



# Antiviral phytochemicals “ellagic acid” and “(+)-sesamin” of *Bridelia retusa* identified as potential inhibitors of SARS-CoV-2 3CL pro using extensive molecular docking, molecular dynamics simulation studies, binding free energy calculations, and bioactivity prediction

Abd. Kakhar Umar<sup>1</sup> · James H. Zothantluanga<sup>2</sup> · Keerthic Aswin<sup>3</sup> · Saipul Maulana<sup>4</sup> · Muhammad Sulaiman Zubair<sup>5</sup> · H. Lahlhenmawia<sup>6</sup> · Mithun Rudrapal<sup>7</sup> · Dipak Chetia<sup>2</sup>

Received: 23 November 2021 / Accepted: 28 April 2022 / Published online: 10 May 2022  
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

## Abstract

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has infected billions and has killed millions to date. Studies are being carried out to find therapeutic molecules that can potentially inhibit the replication of SARS-CoV-2. 3-chymotrypsin-like protease (3CL pro) involved in the polyprotein cleavage process is believed to be the key target for viral replication, and hence is an attractive target for the discovery of antiviral molecules. In the present study, we aimed to identify natural phytochemicals from *Bridelia retusa* as potential inhibitors of SARS-CoV-2 3CL pro (PDB ID: 6M2N) using in silico techniques. Molecular docking studies conducted with three different tools in triplicates revealed that ellagic acid (BR6) and (+)-sesamin (BR13) has better binding affinity than the co-crystal inhibitor “3WL” of 6M2N. BR6 and BR13 were found to have a high LD<sub>50</sub> value with good bioavailability. 3WL, BR6, and BR13 bind to the same active binding site and interacted with the HIS41-CYS145 catalytic dyad including other crucial amino acids. Molecular dynamics simulation studies revealed stability of protein–ligand complexes as evidenced from root-mean-square deviations, root-mean-square fluctuations (RMSF), protein secondary structure elements, ligand-RMSF, protein–ligand contacts, ligand torsions, and ligand properties. BR6 (−22.3064 kcal/mol) and BR13 (−19.1274 kcal/mol) showed a low binding free energy value. The Bayesian statistical model revealed BR6 and BR13 as better protease inhibitors than 3WL. Moreover, BR6 and BR13 had already been reported to elicit antiviral activities. Therefore, we conclude that ellagic acid and (+)-sesamin as natural antiviral phytochemicals with inhibitory potential against SARS-CoV-2 3CL pro.

**Keywords** SARS-CoV-2 · *Bridelia retusa* · Ellagic acid · Sesamin · Molecular docking · Molecular dynamics

✉ James H. Zothantluanga  
jameshzta@gmail.com

Abd. Kakhar Umar  
abdulkaharumar@gmail.com

Keerthic Aswin  
keerthicaswin.s@gmail.com

Saipul Maulana  
saifulmaulana011@gmail.com

Muhammad Sulaiman Zubair  
sulaiman\_zubair80@yahoo.co.id

H. Lahlhenmawia  
hlenmawia@gmail.com

Mithun Rudrapal  
rsmrpal@gmail.com

Dipak Chetia  
dchetia@dibru.ac.in

- 1 Department of Pharmaceutics and Pharmaceutical Technology, Faculty of Pharmacy, Universitas Padjadjaran, Jatinangor, 45363 Bandung, Indonesia
- 2 Department of Pharmaceutical Sciences, Faculty of Science and Engineering, Dibrugarh University, Dibrugarh, 786004 Assam, India
- 3 Council of Scientific and Industrial Research-Institute of Genomics and Integrative Biology, New Delhi 110020, India
- 4 Department of Pharmacognosy and Phytochemistry, Faculty of Pharmacy, Airlangga University, Surabaya, Indonesia
- 5 Department of Pharmacy, Faculty of Math and Natural Sciences, Universitas Tadulako, Palu City, Indonesia
- 6 Department of Pharmacy, Regional Institute of Paramedical and Nursing Sciences, Aizawl 796001, Mizoram, India
- 7 Department of Pharmaceutical Chemistry, Rasiklal M. Dhariwal Institute of Pharmaceutical Education & Research, Pune 411019, Maharashtra, India

## Introduction

Coronavirus disease 2019 (COVID-19) has spread to all parts of the world. The total number of COVID-19 cases has reached 2.5 billion cases, with the death toll reaching over 5 million [1]. The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the causative agent of COVID-19 [2–4]. During viral replication, SARS-CoV-2 produces large viral polyproteins which exhibit excessive variability. These polyproteins undergo proteolytic cleavage which is mediated by the 3-chymotrypsin-like protease (3CL pro) enzyme at 11 different sites [5, 6]. This enzyme (3CL pro) breaks down polyproteins into a variety of non-structural proteins that are essential for viral replication [6]. As 3CL pro is essential for viral replication of SARS-CoV-2, the protein was investigated as a potential drug target for the development of anti-SARS-CoV-2 agents [7, 8]. Several studies have used 3CL pro as the target protein to identify a potential inhibitor of SARS-CoV-2 [9–11].

Although modern medicines and vaccines are available to treat COVID-19, the emergence of new variants of SARS-CoV-2 might weaken the efficacy of the present vaccines and antiviral medications [12, 13]. Adverse drug reactions are sometimes reported against COVID-19 vaccines resulting in vaccine hesitancy [14–16]. SARS-CoV-2 continues to mutate potentially leading to the emergence of deadlier variants with higher transmissibility [17]. Therefore, the development of new/adjuvant/alternative drugs against SARS-CoV-2 is the need of the hour. Among different approaches that are available for the discovery and development of drugs for COVID-19, investigating the bioactive molecules of plant origin seems to be a promising approach [6, 10, 11]. Many plants such as *Lycoris radiata*, *Artemisia annua*, *Pyrrhosia lingaua*, *Lindera aggregata*, *Isatis indigotica*, *Torreya nucifera*, *Houttuynia cordata*, *Echinacea*, *Cinchona*, *Curcuma longa*, and *Curcuma xanthorrhiza* have been reported to contain antiviral phytoconstituents that might be effective against SARS-CoV-2 [18–20]. A study reported that selected herbal extracts have the potential to inhibit SARS-CoV-2 [21]. Another study reported that the polyphenolic leaf extract of *Vitis vinifera* inhibited in vitro replication of SARS-CoV-2 [22]. Furthermore, several in silico investigations have revealed many phytochemicals as having the potential to inhibit the replication of SARS-CoV-2 by binding with target proteins such as 3CL pro [23–25]. These studies highlight the significance of phytochemicals from medicinal plants in curtailing the viral replication of SARS-CoV-2 by 3CL pro inhibition.

*Bridelia retusa* has been reported as an antiviral medicinal plant of Northeast India [26]. *B. retusa* is also

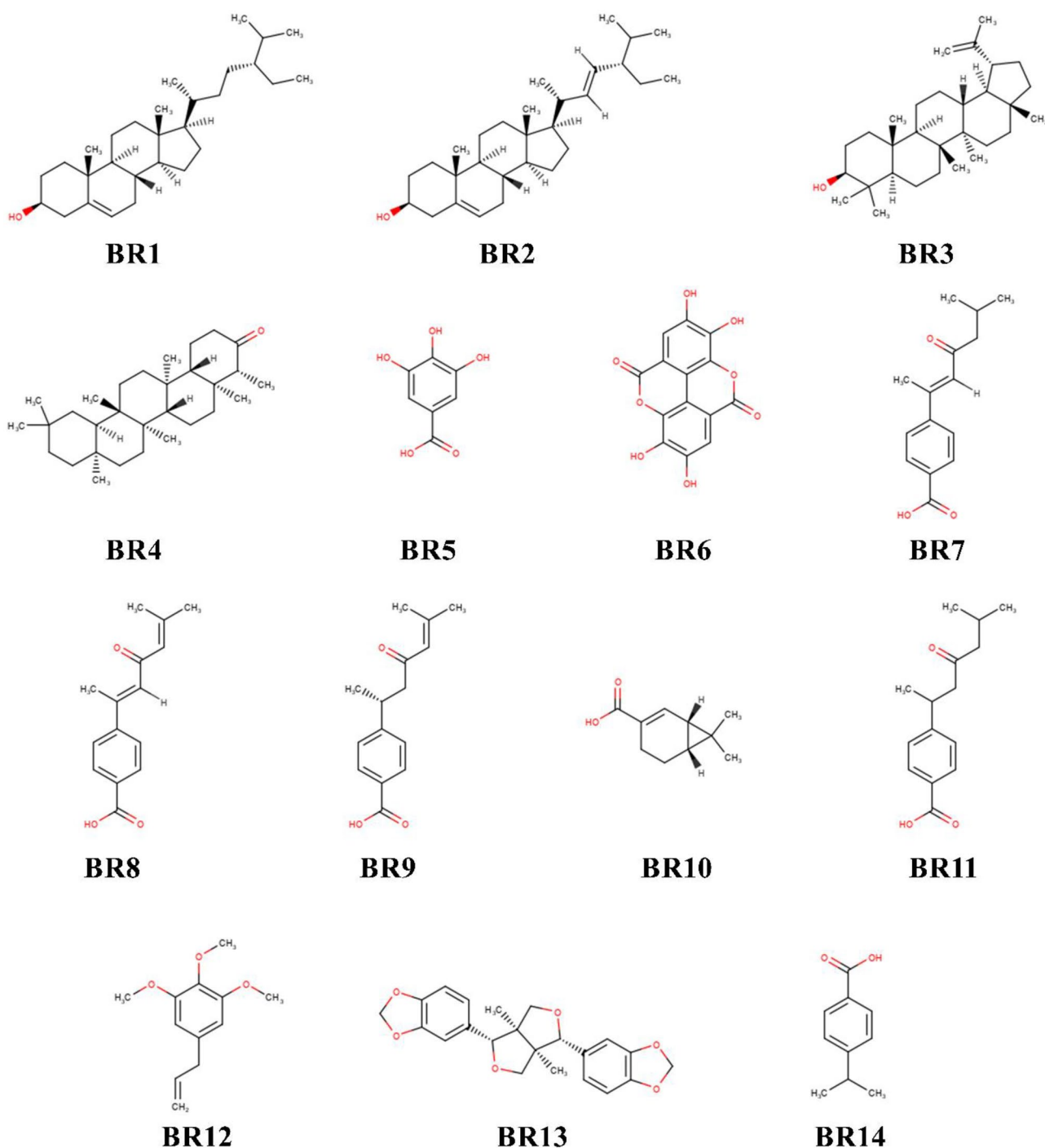
distributed in different parts of the Asian continent. In China, *B. retusa* has been traditionally used as a herbal medicine [27]. According to reported in vivo studies, *B. retusa* elicited antinociceptive, anti-inflammatory, anti-fungal, and antimicrobial activities [28–31]. In this study, fourteen phytochemicals (Fig. 1) namely  $\beta$ -sitosterol (BR1), stigmasterol (BR2), lupeol (BR3), friedelin (BR4), gallic acid (BR5), ellagic acid (BR6), (E)-4-(1,5-dimethyl-3-oxo-1-hexenyl) benzoic acid (BR7), (E)-4-(1,5-dimethyl-3-oxo-1,4-hexadienyl) benzoic acid (BR8), (R)-4-(1,5-dimethyl-3-oxo-4-hexenyl) benzoic acid (BR9), (-)-isochamic acid (BR10), (R)-4-(1,5-dimethyl-3-oxohexyl) benzoic acid (ar-todomatuic acid) (BR11), 5-allyl-1,2,3-trimethoxybenzene (elemicin) (BR12), (+)-sesamin (BR13), and 4-isopropylbenzoic acid (cumic acid) (BR14) that have been reported to be present in *B. retusa* were computationally investigated for their inhibitory potential against the SARS-CoV-2 3CL pro [28–30]. Among the phytochemicals present in *B. retusa*, gallic acid and ellagic acid have been reported to possess antiviral activity [32, 33]. Other phytochemicals present in *B. retusa* such as (+)-sesamin and  $\beta$ -sitosterol were also reported to play an important role in ameliorating viral infection [34, 35].

Therefore, molecular docking simulation studies were carried out with three different docking approaches in triplicates. Phytochemicals with better binding affinity than the co-crystal inhibitor (CCI) were investigated for their toxicity, bioavailability, drug-likeness, and other important parameters. Following this, the phytochemicals were subjected to 100 ns molecular dynamics (MD) simulations to determine the stability of the protein–ligand complexes. The molecular mechanics-generalized born surface area (MM-GBSA) binding free energies of the protein–ligand complexes were calculated. The protease inhibitory potential of the phytochemicals was also evaluated. Finally, two phytochemicals of *B. retusa* (ellagic acid and (+)-sesamin) with reported antiviral activities were found as the most promising bioactive molecules for the inhibition of SARS-CoV-2 3CL pro.

## Materials and methods

### Preparation of phytochemical library

The compound library consisting of the structure data file format of 14 phytochemicals of *B. retusa* was manually prepared with MarvinSketch 20.10 software (ChemAxon). The structure of the prepared phytochemicals was checked with the ‘Structure checker’ add-on feature of the MarvinSketch 20.10 software.



**Fig. 1** Chemical structures of phytochemicals of *B. retusa*

### Retrieval of the target protein

The crystal structure of SARS-CoV-2 3CL pro (PDB ID: 6M2N) was retrieved from the RCSB-PDB website (<https://www.rcsb.org/structure/6M2N>). The co-crystallized inhibitor of 6M2N “3WL” was taken and its structure was also manually prepared with MarvinSketch 20.10 software. The

co-crystallized inhibitor (3WL) of 6M2N was used as a positive control throughout the study.

### Pre-processing of protein

The Discovery Studio Visualizer v20.1.0.19295 software was used to pre-process the protein before further studies.

The software was used to remove water and ligands. After this, polar hydrogens were added to the protein structure. The “define and edit binding site” feature of Discovery Studio Visualizer v20.1.0.19295 software was used to obtain the active binding site coordinates ( $x = -33.516547$ ,  $y = -63.976760$ ,  $z = 41.592773$ ) of SARS-CoV-2 3CL pro.

## Molecular docking studies

In the present study, molecular docking studies were carried out with three different in silico tools. Also, molecular docking was carried out in triplicates.

### DockFlin

The default geometry of all the phytocompounds was removed with the “clean geometry” feature in Discovery Studio 2021 Client (DS) (BIOVIA, San Diego, CA, USA). Energy minimization (MM2) was performed with Chem3D. Each ligand was then optimized with AutoDockTools 1.5.6 (ADT) (The Scripps Research Institute, USA) to add Gasteiger charges, set rotatable bonds, and TORSDOF. Furthermore, the ligands are stored in the PDBQT format.

The protein was optimized using ADT to regulate the charges (Kollman charges). The protein was then stored in the PDBQT format. The grid position was arranged based on the active site coordinates previously defined with DS. The grid dimension was set to  $40 \times 40 \times 40$  magnification with a spacing of  $0.375 \text{ \AA}$ .

Molecular docking was performed with DockFlin software (ETFLIN, Palu City, Indonesia). This software is an in silico tool that allows systematic scheduling of multi-ligand and multi-protein docking processes using the AutoDock Vina. Ligands and proteins were added to their respective list panel, then the docking parameters (grid file) per protein were also loaded in the grid list panel. Loading of all necessary files follows the same order. The docking parameters used were energy range of 4 and exhaustiveness of 8.

### PyRx

The in-built energy minimization parameters of the PyRx 0.8 tool were used to minimize the energy of all the ligands. The energy minimization parameters were kept default (force field = universal force field; optimization algorithm = conjugate; total number of steps = 200; number of steps for update = 1; stop is energy difference is less than = 0.1). The ligands were converted into the PDBQT file format.

The pre-processed protein was converted into the PDBQT file format after it was loaded onto the 3D scene of the PyRx tool. In the Vina search space, the 3D affinity grid box was manually adjusted according to the pre-defined active binding site coordinates so that the entire active site residues are

covered. The size of the 3D affinity grid box ( $25 \text{ \AA}$ ) and the exhaustiveness were kept default ( $n = 8$ ).

Molecular docking simulation with the PyRx 0.8 tool. PyRx is a useful tool that uses AutoDock 4 and AutoDock Vina for docking while providing a simple graphical user interface. Also, it uses ADT and Open Babel for preparing all the necessary files required for docking [36].

### AutoDock Vina standalone package

The target protein was prepared for docking by minimizing its energy (force field = OPLS\_2005 and the minimization proceeds until there is no significant change) with Schrödinger Maestro (Schrödinger, LLC, New York, NY). Kollman charges were added to the target protein with ADT followed by its conversion to “pdbqt” file format.

Energy minimization of all the phytocompounds was carried out with the same procedure and parameters of the PyRx 0.8 tool. Polar hydrogens and Gasteiger charges were added with ADT. The ligands were then converted to the PDBQT file format.

The size of the grid box was set ( $X = 18$  points,  $Y = 14$  points, and  $Z = 18$  points). The active binding site coordinates previously defined with DS were used. AutoDock Vina standalone package was used for molecular docking.

## Toxicity study

The toxicity of the phytocompounds with a binding affinity higher than 3WL was subjected to toxicity screening. Data Warrior v4.5.1 software was used to generate the output of different toxicity parameters such as mutagenic, tumorigenic, reproductive effective, and irritant [37]. The overall drug score was also calculated with the Data Warrior v4.5.1 software.

The ProTox-II web tool ([https://tox-new.charite.de/pro-tox\\_II/](https://tox-new.charite.de/pro-tox_II/)) was then used to predict the LD<sub>50</sub> and toxicity class of the phytocompounds. In addition, hepatotoxicity, carcinogenicity, immunotoxicity, mutagenicity, and cytotoxicity of the phytocompounds were also computed.

## Bioavailability study

Pharmacokinetic parameters such as absorption, distribution, metabolism, and excretion were studied with the SwissADME web tool (<http://www.swissadme.ch/index.php#>). Different physicochemical parameters such as molecular weight, heavy atoms, rotatable bonds, hydrogen bond donors, hydrogen bond acceptors, molar refractivity, topological surface area, lipophilicity, and water solubility were also computed. Then, the bioavailability of the phytocompounds was computed.

## Visualization and analysis of ligand interactions

The protein–ligand interactions shown by the phytochemicals that pass through the ADMET screening were visualized and analyzed with the Discovery Studio Visualizer v20.1.0.19295 software.

## Molecular dynamic simulation studies

Molecular dynamics study was carried out on Linux Ubuntu 20.04.1 LTS 64-bit, NVIDIA Quadro P2200 graphic card, 16 GB RAM, and Intel Xeon(R) W-2223 @3.60 GHz octa-core processor with Desmond on Maestro Schrödinger 2021–2 [38–40]. The molecular dynamics protocol described by Zubair et al. was followed [41]. MD process was commenced by generating complexes of the ligand–protein system from docking results and placing them in a cubic box at 10 Å solvating with SPC (simple point charge) water model. To simulate under physiological conditions, salts ion (sodium and chloride) was set at 0.15 M and added counter ions consisting of 31Na<sup>+</sup> and 29 Cl<sup>−</sup> for original and sesamin, 41Na<sup>+</sup> and 29 Cl<sup>−</sup> for ellagic to neutralize the charge. The MD simulation was performed for 100 ns in NPT conditions with temperature 300 K, pressure 1.63 bar, OPLS\_2005 force field, recording intervals set to 1.2 ps for energy and 20 ps for trajectory.

## MM-GBSA binding free energy calculations

The MD trajectories were generated by the Desmond module from Maestro-Schrodinger 2021–3. As described by Duarte et al. (2021) and Mahmoud et al. (2021), the calculations of the average MM-GBSA binding free energy in Desmond MD trajectory (\*.cms.out) was carried out by executing the thermal\_mmgbsa.py script of Schrodinger [42, 43]. The thermal script processed the MD trajectory of Desmond and splits it into 1001 individual frame snapshots that were used as input to compute the average binding free energy and each individual snapshot runs through MM-GBSA analysis. The MM-GBSA binding free energy was estimated according to the formula:  $\Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{receptor}} - G_{\text{ligand}}$ ,

where  $\Delta G_{\text{bind}}$  is the binding free energy and  $G_{\text{complex}}$ ,  $G_{\text{receptor}}$ , and  $G_{\text{ligand}}$  are the free energies of the complex, receptor, and ligand, respectively. MM-GBSA values were obtained from the script of thermal MM-GBSA provided by Schrödinger (<https://www.schrodinger.com/scriptcenter>) and were used to calculate the  $\Delta G_{\text{bind}}$  for the protein–ligand complexes [44, 45].

## Bioactivity prediction using Bayesian statistical model

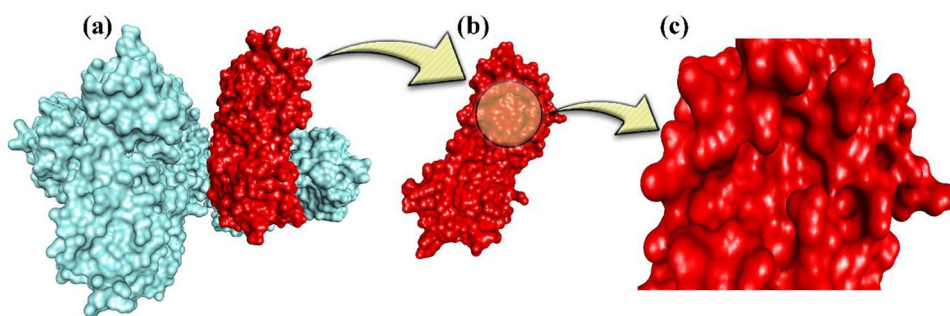
The SMILES ID of the phytochemicals was generated with MarvinSketch 20.10 software. The protease inhibitory activity of the phytochemicals was studied with the Molinspiration Chemoinformatics web tool (<https://www.molinspiration.com/>) which uses a sophisticated Bayesian statistical model to predict the bioactivity of a molecule. Initially, a Bayesian statistical model was generated by fragmenting a library of active molecules and inactive molecules. After the SMILES ID of the phytochemical was entered into the Molinspiration Chemoinformatics web tool, the phytochemicals were also fragmented into ~80,000 fragments. The contribution of each fragment toward the bioactivity was studied to determine the protease inhibitory potential of the phytochemicals.

## Results and discussion

### Features of SARS-CoV-2 3CL pro

The target protein (SARS-CoV-2 3CL pro) used in the present study falls under the classification of viral proteins. The crystal structure of the target protein was obtained from the Research Collaboratory for Structural Bioinformatics-Protein Data Bank (RCSB-PDB) (<https://www.rcsb.org/structure/6M2N>) and has a resolution of 2.20 Å. The viral protein is made up of four chains (A, B, C, and D) and it has a sequence length of 306 amino acids. A snapshot of SARS-CoV-2 3CL pro is given in Fig. 2. The target protein is complexed with a co-crystallized inhibitor “3WL.” Chemically,

**Fig. 2** Snapshot of **a** SARS-CoV-2 3CL pro; **b** chain A of SARS-CoV-2 3CL pro; **c** active binding pocket at chain A of SARS-CoV-2 3CL pro



baicalein or 3WL is a trihydroxyflavone with many pharmacological activities including antiviral activity [46, 47].

### Molecular docking simulation studies

Molecular docking simulation studies (MDSS) is an effective technique to study the binding potential of any compound against a particular target protein. MDSS is now widely recognized and used by many in their drug discovery endeavors [48]. Interestingly, to find out new potential drug candidates against the ongoing pandemic, many scholars have also used the technique of molecular docking [49, 50]. In the present study, MDSS was carried out with three different approaches in triplicates. This initiative was taken to minimize biased outcomes. Although DockFlin, PyRx, and AutoDock Vina standalone packages used the same algorithm (i.e., AutoDock Vina) to carry out molecular docking, preparation of proteins and ligands were carried out with multiple software and different techniques.

Generally, a certain amount of numerical value (scoring function) is given for a ligand. Based on this, the ligands can be ranked in the order of their binding affinity for a target protein [51]. This output of MDSS is useful to filter out potent candidates for a target from ligands with low binding affinity. Higher binding energy means low binding affinity while lower binding energy correlates to a high binding affinity. The binding energy (kcal/mol) of the phyto-compounds of *B. retusa* toward the active binding site of SARS-CoV-2 3CL pro is given in Table 1. The mean of the binding energies calculated from the three simulations was used to prioritize and obtain ligands with the highest binding affinities.

According to DockFlin, BR6 (−8.0 kcal/mol) has a better binding affinity than 3WL (−7.66 kcal/mol). According to PyRx, BR6 (−7.7 kcal/mol) and BR13 (−7.7 kcal/mol) have a better binding affinity than 3WL (−7.3 kcal/mol). According to AutoDock Vina standalone package, BR6 (−8.0 kcal/mol) and BR13 (−7.86 kcal/mol) has a better binding affinity than 3WL (−7.1 kcal/mol). The MDSS conducted with three different approaches in triplicates revealed that BR6 and BR13 can be potential candidates' worthy of further investigations as they had shown a remarkably high binding affinity toward the active binding site of SARS-CoV-2 3CL pro. Thus, BR6 and BR13 were subjected to further analysis. In addition to the binding affinity, ligand interaction is an equally important parameter that can determine a drug's suitability for further studies. To inhibit an enzyme, a drug must interact with the crucial amino acids of an enzyme through hydrogen bonds and hydrophobic interactions [11]. Thus, ligand interaction can also be used to filter out promising drug candidates. As it is industrious to check for protein–ligand interactions for all the ligands used in large-scale drug screening, only the binding affinity was taken

into consideration at this preliminary stage to filter out low-affinity candidates.

### Toxicity study

Toxicity is a serious issue when it comes to drug development. For instance, many drugs had been withdrawn from the market due to toxicity issues. Due to this, several years of hard work, labor, and money that is invested in a drug development program can go to waste [52]. Therefore, toxicity study constitutes an important part of a drug development program. In the present study, different toxicity parameters of BR6 and BR13 were studied with Data Warrior v.4.5.1 software (Table 2) and the ProTox-II web tool (Table 3).

According to Data Warrior v.4.5.1 software (Table 2), both the phyto-compounds and the positive control (3WL) were found to be non-toxic. This is a good indicator that both the phyto-compounds exhibit a good safety profile. Data Warrior v.4.5.1 software was also used to calculate the drug score which correlates to the quality of a compound as a potential drug candidate [53]. The co-crystallized inhibitor (3WL) has the highest drug score (0.6981475) followed by BR6 (0.5104488) and BR13 (0.4094924).

The toxicity class of each compound is presented in Table 2. The ProTox-II web tool has categorized toxicity into different classes which determines the risk associated with each class. 3WL belonged to class 5 which means that 3WL may be harmful if swallowed. BR6 falls under the category of class 4 which indicates that BR6 might be harmful if swallowed. BR13 belonged to class 3 which means that it might be toxic if swallowed.

The LD<sub>50</sub> (median lethal dose) is the amount of a substance that is required to kill half of the participants in a research study [54]. Although 3WL showed signs of carcinogenicity and mutagenicity, it was computed to have a high LD<sub>50</sub> value of 3919 mg/kg body weight (Table 2). According to ProTox-II, BR6 also showed signs of carcinogenicity. However, it has a high LD<sub>50</sub> value of 2991 mg/kg body weight (Table 2). Signs of carcinogenicity and immunotoxicity were observed in BR13. Also, BR13 has the lowest LD<sub>50</sub> (1500 mg/kg body weight) among the three compounds. The experimental LD<sub>50</sub> of medicinal plants and polyherbal formulations showed a lower LD<sub>50</sub> value than BR6 and BR13 [55–57]. BR6 and BR13 can be considered safe and therefore, they are subjected to further studies.

### Physicochemical properties, pharmacokinetics, drug-likeness, and medicinal chemistry

In computational studies such as MDSS, any compound that has a high binding affinity for a target protein does not mean that it will become the potential lead. A compound must also have suitable drug-like properties. Therefore, it

**Table 1** Binding energy (kcal/mol) of the phytochemicals of *B. retusa* toward the active binding site of SARS-CoV-2 3CL pro

Comp	DockFlin				PyRx				AutoDock Vina			
	1st	2nd	3rd	Mean	1st	2nd	3rd	Mean	1st	2nd	3rd	Mean
3WL	-7.7	-7.6	-7.7	-7.66	-7.3	-7.3	-7.3	-7.3	-7.1	-7.1	-7.1	-7.1
BR1	-6.3	-6.9	-6.8	-6.66	-6.8	-7.0	-6.9	-6.9	-6.8	-6.8	-6.9	-6.83
BR2	-7.1	-7.0	-7.1	-7.06	-6.7	-6.7	-6.7	-6.7	-6.7	-6.7	-6.7	-6.7
BR3	-7.6	-7.6	-7.6	-7.6	-7.3	-7.3	-7.3	-7.3	-4.6	-4.6	-4.6	-4.6
BR4	-7.7	-7.5	-7.5	-7.56	-7.0	-7.0	-7.0	-7	-4.0	-4.0	-4.1	-4.03
BR5	-5.1	-5.1	-5.1	-5.1	-5.3	-5.3	-5.3	-5.3	-5.3	-5.3	-5.3	-5.3
BR6	-8.0	-8.0	-8.0	-8.0	-7.7	-7.7	-7.7	-7.7	-8.0	-8.0	-8.0	-8.0
BR7	-6.0	-6.1	-6.1	-6.06	-6.3	-6.3	-6.4	-6.3	-6.0	-6.0	-6.1	-6.03
BR8	-6.2	-6.3	-6.1	-6.2	-6.5	-6.4	-6.5	-6.46	-6.3	-6.3	-6.3	-6.3
BR9	-5.7	-5.8	-5.8	-5.76	-6.4	-6.3	-6.4	-6.36	-6.3	-6.3	-6.3	-6.3
BR10	-6.7	-6.7	-6.7	-6.7	-5.4	-5.4	-5.4	-5.4	-5.4	-5.4	-5.3	-5.36
BR11	-5.6	-5.6	-5.6	-5.6	-6.2	-6.3	-6.1	-6.2	-6.1	-6.1	-6.0	-6.06
BR12	-4.9	-4.9	-4.9	-4.9	-4.9	-4.9	-4.9	-4.9	-5.3	-5.3	-5.3	-5.3
BR13	-6.8	-6.8	-6.8	-6.8	-7.7	-7.7	-7.7	-7.7	-7.9	-7.9	-7.8	-7.86
BR14	-5.1	-5.1	-5.1	-5.1	-5.4	-5.4	-5.4	-5.4	-5.3	-5.3	-5.3	-5.3

Comp compounds, 1st 1st simulation, 2nd 2nd simulation, 3rd 3rd simulation

is important to study different properties and parameters of a compound to make a definitive conclusion. The physico-chemical properties, pharmacokinetics, drug-likeness, and medicinal chemistry of 3WL, BR6, and BR13 are given in Table 4. The graphical representation of the oral bioavailability diagram is also given in Fig. 3.

From Table 4, it can be observed that 3WL, BR6 and BR13 abide by the Lipinski's rule of 5 (molecular weight  $\leq 500$ ,  $cLogP \leq 5$ , hydrogen acceptor  $\leq 10$ , hydrogen donor  $\leq 5$ ), Ghose filter (molecular weight  $< 480$ , molar refractivity  $< 130$ ,  $WLOGP < -0.4$ , number of atoms  $< 70$ ), and Muegge filters (molecular weight  $< 600$ , topological polar surface area  $< 150$ , hydrogen acceptor  $< 10$ , hydrogen donor  $< 5$ ) [58–60]. Also, 3WL and BR13 abide by the Veber filter (rotatable bonds  $\leq 10$ , topological polar surface area  $\leq 140$ ) and the Egan filter ( $WLOGP \leq 5.88$ , topological polar surface area  $\leq 131.6$ ) [61, 62]. However, the topological polar surface area of ellagic acid (BR6) is  $141.34 \text{ \AA}^2$  and therefore it does not completely abide by the Veber and Egan filters. In a drug discovery process, a compound with a lower topological polar surface area is preferred over those having a higher value [63].

Drugs that are intended to be administered orally must be soluble in water. This is because low water solubility can lead to gastrointestinal toxicity. Low water solubility can also lead to low bioavailability [64]. BR6 and BR13 are more soluble in water when compared to 3WL (Table 4). The gastrointestinal absorption of BR6 and BR13 is also predicted to be high (Table 4). Also, BR13 was predicted to be able to permeate through the blood–brain barrier (Table 4). Bioavailability issues can also slow down the pace of a drug discovery process [61, 62]. BR6 and BR13 had a bioavailability score of 0.55 which is also the same as the bioavailability score of 3WL (0.55) (Table 4). 3WL and BR6 showed lead likeness potential while BR13 did not show a lead likeness potential (Table 4). This is because the molecular weight of BR13 is  $354.35 \text{ g/mol}$  which is higher than the acceptable  $350 \text{ g/mol}$ . 3WL, BR6, and BR13 have a synthetic accessibility score of 3.02, 3.17, and 4.12 (Table 4). This means that 3WL will be able to be synthesized easily followed by BR6 while BR13 will be the hardest to be synthesized among the three phytochemicals [65].

According to generated data, 3WL, BR6, and BR13 will be absorbed in the gastrointestinal tract. However, only

**Table 2** Toxicity parameters and drug score of 3WL, BR6, and BR13

Compound	Mutagenic	Tumorigenic	Reproductive effective	Irritant	Drug score
3WL	None	None	None	None	0.6981475
BR6	None	None	None	None	0.5104488
BR13	None	None	None	None	0.4094924

**Table 3** Toxicity parameters, LD<sub>50</sub>, and toxicity class of 3WL, BR6, and BR13

Compound	Hepatotoxicity	Carcinogenicity	Immunotoxicity	Mutagenicity	Cytotoxicity	LD <sub>50</sub> (mg/kg body weight)	Toxicity class
3WL	Inactive	Active	Inactive	Active	Inactive	3919	5
BR6	Inactive	Active	Inactive	Inactive	Inactive	2991	4
BR13	Inactive	Active	Active	Inactive	Inactive	1500	3

**Table 4** Different properties of 3WL, BR6, and BR13

Properties	3WL	BR6	BR13
<b>Physicochemical properties</b>			
Molecular weight	270.24 g/mol	302.19 g/mol	354.35 g/mol
Heavy atoms	20	22	26
Aromatic heavy atoms	16	16	12
Rotatable bonds	1	0	2
H-bond acceptors	5	8	6
H-bond donors	3	4	0
Molar refractivity	73.99	75.31	90.00
TPSA	90.90 Å <sup>2</sup>	141.34 Å <sup>2</sup>	55.38 Å <sup>2</sup>
<b>Lipophilicity</b>			
iLOGP	2.43	0.79	3.46
XLOGP3	3.16	1.10	2.68
WLOGP	2.58	1.31	2.57
MLOGP	0.52	0.14	1.98
SILICOS-IT	2.52	1.67	3.25
<b>Water solubility</b>			
ESOL	Moderately soluble	Soluble	Soluble
Ali	Moderately soluble	Soluble	Soluble
SILICOS-IT	Moderately soluble	Soluble	Moderately soluble
<b>Pharmacokinetics</b>			
GI absorption	High	High	High
BBB permeant	No	No	Yes
P-gp substrate	No	No	No
CYP1A2 inhibitor	Yes	Yes	No
CYP2C19 inhibitor	No	No	Yes
CYP2C9 inhibitor	No	No	No
CYP2D6 inhibitor	Yes	No	Yes
CYP3A4 inhibitor	Yes	No	Yes
Skin permeation	−5.70 cm/s	−7.36 cm/s	−6.56 cm/s
<b>Drug likeness</b>			
Lipinski	Yes	Yes	Yes
Ghose	Yes	Yes	Yes
Veber	Yes	No	Yes
Egan	Yes	No	Yes
Muegge	Yes	Yes	Yes
Bioavailability score	0.55	0.55	0.55
<b>Medicinal chemistry</b>			
PAINS	1 alert	1 alert	0 alert
Brenk	1 alert	3 alerts	0 alert
Leadlikeness	Yes	Yes	No
Synthetic accessibility	3.02	3.17	4.12



BR13 was computed to be able to penetrate the blood–brain barrier. 3WL might interact with CYP1A2, CYP2D6, and CYP3A4. BR6 might interact with CYP1A2. BR13 might interact with CYP2C19, CYP2D6, and CYP3A4. Cytochromes are enzymes produced in the liver that can determine the pharmacokinetic profile of a drug [66]. Overall, 3WL, BR6, and BR13 have an acceptable pharmacokinetic profile (Table 4). From Fig. 3, it can be observed that all the properties of BR13 lie within the physicochemical space that is suitable for oral bioavailability. The majority of the properties of BR6 also lie within the suitable physicochemical space for oral bioavailability. In comparison to 3WL, BR6 and BR13 are also computed to be orally bioavailable. From the evidence available at hand, it can be concluded that BR6 and BR13 possessed suitable properties which qualifies them to be subjected for further studies.

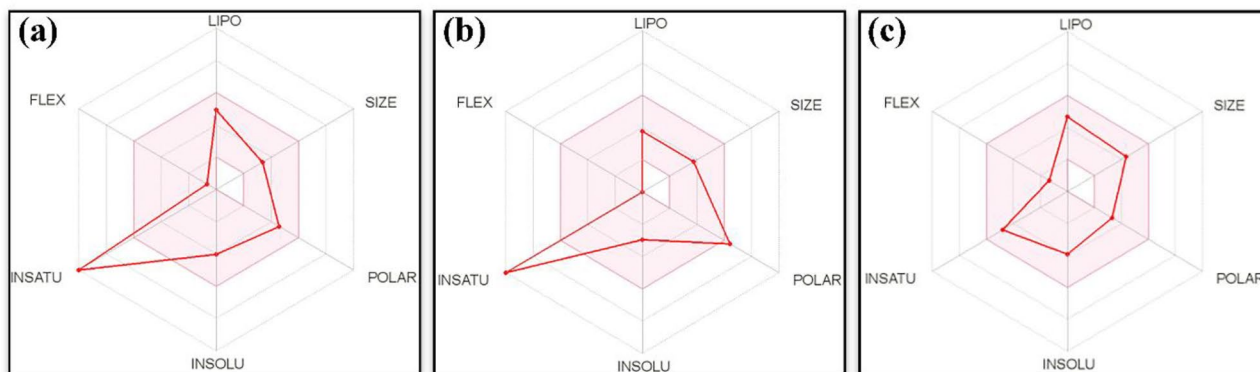
### Visualization and analysis of ligand interactions

In the active binding site of SARS-CoV-2 3CL pro, important catalytic residues of the viral protein such as HIS41 and CYS145 are present. This catalytic dyad is directly involved in the chemical catalysis of the protein. Other amino acid residues such as SER46, LEU141, ASN142, GLU166, PRO168, GLN189, THR190, and ALA191 are also found at the catalytic site cavity [67]. For our phytochemicals to be regarded as a potential inhibitor of SARS-CoV-2 3CL pro, they must bind to the active catalytic site and interact (preferably form a hydrogen bond) with the active site residues. BR6 and BR13 cannot be considered potential inhibitors of SARS-CoV-2 3CL pro simply because they had a high binding affinity for the target protein. It is important that BR6 and BR13 must bind with catalytic dyad and other crucial amino acids present at the active site. Therefore, the ligand interactions of 3WL, BR6, and BR13 were studied with Discovery Studio Visualizer v20.1.0.19295 software, and the output is represented in Fig. 4.

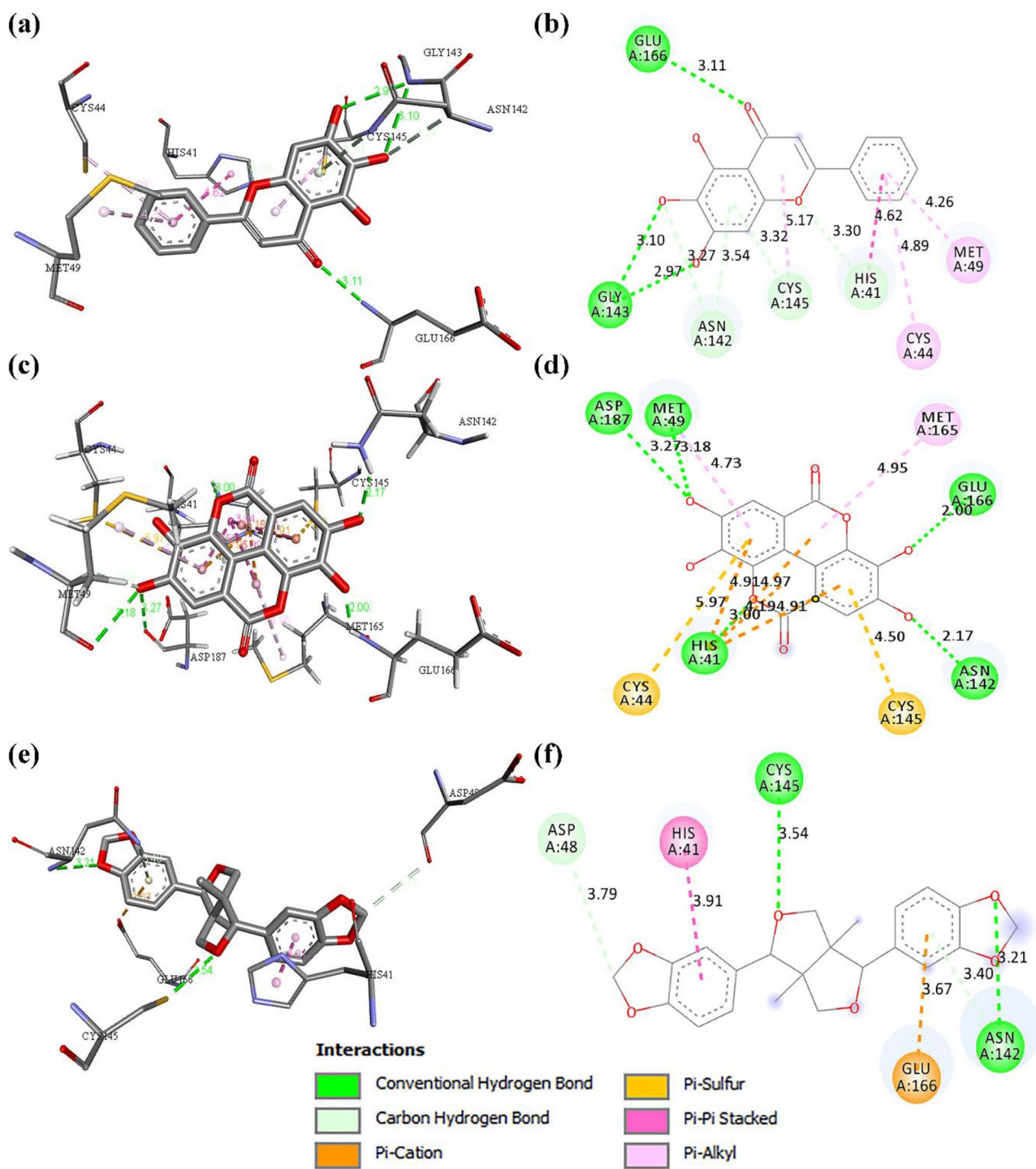
As 3WL is the standard drug, the ligand interactions of the original binding pose of 3WL was also studied (Fig. 4a, b). 3WL formed conventional hydrogen bond with GLU166 (bond length = 3.11 Å); two conventional hydrogen bonds with GLY143 (bond length = 2.97 and 3.10 Å); pi-alkyl interaction with CYS44 (bond length = 4.89 Å), MET49 (bond length = 4.26 Å), CYS145 (bond length = 5.17 Å); pi-pi stacked with HIS41 (bond length = 4.62 Å); carbon-hydrogen bonds with HIS41 (bond length = 3.30 Å), ASN142 (bond length = 3.27 Å and 3.54 Å), and CYS145 (bond length = 3.32 Å). 3WL interacted with 4 amino acid residues (HIS41, CYS145, ASN142, and GLU166) present at the catalytic cavity of SARS-CoV-2 3CL pro.

Ellagic acid (BR6) also showed good ligand interactions with SARS-CoV-2 3CL pro (Fig. 4c, d). BR6 formed conventional hydrogen bonds with HIS41 (bond length = 3.00 Å), MET49 (bond length = 3.18 Å), ASN 142 (bond length = 2.17 Å), GLU166 (bond length = 2.00 Å), ASP187 (bond length = 3.27 Å); pi-alkyl interaction with MET49 (bond length = 4.73 Å), MET165 (bond length = 4.95 Å); pi-cation interactions with HIS41 (bond length = 4.19 Å, 4.91 Å, 4.97 Å, 4.91 Å); pi-sulfur interactions with CYS44 (bond length = 5.97 Å), and CYS145 (bond length = 4.50 Å). BR6 interacted with 4 amino acid residues (HIS41, CYS145, ASN142, and GLU166) present at the catalytic cavity of SARS-CoV-2 3CL pro.

The ligand interactions between (+)-sesamin (BR13) and SARS-CoV-2 3CL pro is given in Fig. 4e, f. BR13 formed conventional hydrogen bond with CYS145 (bond length = 3.54 Å), ASN142 (bond length = 3.40 Å); pi-pi stacked with HIS41 (bond length = 3.91 Å); pi-cation interaction with GLU166 (bond length = 3.67 Å); carbon-hydrogen bond with ASP48 (bond length = 3.79 Å). BR13 also interacted with 4 amino acid residues (HIS41, CYS145, ASN142, and GLU166) present at the catalytic cavity of SARS-CoV-2 3CL pro.



**Fig. 3** Graphical representation of the oral bioavailability of **a** 3WL, **b** ellagic acid, and **c** (+)-sesamin



**Fig. 4** 2D and 3D ligand interactions of **a, b** 3WL, **c, d** ellagic acid, and **e, f** (+)-sesamin

3WL, BR6, and BR13 interacted with a total of 7, 8, and 5 amino acid residues, respectively. A total of 3, 5, and 2 conventional hydrogen bonds were formed by 3WL, BR6, and BR13, respectively. BR6 was able to form a conventional hydrogen bond with HIS41 while BR13 formed a conventional hydrogen bond with CYS145 of the catalytic

dyad of SARS-CoV-2 3CL pro. BR6 also formed a conventional hydrogen bond with two crucial amino acid residues (ASN142 and GLU166) that are found at the catalytic cavity of SARS-CoV-2 3CL pro. BR13 also formed a conventional hydrogen bond with ASN142 that is present at the catalytic cavity of SARS-CoV-2 3CL pro. BR6 and BR13 interacted

with the active site residues (HIS41 and CYS145) of SARS-CoV-2 3CL pro that are directly involved in the catalysis of the protein.

Analysis of the ligand interactions showed that the antiviral phytochemicals of *B. retusa* (ellagic acid and (+)-sesamin) were able to form significant interactions with crucial active site amino acid residues present at the catalytic cavity of SARS-CoV-2 3CL pro. Also, visualization of the 3-dimensional binding pose of BR6 and BR13 (Fig. 5) revealed that they bind to the same position with a similar pose as that of 3WL. Hydrogen bonds are important interactions that occur between a protein and a ligand. Hydrogen bonds are present almost everywhere and they play an important role in catalysis, protein folding, and protein–ligand interactions [68]. Analysis of the ligand interactions of BR6 and BR13 revealed that they formed hydrogen bonds with at least one of the HIS41-CYS145 catalytic dyad in addition to forming hydrogen bonds with other amino acids present at the catalytic cavity. This justifies that BR6 and BR13 have the potential to inhibit SARS-CoV-2 3CL pro by interacting with the important amino acids. After the careful examination of protein–ligand interactions, BR6 and BR13 were subjected for further studies.

### Molecular dynamics simulation studies

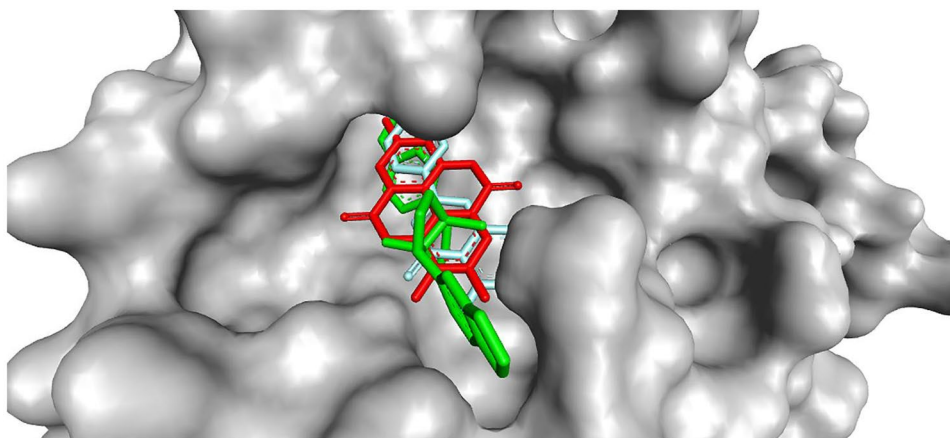
Molecular docking results are validated with MD simulation studies [69]. The conformational stability of the protein–ligand complexes is also evaluated with MD simulation studies [24]. Some images of the MD simulation snapshot over time are given in Fig. 6. The numbers 0, 20, 40, 60, 80, and 100 highlighted in yellow in Fig. 6 represents the time period (nanoseconds) of the MD simulation. The MD simulation snapshot of 3WL is given in Fig. 6a–f. The MD simulation snapshot of ellagic acid (BR6) is given in Fig. 6g–l. The MD simulation snapshot of (+)-sesamin (BR13) is given in Fig. 6m–r. From the MD simulation snapshots, it

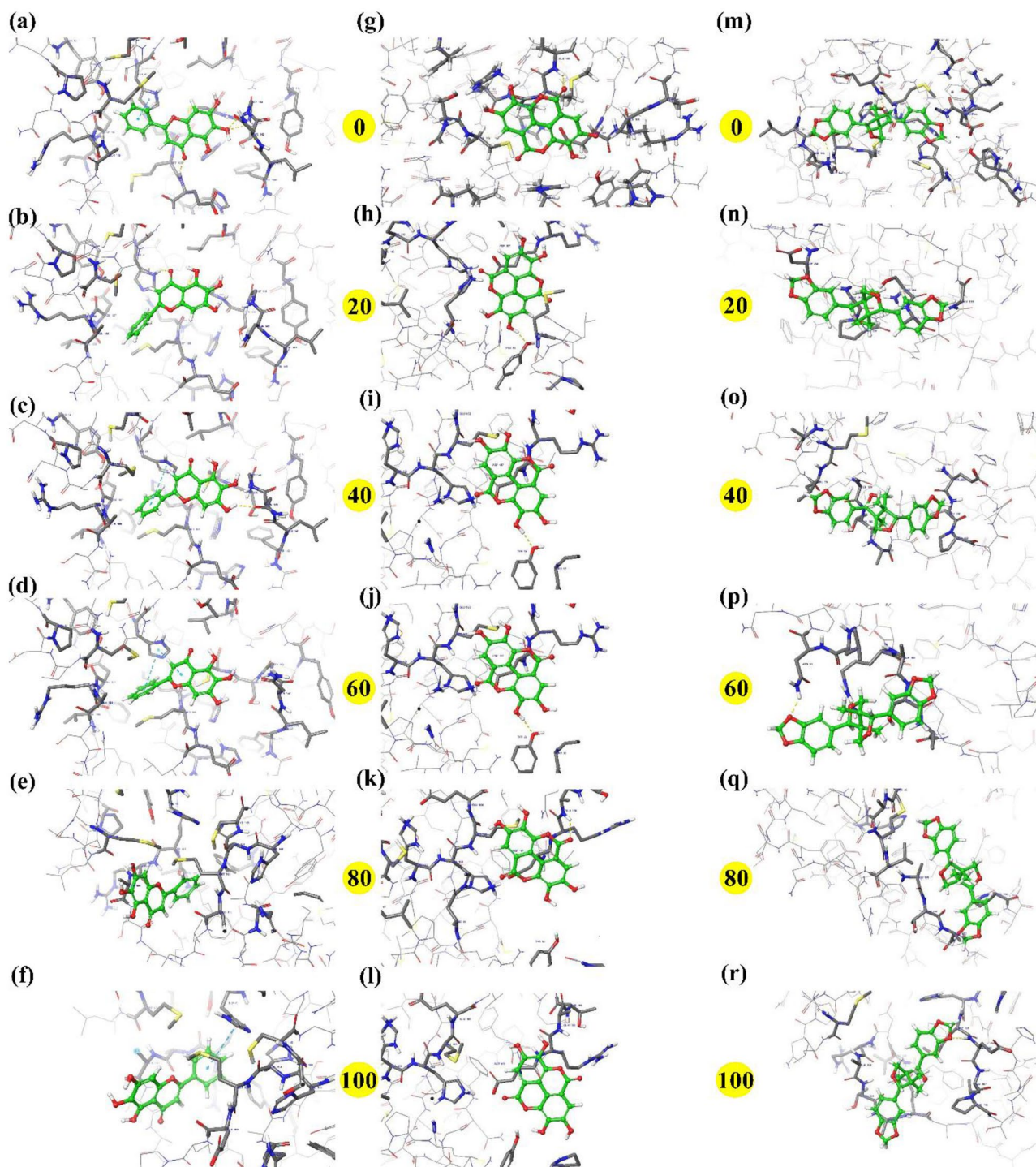
can be observed that during the entire course of 100 ns simulations, each ligands interacts with different active amino acid residues of the target protein through different binding poses. In this way, the ellagic acid and (+)-sesamin were able to interact with the catalytic dyad (HIS41-CYS145) and other important amino acids of SARS-CoV-2 3CL pro.

In the present study, MD simulations were carried out at a timescale of 100 ns and the generated RMSD and RMSF plots are given in Fig. 7. The RMSD of the protein-3WL, protein-ellagic acid, and protein-(+)-sesamin complexes are given in Fig. 7a, c, e respectively. In Fig. 7a, c, e, the protein RMSD trajectory scale is given on the left side while the ligand RMSD trajectory scale is given on the right side of each Y-axis plot. For small globular proteins, RMSD fluctuations between the range of 1 to 3 Å are acceptable [70]. The ligand RMSD determines the stability of the original docked ligand at the active binding site of the protein. A ligand RMSD value slightly higher than that of the protein RMSD is acceptable. If the RMSD of a ligand fluctuates significantly, it means that the ligand is trying to find a new pose as the simulation progress with time [67, 68]. The protein RMSF can be used to evaluate local changes that took place for each amino acid during the simulations as it interacts with a ligand. In the RMSF plot, the peak of each residue determines the degree of changes occurring for each amino acid. The lower the RMSF value, the more stable are the amino acids [24]. Loops or terminal residues often show high RMSF values, and  $\alpha$ -helices and  $\beta$ -sheets are associated with low RMSF values [71]. A low RMSF value of a protein indicates the structural stability of the interacting amino acids during the MD simulation. The result suggests that the protein–ligand complex showed equilibrium fluctuation [71].

In Fig. 7a, the fluctuations of the protein RMSD trajectory in the protein-3WL complex ranged between 1.2 and 2.8 Å. Large deviations were observed for the protein RMSD between 0 and 10 ns. The protein RMSD remained stable between 10 and 50 ns followed by a sharp drift. After

**Fig. 5** Binding pose of 3WL (sky blue), ellagic acid (red), and (+)-sesamin (green) at the active binding pocket of SARS-CoV-2 3CL pro

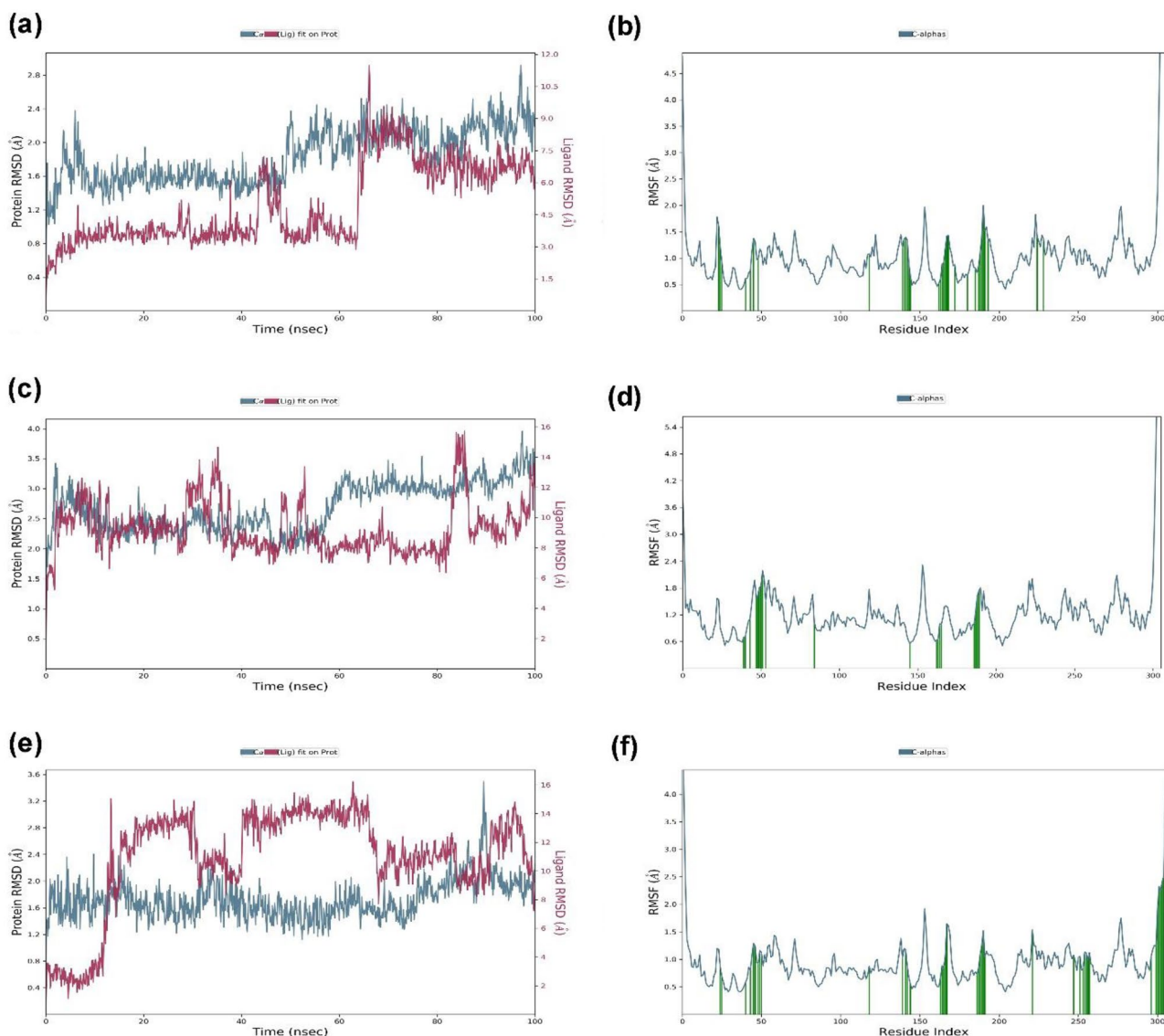




**Fig. 6** MD simulation snapshots of 3WL (a, b, c, d, e, f), ellagic acid (g, h, i, j, k, l), and (+)-sesamin (m, n, o, p, q, r) at various time points (0 ns, 20 ns, 40 ns, 60 ns, 80 ns, and 100 ns)

50 ns, the protein RMSD reached equilibrium and stabilizes around an RMSD value of 2.0 Å. The ligand RMSD of the protein-3WL complex remained stable at around 4.0 Å for up to 60 ns followed by a sharp rise in the RMSD value. After 70 ns, the ligand RMSD trajectory gradually drops,

attained equilibrium, and stabilizes around an RMSD value of 7.0 Å till the end of the simulation. The stable trajectory of the protein–ligand complex after 60 ns indicates that the complex takes more time to attain equilibrium. The stable trajectory of the complex after 60 ns suggests that 3WL is



**Fig. 7** RMSD (Y-axis, left side: protein RMSD, right side: ligand RMSD) of **a** protein-3WL complex, **c** protein-ellagic acid complex, **e** protein-(+)-sesamin complex; RMSF of **b** protein-3WL complex, **d** protein-ellagic acid complex, **f** protein-(+)-sesamin complex

stabilized and spatially occupied at the binding pocket of SARS-CoV-2 3CL pro [71–73]. In Fig. 7b, the protein residues that interacted with 3WL are marked in green color. The RMSF of all the amino acid residues that interacted with 3WL remained below 2.0 Å. The protein RMSD, the ligand RMSD, and the protein RMSF data suggest that 3WL formed a stable complex with SARS-CoV-2 3CL pro during the 100 ns MD simulations. For the protein-3WL complex, the 100 ns MD simulations is sufficient for further analysis.

In Fig. 7c where the protein-ellagic acid complex is depicted, the protein RMSD trajectory initially fluctuated between 2.0 and 3.5 Å (0–10 ns). Between 10 and 50 ns, the RMSD trajectory remained stable. The RMSD gradually rises between 50 and 60 ns. After 60 ns, the protein

attained equilibrium, showed stability in conformation and the protein RMSD stabilizes around 3.0 Å until around 90 ns. A small fluctuation is observed in the protein RMSD trajectory just before the end of the 100 ns simulation. In the protein-ellagic acid complex, the ligand RMSD showed sharp fluctuations at around 10 ns (6 to 12 Å), 30 ns (8 to 14 Å), 50 ns (8 to 14 Å), and 85 ns (8 to 16 Å). The majority of the ligand RMSD fluctuates between the range of 8 to 11 Å. After 85 ns, the ligand RMSD trajectory slowly rises until the end of the MD simulation. The ligand RMSD value (Å) is slightly higher than the protein RMSD value (Å) and can be considered to lie within the acceptable range [67, 68]. In Fig. 7d, the RMSF value of all the amino acid residues of SARS-CoV-2 3CL pro that interacted with ellagic

acid remained below 2.0 Å. A low RMSF value of a protein indicates the structural stability of the interacting amino acids during the MD simulation. The result suggests that the protein–ligand complex showed equilibrium fluctuation [71]. The protein RMSD, the ligand RMSD, and the protein RMSF data suggest that ellagic acid tends to form a stable complex with SARS-CoV-2 3CL pro during the 100 ns MD simulations. For the protein–ellagic acid complex, the 100 ns MD simulations is sufficient for further analysis.

In Fig. 7e, the majority of the RMSD value in the trajectory of the protein-(+)-sesamin complex stabilized between 1.2 and 2.0 Å. Small fluctuations in the protein RMSD trajectory were observed at around 15 ns, 35 ns, and 60 ns. The protein RMSD trajectory gradually rises after 70 ns. At around 90 ns, the protein RMSD fluctuated sharply and then the protein RMSD trajectory gradually decreases to stabilize between 1.6 and 2.0 Å till the end of the simulation. For the majority of the 100 ns MD simulation, the ligand-RMSD of the protein-(+)-sesamin complex stabilizes between 10 and 14 Å. The ligand RMSD trajectory fluctuated sharply at around 10 ns, gradually increases from 20 up to 30 ns, dropped sharply at 30 ns, increased sharply at 40 ns, stabilizes from 40 up to 65 ns, dropped sharply at around 65 ns, stabilizes from 65 up to 85 ns, dipped slightly at 85 ns, stabilized up to 90 ns, increased sharply at 90 ns, gradually increased till 95 ns, and then the ligand RMSD began to drop again till the end of the MD simulations. The ligand RMSD value (Å) is slightly higher than the protein RMSD value (Å) and can be considered to lie within the acceptable range [67, 68]. In Fig. 7f, the majority of the RMSF of the amino acids that interacted with (+)-sesamin remained below 2.0 Å. Few amino acid residues have an RMSF value above 2.0 Å. A low RMSF value of a protein indicates the structural stability of the interacting amino acids during the MD simulation. The result suggests that the protein–ligand complex showed equilibrium fluctuation [71]. The protein RMSD, ligand RMSD, and the protein RMSF data suggest that (+)-sesamin formed a stable complex with SARS-CoV-2 3CL pro during the 100 ns MD simulations. For the protein-(+)-sesamin complex, the 100 ns MD simulations is sufficient for further analysis.

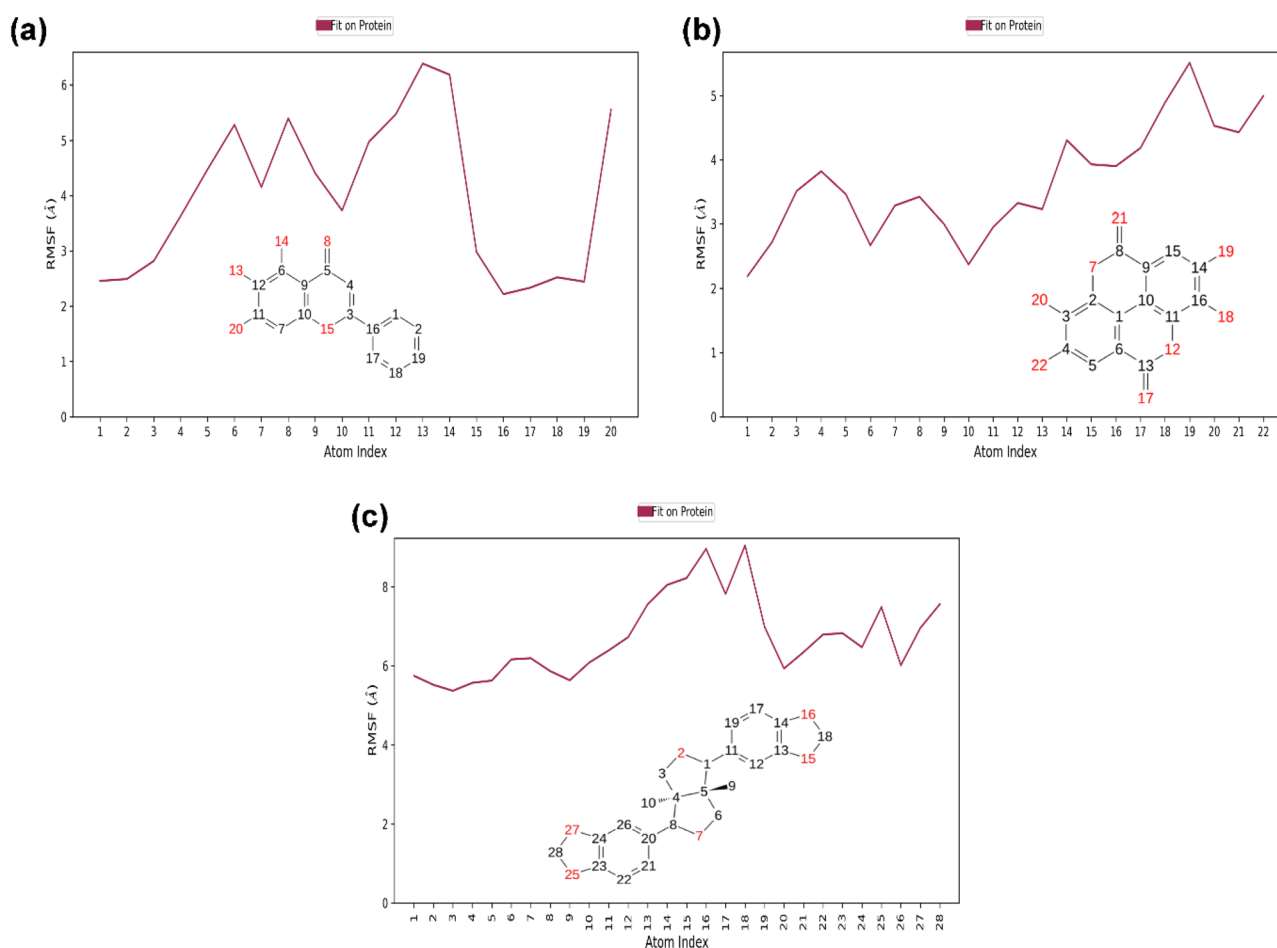
Protein secondary structure elements (SSE) such as  $\alpha$ -helices and  $\beta$ -strands are monitored throughout the 100 ns MD simulations. SSE plays an important role in the overall protein folding and protein structure [74]. The MD simulations analyze different protein SSE data such as the %SSE distribution by residue index throughout the protein structure, SSE composition for each trajectory frame throughout the simulation, monitoring of each residue, and its SSE assignment over time for each protein–ligand complex. The  $\alpha$ -helices and  $\beta$ -strands are represented in red and blue color, respectively. During the MD simulation, the protein that is complexed with 3WL highlighted in Fig. s1a

has a total of 43.37% protein SSE (19.61%  $\alpha$ -helices and 23.76%  $\beta$ -strands). The protein complexed with ellagic acid highlighted in Fig. s1b has a total of 41.41% protein SSE (19.32%  $\alpha$ -helices and 22.09%  $\beta$ -strands). The protein in complex with (+)-sesamin highlighted in Fig. s1c has a total of 43.58% protein SSE (19.47%  $\alpha$ -helices and 24.11%  $\beta$ -strands).

The ligand-RMSF depicted in Fig. 8a–c showed atom-specific fluctuations. The ligand RMSF may give insights into how ligand fragments interact with the protein and their entropic role in the binding event. In the bottom panel, the “fit ligand on protein” line shows the ligand fluctuations, with respect to the protein. The protein–ligand complex is first aligned on the protein backbone, and then the ligand RMSF is measured on the ligand heavy atoms. The RMSF of each atom of the CCI “3WL” complexed with the target protein ranged between 2 and 6 Å (Fig. 8a). The RMSF of each atom of the ellagic acid complexed with the target protein ranged between 2 and 5 Å (Fig. 8b). The RMSF of each atom of the (+)-sesamin complexed with the target protein ranged between 5 and 9 Å (Fig. 8c). Ellagic acid has the lowest RMSF value followed by 3WL and (+)-sesamin.

Protein interactions with the ligand can be monitored throughout the simulation. These interactions can be categorized by type and summarized, as shown in Fig. 9a, c, e. Protein–ligand interactions (or “contacts”) are categorized into four types: hydrogen bonds (green), hydrophobic (purple), ionic (pink), and water bridges (blue). Each interaction type contains more specific subtypes, which can be explored through the “simulation interactions diagram” panel given in Fig. 9a, c, e. The stacked bar charts are normalized throughout the trajectory: for example, a value of 0.7 suggests that 70% of the simulation time the specific interaction is maintained. Values over 1.0 are possible as some protein residue may make multiple contacts of the same subtype with the ligand. In Fig. 9b, d, f, a timeline representation of the interactions and contacts (H-bonds, hydrophobic, ionic, water bridges) are summarized. The total number of specific contacts the protein makes with the ligand throughout the trajectory is given in the top panel (Fig. 9b, d, f). The bottom panel showed which residues interact with the ligand in each trajectory frame (Fig. 9b, d, f). Some residues make more than one specific contact with the ligand, which is represented by a darker shade of orange, according to the scale to the right of the plot.

MDSS revealed that 3WL interacted with HIS41, CYS44, MET49, ASN142, GLY143, CYS145, and GLU166 (Fig. 4a, b). Simulation interactions diagram and timeline representation of interactions generated from MD simulations validate the MDSS by confirming that 3WL interacted with HIS41, CYS44, MET49, ASN142, GLY143, CYS145, and GLU166 at certain time points (0–100 ns) during the 100 ns



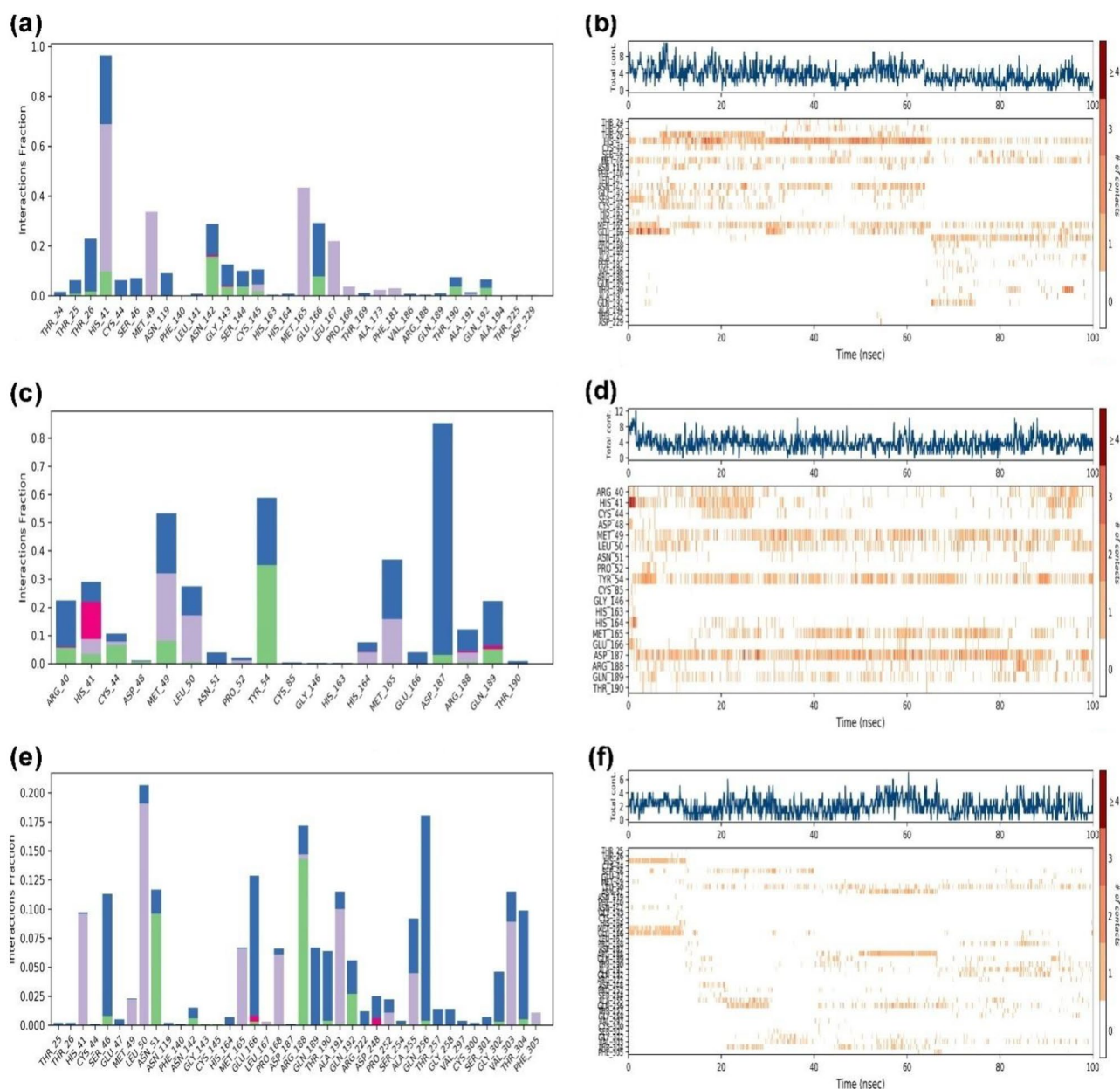
**Fig. 8** Ligand-RMSF plot of **a** 3WL, **b** ellagic acid, and **c** (+)-sesamin

simulations (Fig. 9a, b). MDSS revealed that ellagic acid interacted with HIS41, CYS44, MET49, ASN142, CYS145, MET165, GLU166, and ASP187 (Fig. 4c, d). MD simulations validate the docking study by confirming that ellagic acid interacted with HIS41, CYS44, MET49, ASN142, CYS145, MET165, GLU166, and ASP187 at certain time points (0–100 ns) during the 100 ns MD simulations (Fig. 9c, d). MDSS revealed that (+)-sesamin interacted with HIS41, ASP48, ASN142, CYS145, and GLU166 (Fig. 4e, f). MD simulations validate the docking study by confirming that (+)-sesamin interacted with HIS41, ASP48, ASN142, CYS145, and GLU166 at certain time points (0–100 ns) during the 100 ns MD simulations (Fig. 9e, f).

The ligand torsions plot summarizes the conformational evolution of every rotatable bond (RB) in the ligand throughout the simulation trajectory (0.00 through 100.00 ns). The top panel shows the 2D schematic of a ligand with color-coded rotatable bonds. Each rotatable bond torsion is accompanied by a dial plot and bar plots of the same color (Fig. 10a–e). Dial (or radial) plots describe the

conformation of the torsion throughout the simulation. The beginning of the simulation is in the center of the radial plot and the time evolution is plotted radially outwards. The bar plots summarize the data on the dial plots, by showing the probability density of the torsion. The data provided on the dial plots and the bar plots of 3WL (Fig. 10a), ellagic acid (Fig. 10c), and (+)-sesamin (Fig. 10e) revealed that these ligands underwent a conformational strain to maintain the protein-bound conformation.

Different ligand properties such as the RMSD, radius of gyration (rGyr), intramolecular hydrogen bonds, molecular surface area (MolSA), solvent accessibility surface area (SASA), and polar surface area are given in Fig. 10b, d, f. The RMSD of a ligand is depicted with respect to the reference conformation wherein the first frame is used as the reference and it is considered time  $t=0$ . An RMSD trajectory with minimum fluctuations indicates greater stability [75]. rGyr measures the “extendedness” of a ligand and is equivalent to its principal moment of inertia [76]. A small fluctuation in the rGyr trajectory is acceptable and denotes



**Fig. 9** Simulation interactions diagram (hydrogen bonds (green), hydrophobic (purple), ionic (pink), and water bridges (blue)) between the protein and **a** 3WL, **c** ellagic acid, and **e** (+)-sesamin; timeline

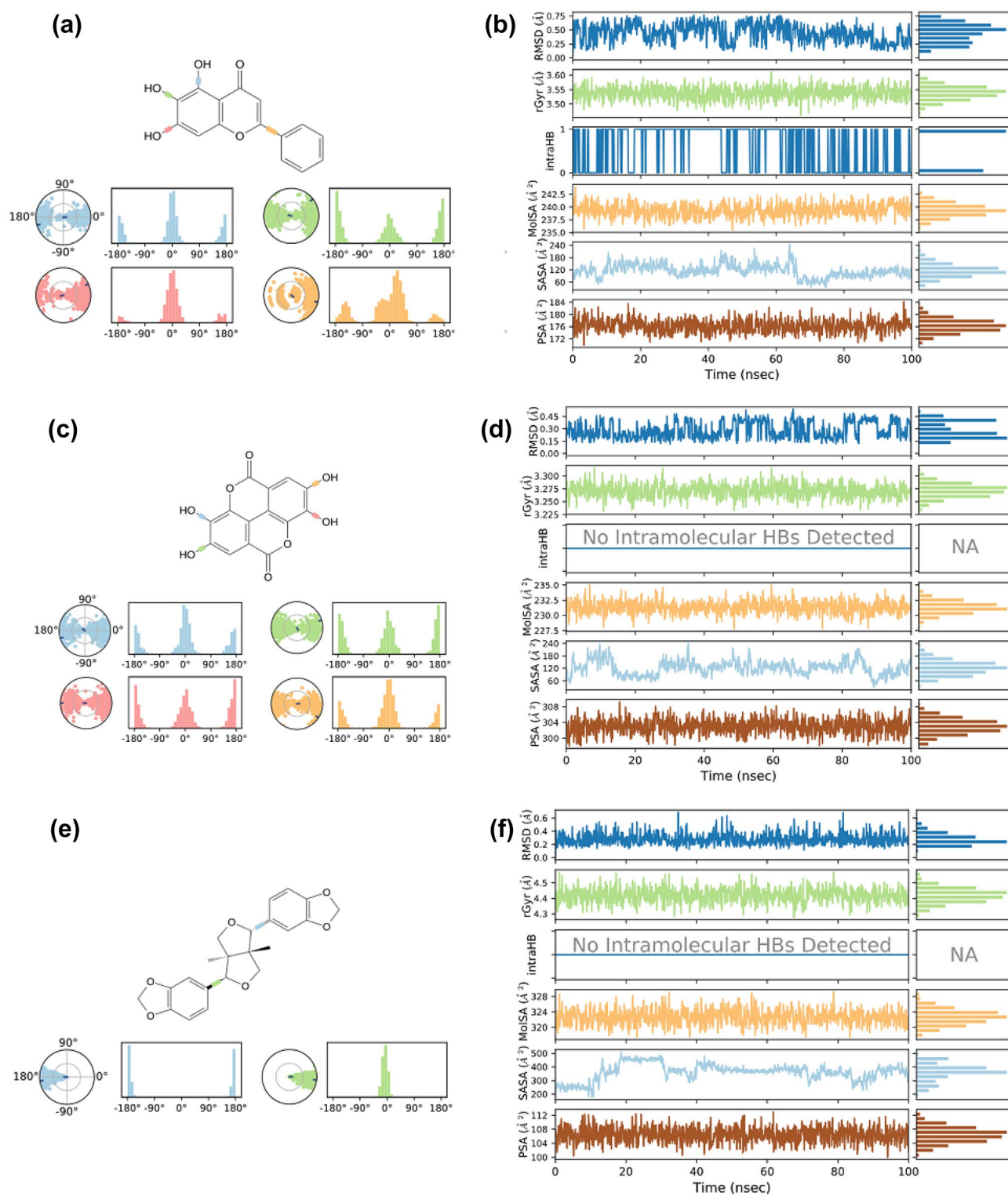
representation of interactions and contacts between the protein and **b** 3WL, **d** ellagic acid, and **f** (+)-sesamin

the stability of the ligand [75]. Intramolecular hydrogen bonds refer to the number of internal hydrogen bonds within a ligand molecule. MolSA is calculated within a 1.4 Å probe radius and the calculated value is equivalent to a van der Waals surface area [75]. SASA refers to the surface area of a molecule accessible by a water molecule [75]. PSA refers to the solvent-accessible surface area in a molecule contributed only by oxygen and nitrogen atoms [77].

The ligand properties of 3WL bound to the target protein during the 100 ns MD simulations are given in Fig. 10b.

The RMSD of 3WL ranged between 0.25 and 0.75 Å, and the equilibrium was observed at around 0.50 Å. The rGyr of 3WL ranged between 3.50 and 3.60 Å, and the equilibrium was observed at around 3.55 Å. Around one intramolecular hydrogen bond was found to be present in the ligand molecule. The RMSD and rGyr trajectory of 3WL remained stable during the entire simulation period. The MolSA of 3WL ranged between 237.5 and 242.5 Å<sup>2</sup>, and the equilibrium was observed at around 240.0 Å<sup>2</sup>. The SASA of 3WL ranged between 60 and 240 Å<sup>2</sup>, and the equilibrium





**Fig. 10** Ligand torsion plot (a, c, e) and ligand properties (b, d, f) of **a, b** 3WL; **c, d** ellagic acid; **e, f** (+)-sesamin

was observed at around  $120 \text{ \AA}^2$ . The PSA of 3WL ranged between  $172$  and  $184 \text{ \AA}^2$ , and the equilibrium was observed at around  $176 \text{ \AA}^2$ .

The ligand properties of ellagic acid bound to the target protein during the 100 ns MD simulations are given in Fig. 10d. The RMSD of ellagic acid ranged between 0.15

and 0.45 Å, and the equilibrium was observed at around 0.15 Å. The rGyr of ellagic acid ranged between 3.250 and 3.300 Å, and the equilibrium was observed at around 3.275 Å. The RMSD and rGyr trajectory of ellagic acid remained stable during the entire simulation period. No intramolecular hydrogen bond was found to be present in the ligand molecule. The MolSA of ellagic acid ranged between 227.5 and 235.0 Å<sup>2</sup>, and the equilibrium was observed at around 231.0 Å<sup>2</sup>. The SASA of ellagic acid ranged from 60 to 240 Å<sup>2</sup>, and the equilibrium was observed at around 120 Å<sup>2</sup>. The PSA of 3WL ranged between 300 and 308 Å<sup>2</sup>, and the equilibrium was observed at around 304 Å<sup>2</sup>.

The ligand properties of (+)-sesamin bound to the target protein during the 100 ns MD simulations is given in Fig. 10f. The RMSD of (+)-sesamin ranged between 0.2 and 0.6 Å, and the equilibrium was observed at around 0.2 Å. The rGyr of (+)-sesamin ranged between 4.3 and 4.5 Å, and the equilibrium was observed at around 4.4 Å. The RMSD and rGyr trajectory of (+)-sesamin remained stable during the entire simulation period. No intramolecular hydrogen bond was found to be present in the ligand molecule. The MolSA of (+)-sesamin ranged between 320 and 328 Å<sup>2</sup>, and the equilibrium was observed at around 324 Å<sup>2</sup>. The SASA of (+)-sesamin ranged between 200 to 500 Å<sup>2</sup>, and the equilibrium was observed at around 400 Å<sup>2</sup>. The PSA of (+)-sesamin ranged between 100 and 112 Å<sup>2</sup>, and the equilibrium was observed at around 108 Å<sup>2</sup>.

### MM-GBSA binding free energy calculations

The estimation of the binding free energies of the protein–ligand complexes was performed with MM-GBSA analysis. A low binding free energy value is preferred over a higher value. When compared to the scoring function given by molecular docking simulation studies, MM-GBSA is more accurate and reliable for binding energy calculations [78, 79]. The binding free energy for all the protein–ligand complexes during the 100 ns MD simulation was carried out to have a better estimate of the binding strength and potency of ellagic acid and (+)-sesamin against SARS-CoV-2 3CL pro. The MM-GBSA binding free energies of protein-3WL, protein-ellagic acid, and protein-(+)-sesamin complexes are given in Table 5. Among the protein–ligand complex, protein-3WL showed the lowest binding free energy value (−23.2356 kcal/mol) followed by protein-ellagic acid (−22.3064 kcal/mol) and protein-(+)-sesamin (−19.1274 kcal/mol). Among the two phytochemicals, ellagic acid has a better binding energy value than (+)-sesamin. The MM-GBSA binding free energy calculations reveal that BR6 and BR13 will bind to the active binding pocket of SARS-CoV-2 3CL pro with a significantly low energy involvement.

**Table 5** MM-GBSA binding free energies of each protein–ligand complexes

Protein–ligand complex	ΔG (kcal/mol)
SARS-CoV-2 3CLpro-3WL complex	−23.2356
SARS-CoV-2 3CLpro-ellagic acid complex	−22.3064
SARS-CoV-2 3CLpro-(+)-sesamin complex	−19.1274

### Bioactivity prediction using Bayesian statistical model

Proteases constitute an important target for drugs that are intended for antiviral therapy. Proteases are generally associated with viral diseases. When a viral protease is inhibited, the replication of the viral genome is hindered and this prevents transcription and replication of the viral RNA [80]. Since SARS-CoV-2 3CL pro is a protease, it will be interesting to study the general protease inhibitory potential of BR6 and BR13. Therefore, for the present study, the protease inhibitory potential of the phytochemicals was prioritized and studied with the sophisticated Bayesian statistical model. This statistical model can be accessed through the Molinspiration Chemoinformatics web tool (<https://www.molinspiration.com/>). The predicted bioactivity of the phytochemicals is given in Table 6.

The Bayesian statistical model generally provides an idea regarding the bioactivity of a molecule within the numerical range of −3 to +3 (<https://www.molinspiration.com/docu/miscreen/virtualsecreening.html>). A higher score indicates a higher activity and vice versa. 3WL has a bioactivity score of −0.35, BR6 has a bioactivity score of −0.18 and BR13 has a bioactivity score of −0.15. The predicted protease inhibitory potential of BR6 and BR13 is higher than that of 3WL (Table 6). From the numeral range provided by the statistical model, it can be observed that BR6 and BR13 might be active against viral protease enzymes. Interestingly, ellagic acid had been reported to exhibit antiviral activity against the Zika virus and human rhinoviruses [33, 81]. Although (+)-sesamin was not directly reported to show antiviral activity, it was demonstrated that (+)-sesamin could reduce cytokine production in influenza type A H1N1 infection [34].

In addition to the protease inhibition, the Molinspiration Chemoinformatics web tool also generates other relevant

**Table 6** Predicted protease inhibitory potential of 3WL, BR6, and BR13

Activity	Activity range	3WL	BR6	BR13
Protease inhibition	−3 to +3	−0.35	−0.18	−0.15

data such as GPCR ligand, ion channel modulator, a kinase inhibitor, nuclear receptor ligand, and enzyme inhibitor. As the present study deals with a protease (SARS-CoV-2 3CL pro), only the protease inhibitory potential of the compounds was highlighted. The other data can be found in Table s1.

## Conclusion

The computational studies on the phytochemicals of the traditional medicinal herb *B. retusa* are reported for the first time in this paper. Ellagic acid and (+)-sesamin has a high binding affinity for the catalytic cavity and they interact with catalytic amino acid residues of SARS-CoV-2 3CL pro. Based on RMSD, RMSF, protein SSE, ligand-RMSF, protein–ligand contacts, ligand torsions properties, and ligand properties generated for 3WL, ellagic acid, and (+)-sesamin, it was observed that all the ligands formed stable conformations with the active binding site of SARS-CoV-2 3CL Pro. The MM-GBSA binding free energy calculations showed that ellagic acid and (+)-sesamin bind to the active pockets of SARS-CoV-2 3CL pro with a significantly low energy involvement. The MM-GBSA values of ellagic acid and (+)-sesamin also complement the low binding energy (high binding affinity) values obtained from molecular docking simulation studies. Ellagic acid and (+)-sesamin have a high LD<sub>50</sub> value, a good safety profile, and good bioavailability. Ellagic acid and (+)-sesamin were also computed to have more activity against protease enzyme than the CCI (3WL) of SARS-CoV-2 3CL pro. Based on our computational findings, we conclude that the antiviral phytochemicals (ellagic acid and (+)-sesamin) of *B. retusa* has the potential to inhibit the chymotrypsin-like protease (3CL Pro) of SARS-CoV-2. Further studies such as in vitro assays are recommended for ellagic acid and (+)-sesamin to confirm their anti-SARS-CoV-2 activity.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1007/s11224-022-01959-3>.

**Acknowledgements** The corresponding author thanks Airlangga University and Universitas Tadulako, Indonesia, for providing computational facilities to carry out MD simulations.

**Author contribution** Study design: James H. Zothantluanga; computational work: Abd. Kakhar Umar, James H. Zothantluanga, S Keerthi Aswin, Saipul Maulana, and Muhammad Sulaiman Zubair; paper drafting: Abd. Kakhar Umar and James H. Zothantluanga; critical review and additional inputs: H. Lahlhenmawia, Mithun Rudrapal, and Dipak Chetia.

**Availability of data and material** All data associated with the paper shall be made available upon request to the corresponding author.

## Declarations

**Conflict of interest** The authors declare no competing interests.

## References

- World Health Organization (2021) WHO coronavirus (COVID-19) Dashboard. <https://covid19.who.int/>. Accessed 10 Nov 2021
- Tang Y, Liu J, Zhang D, Xu Z, Ji J, Wen C (2020) Cytokine storm in COVID-19: the current evidence and treatment strategies. *Front Immunol* 11:1708
- Melo AKG, Milby KM, Caparroz ALMA, Pinto ACPN, Santos RRP, Rocha AP, Ferreira GA, Souza VA, Valadares LDA, Vieira RMRA, Pileggi GS, Trevisani VFM (2021) Biomarkers of cytokine storm as red flags for severe and fatal COVID-19 cases: a living systematic review and meta-analysis. *PLoS ONE* 16:335–341
- Ruscitti P, Berardicurti O, Iagnocco A, Giacomelli R (2020) Cytokine storm syndrome in severe COVID-19. *Autoimmun Rev* 19:102562
- Anand K, Ziebuhr J, Wadhwani P, Mesters JR, Hilgenfeld R (2003) Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* 300:1763–1767
- ul Qamar MT, Alqahtani SM, Alamri MA, Chen LL (2020) Structural basis of SARS-CoV-2 3CLpro and anti-COVID-19 drug discovery from medicinal plants. *J Pharm Anal* 10:313–319
- Skariyachan S, Challapilli SB, Packirisamy S, Kumargowda ST, Sridhar VS (2019) Recent aspects on the pathogenesis mechanism, animal models and novel therapeutic interventions for middle east respiratory syndrome coronavirus infections. *Front Microbiol* 10:569
- Tutura AL, Bavari S (2019) Broad-spectrum coronavirus antiviral drug discovery. *Expert Opin Drug Discov* 14:397–412
- Zothantluanga JH (2021) Molecular docking simulation studies, toxicity study, bioactivity prediction, and structure-activity relationship reveals rutin as a potential inhibitor of SARS-CoV-2 3CL pro. *J Sci Res* 65:96–104
- Zothantluanga JH, Gogoi N, Shakya A, Chetia D, Lalthanzara H (2021) Computational guided identification of potential leads from *Acacia pennata* (L.) Willd. as inhibitors for cellular entry and viral replication of SARS-CoV-2. *Futur J Pharm Sci* 7:201
- Rudrapal M, Celik I, Khan J, Ansari MA, Alarousy RMI, Yadav R, Sharma T, Tallei TE, Pasala PK, Sahoo RK, Khairnar SJ, Bendale AR, Zothantluanga JH, Chetia D, Walode SG (2022) Identification of bioactive molecules from *Triphala* (Ayurvedic herbal formulation) as potential inhibitors of SARS-CoV-2 main protease (Mpro) through computational investigations. *J King Saud Univ - Sci*. <https://doi.org/10.1016/j.jksus.2022.101826>
- Kennedy DA, Read AF (2020) Monitor for COVID-19 vaccine resistance evolution during clinical trials. *PLOS Biol* 18:e3001000
- Singh DD, Parveen A, Yadav DK (2021) SARS-CoV-2: emergence of new variants and effectiveness of vaccines. *Front Cell Infect Microbiol* 11:777212
- Dror AA, Eisenbach N, Taiber S, Morozov NG, Mizrahi M, Zigron A, Srouji S, Sela E (2020) Vaccine hesitancy: the next challenge in the fight against COVID-19. *Eur J Epidemiol* 35:775–779
- Lucia VC, Kelekar A, Afonso NM (2021) COVID-19 vaccine hesitancy among medical students. *J Public Health* 43:445–449
- Kwok KO, Li KK, Wei WI, Tang A, Wong SYS, Lee SS (2021) Influenza vaccine uptake, COVID-19 vaccination intention and vaccine hesitancy among nurses: a survey. *Int J Nurs Stud* 114:103854

17. Kirby T (2021) New variant of SARS-CoV-2 in UK causes surge of COVID-19. *Lancet Respir Med* 9:e20–21
18. Panyod S, Ho CT, Sheen LY (2020) Dietary therapy and herbal medicine for COVID-19 prevention: a review and perspective. *J Tradit Complement Med* 10:420–427
19. Murugesan S, Kottekad S, Crasta I, Sreevathsan S, Usharani D, Perumal MK, Mudliar SN (2021) Targeting COVID-19 (SARS-CoV-2) main protease through active phytocompounds of ayurvedic medicinal plants – *Emblca officinalis* (Amla), *Phyllanthus niruri* Linn. (*Bhumi Amla*) and *Tinospora cordifolia* (Giloy) – a molecular docking and simulation study. *Comput Biol Med* 136:104683
20. Nugraha RV, Ridwansyah H, Ghazali M, Khairani AF, Atik N (2020) Traditional herbal medicine candidates as complementary treatments for COVID-19: a review of their mechanisms, pros and cons. *Evid Based Complement Alternat Med*. <https://doi.org/10.1155/2020/2560645>
21. Guijarro-Real C, Plazas M, Rodríguez-Burruezo A, Prohens J, Fita A (2021) Potential in vitro inhibition of selected plant extracts against SARS-CoV-2 chymotrypsin-like protease (3CLPro) activity. *Foods* 10:1503
22. Zannella C, Giugliano R, Chianese A, Buonocore C, Vitale GA, Sanna G, Sarno F, Manzin A, Nebbioso A, Termolino P, Altucci L, Galdiero M, de Pascale D, Franci G (2021) Antiviral activity of *Vitis vinifera* leaf extract against SARS-CoV-2 and HSV-1. *Viruses* 13:1263
23. Rolta R, Salaria D, Sharma P, Sharma B, Kumar V, Rathi B, Verma M, Sourirajan A, Baumler DJ, Dev K (2021) Phytocompounds of *Rheum emodi*, *Thymus serpyllum*, and *Artemisia annua* inhibit spike protein of SARS-CoV-2 binding to ACE2 receptor: in silico approach. *Curr Pharmacol Reports* 7:135–149
24. Kar P, Sharma NR, Singh B, Sen A, Roy A (2021) Natural compounds from *Clerodendrum* spp. as possible therapeutic candidates against SARS-CoV-2: an in silico investigation. *J Biomol Struct Dyn* 39:4774–4785
25. Ogunyemi OM, Gyebe GA, Elfiky AA, Afolabi SO, Ogunro OB, Adegunloye AP, Ibrahim IM (2020) Alkaloids and flavonoids from African phytochemicals as potential inhibitors of SARS-CoV-2 RNA-dependent RNA polymerase: an in silico perspective. *Antivir Chem Chemother* 28:204020662098407
26. Ministry of Science and Technology (2020) Compendium of antiviral medicinal plants. Mukherjee PK, Mao AA, editors. New Delhi, India
27. Kai F, Pei H, Qiong G, Lin H, XiFeng T, LingJie Z (2019) Chemical constituents of *Bridelia retusa* and their anti-neuroinflammatory activity. *Nat Prod Res Dev* 31:264–268
28. Tatiya AU, Saluja AK, Kalaskar MG, Surana SJ, Patil PH (2017) Evaluation of analgesic and anti-inflammatory activity of *Bridelia retusa* (Spreng) bark. *J Tradit Complement Med* 7:441–751
29. Kumar T, Jain V (2014) Antinociceptive and anti-inflammatory activities of *Bridelia retusa* methanolic fruit extract in experimental animals. *Sci World J* 2014:890151
30. Jayasinghe L, Kumarihamy BMM, Jayarathna KHRN, Udishani NWMG, Bandara BMR, Hara N, Fujimoto Y (2003) Antifungal constituents of the stem bark of *Bridelia retusa*. *Phytochemistry* 62:637–641
31. Jadhav N, Kulkarni S, Mane A, Kulkarni R, Palshetker A, Singh K, Joshi S, Risbud A, Kulkarni S (2015) Antimicrobial activity of plant extracts against sexually transmitted pathogens. *Nat Prod Res* 29:1562–1566
32. Govea-Salas M, Am R-E, Rodríguez-Herrera R, Sa L-S, Cn A-G, Zugasti-Cruz A, Salas-Villalobos TB, Morlett-Chávez JA (2016) Gallic acid decreases hepatitis C virus expression through its antioxidant capacity. *Exp Ther Med* 11:619–624
33. Park SW, Kwon MJ, Yoo JY, Choi HJ, Ahn YJ (2014) Antiviral activity and possible mode of action of ellagic acid identified in *Lagerstroemia speciosa* leaves toward human rhinoviruses. *BMC Complement Altern Med* 14:171
34. Fahnchaksai K, Kodchakorn K, Pothacharoen P, Kongtawelert P (2016) Effect of sesamin against cytokine production from influenza type A H1N1-induced peripheral blood mononuclear cells: computational and experimental studies. *Vitr Cell Dev Biol - Anim* 52:107–119
35. Zhou B, Li J, Liang X, Pan X, Hao Y, Xie PF, Jiang HM, Yang ZF, Zhong NS (2020)  $\beta$ -sitosterol ameliorates influenza A virus-induced proinflammatory response and acute lung injury in mice by disrupting the cross-talk between RIG-I and IFN/STAT signaling. *Acta Pharmacol Sin* 41:1178–1196
36. Dallakyan S, Olson AJ (2015) Small-molecule library screening by docking with PyRx. *Methods Mol Biol* 1263:243–250
37. Sander T, Freyss J, von Korff M, Rufener C (2015) DataWarrior: an open-source program for chemistry aware data visualization and analysis. *J Chem Inf Model* 55:460–473
38. Schrodinger LLC (2020a) Schrödinger Release 2020–1: Maestro Desmond interoperability tools; Schrödinger. New York, New York, USA; 2020
39. Schrodinger LLC (2020b) Release 2020–4: Desmond molecular dynamics system, D.E. Shaw Research, New York, NY. New York, New York, USA; 2020
40. Bowers KJ, Chow DE, Xu H, Dror RO, Eastwood MP, Gregersen BA, Klepeis JL, Kolossvary I, Moraes MA, Sacerdoti FD, Salmon JK, Shan Y, Shaw DE (2006) Scalable algorithms for molecular dynamics simulations on commodity clusters. In: ACM/IEEE SC 2006 Conference (SC'06) IEEE. <https://doi.org/10.1109/SC.2006.54>
41. Zubair MS, Maulana S, Widodo A, Pitopang R, Arba M, Hariono M (2021) GC-MS, LC-MS/MS, docking and molecular dynamics approaches to identify potential SARS-CoV-2 3-chymotrypsin-like protease inhibitors from *Zingiber officinale* Roscoe. *Molecules* 26:5230
42. Duarte Y, Rojas M, Canan J, Pérez EG, González-Nilo F, García-Colunga J (2021) Different classes of antidepressants inhibit the rat  $\alpha 7$  nicotinic acetylcholine receptor by interacting within the ion channel: a functional and structural study. *Molecules* 26:998
43. Mahmoud A, Mostafa A, Al-Karmalawy AA, Zidan A, Abulkhair HS, Mahmoud SH, Shehata M, Elhefnawi MM, Ali MA (2021) Telaprevir is a potential drug for repurposing against SARS-CoV-2: computational and in vitro studies. *Heliyon* 7:e07962
44. Li J, Abel R, Zhu K, Cao Y, Zhao S, Friesner RA (2011) The VSGB 2.0 model: a next generation energy model for high resolution protein structure modeling. *Proteins Struct Funct Bioinforma* 79:2794–2812
45. de Campos LJ, Palermo NY, Conda-Sheridan M (2021) Targeting SARS-CoV-2 receptor binding domain with stapled peptides: an in silico study. *J Phys Chem B* 125:6572–6586
46. Johari J, Kianmehr A, Mustafa MR, Abubakar S, Zandi K (2012) Antiviral activity of baicalein and quercetin against the Japanese encephalitis virus. *Int J Mol Sci* 13:16020–16045
47. Moghaddam E, Teoh BT, Sam SS, Lani R, Hassandarvish P, Chik Z, Yueh A, Abubakar S, Zandi K (2014) Baicalin, a metabolite of baicalein with antiviral activity against dengue virus. *Sci Rep* 4:5452
48. Pinzi L, Rastelli G (2019) Molecular docking: shifting paradigms in drug discovery. *Int J Mol Sci* 20:4331
49. Elfiky AA (2020) Ribavirin, remdesivir, sofosbuvir, galidesivir, and tenofovir against SARS-CoV-2 RNA dependent RNA polymerase (RdRp): a molecular docking study. *Life Sci* 253:117592
50. Dahab MA, Hegazy MM, Abbass HS (2020) Hordatinines as a potential inhibitor of COVID-19 main protease and rna polymerase: an in-silico approach. *Nat Products Bioprospect* 10:453–462

51. Ferreira L, dos Santos R, Oliva G, Andricopulo A (2015) Molecular docking and structure-based drug design strategies. *Molecules* 20:13384–13421
52. Onakpoya JJ, Heneghan CJ, Aronson JK (2016) Post-marketing withdrawal of 462 medicinal products because of adverse drug reactions: a systematic review of the world literature. *BMC Med* 14:10
53. López-López E, Naveja JJ, Medina-Franco JL (2019) DataWarrior: an evaluation of the open-source drug discovery tool. *Expert Opin Drug Discov* 14:335–341
54. Morris-Schaffer K, McCoy MJ (2021) A review of the LD 50 and its current role in hazard communication. *ACS Chem Heal Saf* 28:25–33
55. Adebayo AH, Ishola TA, Yakubu OF (2021) Acute toxicity and antimalarial studies of extract of *Allophylus spicatus* in animals. *Toxicol Res* 37:345–354
56. Ishtiaq S, Akram M, Kamran SH, Hanif U, Afridi MSK, Sajid-ur-Rehman AA, Asif A, Younus M, Akbar S (2017) Acute and sub-acute toxicity study of a Pakistani polyherbal formulation. *BMC Complement Altern Med* 17:387
57. Paul A, Shakya A, Zaman MK (2020) Assessment of acute and sub-chronic neurotoxicity of *Morus alba* L. fruits in rodents. *Futur J Pharm Sci* 6:88
58. McKerrow JH, Lipinski CA (2017) The rule of five should not impede anti-parasitic drug development. *Int J Parasitol Drugs Drug Resist* 7:248–249
59. Ghose AK, Viswanadhan VN, Wendoloski JJ (1999) A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. *J Comb Chem* 1:55–68
60. Muegge I, Heald SL, Brittelli D (2001) Simple selection criteria for drug-like chemical matter. *J Med Chem* 44:1841–1846
61. Veber DF, Johnson SR, Cheng H-Y, Smith BR, Ward KW, Kopple KD (2002) Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem* 45:2615–2623
62. Egan WJ, Merz KM, Baldwin JJ (2000) Prediction of drug absorption using multivariate statistics. *J Med Chem* 43:3867–3877
63. Zhong H, Mashinson V, Woolman T, Zha M (2013) Understanding the molecular properties and metabolism of top prescribed drugs. *Curr Top Med Chem* 13:1290–1307
64. Pal S, Mehta D, Dasgupta U, Bajaj A (2021) Advances in engineering of low molecular weight hydrogels for chemotherapeutic applications. *Biomed Mater* 16:024102
65. Daina A, Michielin O, Zoete V (2017) SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep* 7:42717
66. Zanger UM, Schwab M (2013) Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol Ther* 138:103–141
67. Kneller DW, Phillips G, O'Neill HM, Jedrzejczak R, Stols L, Langan P, Joachimiak A, Coates L, Kovalevsky A (2020) Structural plasticity of SARS-CoV-2 3CL Mpro active site cavity revealed by room temperature X-ray crystallography. *Nat Commun* 11:3202
68. Chen D, Oezguen N, Urvil P, Ferguson C, Dann SM, Savidge TC (2016) Regulation of protein-ligand binding affinity by hydrogen bond pairing. *Sci Adv* 2:e1501240
69. Khan MT, Ali A, Wang Q, Irfan M, Khan A, Zeb MT, Zhang YJ, Chinnasamy S, Wei DQ (2021) Marine natural compounds as potent inhibitors against the main protease of SARS-CoV-2—a molecular dynamic study. *J Biomol Struct Dyn* 39:3627–3736
70. Junejo JA, Zaman K, Rudrapal M, Celik I, Attah EI (2021) Antidiabetic bioactive compounds from *Tetrastigma angustifolia* (Roxb.) Deb and *Oxalis debilis* Kunth.: validation of ethnomedicinal claim by in vitro and in silico studies. *South African J Bot* 143:164–175
71. Mishra CB, Pandey P, Sharma RD, Malik MZ, Mongre RK, Lynn AM, Prasad R, Jeon R, Prakash A (2021) Identifying the natural polyphenol catechin as a multi-targeted agent against SARS-CoV-2 for the plausible therapy of COVID-19: an integrated computational approach. *Brief Bioinform* 22:1346–1360
72. Bhowmik D, Sharma RD, Prakash A, Kumar D (2021) Identification of nafamostat and VR23 as COVID-19 drug candidates by targeting 3CLpro and PLpro. *J Mol Struct* 1233:130094
73. Khater S, Kumar P, Dasgupta N, Das G, Ray S, Prakash A (2021) Combining SARS-CoV-2 proofreading exonuclease and RNA-dependent RNA polymerase inhibitors as a strategy to combat COVID-19: a high-throughput in silico screening. *Front Microbiol* 12:647693
74. Ji YY, Li YQ (2010) The role of secondary structure in protein structure selection. *Eur Phys J E* 32:103–107
75. Prajapati J, Patel R, Goswami D, Saraf M, Rawal RM (2021) Sterenin M as a potential inhibitor of SARS-CoV-2 main protease identified from MeFSAT database using molecular docking, molecular dynamics simulation and binding free energy calculation. *Comput Biol Med* 135:104568
76. Yadav DK, Kumar S, Saloni MS, Yadav L, Teli M, Sharma P, Chaudhary S, Kumar N, Choi EH, Kim HS, Kim MH (2018) Molecular insights into the interaction of RONS and thieno[3,2-c]pyran analogs with SIRT6/COX-2: a molecular dynamics study. *Sci Rep* 8:4777
77. Ertl P, Rohde B, Selzer P (2000) Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *J Med Chem* 43:3714–3717
78. Pattar SV, Adhoni SA, Kamanavalli CM, Kumbar SS (2020) In silico molecular docking studies and MM/GBSA analysis of coumarin-carbonodithioate hybrid derivatives divulge the anticancer potential against breast cancer. *Beni-Suef Univ J Basic Appl Sci* 9:36
79. Wang E, Sun H, Wang J, Wang Z, Liu H, Zhang JZH, Hou T (2019) End-point binding free energy calculation with MM/PBSA and MM/GBSA: strategies and applications in drug design. *Chem Rev* 119:9478–9508
80. López-Otín C, Bond JS (2008) Proteases: multifunctional enzymes in life and disease. *J Biol Chem* 283:30433–30437
81. Acquadro S, Civra A, Cagliero C, Marengo A, Rittà M, Francese R, Sanna C, Bertea C, Sgorbini B, Lembo D, Donalisio M, Rubiolo P (2020) Punica granatum leaf ethanolic extract and ellagic acid as inhibitors of Zika virus infection. *Planta Med* 86:1363–1374

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.