol.* 72, 497–503, (1998).
- Seeram, N. P., Lee, R. & Heber, D. Bioavailability of ellagic acid in human plasma after consumption of ellagitannins from pomegranate (*Punica granatum* L.) juice. *Clin. Chim. Acta* 348, 63–68, <https://doi.org/10.1016/j.cccn.2004.04.029> (2004).
- Seeram, N. P. *et al.* Pomegranate juice and extracts provide similar levels of plasma and urinary ellagitannin metabolites in human subjects. *J. Med. Food* 11, 390–394, <https://doi.org/10.1089/jmf.2007.650> (2008).
- Serra, A. *et al.* Bioavailability of procyanidin dimers and trimers and matrix food effects in *in vitro* and *in vivo* models. *Br. J. Nutr.* 103, 944–952, <https://doi.org/10.1017/S0007114509992741> (2010).
- Shi, J. & Song, J. The catalysis of the SARS 3C-like protease is under extensive regulation by its extra domain. *FEBS J.* 273, 1035–1045, <https://doi.org/10.1111/j.1742-4658.2006.05130.x> (2006).
- Shih, C. M., Wu, C. H., Wu, W. J., Hsiao, Y. M. & Ko, J. L. Hypericin inhibits hepatitis C virus replication via deacetylation and down-regulation of heme oxygenase–1. *Phytomedicine* 46, 193–198, <https://doi.org/10.1016/j.phymed.2017.08.009> (2018).
- Shulla, A. *et al.* A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry. *J. Virol.* 85, 873–882, <https://doi.org/10.1128/JVI.02062-10> (2011).
- Smith, M. & Smith, J. C. Repurposing therapeutics for COVID–19: Supercomputer-based docking to the SARS-CoV–2 viral spike protein and viral spike protein-human ACE2 interface. *ChemRxiv* <https://doi.org/10.26434/chemrxiv.11871402.v4> (2020).

- Subissi, L. *et al.* One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. *Proc. Natl. Acad. Sci. U.S. A.* *111*, E3900-E3909, <https://doi.org/10.1073/pnas.1323705111> (2014).
- Takeda, S. *et al.* *In-vivo* assessment of extrahepatic metabolism of paeoniflorin in rats: relevance to intestinal floral metabolism. *J. Pharm. Pharmacol.* *49*, 35–39, <https://doi.org/10.1111/j.2042-7158.1997.tb06748.x> (1997).
- Tortorici, M. A. & Veessler, D. Structural insights into coronavirus entry. *Adv. Virus Res.* *105*, 93–116, <https://doi.org/10.1016/bs.aivir.2019.08.002> (2019).
- Trott, O. & Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* *31*, 455–461, <https://doi.org/10.1002/jcc.21334> (2010).
- Varshney, I. P., Jain, D.C. & Srivastava, H. C. Saponins from *Trigonella foenum-graecum* leaves. *J. Nat. Prod.* *47*, 44–46, <https://doi.org/10.1021/np50031a002> (1984).
- Vilhelmova, N., Jacquet, R., Quideau, S., Stoyanova, A. & Galabov, A. S. Three-dimensional analysis of combination effect of ellagitannins and acyclovir on herpes simplex virus types 1 and 2. *Antiviral Res.* *89*, 174–181, <https://doi.org/10.1016/j.antiviral.2010.11.014> (2011).
- Vincken, J. P., Heng, L., de Groot, A. & Gruppen, H. Saponins, classification and occurrence in the plant kingdom. *Phytochemistry* *68*, 275–297, <https://doi.org/10.1016/j.phytochem.2006.10.008> (2007).
- Vlasak, R., Luytjes, W., Spaan, W. & Palese, P. Human and bovine coronaviruses recognize sialic acid-containing receptors similar to those of influenza C viruses. *Proc. Natl. Acad. Sci. U.S. A.* *85*, 4526–4529, <https://doi.org/10.1073/pnas.85.12.4526> (1988).
- Walle, T. Bioavailability of resveratrol. *Ann. N. Y. Acad. Sci.* *1215*, 9–15, <https://doi.org/10.1111/j.1749-6632.2010.05842.x> (2011).
- Walle, T., Hsieh, F., DeLegge, M. H., Oatis, J. E. & Walle, U. K. High absorption but very low bioavailability of oral resveratrol in humans. *Drug Metab. Dispos.* *32*, 1377–1382, <https://doi.org/10.1124/dmd.104.000885> (2004).
- Walls, A. C. *et al.* Structure, function, and antigenicity of the SARS-CoV-2 Spike glycoprotein. *Cell* *181*, 281–292, <https://doi.org/10.1016/j.cell.2020.02.058> (2020).
- Wang, L., Wang, Y., Ye, D. & Liu, Q. Review of the 2019 novel coronavirus (SARS-CoV-2) based on current evidence. *Int. J. Antimicrob. Agents* *105948*, <https://doi.org/10.1016/j.ijantimicag.2020.105948> (2020).
- Wu, F. *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* *579*, 265–269, <https://doi.org/10.1038/s41586-020-2008-3> (2020).

- Xu, X. *et al.* 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.* 63, 457–460, <https://doi.org/10.1007/s11427-020-1637-5> (2020).
- Yang, H. *et al.* The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proc. Natl. Acad. Sci. U.S. A.* 100, 13190–13195, <https://doi.org/10.1073/pnas.1835675100> (2003).
- Yin, M. C., Lin, M. C., Mong, M. C. & Lin, C. Y. Bioavailability, distribution, and antioxidative effects of selected triterpenes in mice. *J. Agric. Food Chem.* 60, 7697–7701, <https://doi.org/10.1021/jf302529x> (2012).
- Yin, Y. & Wunderink, R. G. MERS, SARS and other coronaviruses as causes of pneumonia. *Respirology* 23, 130–137, <https://doi.org/10.1111/resp.13196> (2018).
- Yu, L., Sun, Y., Su, J. & Li, X. Bismahanine exerts anticancer effects on human cervical cancer cells by inhibition of growth, migration and invasion via suppression of NF- κ B signalling pathway. *J. BUON* 25, 93–98, (2020).
- Yu, S. *et al.* A review on the phytochemistry, pharmacology, and pharmacokinetics of amentoflavone, a naturally-occurring biflavonoid. *Molecules* 22, 299, <https://doi.org/10.3390/molecules22020299> (2017).
- Zhang, L. *et al.* Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α -ketoamide inhibitors. *Science* 368, 409–412, <https://doi.org/10.1126/science.abb3405> (2020).
- Zhao, L. *et al.* Effect of (E)-2-isopropyl-5-methylcyclohexyl octadec-9-enoate on transdermal delivery of Aconitum alkaloids. *Drug Dev. Ind. Pharm.* 37, 290–299, <https://doi.org/10.3109/03639045.2010.510141> (2011).
- Zheng, S. *et al.* Synthesis and biological profiling of tellimagrandin I and analogues reveals that the medium ring can significantly modulate biological activity. *Org. Biomol. Chem.* 10, 2590–2593. <https://doi.org/10.1039/c2ob25065a> (2012).
- Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579, 270–273, <https://doi.org/10.1038/s41586-020-2012-7> (2020).
- Zhu, N. *et al.* A novel coronavirus from patients with Pneumonia in China, 2019. *N. Engl. J. Med.* 382, 727–733, <https://doi.org/10.1056/NEJMoa2001017> (2020).

Figures

A

2

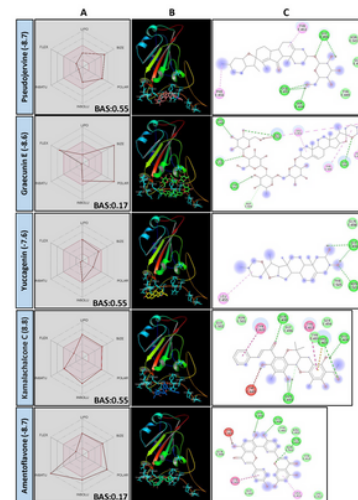


Figure 1

1a. Structural activity relationship : Correlation of canonical smiles strcutre similarity and binding energy of selected Plant secondary metabolites against SARS-CoV-2 Spike protein. Structurally similar molecules are grouped in dotted lines and a representative molecule with low binding energy (Kcal/mol) (values in perianthesis) is represned in box. 1b. Data analysis of selected plant secondary metabolites against SARS-CoV-2 Spike protein. A. Bioavailability radar chart with Bioavailabilty (BAS) score of selected molecules, B. 3D vizualization of protein-ligand interaction using pymol (selected aminoacid residue of target site of protein are colored in cyan, and C. 2D Vizualization of different types of interaction between ligand and target site of protein using Discovery studio (different types of interactions are represened in color codes).

A

1

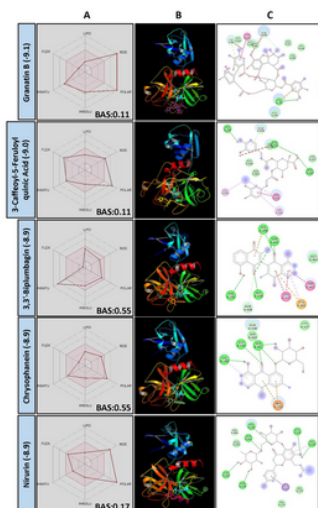


Figure 2

2a. Structural activity relationship : Correlation of canonical simles strucutre similarlity and binding energy of selected Plant secondary metabolites against SARS-CoV-2 TMPRSS2. Structurally similar molecules are grouped in dotted lines and a representative molecule with low binding energy (Kcal/mol) (values in perianthesis) is represned in box. 2b. Data analysis of selected plant secondary metabolites against SARS-CoV-2 TMPRSS2.. A. Bioavailability radar chart with Bioavailabilty (BAS) score of selected molecules, B. 3D vizualization of protein-ligand interaction using pymol (selected aminoacid residue of target site of protein are colored in cyan, and C. 2D Vizualization of different types of interaction between ligand and target site of protein using Discovery studio (different types of interactions are represented in color codes).

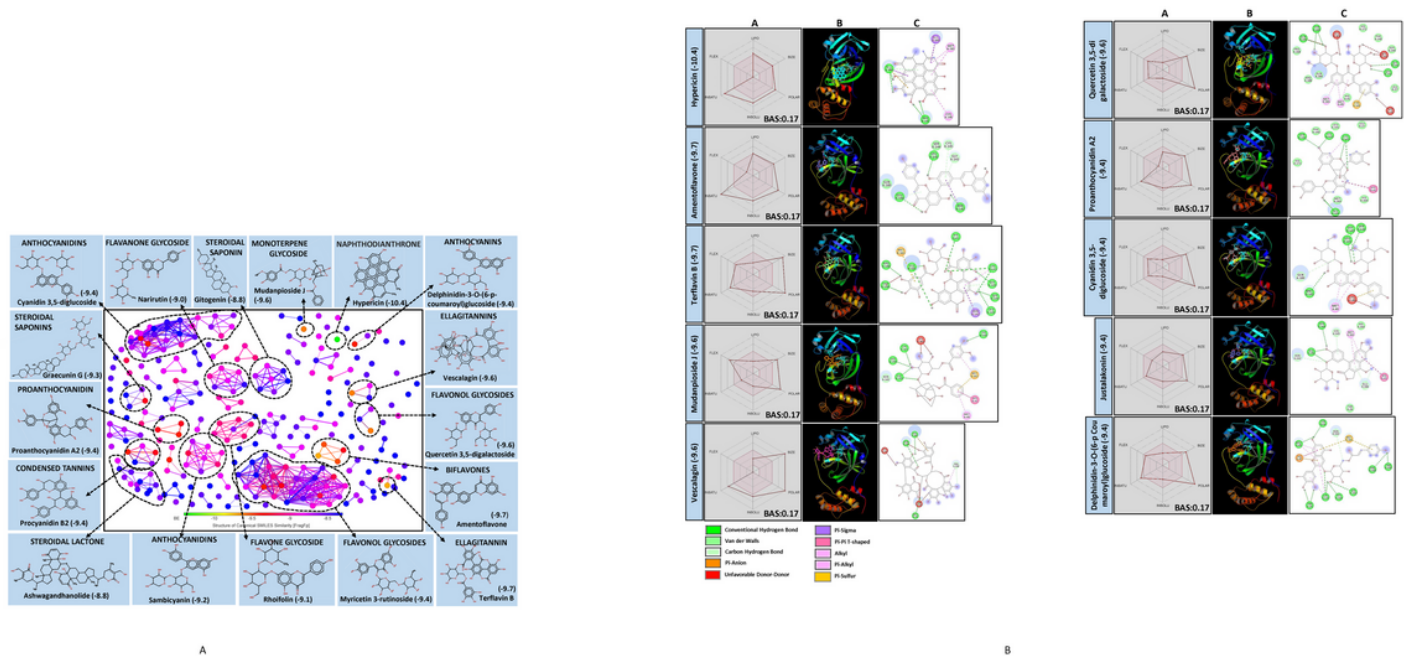
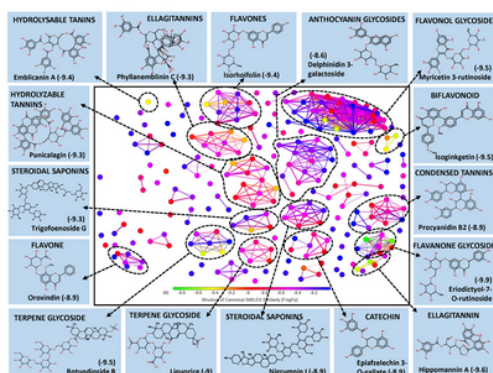
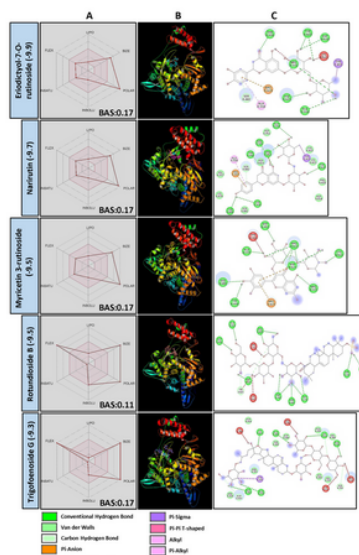


Figure 3

a. Structural activity relationship : Correlation of canonical simles strucutre similarlity and binding energy of selected Plant secondary metabolites against SARS-CoV-2 Mpro. Structurally similar molecules are grouped in dotted lines and a representative molecule with low binding energy (Kcal/mol) (values in perianthesis) is represned in box. Figure 3b. Data analysis of selected plant secondary metabolites against SARS-CoV-2 Mpro.. A. Bioavailability radar chart with Bioavailabilty (BAS) score of selected molecules, B. 3D vizualization of protein-ligand interaction using pymol (selected aminoacid residue of target site of protein are colored in cyan, and C. 2D Vizualization of different types of interaction between ligand and target site of protein using Discovery studio (different types of interactions are represented in color codes).



A



B

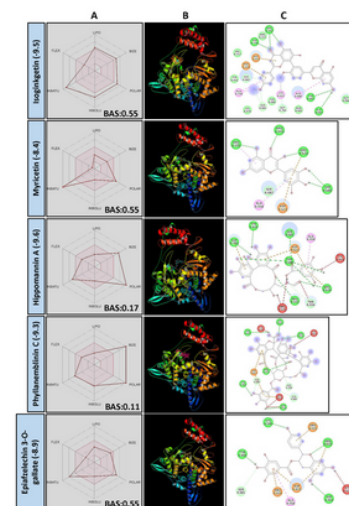


Figure 4

a. Structural activity relationship : Correlation of canonical simles strucutre similrity and binding energy of selected Plant secondary metabolites against SARS-CoV-2 RdRp. Structurally similar molecules are grouped in dotted lines and a representative molecule with low binding energy (Kcal/mol) (values in perianthesis) is represned in box. 4b. Data analysis of selected plant secondary metabolites against SARS-CoV-2 RdRp. A. Bioavailability radar chart with Bioavailibilty (BAS) score of selected molecules, B. 3D vizualization of protein-ligand interaction using pymol (selected aminoacid residue of target site of protein are colored in cyan, and C. 2D Vizualization of different types of interaction between ligand and target site of protein using Discovery studio (different types of interactions are represented in color codes).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SUPPLEMENTARYFILE3.xlsx](#)
- [SUPPLEMENTARYFILE1.xlsx](#)
- [SUPPLEMENTARYFILE2.xlsx](#)