In silico-based studies on phytochemicals from native Indian plants as potential inhibitors of SARS-CoV-2

Apoorva Sharma*

Departments of Biotechnology, Delhi Technological University, Rohini-110 042, Delhi, India

This study aims to analyze the AntiCovid effect of Phytocompounds extracted from Native Indian Plant species by computational methods such as Molecular Docking. Through this study keeping the Indian Heritage alive we characterized the ability of these phytochemicals as inhibiting agents of the Main Protease enzyme of this Virus. The lack of any effective treatment and the reoccurrence of cases despite Vaccination necessitates the quick provision of anti-SARS-CoV-2 drugs. Natural substances are getting a lot of attention for SARS-CoV-2 therapy as they have proven antimicrobial activities and are a key source for numerous antiviral drugs. Despite the fact that this virus has several identified target receptors, Main Protease (M^{pro}) is crucial for viral replication. In this study, 26 phytochemicals from 10 native Indian plant species were studied. Our docking studies demonstrated that compounds Quercetin, Withaferin A, Sominone, and Nimbin were likely to be more favorable than the natural inhibitor N3, with binding energies of−8.42, −9.21, −9.95, and −8.88 kcal/mol, respectively. These four candidate natural compounds were further examined for their bioavailability scoresthrough ADMET analysis to prove the safety of these compounds as well as their drug likeliness.Through the results it was indicated that these natural phytochemicals have a significant potential of inhibiting the SARS-CoV-2 M^{pro} enzyme and might be utilized to treat SARS-CoV-2 and manage public health, subject to *in vitro* validation in the future.

Keywords: Bioavailability, COVID-19, Molecular docking, M^{pro}, SARS-CoV-2

The Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection is unquestionably one of the most critical public health threats of this century. It has resulted in millions of mortalities and sparked healthcare alerts globally. The virus has a tendency to aggressively mutate into numerous deadly strains and thus spread widely among populations. December 2019 saw the outbreak of this deadly respiratory and enteric disease[1].

SARS Cov-2 belongs to the beta lineage of coronaviruses and is closely related to the SARS CoV virus[2]. It has already affected more than 200 countries and caused millions of casualties with significant post-illness abnormalities in individuals around the globe[3].The patients affected by this virus experience symptoms like cough, runny nose, headache, fever, chills, and in some cases hemoptysis and diarrhea[4]. The virus most troubled the patients already diagnosed with comorbid conditions (with a pre-resperatorial or diabetic condition) and improper treatment regimen[5]. The high cost of synthetic drugs combined with observable side effects post treatment led to the investigation of their natural counterparts. India was recently hit by the second wave of the Pandemic far worse than the first one with statistical analysis showing the high number of reported cases assessing the need of ongoing significant research on the development of targeted inhibitors.

The causative agent of the pandemic belongs to the family of enveloped positive-single-stranded RNA viruses, Coronaviridae with a very high infection rate. Various studies have targeted alternative proteins and enzymes such as Helicase, Spike Protein, RNA Dependent RNA polymerase, Main proteases (M^{pro}, $3CL^{pro}$) for drug designing and targeting[7]. Yet one of the most favored drug designing target with specific eleven maturation cleavage sites of the large polyprotein was M^{pro}[7]. This releases the replicase enzyme which is essential to the viral genome for transcription and regulatory mechanisms, this cleavage action performed by the enzyme plays a pivotal role in the viral life cycle8. The SARS M^{pro} is 33.8 kDa with three domains, the first two being antiparallel beta barrels and the third being a cluster of helices. The substrate generally attaches at the junction of domains 1 and 2, domain 3 acts as the catalytic regulator.

*Correspondence:
E-mail: apoorva_2k19bt009@dtu.ac.in

India has an old History of its advocacy on the medicinal importance of its Native plants with rich ingredients. The antiviral and antiseptic properties of Neem leaves have been known to us centuries before. Tulsi is a common plant in Indian households and is known to have antipyretic, antibacterial, and antifungal properties[9]. Ursolic acid is a phytocompound present in Tulsi which has shown to have high binding efficiency against RNA Polymerase and spike protein of SARS[10]. Ginger is an indigenous plant to southeast Asian countries and has been an indispensable part of Unani medicine. It has also been found to be bioactive against the Influenza virus[11], Herpes, and Chikungunya virus by various studies and has a great potential in being a candidate for drug designing[12]. Neem has been an active extensively studied for its antiviral properties against Newcastle disease[13], Herpes simplex virus[14]. The crude acidic extract of neem seeds has a significant amount of virucidal inhibition[15]. Recent studies have also shown the antiviral properties of cinnamon bark extract against the H7N3 Influenza virus[16]. Indian recipes have been rich with the usage of garlic in dishes. It is rich in organosulfur compounds and has been known to levitate viral diseases like Influenza, Herpes simplex, and Coxsackievirus[17]. Aloe also has been reported for the study of Herpes Simplex Virus-2[18], various chemicals of this plant have interacted with the viral enzyme and have caused the breakdown of the viral envelope[19]. Coriander also called Dhaniya in India is known to be constituted of phytochemicals such as Terpenoids, Tannins, Phlobatanins, and alkaloids thus reported to be a significant antioxidant and antiviral source[20]. The antiviral study of these native Indian plants has great potential in discovering new therapeutic techniques for various diseases across the world.

In this study, we have selected ten Indian native plants Tulsi (*Ocimum sanctum*), Neem (*Azadirachta indica*), Ginger (*Zingibere officinale*), Garlic (*Allium sativum*), Cinnamon (*Cinnamonum zeylanicum*), Aloe Vera (*Aloe barbadenis*), Coriander (*Coriander sativum*), Lemongrass (*Cymbopogon citratus*), Ashwagandha (*Withania somnifera*), Giloy (*Tinospora cordifolia*) against the M^pro of SARS CoV-2 as the target for docking procedures. We tried to find a natural compound that can be therapeutic and thus help in the identification of an effective treatment for the same (Fig. 1).

Materials and Methods
Target selection
In this study, the main protease enzyme (M^pro) of SARS CoV-2 was used as the potential drug target.

The protein sequence and structure of the target protein were extracted from the protein data bank (www.rcsb.org) with PDB ID: 6LU7. This enzyme is a homodimer, with each monomer composed of 306 residues and three domains and the N3 molecule acting as its natural inhibitor.

Sequence analyses
Physicochemical parameters of the main protease enzyme of SARS CoV-2 including isoelectric point, instability index, hydropathicity, the atomic composition was computed using the ProtParam tool of ExPASy[21].

Structural analyses
From the data repository of proteins, Protein Data bank (https://rcsb.org/), 2.16 Angstrom Resolution Crystal Structure of the Main Protease from SARS-CoV-2 with PDB ID: 6LU7 was retrieved in the .PDB format. Yet Another Scientific Artificial Reality Application (YASARA)[22] energy minimization server was employed for initial quality assessment, structural refinement, and energy minimization of the target protein structure with its reliability evaluation through ProCheck3, ProSA-web[24], ProQ[25] and ERRAT server[26].

Ligands
During the procedure, a total of 26 phytochemicals were selected for study from 10 native Indian plants. The plants being namely Tulsi (*Ocimum sanctum*), Neem (*Azadirachta indica*), Ginger (*Zingibere officinale*), Garlic (*Allium sativum*), Cinnamon (*Cinnamonum zeylanicum*), Aloe Vera (*Aloe barbadenis*), Coriander (*Coriander sativum*), Lemongrass (*Cymbopogon citratus*), Ashwagandha (*Withania somnifera*), Giloy (*Tinospora cordifolia*). The compounds were obtained from PubChem data bank as.SDF file[27]. The .SDF files were converted to the PDB format using Online Smiles generator by National Cancer Institute (https://cactus.nci.nih.gov/translate/).

ADME Analysis
For the selection purpose, Lipinski's Rule of Five parameters were used. This rule was studied using the online Swiss ADME Tool (http://www.swissadme.ch/). Lipinski's rule states that for a compound to qualify as a ligand it must possess characteristics like H bond acceptors less than 10, H bond donors less than 5, Molecular weight being less than 500, High lipophilicity; the value of Log P less than 5. Any compound which violated more than 2 rules was debarred from the study.
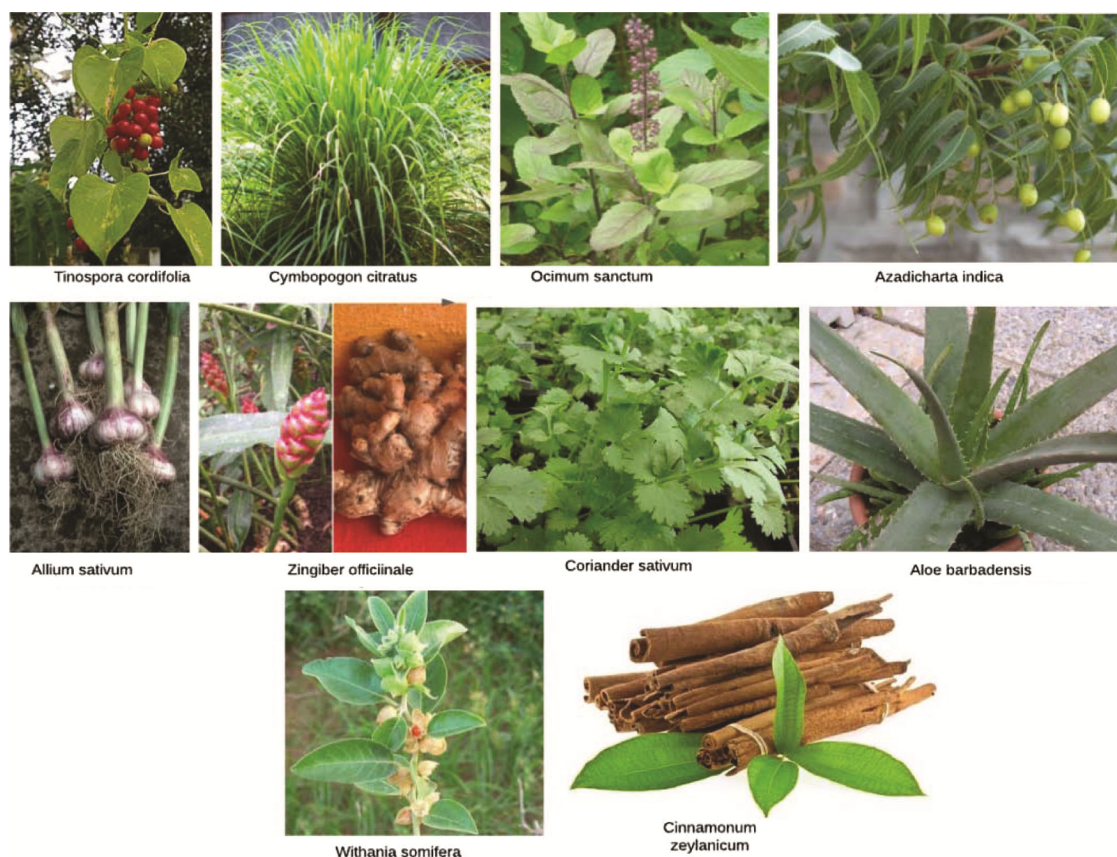
Fig. 1 — Pictorial description of 10 potential Indian native plants Giloy (*Tinospora cordifolia*), Lemongrass (*Cymbopogon citratus*), Tulsi (*Ocimum sanctum*), Neem (*Azadirachta indica*), Garlic (*Allium sativum*), Ginger (*Zingibere officinale*), Coriander (*Coriander sativum*), Aloe Vera (*Aloe barbadenis*), Ashwagandha (*Withania somnifera*), Cinnamon (*Cinnamonum zeylanicum*)

Molecular docking

The procedure of docking was performed through the software Auto Dock 4.2 (http://autodock.scripps. edu/downloads/autodock-registration/autodock-4-2-download-page/)[28]. The protein structure was optimized before being used. Water atoms were removed, Polar hydrogen atoms were added followed by the addition of Kollman charges. N3 hetatm acting as the native inhibitor was removed. Similarly, each ligand was prepared before the docking procedure. A grid box of 60X60X60 was used with a spacing of 0.375 Å. The search parameter set was Genetic Algorithm and the output was procured in Lamarckian GA run. A DLG (docking log file) was studied for further analysis of the binding energy. Each ligand- Protein pair had 10 conformations. Out of these only, the most stable conformation was selected and converted to a 2D structure to examine the chemical interactions present between the both.

Toxicity prediction

Small molecule toxicology prediction is critical for predicting the quantity of tolerance before they are physiologically adapted. Toxicology prediction analysis was performed using the pkCSM online database (http://biosig.unimelb.edu.au/pkcsm/prediction)[29]. Toxicity effects in the categories of AMES Toxicity, Human Maximum Tolerance Dose, hERG-I Inhibitor, hERG-II Inhibitor, LD_{50}, LOAEL, Hepatotoxicity, Skin Toxicity, *T. pyriformis* Toxicity, and Minnow Toxicity were studies.

Target prediction

Molecular Target studies are important to find the phenotypic side effects or potential cross-reactivity caused by the action of small biomolecules. Swiss Target Prediction website (http://www.swisstar-getprediction.ch/)[30] was used for the target prediction analysis.

Bioavailability radar

The drug-like consumption of the studied ligands which showed binding energy less than the controlled setup was performed taking into consideration 6 physicochemical properties. Swiss ADME tool (http://www.swissadme.ch/) was used for the study of parameters like - Solubility, Polarity, Size, Flexibility,
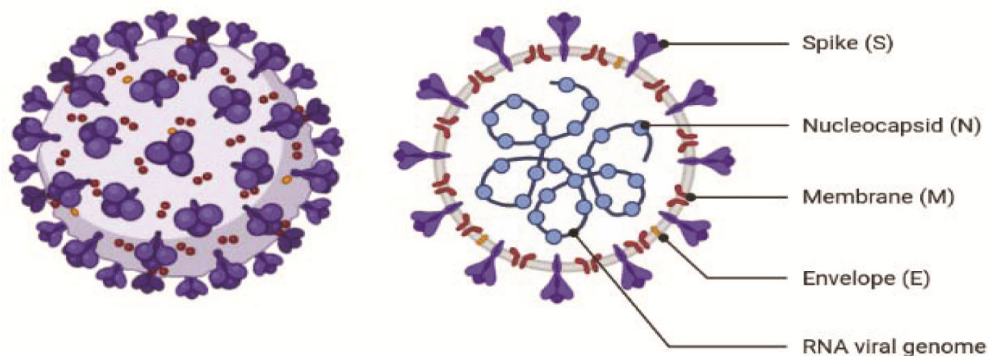
Fig. 2 — SARS-CoV-2 Viral genome structure

Saturation, and Lipophilicity. The pink region shows obedience to all the factors and any violation from them on a large scale indicates the non- bioavailability of the said compound.

Results and Discussion

SARS-CoV-2 viral genome structure

SARS-CoV-2 is a non-segmented, spherical virus encapsulating single-stranded RNA (ssRNA), measuring roughly 30 kb in length as genetic material. The SARS-CoV-2 Wuhan-Hu-1 strain was the first entire viral genome sequenced, comprising 29,903 bp of RNA (GenBank ID: MN908947.3). The genome sequence reveals that the 5' end is capped and the 3' end is polyadenylated, with two non-coding untranslated regions (UTRs), structural proteins (S, E, M, and N), and other non-coding elements and replicase genes for auxiliary proteins (ORF1ab). Several proteins are encoded via several reading frames (Fig. 2). ORF1a/b, which is found at the 5' end of the genome and encodes 15 nsps, is the largest ORF[31].

Target protein sequence and structural analyses

We evaluated the physicochemical characteristics of the M^{pro} of SARS-CoV-2. The stability index of a protein is a measurement of the protein's stability in a test tube. A protein with an instability index of less than 40 is deemed stable, based on the weight values of various dipeptides. With an instability score of 27.65, our findings indicated that this protein complex is quite stable (Table 1).

Structure evaluation and validation of target protein Mpro

The coronavirus M^{pro} enzyme is a single-chain protein of 306 amino acid residues. The Protein Data Bank provided the experimentally determined structure (X-RAY DIFFRACTION) of our target complex with 2.16 Angstrom resolution crystal structure of M^{pro} (PDB ID: 6LU7)[6] (Fig. 3).

Table 1 —Physicochemical parameters of M^{pro} of SARS-CoV-2.

S. No	Parameters	M^{pro} of SARS-CoV-2
1	Mol. Weight	33796.64
2	No. of amino acids	306
3	Theoretical pI	5.95
4	Instability index (II)	27.65
5	No. of Negatively Charged Residues (Asp + Glu)	26
6	No. of Positively Charged Residues (Arg + Lys)	22
7	Aliphatic Index	82.12
8	Grand average of Hydropathicity (GRAVY)	−0.019
9	Atomic Composition	C 1499 H 2318 N 402 O 445 S 22

The energy minimization and structural refinement of the aforementioned structure of the target were done utilizing the YASARA Energy Minimization Server to validate the structure. In the revised model, we were able to reduce the structure's energy from −125921.3 kJ/mol (score, −0.93) to −165435.7.5 kJ/mol (score, 0.26). Following that, the stereochemistry of the improved model of the target M^{pro} was analyzed using ProCheck. The results were plotted on the Ramachandran Plot, with the majority of the residues (89.8%) in the most favorable region (red), allowed zones (yellow) 9.4%, and the remaining 0.4% in the generously allowed region (light yellow), followed by only 0.4% residues in the most unfavorable zone of the disallowed region (white) (Fig. 4).

We then analyzed our protein via the ProSA-web server's protein structure analysis, resulting in a Z score of –7.16 (Fig. 5A). The Levitt-Gerstein (LG) score of 6.388 and Maxus 0.370 retrieved in the Protein Quality Predictor (ProQ) (Fig. 5B) tool indicated the excellent correctness of our structure. A ProQ LG score of > 2.5 indicates that the model structure is of high quality.

In the ERRAT plot (which is used to evaluate and validate the crystal structure of a protein in which the error values are plotted as a function of the sliding 9-residue window location), the Quality Factor for our target protein was 99.6564 (Fig. 6), further assuring the quality and reliability of the structure as the higher quality score indicates higher quality. Yellow bars show parts of the structure that are likely to be rejected at a 95% confidence level. Our findings point to the target protein structure's stability, quality, and reliability.

ADME Analysis

Using Lipinski's rule of 5 parameters. We determined the potential drug candidates out of the ligands. Luckily out of 26 taken phytochemicals none was observed to be violating the rule (Table 2). Hence all the selected ligands were docked for further studies.

Fig. 3 — A ribbon representation of crystal structure of the M^pro of SARS CoV-2 (PDB ID: 6LU7)
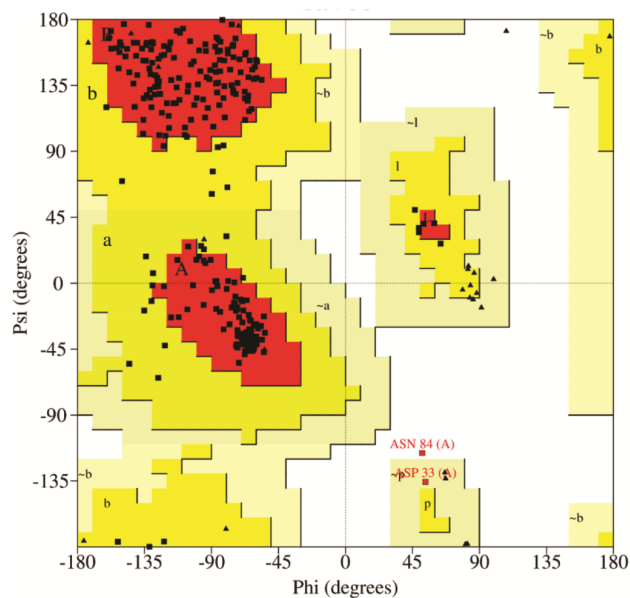

Fig. 4 — Ramachandran plot of target protein Structure (the red, dark yellow, and light-yellow and white regions represent the most favored, allowed, and generously allowed and disallowed regions respectively)
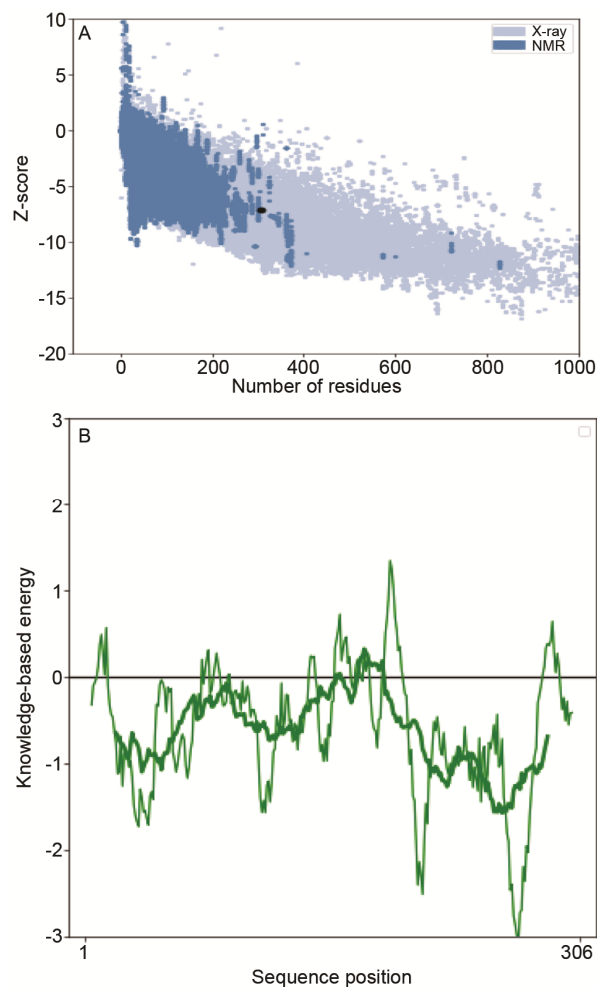

Fig. 5 — (A) ProSA-web Z-scores of target protein (all protein chains in Protein Data Bank [PDB] determined by X-ray crystallography [light blue] and nuclear magnetic resonance spectroscopy [dark blue] with respect to their length). The black dot in the dark blue region represents the Z-score of our target; and (B) Energy plot for the Main Protease of SARS-CoV-2
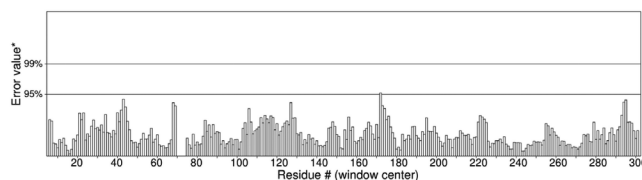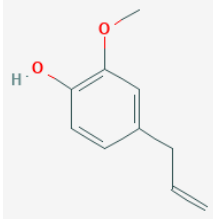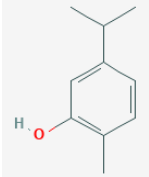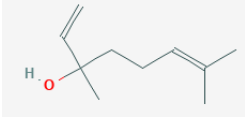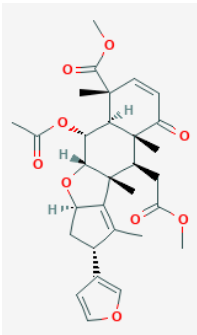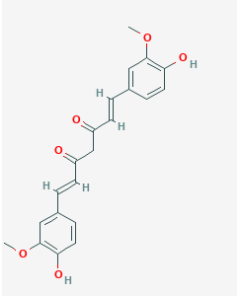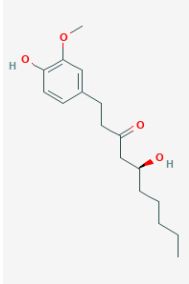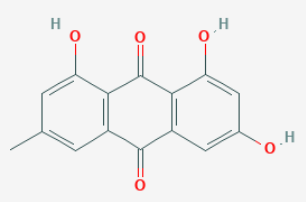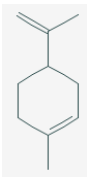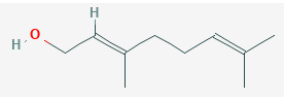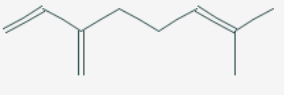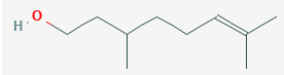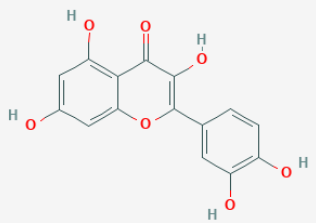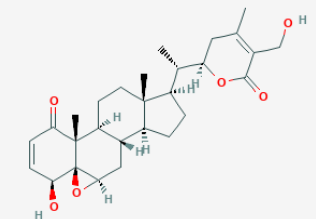

Fig. 6 — ERRAT plot of Main Protease of SARS-CoV-2

Table 2 — ADME Analysis of potential ligands

S.No.	Compound name	PubChem ID	Compound structure	Parameter characteristics	
1	Eugenol	3314		Molecular Weight (<500Da)	164.20 g/ml
				Lipophilicity (Log P <5)	2.01
				H Bond Donor (<5)	1
				H Bond acceptor (<10)	2
				Violations	0
2	Carvacrol	10364		Molecular Weight (<500Da)	150.22g/mol
				Lipophilicity (Log P <5)	2.76
				H Bond Donor (<5)	1
				H Bond acceptor (<10)	1
				Violations	0
3	Linalool	6549		Molecular Weight (<500Da)	154.25g/mol
				Lipophilicity (Log P <5)	2.59
				H Bond Donor (<5)	1
				H Bond acceptor (<10)	1
				Violations	0
4	Nimbin	108058		Molecular Weight (<500Da)	540.60g/mol
				Lipophilicity (Log P <5)	2.04
				H Bond Donor (<5)	0
				H Bond acceptor (<10)	9
				Violations	1
5	Curcumin	969516		Molecular Weight (<500Da)	368.38g/mol
				Lipophilicity (Log P <5)	1.47
				H Bond Donor (<5)	2
				H Bond acceptor (<10)	6
				Violations	0
6	Gingerol	442793		Molecular Weight (<500Da)	294.39g/mol
				Lipophilicity (Log P <5)	2.14
				H Bond Donor (<5)	2
				H Bond acceptor (<10)	4
				Violations	0

(Contd.)

Table 2 — ADME Analysis of potential ligands— (*Contd.*)

S.No.	Compound name	PubChem ID	Compound structure	Parameter characteristics	
7	Zingerone	31211		Molecular Weight (<500Da)	194.23g/mol
				Lipophilicity (Log P <5)	1.42
				H Bond Donor (<5)	1
				H Bond acceptor (<10)	3
				Violations	0
8	Allicin	65036		Molecular Weight (<500Da)	162.27g/mol
				Lipophilicity (Log P <5)	1.18
				H Bond Donor (<5)	0
				H Bond acceptor (<10)	1
				Violations	0
9	Ajoene	5386591		Molecular Weight (<500Da)	234.40g/mol
				Lipophilicity (Log P <5)	2.10
				H Bond Donor (<5)	0
				H Bond acceptor (<10)	1
				Violations	0
10	Cinnamaldehyde	637511		Molecular Weight (<500Da)	132.16g/mol
				Lipophilicity (Log P <5)	2.01
				H Bond Donor (<5)	0
				H Bond acceptor (<10)	1
				Violations	0
11	Alpha thujene	17868		Molecular Weight (<500Da)	136.23g/mol
				Lipophilicity (Log P <5)	4.29
				H Bond Donor (<5)	0
				H Bond acceptor (<10)	0
				Violations	0
12	Terpineol	17100		Molecular Weight (<500Da)	154.25g/mol
				Lipophilicity (Log P <5)	2.30
				H Bond Donor (<5)	1
				H Bond acceptor (<10)	1
				Violations	0
13	Barbaloin	12305761		Molecular Weight (<500Da)	418.39g/mol
				Lipophilicity (Log P <5)	−1.59
				H Bond Donor (<5)	7
				H Bond acceptor (<10)	9
				Violations	1

(*Contd.*)

Table 2 — ADME Analysis of potential ligands— (*Contd.*)

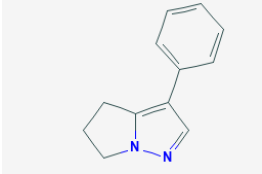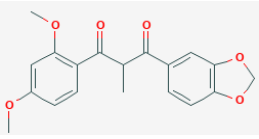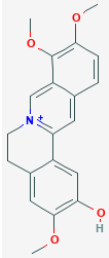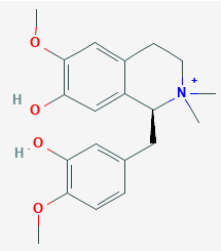| S.No.	Compound name	PubChem ID	Compound structure	Parameter characteristics	
14	Emodin	3220		Molecular Weight (<500Da) Lipophilicity (Log P <5) H Bond Donor (<5) H Bond acceptor (<10) Violations	270.24g/mol 0.36 3 5 0
15	Limonene	22311		Molecular Weight (<500Da) Lipophilicity (Log P <5) H Bond Donor (<5) H Bond acceptor (<10) Violations	136.23g/mol 3.27 0 0 0
16	Camphor	2537		Molecular Weight (<500Da) Lipophilicity (Log P <5) H Bond Donor (<5) H Bond acceptor (<10) Violations	152.23g/mol 2.30 0 1 0
17	Geraniol	637566		Molecular Weight (<500Da) Lipophilicity (Log P <5) H Bond Donor (<5) H Bond acceptor (<10) Violations	154.25g/mol 2.59 1 1 0
18	Myrcene	31253		Molecular Weight (<500Da) Lipophilicity (Log P <5) H Bond Donor (<5) H Bond acceptor (<10) Violations	136.23g/mol 3.56 0 0 0
19	Citronellol	8842		Molecular Weight (<500Da) Lipophilicity (Log P <5) H Bond Donor (<5) H Bond acceptor (<10) Violations	156.27g/mol 2.70 1 1 0
20	Quercetin	5280343		Molecular Weight (<500Da) Lipophilicity (Log P <5) H Bond Donor (<5) H Bond acceptor (<10) Violations	302.24g/mol −0.56 5 7 0
21	Withaferin A	265237		Molecular Weight (<500Da) Lipophilicity (Log P <5) H Bond Donor (<5) H Bond acceptor (<10) Violations	470.6g/mol 2.75 2 6 0

(*Contd.*)

Table 2 — ADME Analysis of potential ligands— (*Contd.*)

S.No.	Compound name	PubChem ID	Compound structure	Parameter characteristics	
22	Sominone	44249449		Molecular Weight (<500Da)	458.63g/mol
				Lipophilicity (Log P <5)	3.66
				H Bond Donor (<5)	3
				H Bond acceptor (<10)	5
				Violations	0
23	Withasominone	442877		Molecular Weight (<500Da)	184.24g/mol
				Lipophilicity (Log P <5)	2.34
				H Bond Donor (<5)	0
				H Bond acceptor (<10)	1
				Violations	0
24	Tinosporinine	42607646		Molecular Weight (<500Da)	342.3g/mol
				Lipophilicity (Log P <5)	1.19
				H Bond Donor (<5)	0
				H Bond acceptor (<10)	6
				Violations	0
25	Columbamine	72310		Molecular Weight (<500Da)	338.38g/mol
				Lipophilicity (Log P <5)	1.78
				H Bond Donor (<5)	1
				H Bond acceptor (<10)	4
				Violations	0
26	Tembetarine	167718		Molecular Weight (<500Da)	344.42g/mol
				Lipophilicity (Log P <5)	−1.71
				H Bond Donor (<5)	2
				H Bond acceptor (<10)	4
				Violations	0

Molecular Docking

All the selected ligands were used for molecular docking. Molecular docking undoubtedly remains an essential tool in computational biology with a diverse usage in drug designing and delivery. After the docking procedure, potentially stable ligand-protein complexes were selected for further bond formation study. There are 10 residues present in the active binding site of protein 6LU7 namely THR24, THR26, ASN142, CYS145, PHE140, HIS163, HIS164, GLY143, GLU166, HIS172. GLU166 is involved in the homodimerization of the protease enzyme and also plays a key role in the creation of a binding pocket. Moreover, CYS141 and HIS41 form the catalytic dyad on the third domain. To consider the ligands the native inhibitor N3 was taken as the comparative analyzer for all the 26 phytochemicals. The binding energy of N3 is (−8.15 kcal/mol). Our analysis led to the result of four compounds having a more stable interaction with the protease enzyme than the native inhibitor. The phytochemicals were as followed Nimbin −8.88 kcal/mol, Quercetin −8.42 kcal/mol, Withaferin A −9.21 kcal/mol and Sominone −9.95 kcal/mol (Table 3 & Fig. 7).

Study of Interactions

The interactions between the best four binding ligands were studied (Fig. 8). Interactions between 6LU7 and Quercetin (Fig. 8A) showed chemical

Table 3 —Docking results revealing Polar contact information and binding energy of different ligands with Protease of SARS CoV-2

| S.No	Ligands	Binding Energy (Delta G) (kcal/mol)	Ligand Efficiency	Inhibition constant	Inter-molecular "Energy (kcal/mol)	Vdw H-bond "desolvation (kcal/mol)
1	Eugenol	−5.02	−0.42	208.49 μM	−6.21	−6.19
2	Carvacrol	−5.23	−0.48	146.47 μM	−5.83	−5.77
3	Linalool	−5.29	−0.48	133.1 μM	−6.78	−6.76
4	Nimbin	−8.88	−0.23	310.06 nM	−11.27	−11.24
5	Curcumin	−8.07	−0.30	1.21 μM	−11.06	−11.04
6	Gingerol	−6.55	−0.31	15.87 μM	−10.13	−10.11
7	Zingerone	−5.64	−0.40	73.42 μM	−7.13	−7.03
8	Allicin	−3.86	−0.43	1.48 mM	−5.35	−5.33
9	Ajoene	−5.56	−0.43	83.46 μM	−7.95	−7.93
10	Cinnamaldehyde	−4.79	−0.48	309.93 μM	−5.38	−5.37
11	Alpha Thujene	−5.18	−0.52	160.69 μM	−5.47	−5.47
12	Terpineol	−5.49	−0.50	93.88 μM	−6.09	−6.02
13	Barbaloin	−5.17	−0.52	163.39 μM	−5.46	−5.46
14	Emodin	−6.63	−0.33	13.83 μM	−7.52	−7.37
15	Limonene	−5.23	−0.52	147.59 μM	−5.52	−5.53
16	Camphor	−5.20	−0.52	154.83 μM	−5.50	−5.49
17	Geraniol	−5.24	−0.48	144.42 μM	−6.73	−6.67
18	Myrcene	−4.67	−0.47	376.15 μM	−5.87	−5.86
19	Citronellol	−4.91	−0.45	252.45 μM	−6.70	−6.60
20	Quercetin	−8.42	−0.38	669.45 nM	−10.21	−10.01
21	Withaferin A	−9.21	−0.27	177.63 nM	−10.70	−10.55
22	Sominone	−9.95	−0.30	50.55 nM	−11.74	−11.49
23	Withasominone	−6.00	−0.43	39.77 μM	−6.30	−6.28
24	Tinosporanine	−7.50	−0.30	3.17 μM	−9.29	−8.99
25	Columbamine	−7.93	−0.32	1.54 μM	−9.12	−9.04
26	Tembetarine	−7.55	−0.30	2.9 μM	−9.34	−8.98
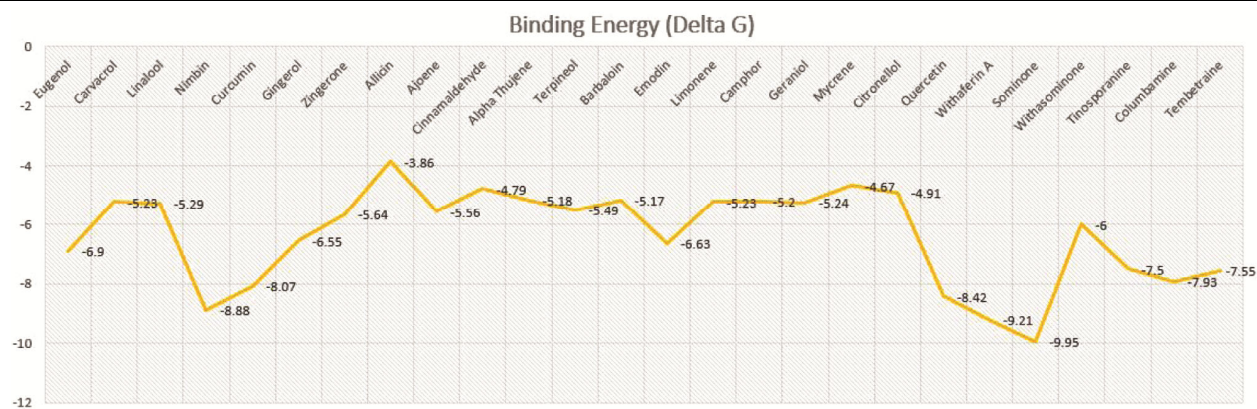
Fig. 7 — Graphical representation of the binding energy of the phytochemicals

interaction by formin conventional Hydrogen bonds withamino acids positioned asHIS164, ASP187, THR190, and GLN192, pi sulfur bond was formed with MET165, pi-pi T shaped bond was formed with HIS41 and it also formed the conventional Van der Waal forces.

Interaction between 6LU7 and Nimbin (Fig. 8B). The ligand formed conventional hydrogen bonds with GLU166 and HIS163, conventional carbon-hydrogen bond with GLU166 and GLN189, pi sulfur bond with MET165, alkyl, and pi-alkyl bonds with PRO168 and it also showed a significant amount of van der waal interactions with more than ten nearby amino acids.

The next studied interaction was Withaferin A and the Mpro (Fig. 8C) of the SARS Cov-2. Pi-sulfur bond was observed with MET165, pi-pi T shaped bond with HIS41, conventional Hydrogen bonds with
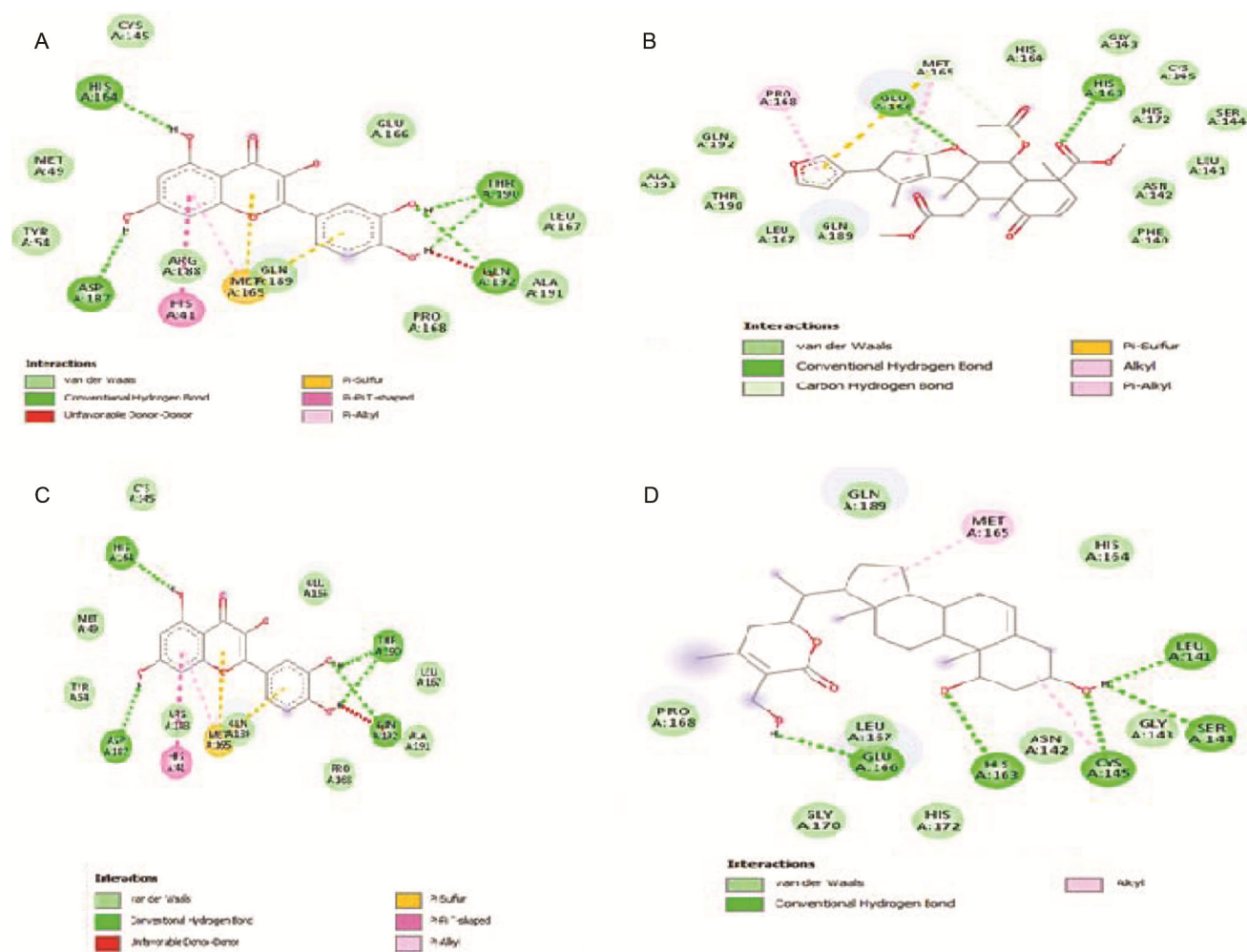
Fig. 8 — (A) Interactions between 6LU7 and Quercetin; (B) Interactions between Nimbin and 6LU7; (C) Interactions between Withaferin A and 6LU7; and (D) Interactions between Sominone and 6LU7

HIS164, ASP187, THR190, and GLN192. Moreover, it also showed the van der waal interaction with about ten more nearby amino acids.

The most negative binding energy was showcased by the interaction of Sominone and 6LU7 (Fig. 8D). It showed three types of diverse interactions namely conventional Hydrogen bond with LEU141, SER144, CYS145, HIS163, GLU166, and Van der waal interaction with closely associated multiple amino acids and it also formed a conventional alkyl bond with MET165.

Toxicity prediction

None of the substances tested in the toxicity studies crossed the blood-brain barrier, decreasing the likelihood of neurotoxicity. The *in silico* toxicity scores obtained by the pKCSM algorithm demonstrated that none of the four natural chemicals displayed AMES toxicity. Similarly, none of the substances inhibited hERG-I and hERG-II. Sominone was anticipated to have harmful potential in terms of hepatotoxicity. A high range of *T. pyriformis* toxicity was found in all of the substances (0.288 to 0.317 log mg/L) (Table 4).

Target prediction

The *in silico* target prediction of four potential natural compounds in the human proteome is shown in (Fig. 9). With a probability of 1%, the family kinase (33.3%), oxidoreductase (20%), AG protein-coupled receptors (13.3%), and enzymes (13.3%) were predicted to be the primary targets of Quercetin. This suggests that there is a possibility that Quercetin will have off-target activity in humans. Withaferin A revealed kinase (33.3%) and enzymes (20%) as possible targets, but with a very low probability of 0.12. With a probability of 0.1 percent, kinase (20%), oxidoreductases (13.3%), and secreted
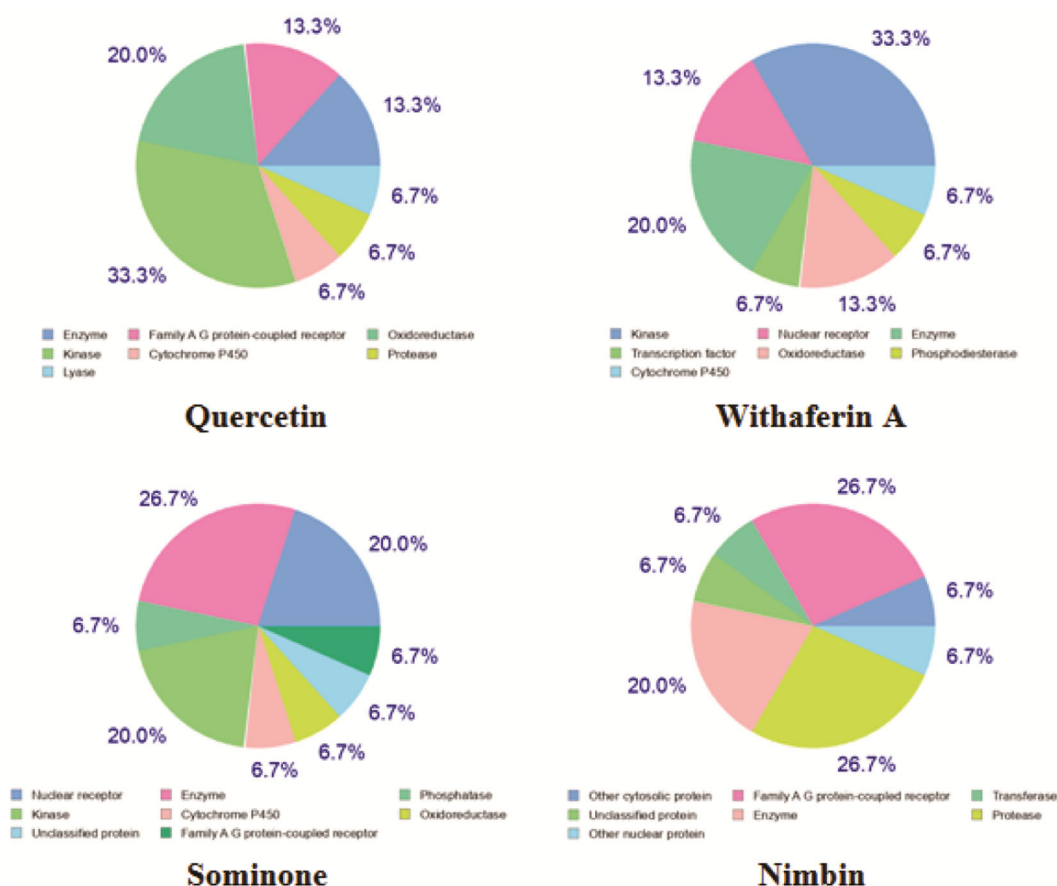
Fig. 9 — The Pie-chart representing top-15 target prediction of four potential natural compounds in the human proteome

Table 4 — *In silico* toxicity prediction of 4 selected compounds

Com-pounds	AMES toxicity Categorical (Yes/No)	Max. tolerated dose (human) Numeric(log mg/kg/day)	hERG I* inhibitor Categorical (Yes/No)	hERG II* inhibitor Categorical (Yes/No)	Oral rat acute toxicity (LD50*) Numeric (log mg/kg)	Oral rat chronic toxicity (LOAEL*) Numeric (log mg/kg_bw /day)	Hepatotoxicity Categorical (Yes/No)	Skin sensitization Categorical (Yes/No)	T. pyriformis toxicity Numeric (log µg/L)	Fathead minnow toxicity Numeric (log mM)
Quercetin	No	0.499	No	No	2.471	2.612	No	No	0.288	3.721
Withaferin A	No	−0.695	No	No	2.799	0.918	No	No	0.299	0.738
Sominone	No	−0.837	No	No	2.225	0.14	Yes	No	0.317	−0.46
Nimbin	No	−0.371	No	No	2.48	1.57	No	No	0.295	1.269

hERG*: human Ether-à-go-go-Related Gene; LD$_{50}$*: lethal dose of 50%; LOAEL*: lowest observed adverse effect level.

proteins (20%) were possible targets of Sominone. Nimbin, with a probability of 0.095, revealed a protease (26.7%), family AG protein-coupled receptors (26.7%), and enzymes (20%) as potential off-targets. These findings suggest that Nimbin has a negligible risk of causing off-target effects in humans.

Bioavailability Radar

The bioavailability radar is based on 6 characteristics – Lipophilicity XLOGP3 (in the range of 0.7to +5.0), Molecular weight (between 150-500 g/mol), Polarity TPSA (between 20 to 130 Å2), Solubility(insolubility) log S not greater than 6, Saturation (insatu) sp3 hybridized carbon not less than 0.25 in fraction, flexibility not more than 9 rotatable bonds.

Thus, Nimbin and Quercetin failed the parameters set (Table 5 & Fig. 10). Nimbin violated the criteria of molecular weight. Quercetin violated the criteria of TPSA polarity being greater than the limit and has the saturation of sp3 hybridized carbon less than 0.25. Through the analysis of bioavailability of the four

Table 5 —Radar Plot statistics of oral bioavailability of the shortlisted natural compounds. Any deviation represents a suboptimal physicochemical property for oral bioavailability

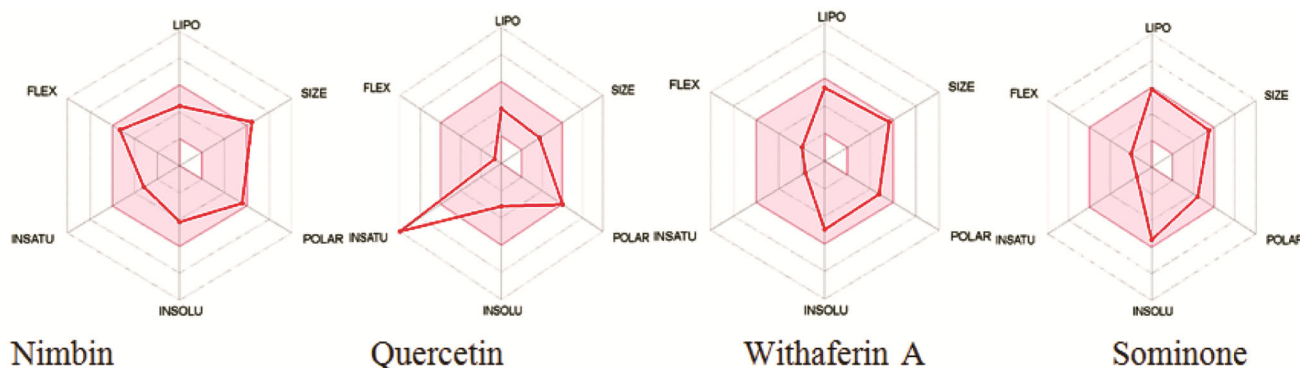| Ligand	Lipophilicity (XLOGP)	Mol. Weight	Polarity (TPSA) Å^2	Solubility (Log S)	Saturation	Flexibility
Nimbin	2.28	540.6 g/mol	118.34	−4.20	0.60	8
Quercetin	1.54	302.24 g/mol	131.36	−3.16	0.00	1
Withaferin A	3.83	470.60 g/mol	96.36	−4.97	0.79	3
Sominone	4.72	458.63 g/mol	86.99	−5.46	0.82	3

Fig. 10 — Radar plots for oral bioavailability of the compounds (Nimbin, Quercetin, Withaferin A and Sominone). Pink area represents the range of optimal values. The colored zone is a suitable physicochemical space for oral bioavailability

selected ligands *i.e.*, Nimbin, Quercetin, Withaferin A, and Sominone. The study deduced that Withaferin A and Sominone are orally bioavailable.

Conclusion

In this study, we took 26 phytochemicals that were selected from 10 native Indian plant species. Indian plants are a part of the rich legacy of the Indian ayurvedic and historic medicinal scienceswith potential to replace the synthetic drugs and cause no side effect. These 26 compounds were filtered using Lipinski's rule of five parameters and were determined for their probable use as drugs. All the phytochemicals were successful in this filtration process and were docked. While docking the native inhibitor N3 was taken as the reference and our studies revealed the presence of four ligands being more potent.

The four resulting compounds were Nimbin (−8.88 kcal/mol), Quercetin (−8.42 kcal/mol), Withaferin A (−9.21 kcal/mol), Sominone (−9.95 kcal/mol). These four compounds were further analyzed for their Toxicity, Target Prediction and Bioavailability for estimating their conventional use as a medication.

In the toxicity analysis none of these compounds demonstrated Blood Brain barrier toxicity, however Sominone was observed to show hepatotoxicity. Through the target prediction analysis Quercetin was seen to show a potential for causing off target effects with null chance of Nimbin and only 0.1% chances of Sominone, 0.12% of Withaferin A. Following these

studies Quercetin and Nimbin failed the Bioavailability test. In the course of these analytical procedures Withaferin A was found to be the ideal Phytocompound with characteristic results in each analysis. This compound can be analyzed further *in vitro* for its promising scope as an Anti- SARS medicine and drug.

Conflict of interest

All authors declare no conflict of interest.

References

1 Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, Hu Y, Tao ZW, Tian JH, Pei YY & Yuan ML, A new coronavirus associated with human respiratory disease in China. *Nature*, 579 (2020) 265.

2 Shu Y & McCauley J, GISAID: Global initiative on sharing all influenza data - from vision to reality. *Euro Surveill*, 22 (2017):30494.

3 Zhao S, Lin Q, Ran J, Musa SS, Yang G, Wang W, Lou Y, Gao D, Yang L, He D & Wang MH,. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. *Int J Infect Dis*, 92 (2020) 214.

4 Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX & Zhong NS, Clinical characteristics of coronavirus disease, 2019 (2020) 1708.

5 Knight TE, Severe acute respiratory syndrome coronavirus 2 and coronavirus disease 2019: a clinical overview and primer. *Biopreserv Biobank*, 18 (2020) 492.

6 Wu C, Liu Y, Yang Y, Zhang P, Zhong W, Wang Y, Wang Q, Xu Y, Li M, Li X & Zheng M, Analysis of therapeutic targets for SARS-CoV-2 and discovery of

potential drugs by computational methods. *Acta Pharmaceutica Sinica B*, 10 (2020) 766.

7 Ma C, Sacco MD, Hurst B, Townsend JA, Hu Y, Szeto T, Zhang X, Tarbet B, Marty MT, Chen Y & Wang J, Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. *Cell Res*, 30 (2020) 678.

8 Mengist HM, Fan X & Jin T, Designing of improved drugs for COVID-19: Crystal structure of SARS-CoV-2 main protease M pro. *Signal Transduct Target Ther*, 5 (2020) 1.

9 Mohan L, Amberkar MV & Kumari M, *Ocimum sanctum* linn. (TULSI)-an overview. *Int J Pharm Sci Rev Res*, 7 (2011) 51.

10 Kumar AHS, Molecular docking of natural compounds from tulsi (*Ocimum sanctum*) and neem (*Azadirachta indica*) against SARS-CoV-2 protein targets. *Biol Eng Med Sci Rep*, 6 (2020) 11.

11 Singh NA, Kumar P, Jyoti & Kumar N, Spices and herbs: Potential antiviral preventives and immunity boosters during COVID-19. *Phyther Res*, 35 (2021) 2745.

12 Dorra N, El-Berrawy M, Sallam S & Mahmoud R, Evaluation of Antiviral and Antioxidant Activity of Selected Herbal Extracts. *J High Inst Public Heal*, 49 (2019) 36.

13 Mahmood MS, Amir HW, Abbas RZ, Aslam B & Rafique A. Evaluation of antiviral activity of *Azadirachta indica* (Neem) bark extract against Newcastle disease virus. *Pak Vet J*, 38 (2018) 25.

14 Tiwari V, Darmani NA, Yue BYJT & Shukla D, *In vitro* antiviral activity of neem (*Azardirachta indica* L.) bark extract against herpes simplex virus type-1 infection. *Phyther Res*, 24 (2010) 1132.

15 Amer H, Helmy WA & Taie HAA, *In vitro* antitumor and antiviral activities of seeds and leaves Neem (*Azadirachta indica*) extracts. *Int J Acad Res*, 2 (2010) 47.

16 Fatima M, Sadaf Zaidi NS, Amraiz D & Afzal F, *In vitro* antiviral activity of Cinnamomum cassia and its nanoparticles against H7N3 influenza a virus. *J Microbiol Biotechnol*, 26 (2016) 151.

17 Tsai Y, Cole LL, Davis LE, Lockwood SJ, Simmons V, Wild GC. Antiviral properties of garlic: *in vitro* effects on influenza B, herpes simplex and coxsackie viruses. *Planta Med*, 51 (1985) 460.

18 Zandi K, Zadeh MA, Sartavi K & Rastian Z, Antiviral activity of Aloe vera against herpes simplex virus type 2: An *in vitro* study. *Afr J Biotechnol*, 6(2007).

19 Mpiana P, Ngbolua K, Tshibangu DS, Kilembe JT, Gbolo BZ, Mwanangombo DT, Inkoto CL, Lengbiye EM, Mbadiko CM, Matondo A, Bongo GN & Tshilanda DD, Aloe vera (L.) Burm. F. as a Potential Anti-COVID-19 Plant: A Mini-review of Its Antiviral Activity. *Eur J Med Plants*, 31 (2020) 86.

20 Nathenial S, Fatima A, Fatima R, Ijaz N, Saeed N, Shafqat A & Leghari L, Phytochemical study of acetone solvent extract of Coriander sativum. *J Pharmacogn Phytochem*, 8 (2019) 136.

21 Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD & Bairoch A, ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res*, 31(2003):3784.

22 Krieger E, Joo K, Lee J, Lee J, Raman S, Thompson J, Tyka M, Baker D & Karplus K, Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins*, 77 (2009) 114.

23 Laskowski RA, MacArthur MW, Moss DS & Thornton JM, PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr*, 26(1993) 283.

24 Wiederstein M & Sippl MJ, ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res*, 35 (2007) W407.

25 Wallner B & Elofsson A, Can correct protein models be identified? *Protein Sci*, 12 (2003) 1073.

26 Colovos C & Yeates TO, Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci*, 2 (1993) 1511.

27 Wang Y, Xiao J, Suzek TO, Zhang J, Wang J & Bryant SH, PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res*, 37 (2009) W623.

28 Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS & Olson AJ, AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem*, 30 (2009) 2785.

29 Pires DE V, Blundell TL & Ascher DB, pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J Med Chem*, 58 (2015) 4066.

30 Daina A, Michielin O, Zoete V, SwissTarget Prediction: updated data and new features for efficient prediction of protein targets of small molecules. *Nucleic Acids Res*, 47 (2019) W357.

31 Klein S, Cortese M, Winter SL, Wachsmuth-Melm M, Neufeldt CJ, Cerikan B, Stanifer ML, Boulant S, Bartenschlager R & Chlanda P, SARS-CoV-2 structure and replication characterized by *in situ* cryo-electron tomography. *Nat Commun*, 11(2020) 10.