



The Open COVID Journal

Content list available at: <https://opencovidjournal.com>



RESEARCH ARTICLE

Identification of Naturally Occurring Antiviral Molecules for SARS-CoV-2 Mitigation

Shiwani Rana¹, Prateek Kumar², Anchal Sharma¹, Sanjay Sharma¹, Rajanish Giri^{2*} and Kalyan S. Ghosh^{1*}

¹Department of Chemistry, National Institute of Technology Hamirpur, Hamirpur, H.P. 177005, India

²School of Basic Sciences, Indian Institute of Technology Mandi, Mandi, H.P. 175005, India

Abstract:

Aim:

This study aimed to virtually screen the naturally occurring antiviral molecules for SARS-CoV-2 mitigation based on multiple molecular targets using docking and molecular dynamics simulations.

Background:

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 a pandemic by the World Health Organization. It has resulted in a huge number of infections as well as fatalities due to severe lower respiratory tract sickness.

Objective:

The objective of this study was the identification of inhibitors against the crucial molecular targets linked with viral infection caused by SARS-CoV-2.

Materials and Methods:

In silico screening of twenty naturally occurring antiviral molecules was performed using the Autodock docking tool. Further, molecular dynamics (MD) simulations were performed on the most stable docked complex between cysteine-like protease or 3CL protease (3CLpro) and the best-identified inhibitor (bilobetin).

Results:

3CLpros is one of the very important molecular targets as it is involved in the replication process of the virus. In the present study, we have initially investigated the inhibitory potential of naturally occurring antiviral molecules against the activity of main viral protease (3CLpro) to put a halt to viral replication. The investigation had been carried out through docking of the molecules with 3CLpro. Based on the results, the three most potential molecules (bilobetin, ginkgetin and sciadopitysin) have been screened. Further, these molecules were subjected to checking their activity on other molecular targets like papain-like protease (PLpro), spike protein S1, RNA dependent RNA polymerase (RdRp), and Angiotensin-Converting Enzyme 2 (ACE2) receptor. In addition to 3CLpro inhibition, ginkgetin was also predicted as an inhibitor of PLpro. However, none of these three compounds was found to be effective on the rest of the molecular targets. Molecular Dynamics (MD) simulations of the most stable docked complex between 3CLpro and its best inhibitor (bilobetin) confirmed notable conformational stability of the docked complex under a dynamic state.

Conclusion:

Bilobetin alone or a combination of bilobetin and ginkgetin may be used to impede viral replication. These observations are solely based on the results from blind docking with protein molecules and need to be further corroborated with experimental results.

Keywords: SARS-Cov-2, Cysteine like protease inhibition, Papain-like protease inhibition, Bilobetin, Ginkgetin, Sciadopitysin.

Article History

Received: December 22, 2020

Revised: March 10, 2021

Accepted: April 13, 2021

1. INTRODUCTION

Novel coronavirus (SARS-CoV-2) was probably surfaced in a seafood market in Wuhan city of China at the end of 2019. This virus had engulfed almost the entire globe in 2020, and it still continues to do so. On February 11, 2020, it was termed COVID-19 and recognized as a cause of viral pneumonia, which had made a huge population sick. It has been declared a global health emergency and a pandemic by World Health Organization (WHO) on January 30, 2020. Since its emergence, around ~2.8 million people have died, and ~127 million have been reported being infected. Coronavirus belongs to *coronaviridae* family (*nidovirales* order) and can cause respiratory illness ranging from the common cold to severe acute respiratory syndrome (SARS-CoV) with symptoms like fever, cough, and shortness of breath [1, 2]. Different coronaviruses were earlier detected in different animals like pigs, camel, bat, *etc.*, and they may be transmitted from those animals to humans. This phenomenon is commonly termed a spillover event. Out of seven known coronaviruses, four have the ability to cause mild to moderate respiratory tract disease, and the rest of them can cause severe and lethal illnesses [3]. HCoV-OC43, HCoV-OC63, HCoV-OC229E, and HKU1 fall in the category of alpha-coronavirus, which can cause modest respiratory illnesses. The beta-coronavirus such as SARS-CoV and MERS-CoV have the potency to cause severe and fatal respiratory lower tract infection [4]. The novel coronavirus associated with COVID-19 is also a beta-coronavirus and has similarities with SARS-CoV. For that reason, it has been termed as SARS-CoV-2. In November 2002, SARS-CoV was first detected in Asia, and its further spread occurred in twenty-six countries. In September 2012, another respiratory syndrome, *i.e.*, Middle East Respiratory Syndrome (MERS-CoV) caused by coronavirus had been reported. The coronaviruses have a large genome sequence of 30 kb in length with 5' cap and 3' poly-A tail [5]. The spherical virion SARS-CoV-2 has a diameter of about 60-140 nm and is constituted with peplomers of crown shape [6]. The structure mainly consists of membrane proteins, nucleoplasmid (enclosed RNA), lipid membrane, spike protein, and envelope proteins [7, 8]. The spike glycoproteins on the viral capsid play a crucial role during the internalization of the virus into the host cells. These proteins bind with the angiotensin-converting enzyme 2 (ACE2) receptor present on the surface of host cells and assist the virus in injecting RNA into cells [9]. Upon viral infection, the RNA is processed to synthesize two polyproteins (pp1a/pp1ab) [10]. The transcription process occurs through the formation of the Replication-Transcription Complex (RCT). In a typical RNA genome, there are a minimum of six Open Reading Frames (ORFs) that function as templates to produce subgenomic mRNAs. The frameshift mutation between ORF1a and ORF1b [11] encodes both pp1a/pp1ab polyproteins. The polyprotein pp1ab contains more than 7,000 residues and also possesses putative RNA-dependent RNA polymerase (RdRp) and RNA helicase activities [10, 12].

Further, these polyproteins are cleaved to form functional proteins by two proteases encoded in the virus, namely 3-Chymotrypsin-Like protease (3CLpro) or main protease and Papain-Like protease (PLpro) [13]. 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. PLpro assists the proteolytic process and also removes ubiquitin to protect viruses from immune responses [14]. Therefore, the prime protease 3CLpro, which basically generates the functional proteins required for viral replication, has attracted the attention of many researchers as a potential drug target against SARS-CoV-2 [15 - 20].

To identify potential drug candidates against SARS-CoV-2, we have adopted computational tools to screen out some inhibitor molecules against 3CLpro. Several reports on the identification of different potent therapeutics against SARS-CoV-2 using computational tools were published in the recent past [15 - 24]. In the present work, we have searched for the inhibitors in nature as several antiviral compounds were found to be present in different medicinal plants [25]. Moreover, plant-derived naturally occurring compounds play a significant role in the discovery of many effective drugs in the past, and they were approved further [26, 27].

2. MATERIALS AND METHODS

2.1. Preparation of the Structures of Small Molecules and Proteins for Docking

Three-dimensional structures (as .mol file) of twenty naturally occurring compounds and control drugs were collected from ChemSpider (www.chemspider.com). Their natural source, structures, and previously reported antiviral activities are given in Supplementary materials in Table S1. Geometry and energy optimization of their 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 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 or proteins and the crystallographic water molecules were removed from the structures.

2.2. Molecular Docking

Protein-ligand dockings were performed by using *Autodock 4.2*. Before docking, hydrogens were added, torsion angles were confirmed, and Kollman charges were added to the protein structure. The grid boxes for the blind docking were made in such a way that the whole protein in each case was enclosed within that box. Further, the Lamarckian Genetic Algorithm (LA) protocol was applied in the docking. The lowest energy docked conformation of a ligand obtained from each docking was saved as a .pdb file. That conformation of ligand was merged with the corresponding protein structure, and then that was used for the analysis of protein-ligand

* Address correspondence to these authors at the Department of Chemistry, National Institute of Technology Hamirpur, H.P. 177005, India; Tel: +91-1972-254104; E-mail: kalyan@nith.ac.in and School of Basic Sciences, Indian Institute of Technology Mandi, H.P. 175005, India; Tel: +91-1905-267154; E-mail: rajanishgiri@iitmandi.ac.in

interactions. Interacting residues of the proteins along with the types of interactions were identified using Protein-Ligand Interaction Profiler (<https://projects.biotech.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 predicted using the *SWISSADME* (www.swissadme.ch/index.php) server.

2.4. Molecular Dynamics (MD) Simulation

3CLpro-bilobetin complex (as bilobetin appeared with the highest docking score with 3CLpro) was selected to probe binding stability. For this purpose, MD simulation of the complex was performed over a period of 150 ns. Here, we had used SPC water model using the GROMOS54A7 force field executed in Gromacs 5.1.2 [28]. A prior simulation setup containing protein-ligand complex was developed and pre-equilibrated under biological environment. 0.15 M salt concentration and 4 Na⁺ were used for neutralizing the simulation setup in the cubic simulation box. For energy minimization of the protein-ligand complex in the arranged system, the steepest descent method for 50,000 steps was employed. Using the SHAKE algorithm, the equilibration runs for the system were performed under NPT and NVT conditions for 1 ns each. To maintain the average temperature and pressure at 300K and 1 bar, we used the V-rescale thermostat (a modified algorithm of Berendsen thermostat) and Parrinello-Rahman barostat methods, respectively.

2.5. Analysis of MD Simulation

The Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), and the Radius of gyration (R_g) for C- α atoms were calculated using *gmx*, *rms*, *rmsf*, and *gyrate* commands, respectively. Global conformational analysis as derived from the Principal Components (PC) of the system was used to estimate almost the whole dynamics of the system. It was measured using *covar* and *anaeig* commands in *Gromacs*. Solvation (polar and non-polar) free energy and molecular mechanics (mm) potential energy (electrostatic and Van der Waals energy) were calculated using *g_mmpbsa* tool [29, 30].

3. RESULTS AND DISCUSSION

To restrict the spread of COVID-19 infections, inhibition of 3CLpro seems to be a potential way to discontinue the process of viral replication. Many recent articles have focused their target on 3CLpro to fight SARS-CoV-2, as mentioned earlier in this article. 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 preclinical and clinical trials against SARS-CoV-2. Based on the clinical and preclinical reports as obtained from *in vitro* animal models as well as clinical studies of the drugs such as remdesivir, hydroxychloroquine, lopinavir, and ritonavir, their emergency

use was approved in 2020 for the treatment of COVID-19 [31].

Further clinical studies reported mixed outcomes on the effectiveness of these drugs [31]. Their efficiency also depends on the clinical protocol frameworks. In the present study, we have chosen those drug molecules as control drugs to have a comparative idea about the effectiveness of our selected test molecules. The free energy of binding all these molecules with 3CLpro as estimated by *Autodock* is given in Table 1.

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

S. No.	Compound	Estimated Free Energy of Binding with 3CLpro (kcal/mol)	Estimated Inhibition Constant for 3CLpro (μ M)	<i>logP_{ow}</i> (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.	Pelargonidin	-6.73	11.60	2.70
13.	Harmine	-6.02	38.79	2.78
14.	Harmene	-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.	Polyandrocarpine B	-5.49	94.61	2.76
19.	Polyandrocarpine 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	-4.38	373.99	-2.05

We also calculated the *logP* value of these compounds to check their drug likeliness. Basically, when a solute is distributed between two solvents, the ratio of the concentration of that solute in those solvents is termed partition coefficient (*P*). If a compound is partitioned between water and a nonpolar solvent, the *logP* value is considered as a measure of lipophilicity or hydrophobicity of that compound. In general, the concentrations of the solute in the nonpolar and polar phase are placed in the numerator and denominator, respectively, to determine the value of *P*. 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 [32]. The *logP* values of all these compounds are enlisted in Table 1. Except for narasin, all of our selected molecules have an estimated *logP* value of less than 5.

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). 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 remdesivir, ritonavir, lopinavir, and ribavirin are -4.35, -3.26, -4.14, and -4.38 kcal/mol, respectively. Based on these values, three molecules from our series, namely bilobetin, ginkgetin, and sciadopitysin were found to be promising inhibitors of 3CLpro. We have further extended our study to trace the nature of interactions between 3CLpro and these three molecules. The lowest energy docked conformation of these compounds with 3CLpro is shown in Fig. (1). The residues of 3CLpro, which strongly interact with bilobetin, ginkgetin, and sciadopitysin are also marked in Fig. (1). Different non-covalent forces such as hydrogen-bonding, hydrophobic, Van der Waals, π -stacking interactions, etc., were found to be involved in protein-ligand interactions. 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 [33]. The importance of these residues in the design of antiviral compounds as inhibitors of 3CLpro was also supported in another recent publication [34]. 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 with Pro 168 through hydrophobic interaction. Sciadopitysin (Fig. 1C) is involved in three H-bonding with His 41 and Gln 166 and five hydrophobic interactions with Glu 166, Pro 168, and Gln 192. The catalytic residue His 41 is 4.13 and 2.92 Å away from bilobetin and sciadopitysin molecules, respectively. In contrast, the distance between ginkgetin and the other catalytic residue Cys 145 is 2.92 Å. In addition to these interactions, the inhibitor molecules may also have π -cation, π -stacking, and Van der Waals interactions.

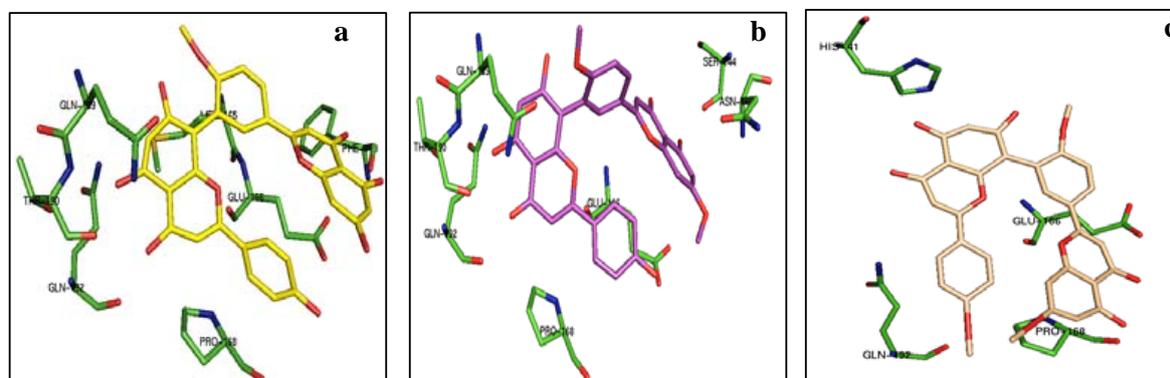


Fig. (1). Lowest energy docked conformation of (a) bilobetin (yellow), (b) ginkgetin (magenta) and (c) sciadopitysin (wheat) with 3CLpro. Interacting protein residues are shown in green color.

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 the case of four control drugs, 3CLpro was not predicted as a 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. Among them, only ribavirin was found to bind at the catalytic site of 3CLpro.

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

Compound	Residues of 3CLpro Interacting with the Molecule
Remdesivir	<i>Hydrophobic interactions:</i> Arg 4, Lys 5, Met 6, Ala 7, Val 125, Tyr 126 <i>Hydrogen bonding:</i> Lys 5, Ala 7, Gly 127, Arg 298
Ritonavir	<i>Hydrophobic interactions:</i> Lys 5, Ala 7, Val 125, Tyr 126, Gln 127, Glu 288, Phe 291 <i>Hydrogen bonding:</i> Lys 5
Lopinavir	<i>Hydrophobic interactions:</i> Tyr 239, Met 276, Ala 285 <i>Hydrogen bonding:</i> Leu 271, Gly 278, Ala 285
Ribavirin	<i>Hydrogen bonding:</i> His 164, Gln 189, Thr 190, Gln 192

3.1. Molecular Dynamics Simulations

MD simulations were performed to investigate the binding stability of topmost screened compound (bilobetin) with 3CLpro of SARS-CoV-2. Bilobetin, out of screened 24 compounds, had shown a significantly high binding energy with 3CLpro. The extensive MD simulation up to 150 ns had been analyzed in terms of mean deviation, fluctuation, and radius of gyration of the system. As observed in the trajectory, the 3CLpro complex with bilobetin had shown an increasing trend (0.2-0.3 nm) in RMSD values up to 60 ns while it has been stabilized after 100 ns with slight fluctuations in the middle (Fig. 2A). The average RMSD of the complex was calculated as ~0.28 nm for 150 ns simulation time which is in a favorable range. The mean fluctuation (measured as RMSF) was observed to be in good correlation with the RMSD of the complex.

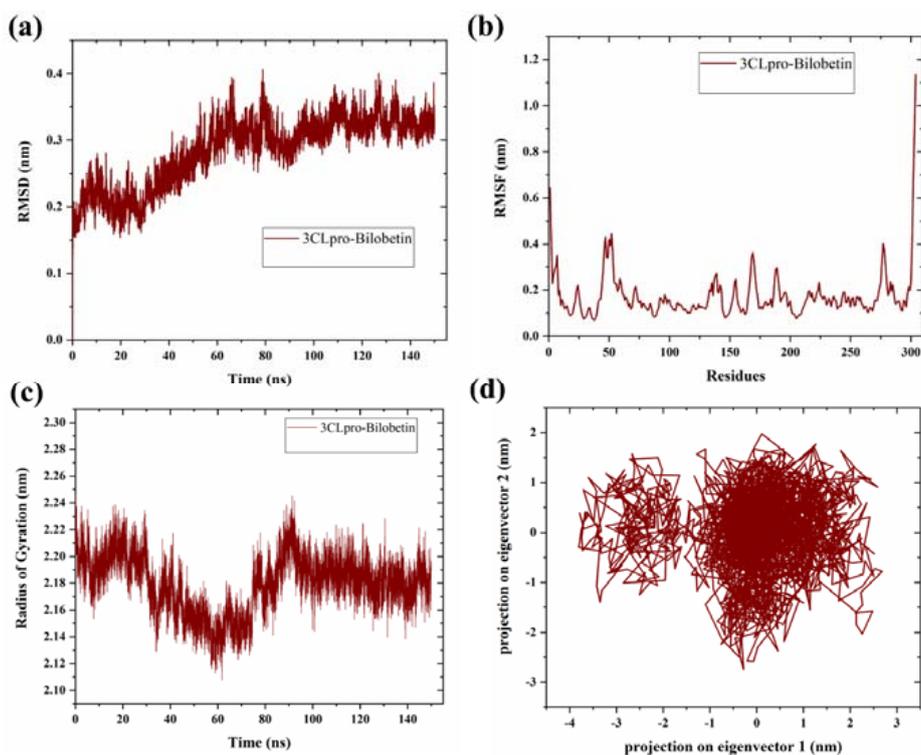


Fig. (2). Molecular dynamics simulation analysis of bilobetin bound 3CLpro of SARS-CoV-2 for 150 ns time period— (a) root mean square deviation (RMSD); (b) root mean square fluctuation (RMSF); (c) radius of gyration and (d) principal component analysis (PCA).

3CLpro-bilobetin complex revealed slightly higher fluctuation upto 0.4 nm near residues 46-53 and the rest of the protease did not show such fluctuations (Fig. 2B). An intense peak was also detected at C-terminal, which may be due to the presence of disordered residues at the terminal [35]. The radius of gyration (a parameter to measure biomolecule compactness) was also determined to study the effect of the binding of bilobetin on the structure of 3CLpro. Like RMSD and RMSF, bilobetin-3CLpro complex had shown fluctuations till nearly half time of the simulation and then attained stability after 85 ns with a lesser radius. The average radius of gyration of bilobetin bound complex was calculated to be approximately 2.21 nm for 150 ns long MD simulations (Fig. 2C). Lesser value of the radius of gyration suggests more compact structure of the complex. This also indicates that bilobetin could bind to 3CLpro with significant binding efficacy. As a conclusion, these three parameters had suggested quite stable binding of bilobetin with 3CLpro but that needs to be validated through experimental studies.

3.2. Principal Component Analysis (PCA)

Principal components or principal motions are the main and responsible components of a trajectory to determine the overall variations that occurred during the simulation time. The conformational alterations can be identified through the PCA for the C- α atoms of the protein. To investigate the overall change in the structural conformations, we have analyzed the

simulation trajectory through these principal motions. Using the normalized data from the trajectory, a co-variance matrix of eigenvalues and eigenvectors was created from first two components that define the overall motion in the system [24061923, 32457393, 31443266]. PCA generates eigenvectors which represent the reduced collective atomic motions in a protein structure [36]. The last 25 ns of the simulation trajectory was analyzed and it was found that the bilobetin bound 3CLpro complex had occupied a less conformational space, as projected with two principal components (Fig. 2D). A cluster is formed by eigenvectors with a distance range -1 to 2 nm and a scattered region is shown at -3.5 to -2 nm. Overall, significantly effective binding of bilobetin indicates that it could act as a potential druggable molecule against SARS-CoV-2.

3.3. Binding Energy Estimation

The binding energy was estimated for the whole trajectory. Using *g_mmpbsa* tool, we had calculated the solvation (polar and non-polar) free energy and molecular mechanics (mm) potential energy (electrostatic and van der waals energy) [29, 30]. The average binding energy was estimated to be -141.38 ± -23.91 kJ/mol, which is quite acceptable (Fig. 3). The solvent accessible surface area (SASA) energy was found to be -21.15 ± -1.58 kJ/mol. The estimated binding energy value also supports the observations from MD simulation and strongly suggests further validation of the protein-ligand binding through *in vitro* experiments.

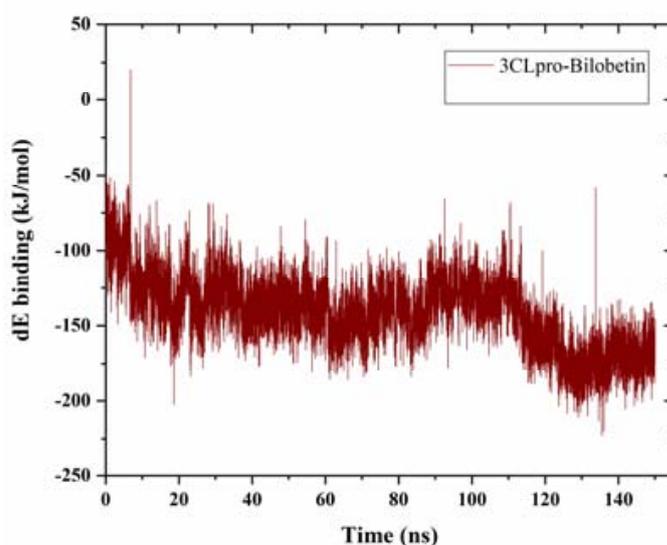


Fig. (3). Binding energy estimation for 150 ns MD simulation of 3CLpro-bilobetin complex using g_mmpbsa tool in Gromacs simulation package.

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. The substrate binding site is most probably located in the S3/S4 pockets, which are much more spacious

than the S1/S2 pockets situated very close to the catalytic residues [37]. 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 substrate binding region of PLpro [37]. Considering the residues of PLpro interacting with these three molecules (Table 4), it was noticed that only ginkgetin can bind in the S3/S4 pockets (Fig. 4). This molecule interacts 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 can inhibit the enzymatic activity of PLpro.

Table 3. Docking results associated with the lowest energy docked conformation of the three screened compounds with PLpro, spike protein S1, ACE2 receptor and RdRp.

S. No.	Compound	Docking with PLpro		Docking with Spike protein S1		Docking with ACE2		Docking with RdRp	
		Binding energy (kcal/mol)	Inhibition constant (μ M)	Binding energy (kcal/mol)	Inhibition constant (μ M)	Binding energy (kcal/mol)	Inhibition constant (μ M)	Binding energy (kcal/mol)	Inhibition constant (μ M)
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.

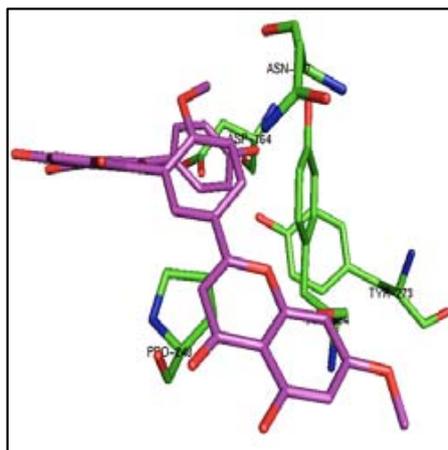


Fig. (4). Lowest energy docked conformation of ginkgetin (magenta) with PLpro. Interacting protein residues are shown in green color.

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

A major hotspot is recently identified in the spike protein S1 of SARS-CoV-2 for its binding with ACE2 receptor [38]. This binding region in the spike protein is composed of Lys 417, Asn 487, Gln 493, Gln 498 and Tyr 505. Another recent article [23] also identified the residues Tyr 453, Arg 454, Leu 455, Lys 458, Ser 459, Ser 469, Glu 471, Pro 491, Leu 492, Gln 493 and Tyr 489 as a part of binding site. The values of estimated free energy of binding with spike protein S1 are highly negative in case of these three molecules (Table 3). However, the binding site for bilobetin, ginkgetin and sciadopitysin (interacting residues enlisted in Table 4) in the spike protein is quite different than the predicted hotspot for receptor binding. In this case, these molecules probably will not be effective in preventing the binding of the spike protein with its receptor on host cells. Similarly, the binding hotspot in ACE2 receptor is composed of Lys 31, His 34, Glu 35, Glu 37, Asp 38 and Tyr 83 [38]. In this case also, none of the three

molecules bind in that region of ACE2 to prevent its binding with 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 [22, 39]. From Table 4, it is also clear that these three molecules do not bind to the active site of RdRp.

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. Molecular dynamics simulation analysis revealed that the highest scoring inhibitor bilobetin can form a complex with 3CLpro with high conformational stability. Among these molecules, ginkgetin was also identified as an inhibitor of Papain-Like protease (PLpro) of SARS-CoV-2. 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 do not bind to the active sites or hotspots on those protein targets. These observations are solely based on the results from blind docking with protein molecules and they need to be further corroborated with experimental results for a fruitful conclusion.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are the basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The data supporting the findings of this study are available within the article.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

KSG is grateful to the Director, NIT Hamirpur for research supports. Authors are thankful to Prashant Yadav, Navdeeshwar Suman, Ayushi Aggarwal, and Khusboo Kumari (UG students from NIT Hamirpur) for their support in some docking experiments. AS is thankful to NIT Hamirpur for her fellowship.

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publisher's website along with the published article.

REFERENCES

- Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol* 2019; 17(3): 181-92. [http://dx.doi.org/10.1038/s41579-018-0118-9] [PMID: 30531947]
- Chary MA, Barbuto AF, Izadmeh S, Hayes BD, Burns MM. COVID-19: Therapeutics and their toxicities. *J Med Toxicol* 2020; 16(3): 284-94. [http://dx.doi.org/10.1007/s13181-020-00777-5] [PMID: 32356252]
- Su S, Wong G, Shi W, *et al.* Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends Microbiol* 2016; 24(6): 490-502. [http://dx.doi.org/10.1016/j.tim.2016.03.003] [PMID: 27012512]
- Lee PI, Hsueh PR. Emerging threats from zoonotic coronaviruses—from SARS and MERS to 2019-nCoV. *J Microbiol Immunol Infect* 2020; 53(3): 365-7. [http://dx.doi.org/10.1016/j.jmii.2020.02.001] [PMID: 32035811]
- Lu R, Zhao X, Li J, *et al.* Genomic characterisation and epidemiology of 2019 novel coronavirus: Implications for virus origins and receptor binding. *Lancet* 2020; 395(10224): 565-74. [http://dx.doi.org/10.1016/S0140-6736(20)30251-8] [PMID: 32007145]
- Wu F, Zhao S, Yu B, *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* 2020; 579(7798): 265-9. [http://dx.doi.org/10.1038/s41586-020-2008-3] [PMID: 32015508]
- Zhang YZ, Holmes EC. A Genomic Perspective on the Origin and Emergence of SARS-CoV-2. *Cell* 2020; 181(2): 223-7. [http://dx.doi.org/10.1016/j.cell.2020.03.035] [PMID: 32220310]
- Jiang S, Hillier C, Du L. Neutralizing antibodies against SARS-CoV-2 and other human coronaviruses. *Trends Immunol* 2020; 41(5): 355-9. [http://dx.doi.org/10.1016/j.it.2020.03.007] [PMID: 32249063]
- Hoffmann M, Kleine-Weber H, Schroeder S, *et al.* SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 2020; 181(2): 271-280.e8. [http://dx.doi.org/10.1016/j.cell.2020.02.052] [PMID: 32142651]
- Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM. Coronavirus genome: Prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis. *Nucleic Acids Res* 1989; 17(12): 4847-61. [http://dx.doi.org/10.1093/nar/17.12.4847] [PMID: 2526320]
- Brierley I, Bournsnel ME, Binns MM, *et al.* An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. *EMBO J* 1987; 6(12): 3779-85. [http://dx.doi.org/10.1002/j.1460-2075.1987.tb02713.x] [PMID: 3428275]
- Lee HJ, Shieh CK, Gorbalenya AE, *et al.* The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. *Virology* 1991; 180(2): 567-82. [http://dx.doi.org/10.1016/0042-6822(91)90071-1] [PMID: 1846489]
- Ziebuhr J, Snijder EJ, Gorbalenya AE. Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J Gen Virol* 2000; 81(Pt 4): 853-79. [http://dx.doi.org/10.1099/0022-1317-81-4-853] [PMID: 10725411]
- Békés M, Rut W, Kasperkiewicz P, *et al.* SARS hCoV papain-like protease is a unique Lys48 linkage-specific di-distributive deubiquitinating enzyme. *Biochem J* 2015; 468(2): 215-26. [http://dx.doi.org/10.1042/BJ20141170] [PMID: 25764917]
- Cardoso WB, Mendanha SA. Molecular dynamics simulation of docking structures of SARS-CoV-2 main protease and HIV protease inhibitors. *J Mol Struct* 2021; 1225: 129143. [http://dx.doi.org/10.1016/j.molstruc.2020.129143] [PMID: 32863430]
- Ghosh R, Chakraborty A, Biswas A, Chowdhuri S. Identification of alkaloids from *Justicia adhatoda* as potent SARS CoV-2 main protease inhibitors: An *in silico* perspective. *J Mol Struct* 2021; 1229: 129489. [http://dx.doi.org/10.1016/j.molstruc.2020.129489] [PMID: 33100380]
- Elzupir AO. Inhibition of SARS-CoV-2 main protease 3CL^{pro} by means of α -ketoamide and pyridone-containing pharmaceuticals using *in silico* molecular docking. *J Mol Struct* 2020; 1222: 128878. [http://dx.doi.org/10.1016/j.molstruc.2020.128878] [PMID: 32834113]
- Chhetri A, Chhetri S, Rai P, Sinha B, Brahman D. Exploration of inhibitory action of Azo imidazole derivatives against COVID-19 main protease (M^{pro}): A computational study. *J Mol Struct* 2021; 1224: 129178. [http://dx.doi.org/10.1016/j.molstruc.2020.129178] [PMID: 32904625]
- Baidya N, Ghosh NN, Chattopadhyay AP. Inhibitory activity of hydroxychloroquine on COVID-19 main protease: An insight from MD-simulation studies. *J Mol Struct* 2020; 1219: 128595. [http://dx.doi.org/10.1016/j.molstruc.2020.128595] [PMID: 32834108]
- Bolelli K, Ertan-Bolelli T, Unsalan O, Altunayar-Unsalan C. Fenoterol and dobutamine as SARS-CoV-2 main protease inhibitors: A virtual screening study. *J Mol Struct* 2021; 1228: 129449. [http://dx.doi.org/10.1016/j.molstruc.2020.129449] [PMID: 33071354]
- Baidya N, Khan AA, Ghosh NN, Dutta T, Chattopadhyay AP. Screening of potential drug from *Azadirachta Indica* (Neem) extracts for SARS-CoV-2: An insight from molecular docking and MD-simulation studies. *J Mol Struct* 2021; 1227: 129390. [http://dx.doi.org/10.1016/j.molstruc.2020.129390] [PMID: 33041371]
- Venkateshan M, Muthu M, Suresh J, Ranjith Kumar R. Azafluorene derivatives as inhibitors of SARS CoV-2 RdRp: Synthesis, physicochemical, quantum chemical, modeling and molecular docking analysis. *J Mol Struct* 2020; 1220: 128741. [http://dx.doi.org/10.1016/j.molstruc.2020.128741] [PMID: 32834110]
- Kulkarni SA, Nagarajan SK, Ramesh V, Palaniyandi V, Selvam SP, Madhavan T. Computational evaluation of major components from

- plant essential oils as potent inhibitors of SARS-CoV-2 spike protein. *J Mol Struct* 2020; 1221: 128823. [http://dx.doi.org/10.1016/j.molstruc.2020.128823] [PMID: 32834111]
- [24] Sepay N, Sekar A, Halder UC, Alarif A, Afzal M. Anti-COVID-19 terpenoid from marine sources: A docking, admet and molecular dynamics study. *J Mol Struct* 2021; 1228: 129433. [http://dx.doi.org/10.1016/j.molstruc.2020.129433] [PMID: 33071352]
- [25] Akram M, Tahir IM, Shah SMA, *et al.* Antiviral potential of medicinal plants against HIV, HSV, influenza, hepatitis, and coxsackievirus: A systematic review. *Phytother Res* 2018; 32(5): 811-22. [http://dx.doi.org/10.1002/ptr.6024] [PMID: 29356205]
- [26] Patridge E, Gareiss P, Kinch MS, Hoyer D. An analysis of FDA-approved drugs: Natural products and their derivatives. *Drug Discov Today* 2016; 21(2): 204-7. [http://dx.doi.org/10.1016/j.drudis.2015.01.009] [PMID: 25617672]
- [27] Thomford NE, Senthane DA, Rowe A, *et al.* Natural products for drug discovery in the 21st century: Innovations for novel drug discovery. *Int J Mol Sci* 2018; 19(6): 1578. [http://dx.doi.org/10.3390/ijms19061578] [PMID: 29799486]
- [28] Berendsen HJC, Spoel DV, Drunen RV. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput Phys Commun* 1995; 91: 43-56. [http://dx.doi.org/10.1016/0010-4655(95)00042-E]
- [29] Kumari R, Kumar, R. Open Source Drug Discovery Consortium, & A. Lynn, g_mmpbsa—A GROMACS tool for high-throughput MM-PBSA calculations. *J Chem Inf Model* 2014; 54: 1951-62. [http://dx.doi.org/10.1021/ci500020m] [PMID: 24850022]
- [30] Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 2001; 98(18): 10037-41. [http://dx.doi.org/10.1073/pnas.181342398] [PMID: 11517324]
- [31] Maciorowski D, Idrissi SZE, Gupta Y, *et al.* A Review of the Preclinical and Clinical Efficacy of Remdesivir, Hydroxychloroquine, and Lopinavir-Ritonavir Treatments against COVID-19. *SLAS Discov* 2020; 25(10): 1108-22. [http://dx.doi.org/10.1177/2472555220958385] [PMID: 32942923]
- [32] Ditzinger F, Price DJ, Ilie AR, *et al.* Lipophilicity and hydrophobicity considerations in bio-enabling oral formulations approaches - A Pearn review. *J Pharm Pharmacol* 2019; 71(4): 464-82. [http://dx.doi.org/10.1111/jphp.12984] [PMID: 30070363]
- [33] Zhang L, Lin D, Sun X, *et al.* Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α -ketoamide inhibitors. *Science* 2020; 368(6489): 409-12. [http://dx.doi.org/10.1126/science.abb3405] [PMID: 32198291]
- [34] Dai W, Zhang B, Jiang XM, *et al.* Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* 2020; 368(6497): 1331-5. [http://dx.doi.org/10.1126/science.abb4489] [PMID: 32321856]
- [35] Giri R, Bhardwaj T, Shegane M, *et al.* Understanding COVID-19 via comparative analysis of dark proteomes of SARS-CoV-2, human SARS and bat SARS-like coronaviruses. *Cell Mol Life Sci* 2021; 78(4): 1655-88. [http://dx.doi.org/10.1007/s00018-020-03603-x] [PMID: 32712910]
- [36] David CC, Jacobs DJ. Principal component analysis: A method for determining the essential dynamics of proteins. *Methods Mol Biol* 2014; 1084: 193-226. [http://dx.doi.org/10.1007/978-1-62703-658-0_11] [PMID: 24061923]
- [37] Gao X, Qin B, Chen P, *et al.* Crystal structure of SARS-CoV-2 papain-like protease. *Acta Pharm Sin B* 2021; 11(1): 237-45. [http://dx.doi.org/10.1016/j.apsb.2020.08.014] [PMID: 32895623]
- [38] Veeramachaneni GK, Thunuguntla VBSC, Bobbillapati J, Bondili JS. Structural and simulation analysis of hotspot residues interactions of SARS-CoV 2 with human ACE2 receptor. *J Biomol Struct Dyn* 2020; 1-11. [http://dx.doi.org/10.1080/07391102.2020.1773318] [PMID: 32448098]
- [39] Elfiky AA. SARS-CoV-2 RNA dependent RNA polymerase (RdRp) targeting: An *in silico* perspective. *J Biomol Struct Dyn* 2020; 1-9. [http://dx.doi.org/10.1080/07391102.2020.1761882] [PMID: 32338164]