Virtual screening of naturally occurring antiviral molecules for SARS-CoV-2 mitigation using docking tool on multiple molecular targets

Shiwani Rana, Sanjay Sharma and Kalyan Sundar Ghosh*

Department of Chemistry, National Institute of Technology Hamirpur, H.P. 177005, India

Corresponding Author

Tel: +91-1972-254104; Fax: +91-1972-223834; e-mail: kalyan@nith.ac.in

Abstract

The coronavirus catastrophe (COVID-19) caused by a novel strain of coronavirus (SARS-CoV-2) has turned the world upside down at an unprecedented level and has been declared as a pandemic by World Health Organization (WHO). It has resulted huge number of fatalities and infections due to the severe lower respiratory tract sickness in the infected people. Research across the world is in progress to identify inhibitors against various molecular targets associated with this viral infection. Among these targets, a very important one is a cysteine like protease or 3CL protease (3CLpro) and that is required for the replication of the virus. In the present study, initially we have investigated the potential of twenty naturally occurring antiviral molecules to function as inhibitors against the activity of main viral protease (3CLpro) so as to put a halt on viral replication. The investigation has been carried out through docking of the molecules with 3CLpro. Based on the results, three most potential molecules (bilobetin, ginkgetin and sciadopitysin) have been screened. Further these molecules were subjected for checking their activity on other molecular targets like a papain like protease (PLpro), spike protein S1, RNA dependent RNA polymerase (RdRp), angiotensin converting enzyme 2 (ACE2) receptor. In addition to 3CLpro inhibition, ginkgetin was predicted as an inhibitor of PLpro also. But none of these three compounds was found effective on rest other molecular targets.

Keywords: SARS-Cov-2, cysteine like protease inhibition, papain like protease inhibition, bilobetin, ginkgetin, sciadopitysin

1 Introduction

The outbreak of the deadly novel coronavirus (SARS-CoV-2) that surfaced probably in the seafood market of Wuhan city of China at the end of year 2019 has brought the world to its knees. This virus has now engulfed almost the whole globe. It was later termed as COVID-19 on February 11, 2020 and was recognized to be the source of viral pneumonia that has sicken a huge population globally. It has been declared as a global health emergency and a pandemic by World Health Organization (WHO) on January 30, 2020. So far from its emergence, around ~3,62,000 people had died and ~5.9 million were reported to be infected. Coronavirus belongs to *coronviridae* family with order *nidovirales* that normally causes respiratory illness ranging from common cold to severe acute respiratory syndrome (SARS-CoV) that includes fever, cough and shortness of breath (Cui et al., 2019). There are number of coronaviruses and they exist in different animals such as pig, camel, bat etc. Sometimes these viruses get transmitted from animals to humans that can cause diseases and the phenomenon is commonly termed as spillover event. Out of the seven known coronaviruses, four have the ability to cause mild to moderate respiratory tract disease and rest three of them can cause severe and lethal illness (Su et al., 2016). HCoV-OC43, HCoV-OC63, HCoV-OC229E and HKU1 fall in to the category of alpha-coronavirus that cause modest respiratory illness. In contrast, beta-coronovirus such as SARS-CoV and MERS-CoV have the potency to cause severe and fatal respiratory lower tract infection (Lee et al., 2020). The novel coronavirus (COVID-19) also falls in the category of beta-coronavirus and share similarities with SARS-CoV and for that reason it has been termed as SARS-CoV-2. In November 2002, SARS-CoV was first identified in Asia and it further engulfed 26 countries. In September 2012, another respiratory syndrome i.e. Middle East Respiratory Syndrome (MERS-CoV) caused by coronavirus had popped up. Coronavirus has a large genome

sequence of ~30 kb in length with 5′ cap and 3′ poly-A tail (Lai et al., 1997). The virion SARS-CoV-2 has a diameter of about 60-140 nm in spherical shape which is constituted with peplomers of crown shape (Wu et al., 2020). Its structure mainly consists of (i) membrane protein, (ii) nucleoplasmid (enclosed RNA), (iii) lipid membrane, (iv) spike protein, and (v) envelope protein (Jiang et al., 2020). The spike glycoproteins on the viral capsid play a crucial role to let the virus enter the host cells in living beings. These proteins binds with the angiotensin converting enzyme 2 (ACE2) receptor present on the surface of host cells and thus permit the virus to inject RNA in to cells (Hoffmann et al., 2020). Upon viral infection, RNA is processed to synthesize two polyproteins (pp1a/pp1ab) (Gorbalenya et al., 1989). The transcription process occurs with the formation of replication-transcription complex (RCT). In a typical RNA genome, there are minimum six open reading frames (ORFs) that function as template to produce subgenomic mRNAs. The frame shift mutation between ORF1a and ORF1b (Brierley et al., 1987) encodes both pp1a/pp1ab polyproteins. The polyprotein pp1ab contains more than 7,000 residues and posses putative RNA-dependent RNA polymerase (RdRp) and RNA helicase activities (Gorbalenya et al., 1989, Lee et al., 1991). Further, these polyproteins are cleaved by two proteases encoded in the virus namely, 3-chymotrypsin-like protease (3CLpro) or main protease and papain like protease (PLpro) in to functional proteins (Ziebuhr et al., 2000). The key enzyme 3CLpro is the prime protease responsible for the cleavage of polypeptides into vital functional proteins required for the replication of the virus. Alongside, PLpro also assist in the process with a supplementary role of removing ubiquitin so as to prevent the CoVs from any immune response (Békés et al., 2015). Thus, the prime protease 3CLpro that generates the functional proteins responsible for the replication and translation of this novel virus has attracted attention of many researches as a potential drug target. As this virus has dismayed the entire

population on this planet, researchers across the world are putting their all to combat against COVID-19. Expectantly, to identify potential drug candidates against SARS-CoV-2, we have decided to adopt molecular docking tool to screen out some compounds initially which can act as inhibitors against 3CLpro. For that purpose, we looked into the nature as a source of such compounds because plant derived naturally occurring compounds play a significant role in the discovery of many effective drugs and of they were approved further (Patridge et al., 2016; Thomford et al., 2018). Not only that various derivatives of natural molecules and the lead compounds inspired by nature have demonstrated their potential in drug design. Several antiviral compounds were also isolated from different medicinal plants (Akram et al., 2018). Few recent articles in this journal have also reported different natural molecules having potential to be considered as effective drugs against SARS-CoV-2 (Aanouz et al., 2020; Abdelli et al., 2020; Ahmad et al., 2020; Borkotoky et al., 2020; Enmozhi et al., 2020; Wahedi et al., 2020).

2. Methods

2.1. Preparation of the structures of small molecules and protein for docking

Three dimensional structures (as .mol file) of twenty naturally occurring compounds and control drugs were collected from ChemSpider (http://www.chemspider.com/). Their natural source, structures are the previously reported antiviral activities are given in Supplementary materials Table S1. Geometry and energy optimization of these structures were performed through quantum mechanical calculations using parametric method 3 (PM3) in ArgusLab 4.0 (http://www.arguslab.com). The crystal structure of the SARS-CoV-2 related proteins namely 3CLpro (PDB ID: 6M0K), PLpro (PDB ID: 6W9C), RdRp (in a complex with SARS-CoV-2 NSP7 and SARS-CoV-2 NSP 8, PDB ID: 6M71), spike protein S1 (in complex with human

antibody, PDB ID: 6W41), ACE2 (in a complex with spike glycoprotein, PDB ID: 6LZG) were downloaded from Protein Data Bank (PDB). To refine these protein structures, bound ligands and/or proteins and the crystallographic water molecules were removed from the structure.

2.2 Molecular docking

Protein-ligand dockings were performed by using Autodock 4.2 software. Before docking, hydrogens were added, torsion angles were confirmed and Kollman charges were added in the protein structure. The grid boxes for the blind docking were created in such a way that the whole protein was trapped within that box. Further, Lamarckian Genetic Algorithm (LA) protocol was applied to perform the docking. The lowest energy docked conformation obtained from each docking was saved as .pdb file. That conformation of ligand was merged with the corresponding protein structure and then that was used for the analysis of protein-ligand interactions. Interacting residues of the proteins along with the types of interactions involved were identified using Protein-Ligand Interaction Profiler (https://projects.biotec.tu-dresden.de/plip-web/plip). Molecular visualization and rendering of the structures were done in Pymol.

2.3 Determination of logP value

The *logP* values of the compounds were estimated using SWISSADME (www.swissadme.ch/index.php) server.

3. Results and Discussion

To combat the spreading of COVID-19 infections, inhibition of 3CLpro seems to a potential way out to discontinue the process of viral replication. Many recent articles in this journal have focused their target on 3CLpro to fight SARS-Cov-2 (Kumar et al. 2020; Mittal et al. 2020;

Lobo-Galo et al. 2020; Bhardwaj et al. 2020; Das et al. 2020). Considering the importance of natural compounds, we have selected twenty naturally occurring antiviral molecules and they were docked with 3CLpro using Autodock for screening their potential. In addition to that, we had also docked some control drugs (remdesivir, lopinavir, ritonavir and ribavirin), which are under some clinical trials against SARS-Cov-2. The free energy of binding as estimated by Autodock for all these molecules with 3CLpro is given in Table 1. We had also calculated the *logP* values of these compounds to check their drug likeliness. The *logP* value or lipophilicity is a crucial parameter to understand the cell penetration behavior of a molecule through cell membranes. If *logP* value is more than 5, it suggests reduced absorption and less permeability due to greater molecular hydrophobicity (Ditzinger et al., 2019). The *logP* values of all these compounds are enlisted in the Table 1. Except narasin, all of our selected molecules have the estimated *logP* value less than 5.

Table 1: *logP* values and the docking parameters associated with the lowest energy docked conformation of the compounds with 3CLpro

Sr. No.	Compound	Estimated free energy of binding with 3CLpro (kcal/mol)	Estimated inhibition constant for 3CLpro (μM)	$logP_{O/W}$ (consensus)
1.	**Bilobetin**	**-10.83**	0.011	3.96
2.	**Ginkgetin**	**-10.19**	0.033	4.34
3.	**Sciadopitysin**	**-9.20**	0.180	4.76
4.	Narasin	-5.18	159.86	5.20
5.	Resveratrol	-5.83	53.61	2.48
6.	Esculetin	-5.90	47.30	1.12
7.	Esculin	-5.30	130.02	-0.56
8.	Matrine	-7.41	3.71	1.80
9.	Scutellarin	-5.07	193.03	-0.22
10.	Delphinidin	-6.13	31.88	0.13
11.	Cyanidin	-6.19	28.87	0.56
12.	Pelargodin	-6.73	11.60	2.70
13.	Harmine	-6.02	38.79	2.78
14.	Harmane	-5.47	97.57	2.70
15.	Harmol	-6.35	22.07	2.16

16.	Avarol	-7.35	4.08	4.75
17.	Avarone	-7.93	1.54	4.24
18.	Polyandrocarpidine B	-5.49	94.61	2.76
19.	Polyandrocarpidine D	-5.83	53.60	2.33
20.	Halitunal	-6.30	23.94	4.54
21.	Remdesivir	-4.35	644.64	1.53
22.	Ritonavir	-3.26	4060	5.03
23.	Lopinavir	-4.14	919.24	4.53
24.	Ribavirin	-4.68	373.99	-2.05

From the above table, it was found that the estimated ΔG is very high for bilobetin (-10.83 kcal/mol), ginkgetin (-10.19 kcal/mol) and sciadopitysin (-9.20 kcal/mol). Rest other molecules have the binding energy in the range of -5.07 to -7.93 kcal/mol. The estimated ΔG values for the control drugs like remdesevir, ritonavir, lopinavir and ribavirin are -4.35, -3.26, -4.14 and -4.68 kcal/mol respectively. Based on these values, three molecules from above series namely bilobetin, ginkgetin and sciadopitysin were found as quite promising inhibitors of 3CLpro. We have further extended our study to trace the interactions playing in between 3CLpro and these three molecules. The lowest energy docked conformation of these compounds with 3CLpro along with the major interacting residues from the protein is shown in Figure 1. These residues interact with the molecules using different non-covalent forces such as hydrogen-bonding, hydrophobic, van der Waals, π-alkyl, π-sigma, π-π stacked interactions etc. The substrate binding site of 3CLpro is constituted by the residues Thr 25, Thr 26, His 41, Met 49, Gly 143, Cys 145, Glu 166, Pro 168 etc. A recent report has revealed the role of two catalytic residues namely His 41 and Cys 145 along with some other residues like Gly 143, Cys 145, His 163, His 164, Glu 166, Pro 168 and Gln 189 for effective design of suitable inhibitors with 3CLpro (Zhang et al., 2020). Importance of these residues for the design antiviral compounds as inhibitors of 3CLpro was also supported in another recent publication (Dai et al., 2020). Bilobetin (Fig. 1A) was found to form seven possible H-bonds with the residues Phe 140, Glu 166, Gln 189, Thr 190 and Gln 192. It also has three hydrophobic interactions with Met 165, Glu 166 and Pro 168. Ginkgetin

(Fig. 1B) interacts with the residues Asn 142, Ser 144, Glu 166, Gln 189, Thr 190 and Gln 192 through nine H-bonds and only with Pro 168 through hydrophobic interaction. Sciadopitysin (Fig. 1C) forms three H-bonding with His 41 and Gln 166 and has five hydrophobic interactions with Glu 166, Pro 168 and Gln 192. The catalytic residue His 41 is 4.13 and 2.92Å away from bilobetin and sciadopitysin molecules respectively. Whereas, the distance between ginkgetin and the other catalytic residue Cys 145 is 2.92Å. In addition to these interactions, the inhibitor molecules also have π-cation, π-π stacked and van der Waals interactions. Therefore, blind docking of the naturally occurring antiviral molecules with 3CLpro suggests that the above mentioned three compounds possess excellent inhibitory potential (with nanomolar inhibition constant) towards this protease enzyme. This is because of their strong binding at the catalytic site of the enzyme, which is crucial for viral replication. In case of four control drugs, 3CLpro may not be predicted as suitable molecular target in terms of the binding energy values mentioned in Table 1. The residues of 3CLpro interacting with these control drugs are also mentioned in Table 2.

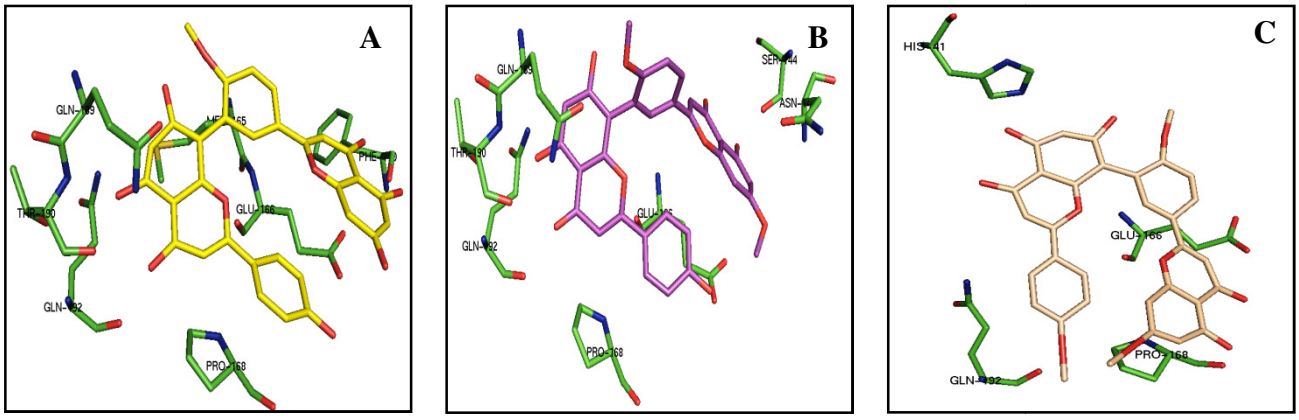


Fig. 1: Lowest energy docked conformation of (A) bilobetin (yellow), (B) ginkgetin (magenta) and (C) sciadopitysin (wheat) with 3CLpro. Interacting protein residues are shown in green color.

Table 2: Residues of 3CLpro interacting with four control drugs

Compound	Residues of 3CLpro interacting with the molecule
Remdesivir	*Hydrophobic interactions:* Arg 4, Lys 5, Met 6, Ala 7, Val 125, Tyr 126 *Hydrogen bonding:* Lys 5, Ala 7, Gly 127, Arg 298
Ritonavir	*Hydrophobic interactions:* Lys 5, Ala 7, Val 125, Tyr 126, Gln 127, Glu 288, Phe 291 *Hydrogen bonding:* Lys 5
Lopinavir	*Hydrophobic interactions:* Tyr 239, Met 276, Ala 285 *Hydrogen bonding:* Leu 271, Gly 278, Ala 285
Ribavirin	*Hydrogen bonding:* Ile 152, Tyr 154, Arg 298

Considering admirable inhibitory capability of these three molecules on 3CLpro, their binding was also studied with another protease PLpro of SARS-CoV-2. The binding of four control drugs (remdesivir, liponavir, ritonavir and ribavirin) with PLpro was also checked. In Table 3, it has been found that the estimated ΔG is very high for bilobetin (-10.83 kcal/mol), ginkgetin (-10.19 kcal/mol) and sciadopitysin (-9.20 kcal/mol). The catalytic residues Cys 111 and His 272 (residue numbering according to the pdb file) of the active site of PLpro are present in S1 pocket. But the substrate binding site is most probably the S3/S4 pockets, which are much more spacious than the S1/S2 pockets situated very close to the catalytic residues (Goswami et al., 2020; Arya et al., 2020). The residues from Asp 164 to Glu 167, Met 208, Cys 217, Ala 246 to Pro 248, Tyr 264, Gly 266 to Gln 269, Gly 271, Tyr 273, Thr 301 and Asp 302 are present in the binding region of PLpro (Goswami et al., 2020; Arya et al., 2020). When we looked into the residues of PLpro interacting with these three molecules (Table 4), it was noticed that only ginkgetin is binding in the S3/S4 pockets (Fig. 2). This molecule is interacting closely with the residues of that pocket as mentioned above. Therefore, ginkgetin is expected to inhibit the proteolytic activity of PLpro as its binding in that region is expected to inhibit the enzymatic activity of PLpro (Arya et al., 2020).

Table 3: Docking parameters associated with the lowest energy docked conformation of the compounds with PLpro, spike protein S1, ACE2 receptor and RdRp

Sr. No.	Compound	Docking with PLpro		Docking with Spike protein S1		Docking with ACE2		Docking with RdRp	
		Binding energy (kcal/mol)	Inhibition constant (µM)	Binding energy (kcal/mol)	Inhibition constant (µM)	Binding energy (kcal/mol)	Inhibition constant (µM)	Binding energy (kcal/mol)	Inhibition constant (µM)
1.	Bilobetin	-9.63	0.088	-11.13	0.007	-9.91	0.054	-9.49	0.11
2.	Ginkgetin	-6.81	10.11	-11.23	0.006	-7.96	1.46	-9.78	0.067
3.	Sciadopitysin	-9.44	0.121	-11.31	0.005	-8.90	0.298	-9.26	0.162
4.	Remdesivir	-2.73	9910	NP	NP	NP	NP	-3.52	2630
5.	Ritonavir	-3.29	3860	NP	NP	NP	NP	NP	NP
6.	Lopinavir	-4.51	497.85	NP	NP	NP	NP	NP	NP
7.	Ribavirin	-3.95	1280	-5.03	205.19	NP	NP	-3.32	3670
8.	Hydroxychloroquine	NP	NP	NP	NP	-5.77	58.63	NP	NP

NP - docking not performed

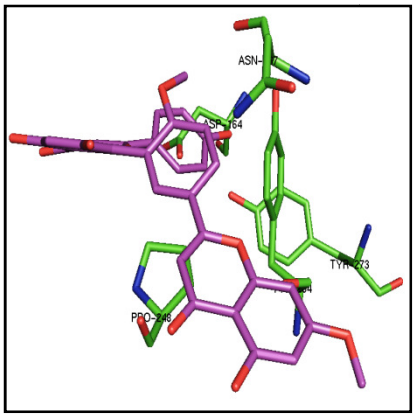


Fig. 2: Lowest energy docked conformation of ginkgetin (magenta) with PLpro. Interacting protein residues are shown in green color.

Table 4: Residues of PLpro, spike protein S1, ACE2 receptor and RdRp interacting with three naturally occurring compounds

Compound	Residues of PLpro interacting with the molecule	Residues of Spike protein S1 interacting with the molecule	Residues of ACE2 receptor interacting with the molecule	Residues of RdRp interacting with the molecule
Bilobetin	*Hydrophobic interactions:* Arg 65, Val 66, Ala 68, Phe 69, Thr 74, Thr 75, Phe 79, Leu 80 *π-stacking:* Phe 69	*Hydrophobic interactions:* Phe 342, Ala 363, Val 367, Leu 368, Phe 374 *Hydrogen bonding:* Cys 336, Glu 340, Asn 343, Asp 364, Tyr 365, Val 367, Leu 368	*Hydrophobic interactions:* Leu 95, Asn 210 *Hydrogen bonding:* Trp 203, Asp 206, Asn 210, Ala 396, Glu 398, Arg 514, Glu 564, Trp 566	*Hydrophobic interactions:* Tyr 273, Leu 329, Val 330, Arg 331, Lys 332, Val 341 *Hydrogen bonding:* Leu 270, Pro 328, Val 330, Arg 331, Thr 344
Ginkgetin	*Hydrophobic interactions:* Asp 164, Pro 248, Tyr 264 *Hydrogen bonding:* Asp 164, Asn 267, Tyr 273	*Hydrophobic interactions:* Pro 337, Phe 342, Ala 363, Val 367, Leu 368, Phe 374 *Hydrogen bonding:* Cys 336, Glu 340, Asp 364, Tyr 365, Val 367, Leu 368, Ser 371	*Hydrophobic interactions:* Asn 290, Ile 291, Leu 370, Phe 438 *Hydrogen bonding:* Asp 292, Asp 367, Phe 438, Gln 442 *π-cation:* Lys 441	*Hydrophobic interactions:* Arg 249, Val 315, Leu 460, Pro461 *Hydrogen bonding:* Ala 250, Leu 251, Ser 255, Thr 319, Arg 349, Phe 396, Asn 628
Sciadopitysin	*Hydrophobic interactions:* Tyr 213, Glu 214, Tyr 305, Lys 306 *Hydrogen bonding:* Lys 217, Glu 307	*Hydrophobic interactions:* Pro 337, Phe 338, Glu 340, Phe 342, Val 367, Leu 368, Phe 374 Hydrogen bonding: Cys 336, Glu 340, Asp 364, Tyr 365, Leu 368 *π-stacking:* Phe 342	*Hydrophobic interactions:* Ala 99, Gln 102, Glu 398, Lys 562 *Hydrogen bonding:* Ala 99, Gln 102, Glu 208, Asn 394 *π-cation:* Lys 562	*Hydrophobic interactions:* Tyr 420, Leu 437, Phe 440, Phe 843 *Hydrogen bonding:* Gly 413, Tyr 420, Glu 436, Leu 437, Phe 441 *π-stacking:* Phe 415, Phe 843

A major hot spot is recently identified in the spike protein S1 of SARS-CoV-2 for its binding with ACE2 receptor (Veeramachaneni et al., 2020). This binding region in the spike protein is composed of Lys 417, Asn 487, Gln 493, Gln 498 and Tyr 505. The values of estimated free energy of binding with spike protein S1 are highly negative in case of these three molecules (Table 3). But the binding site for bilobetin, ginkgetin and sciadopitysin (interacting residues enlisted in Table 4) in the spike protein is quite different than the predicted hot spot for receptor binding. So in this case, these molecules probably will not be effective to prevent the binding of the spike protein with its receptor on host cells. Similarly, the binding hotspot in ACE2 receptor is composed with Lys 31, His 34, Glu 35, Glu 37, Asp 38 and Try 83 (Veeramachaneni et al., 2020). In this case also, none of three molecules binds in that region of ACE2 to prevent the binding of spike protein S1 of SARS-CoV-2 with ACE2. In case of RdRp, two aspartic acid residues namely Asp 760 and Asp 761 (residue numbering as per pdb file) constitute the active site. From Table 4, it is also clear that these three molecules are not binding to the active site of RdRp also.

Conclusion

Using docking tool, three amentoflavone (bilobetin, ginkgetin and sciadopitysin) were predicted to inhibit the main protease (3CLpro) of SARS-CoV-2, which is very important for viral replication. Among them, ginkgetin was also identified as an inhibitor of papain like protease (PLpro) of that virus. When these three promising molecules were docked with other molecular targets associated with SARS-CoV-2 (spike protein S1, RNA dependent RNA polymerase and angiotensin converting enzyme 2 (ACE2) receptor), it was observed that they are not binding to the active sites or hot spots of those targets. These observations are solely based on the results

from blind docking with protein molecules and that need to be further corroborated with experimental results to end up with a fruitful conclusion.

Acknowledgement

KSG is grateful to the Director, NIT Hamirpur for research supports and motivations to work against COVID-19. Authors are thankful to Prashant Yadav, Navdeeshwar Suman, Ayushi Aggarwal and Khusboo Kumari (UG students from NIT Hamirpur) for their supports in some docking experiments. Special thanks to Anchal Sharma, NIT Hamirpur for her assistance in the preparation of the manuscript and Dr. Atanu Singha Roy, NIT Meghalaya for helpful discussions.

References

Aanouz, I., Belhassan, A., El-Khatabi, K., Lakhlifi, T., El-Ldrissi, M., & Bouachrine, M. (2020). Moroccan Medicinal plants as inhibitors against SARS-CoV-2 main protease: Computational investigations. *Journal of Biomolecular Structure and Dynamics,* DOI: 10.1080/07391102.2020.1758790

Abdelli, I., Hassani, F., Bekkel Brikci, S., & Ghalem, S. (2020). In silico study the inhibition of angiotensin converting enzyme 2 receptor of COVID-19 by Ammoides verticillata components harvested from Western Algeria. *Journal of Biomolecular Structure and Dynamics*, DOI: 10.1080/07391102.2020.1763199

Ahmad, S., Abbasi, H. W., Shahid, S., Gul, S., & Abbasi S. W. (2020). Molecular Docking, Simulation and MM-PBSA Studies of Nigella Sativa Compounds: A Computational Quest to identify Potential Natural Antiviral for COVID-19 Treatment. *Journal of Biomolecular Structure and Dynamics*, DOI: 10.1080/07391102.2020.1775129

Akram, M., Tahir, I. M., Shah, S. M. A., Mahmood, Z., Altaf, A., Ahmad, K., Munir, N., Daniyal, M., Nasir, S., & Mehboob, H. (2018). Antiviral potential of medicinal plants against HIV, HSV, influenza, hepatitis, and coxsackievirus: a systematic review. *Phytotherapy Research, 32,* 811-822. DOI: 10.1002/ptr.6024

Arya, R., Das, A., Prashar, V., Kumar, M. (2020). Potential inhibitors against papain-like protease of novel coronavirus (SARS-CoV-2) from FDA approved drugs, *ChemRxiv Preprint*.

Békés, M., Rut, W., Kasperkiewicz, P., Mulder, M. P. C., Ovaa, H., Drag, M., Lima, C. D., & Huang, T. T. (2015). SARS hCoV papain-like protease is a unique Lys48 linkage-specific di-distributive deubiquitinating enzyme. *Biochemical Journal, 468,* 215–226. DOI: 10.1042/BJ20141170

Bhardwaj, V. K., Singh, R., Sharma, J., Rajendran, V., Purohit R., & Kumar, S. (2020). Identification of bioactive molecules from tea plant as SARS-CoV-2 main protease inhibitors. *Journal of Biomolecular Structure and Dynamics*, DOI: 10.1080/07391102.2020.1766572

Borkotoky S., & Banerjee M. (2020). A computational prediction of SARS-CoV-2 structural protein inhibitors from Azadirachta indica (Neem). *Journal of Biomolecular Structure and Dynamics*, DOI: 10.1080/07391102.2020.1774419

Brierley, I., Boursnell, M. E., Binns, M. M., Bilimoria, B., Blok, V. C., Brown, T. D., & Inglis, S. C. (1987). An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. *The EMBO Journal, 6,* 3779-3785. DOI: 10.1002/j.1460-2075.1987.tb02713.x

Cui, J., Li, F., & Shi, Z. L. (2019). Origin and evolution of pathogenic coronaviruses. *Nature Reviews Microbiology, 17,* 181-192.

Dai, W., Zhang, B., Su, H., Li, J., Zhao, Y., Xie, X., et al. (2020). Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. Science, DOI: 10.1126/science.abb4489.

Das, S., Sarmah, S., Lyndem, S. & Singha Roy, A. (2020). An investigation into the identification of potential inhibitors of SARS-CoV-2 main protease using molecular docking study. *Journal of Biomolecular Structure and Dynamics*, DOI: 10.1080/07391102.2020.1763201

Ditzinger, F., Price, D. J., Ilie, A., Köhl, N. J., Jankovic, S., Tsakiridou, G. et al. (2019). Lipophilicity and hydrophobicity considerations in bio-enabling oral formulations approaches - a PEARRL review. *J Pharm Pharmacol* 71, 464–482. DOI:10.1111/jphp.12984.

Elfiky, A. A. (2020). SARS-CoV-2 RNA dependent RNA polymerase (RdRp) targeting: an insilico perspective. *Journal of Biomolecular Structure and Dynamics*, DOI: 10.1080/07391102.2020.1761882

Enmozhi, S. K., Raja, K., Sebastine, I., & Joseph, J. (2020). Andrographolide as a potential inhibitor of SARS-CoV-2 main protease: an in silico approach. *Journal of Biomolecular Structure and Dynamics*, DOI: 10.1080/07391102.2020.1760136

Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., Blinov, V. M. (1989). Coronavirus genome: prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis. *Nucleic Acids Research, 17,* 4847-4861. DOI: 10.1093/nar/17.12.4847

Goswami, D., Kumar, M., Ghosh, S. K., & A. Das (2020). Natural product compounds in Alpinia officinarum and ginger are potent SARS-CoV-2 papain-like protease inhibitors, *ChemRxiv Preprint*.

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T. S., Herrler, G., Wu, N., Nitsche, A., Müller, M. A., Drosten, C., & Pöhlmann, S. (2020). SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell, 181,* 271-280. DOI: 10.1016/j.cell.2020.02.052

Jiang, S., Hillyer, C., Du, L. (2020). Neutralizing antibodies against SARS-CoV-2 and other human Coronaviruses. *Trends in Immunology,* DOI: https://doi.org/10.1016/j.it.2020.03.007

Kumar, A., Choudhir, G., Shukla, S. K., Sharma, M., Tyagi, P., Bhushan, A., & Rathore, M. (2020). Identification of phytochemical inhibitors against main protease of COVID-19 using molecular modeling approaches. *Journal of Biomolecular Structure and Dynamics*, DOI: 10.1080/07391102.2020.1772112

Lai, M. M., & Cavanagh, D. (1997). The molecular biology of coronaviruses. *Advances in Virus Research, 48,* 1-100. DOI: 10.1016/S0065-3527(08)60286-9

Lee, H. J., Shieh, C. K., Gorbalenya, A. E., Koonin, E. V., La Monica, N., Tuler, J., Bagdzhadzhyan, A., & Lai, M. M. (1991). The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. *Virology, 180,* 567-582. DOI: 10.1016/0042-6822(91)90071-i

Lee, P., & Hsueh, P. (2020). Emerging threats from zoonotic coronaviruses-from SARS and MERS to 2019-nCoV, *Journal of Microbiology, Immunology and Infection,* DOI: 10.1016/j.jmii.2020.02.001

Lobo-Galo, N., Terrazas-López, M., Martínez-Martínez A., & Díaz-Sánchez, Á. G. (2020). FDA-approved thiol-reacting drugs that potentially bind into the SARS-CoV-2 main protease, essential for viral replication. *Journal of Biomolecular Structure and Dynamics*, DOI: 10.1080/07391102.2020.1764393.

Mittal, L., Kumari, A., Srivastava, M., Singh M., & Asthana, S. (2020). Identification of potential molecules against COVID-19 main protease through structure-guided virtual screening approach. *Journal of Biomolecular Structure and Dynamics*, DOI: 10.1080/07391102.2020.1768151

Patridge, E., Gareiss, P., Kinch, M. S., & Hoyer, D. (2016). An analysis of FDA-approved drugs: natural products and their derivatives. *Drug Discovery Today, 21,* 204-207. DOI: 10.1016/j.drudis.2015.01.009

Su, S., Wong, G., Shi, W., Liu, J., Lai, A. C., Zhou, J., Liu, W., Bi, Y., & Gao, G. F. (2016). Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends in Microbiology, 24,* 1490–502. DOI: 10.1016/j.tim.2016.03.003

Thomford, N. E., Senthebane, D. A., Rowe, A., Munro, D., Seele, P., Maroyi, A., & Dzobo, K. (2018). Natural Products for Drug Discovery in the 21st Century: Innovations for Novel Drug Discovery. *International Journal of Molecular Sciences, 19,* 1578. DOI: 10.3390/ijms19061578

Veeramachaneni, G. K., Thunuguntla, V. B. S. C., Janaki Ram, B. & Bondili J. S. (2020). Structural and Simulation analysis of hot spot residues interactions of SARS-CoV 2 with Human ACE2 receptor. *Journal of Biomolecular Structure and Dynamics,* DOI: 10.1080/07391102.2020.1773318

Wahedi, H. M., Ahmad S., & Abbasi S. W. (2020). Stilbene-based natural compounds as promising drug candidates against COVID-19. *Journal of Biomolecular Structure and Dynamics*, DOI: 10.1080/07391102.2020.1762743

Wu, F., Zhao, S., Yu, B., Chen, Y., Wang, W., & Song, Z. (2020). A new coronavirus associated with human respiratory disease in China. *Nature, 579,* 265–269. DOI: 10.1038/s41586-020-2008-3

Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., Becker, S., Rox, K., & Hilgenfeld, R. (2020) Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science*, *368*, 409-412.

Ziebuhr, J., Snijder, E. J., & Gorbalenya, A. E. (2000). Virus-encoded proteinases and proteolytic processing in the Nidovirales. *Journal of General Virology, 81,* 853-879. DOI: 10.1099/0022-1317-81-4-853.

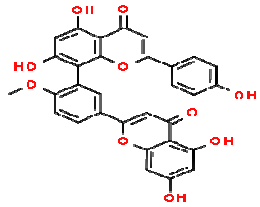
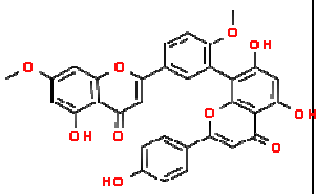
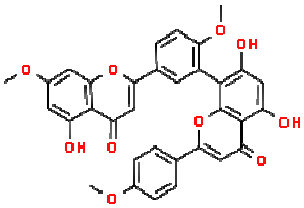
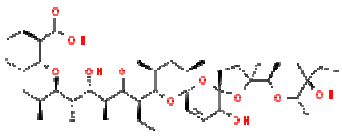
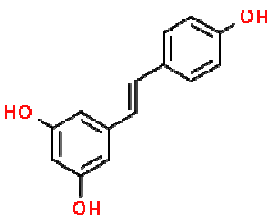
Supplementary Materials

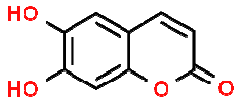
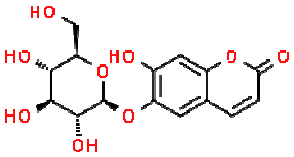
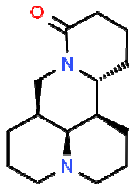
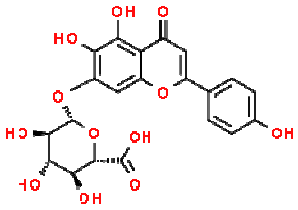
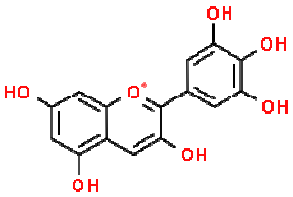
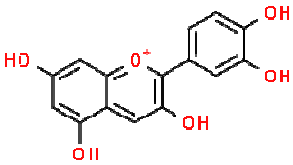
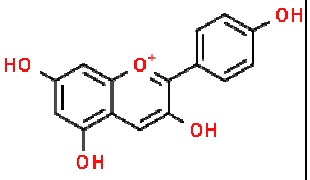
Virtual screening of naturally occurring antiviral molecules for SARS-CoV-2 mitigation using docking tool on multiple molecular targets

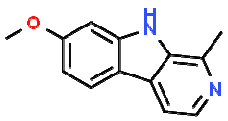
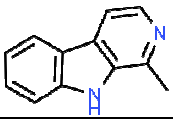
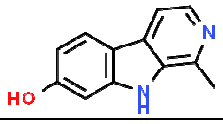
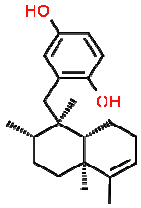
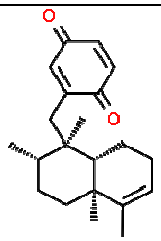
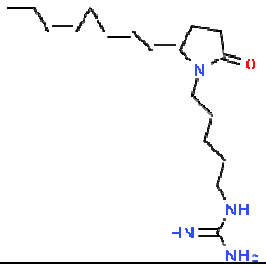
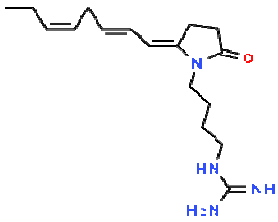
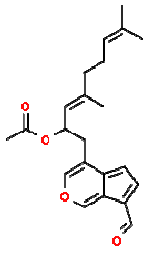
Shiwani Rana, Sanjay Sharma and Kalyan Sundar Ghosh*

Department of Chemistry, National Institute of Technology Hamirpur, H.P. 177005, India

Table S1: Naturally occurring compounds

Sr. No.	Compound	Structure	logP (consensus)	ChemSpider ID	Source	Antiviral activity
1.	Bilobetin		3.96	4474758	*Torreya nucifera*	SARS-CoV [1]
2.	Ginkgetin		4.34	4436579	*Torreya nucifera*	SARS-CoV [1]
3.	Sciadopitysin		4.76	4445013	*Torreya nucifera*	SARS-CoV [1]
4.	Narasin		5.20	58911		Dengue virus (DENV-2) [2]
5.	Resveratrol		2.48	392875	blueberries, cranberries etc.	Respiratory syncytial virus [3]

6.	Esculetin		1.12	4444764	*Lactuca virosa*	Porcine Circovirus Type 2 (PCV2) [4]
7.	Esculin		-0.56	4444765	*Lactuca virosa*	Porcine Circovirus Type 2 (PCV2) [4]
8.	(+)-Matrine		1.80	82591	*Sophora flavescens* and *Sophora tonkinensis*	Human Enterovirus 71 [5]
9.	Scutellarin		-0.22	161366	*Scutellaria barbata*	Porcine Reproductive and Respiratory Syndrome Virus [6]
10.	Delphinidin		0.13	114185	Blueberry	West Nile Virus, Zika Virus, and Dengue Virus [7]
11.	Cyanidin		0.56	114193	Bilberry	H1N1 Influenza Virus [8]
12.	Pelargonidin		2.70	389676	*Solanum tuberosum*	InfV A and B [9]

13.	Harmine		2.78	4444445	*Peganum harmala*	murine cytomegalovirus and Sindbis virus [10]
14.	Harmane		2.70	4444755	*Symplocos setchuensis*	HIV [11]
15.	Harmol		2.16	10296888	*Peganum harmala*	Dengue Virus [12]
16.	Avarol		4.75	65156	Disidea avara	HIV [13]
17.	Avarone		4.24	65157	Disidea avara	HIV [13]
18.	Polyandrocarpidine B		2.76	10469701	Polyandrocarpa sp.	HSV [14]
19.	Polyandrocarpidine D		2.33	10469702	Polyandrocarpa sp.	HSV [14]
20.	Halitunal		4.54	8513688	*Halimeda tuna*	murine coronavirus A59 [15]

References:

[1] Ryu, Y. B., Jeong, H. J., Kim, J. H., Kim, Y. M., Park, J., Kim, D., Naguyen, T. T. H., Park, S., Chang, J. S., Park, K. H., Rho, M., & Lee, W. S. (2010). Biflavonoids from Torreya nucifera displaying SARS-CoV 3CLpro inhibition. *Bioorganic & Medicinal Chemistry, 18,* 7940-7947. DOI: 10.1016/j.bmc.2010.09.035.

[2] Low JS, Wu KX, Chen KC, Ng MM, Chu JJ. Narasin, a novel antiviral compound that blocks dengue virus protein expression. Antivir Ther 2011;16:1203-18.

[3] Zang N, Xie X, Deng Y, Wu S, Wang L, Peng C, et al. Resveratrol-mediated gamma interferon reduction prevents airway inflammation and airway hyperresponsiveness in respiratory syncytial virus-infected immunocompromised mice. J Virol 2011;85:13061-8.

[4] N. Sun, T. Yu, J. Zhao, Y. Sun, J. Jiang, Z. Duan, W. Wang, Y. Hu, H. Lei, and H. Li, Antiviral Activities of Natural Compounds Derived from Traditional Chinese Medicines against Porcine Circovirus Type 2 (PCV2), Biotechnology and Bioprocess Engineering 20 (2015) 180-187

[5] Y. Yang, J. Xiu, X. Zhang, L. Zhang, K. Yan, C. Qin, and J. Liu, Antiviral Effect of Matrine against Human Enterovirus 71, Molecules. 2012, 17, 10370–10376.

[6] J. Cheng, N. Sun, X. Zhao, L. Niu, M. Song, Y. Sun, J. Jiang, J. Guo, Y. Bai, J. He, H. Li, In Vitro Screening for Compounds Derived From Traditional Chinese Medicines With Antiviral Activities Against Porcine Reproductive and Respiratory Syndrome Virus, J Microbiol Biotechnol, 2013, 23, 1076-1083. doi: 10.4014/jmb.1303.03074.

[7] Á. Vázquez-Calvo, N. J. de Oya, M. A. Martín-Acebes, E. Garcia-Moruno, and J. Saiz, Antiviral Properties of the Natural Polyphenols Delphinidin and Epigallocatechin Gallate against the Flaviviruses West Nile Virus, Zika Virus, and Dengue Virus, Front Microbiol. 2017; 8: 1314.

[8] S. Kannan, P. Kolandaivel, The Inhibitory Performance of Flavonoid cyanidin-3-sambubiocide Against H274Y Mutation in H1N1 Influenza Virus, J Biomol Struct Dyn. 2018, 36, 4255-4269.

[9] Hayashi, K., Mori, M., Knox, Y. M., Suzutan, T., Ogasawara, M., Yoshida, I., et al. (2003). Anti influenza virus activity of a red-fleshed potato anthocyanin. Food Sci. Technol. Res. 9 (3), 242–244. doi: 10.3136/fstr.9.242

[10] J. B. Hudson, E. A. Graham, G. H. N. Towers, ANTIVIRAL EFFECT OF HARMINE, A PHOTOACTIVE β-CARBOLINE ALKALOID, 1986, 43, Pages 21-26.

[11] J Ishida, H K Wang, M Oyama, M L Cosentino, C Q Hu, K H Lee, Anti-AIDS Agents. 46. Anti-HIV Activity of Harman, an anti-HIV Principle From Symplocos Setchuensis, and Its Derivatives, J Nat Prod, 2001, 64, 958-960.

[12] V. M Quintana, L. E Piccini, J. D P. Zénere, E. B. Damonte, M. A. Ponce, V. Castilla, Antiviral Activity of Natural and Synthetic β-Carbolines Against Dengue Virus, Antiviral Res, 2016, 134, 26-33.

[13] S. Sagar, M. Kaur, and K. P. Minneman, Antiviral Lead Compounds from Marine Sponges, Mar Drugs. 2010, 8, 2619–2638.

[14] Rinehart K.L. Biologically Active Marine Natural Products. Pure & Appl. Chem. 1989, 61, 525–528.

[15] F. E. Koehn, Sartah, Gunasekera, P., D. Nadine, N. S. Cross, Halitunal, an unusual diterpene aldehyde from the marine alga Halimeda tuna, Tetrahedron Letters, 32, 169-172