Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Computational investigation of natural compounds as potential main protease (M\text{pro}) inhibitors for SARS-CoV-2 virus

Chirag N. Patel\textsuperscript{a,c,1,*}, Siddhi P. Jani\textsuperscript{b,1}, Sivakumar Prasanth Kumar\textsuperscript{c}, Krunal M. Modi\textsuperscript{d,e,**}, Yogesh Kumar\textsuperscript{g}

\textsuperscript{a} Computer-Aided Drug Design Group, Chemical Biology Laboratory, Center for Cancer Research, National Cancer Institute, National Institute of Health, Frederick, MD, 21702, USA
\textsuperscript{b} Institute of Science, Nirma University, Ahmedabad, 382481, Gujarat, India
\textsuperscript{c} Department of Botany, Bioinformatics, and Climate Change Impacts Management, School of Sciences, Gujarat University, Ahmedabad, 380009, Gujarat, India
\textsuperscript{d} Department of Molecular Electrochemistry and Catalysis, J. Heyrovsky Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, Dolanska 2155/3, 182 23 Prague 8, Czech Republic
\textsuperscript{e} Department of Humanities and Science, School of Engineering, Indrashil University, Mehsana, 382740, Gujarat, India
\textsuperscript{f} Department of General Visceral and Thoracic Surgery, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, Hamburg, 20246, Germany

\textbf{ARTICLE INFO}

\textbf{Keywords:}
Main proteases (M\text{pro})
DFT calculation
Dynamics simulation
Binding free energy
Omicron

\textbf{ABSTRACT}

The coronavirus disease 2019 (COVID-19) pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is significantly impacting human lives, overburdening the healthcare system and weakening global economies. Plant-derived natural compounds are being largely tested for their efficacy against COVID-19 targets to combat SARS-CoV-2 infection. The SARS-CoV-2 Main protease (M\text{pro}) is considered an appealing target because of its role in replication in host cells. We curated a set of 7809 natural compounds by combining the collections of five databases viz Dr Duke’s Phytochemical and Ethnobotanical database, IMPPAT, PhytoHub, AromaDb and Zinc. We applied a rigorous computational approach to identify lead molecules from our curated compound set using docking, dynamic simulations, the free energy of binding and DFT calculations. Theaflavin and ginkgetin have emerged as better molecules with a similar inhibition profile in both SARS-CoV-2 and Omicron variants.

1. Introduction

The beginning of 2020 took the planet to a virtual impasse with the outbreak of a novel Severe Acute Respiratory Virus Syndrome Coronavirus-2 (SARS-CoV-2) [1]. The Coronavirus disease 2019 (COVID 2019) pandemic began in Wuhan province, China, in 2019 and has resulted in unprecedented human deaths worldwide [2]. The International Committee on Taxonomy of Viruses (ICTV) recognized this virus as SARS-CoV-2 on 11\textsuperscript{th} March 2020 due to genetic relation to SARS-CoV infection reported in the year 2003 [1,3]. The first SARS-CoV-2 Wuhan Hu-1 strain evolved genetically [4] during worldwide spread leading to five different variants such as B.1.1.7 (Alpha) detected in the United Kingdom (September 2020), B.1.351 (Beta) in South Africa (May 2020), B.1.617.2 (Delta) in India (October 2020) and B.1.1.529 (Omicron) in South Africa (November 2021) [5]. The COVID virus majorly affects the respiratory system with mild to severe symptoms including watery nose, persistent cough, sore throat, loss of smell or taste, frequently with fever and body pain [2,6,7]. Although antiviral medications (e.g., nirmatrelvir, ritonavir) [8] are being prescribed to alleviate the symptoms, the current practices suggest that vaccine administration boosts the immune system to develop antibodies curbing the COVID infection [2,6,7]. However, the side effects and other related ailments associated with the COVID vaccines are largely unknown and necessitate the pursuit of new antivirals [9].

Coronaviruses (CoVs) are enveloped and positive-stranded RNA viruses...
viruses belonging to the Coronaviridae family and are categorized into seven major classes based on their genome sequences and serological reactions [10]. This includes 229E (Alpha coronavirus), NL63 (Alpha coronavirus), OC43 (Beta coronavirus), HUK1 (Beta coronavirus), MERS-CoV (Beta coronavirus, the causative agent of Middle East Respiratory Syndrome, MERS), SARS-CoV (the Beta coronavirus, the agent of severe acute respiratory syndrome, SARS) and SARS-CoV-2 (the novel coronavirus responsible for COVID-19) [11]. Certain CoVs can infect humans (e.g., damage to lungs, difficulty breathing, recurrent fevers, weakness, sadness, anxiety, and chronic memory and attention impairment) while others are only capable of infecting animals (e.g., respiratory tract and olfactory epithelium damage in golden hamster) [12]. Initial reports suggest that bat is the primary host of SARS-CoV-2 transmission in humans while pangolin is an intermediary host [13]. These reports are getting challenged based on emerging pieces of evidence on the genetic and evolutionary relationships among the human, bat and pangolin coronaviruses [14].

In the current situation, it has become necessary to monitor and ultimately eradicate this deadly disease by using an effective vaccine. Several vaccines are available on the market, but it contains some side effects; hence there is a dire need to combat this coronavirus with potential drug or natural inhibitors. CoV infection can be controlled by masking the virus entry into the host cells through the invention of inhibitors protecting the replication and transcription of the virus [15,16].

SARS-CoV-2 encodes both structural and non-structural proteins (NSPs). The structural proteins include envelope (small) membrane protein (E protein), membrane protein (M protein), nucleocapsid protein (N protein) and spike protein (S protein). Some proteases such as main protease (Mpro, NSP5) and Papain-like protease (PLpro, NSP3) are also present [16,17]. A series of NSPs facilitate crucial functions essential for replication, survival and virulence [18]. These include NSP13 (helicase), NSP12 (RNA-dependent RNA polymerase, RdRp), NSP14 (N-terminal exoribonuclease and C-terminal guanine-N’-methyltransferase), NSP15 (uridylicate-specific endoribonuclease), NSP16 (2’-O-methyltransferase) among others [19].

An ideal strategy to develop drugs for treating COVID-19 is targeting the SARS-CoV-2 Mpro protein that plays a prominent role in viral replication and transcription in host cells [6,20]. Inhibition of Mpro protein is actively pursued in the research community by leveraging the techniques of structure-based drug design and partially due to the absence of human homologs that may constitute a high rate of therapeutic efficacy [16]. SARS-CoV-2 Mpro is a cysteine protease (EC 3.4.22.69) belonging to the PA protease clan family. Jin et al., released the first crystal structure of SARS-CoV-2 Mpro (resolution, 2.16 Å) in the Protein Data Bank (PDB, entry: 6LU7) on 5th February 2020 [20]. The functional form of Mpro is a dimer (~33 kDa), and each monomer (306 amino acid residues) consists of three domains (I to III; domain I - residues 8 to 101, II – 102 to 184, and III – 201 to 303) [16,21]. The active site of Mpro is situated between domains I and II and carry out proteolytic function using the Cys-His catalytic dyad [16]. Both domains I and II are composed of antiparallel β-barrels, and domain III consists of five α-helical strands. The residue Glu166 is responsible for bringing the active site close to the dimer interface that promotes the access of the S1 subsite necessary for ligand binding. Mutation of this residue results in the loss of substrate-induced dimerization [22]. The presence of an oxyanion hole, which is composed of residues Gly143, Ser144, and Cys145, is also notable in stabilizing the transition state during catalysis by protecting the negative charge on the ligand’s proximal oxygen atom and enabling hydrogen bond. The active site of Mpro is encompassed by the following residues, Ser46, Gin189, Thr190, Ala191, Pro168, Glu166, Leu141, and Asn142 in addition to the catalytic dyad (Cys145, His41) [16,21].

On 24 November 2021, the World Health Organization (WHO) recognized Omicron as a new SARS-CoV-2 variant (B.1.1.529) of severe concern and highly contagious compared to the Delta variant but possesses relatively less severity [23,24]. Despite increased surveillance and vaccination drive, Omicron has spread to 108 nations. A large-scale study is necessary to investigate vaccine-induced immunity after booster administration and adaptation of acquired immunity in human subjects to circumvent the Omicron variant [25,26]. Researchers obtained the three-dimensional structure of Omicron Mpro harbouring a single mutation P132H to its SARS-CoV-2 counterpart and reported similar catalytic efficiency of antivirals such as GC-376, PF-07321332 (nirmatrelvir), and PF-00835231 [27,28].

Plants have been the primary source of medication since ancient times due to their therapeutic potential and low toxicity. This allowed researchers to use them as initial lead molecules for the drug discovery and development process. Several studies have previously shown that plant-based research might be an effective strategy to identify better leads for COVID-19 medication development [29]. Kumar et al., 2021 tested natural metabolites against COVID-19 Mpro [30]. Pandey et al., 2020 used phytochemicals to target molecules against SARS-CoV-2 spike protein. To promote such lead compounds to the pre-clinical stage of drug development, they must be toxic-free and thereby selecting a library of natural compounds for such assessment is critical [31]. We selected 7809 natural compounds from Zinc [32], IMPMAT [33], PhytoHub [34], AromaDb [35] and Dr. Duke’s Phytochemical and Ethnobotanical databases [36] with emphasis on physicochemical properties, drug performance rates, targeted proteins, business providers and drug similarity profile. In this study, we used an integrative workflow comprising structure-based virtual screening, DFT calculations, dynamic simulations, and binding free energy calculations to find Mpro inhibitors from our compiled library of natural compounds.

2. Materials and methods

2.1. Ligand selection

A total of 7809 natural compounds were chosen based on drug-like properties through Dr. Duke’s Phytochemical and Ethnobotanical database, IMPMAT, PhytoHub, AromaDb and Zinc databases. These compounds have favourable ADMET characteristics. The three-dimensional structures were extracted in sdf format and polar hydrogen atoms were added. We carried out stereochemistry checks and energy was minimized using the Amber 03 forcefield using 1000 steps of the steepest descent technique with no convergence. The minimized molecules were then saved in.sdf format convenient for further calculations in the YASARA software (academic license) [37] [-] [39].

2.2. Protein selection

The crystal structures of SARS-CoV-2 and Omicron Mpro in complex with inhibitors (N3, oxazole-pyridoline molecule, resolution 2.16 Å and K36 (pyrrolidine-sulfonic acid molecule, 2.05 Å) were retrieved from PDB with the entries 6LU7 [16] and 7TOB [40], respectively. Structure files were prepared using the “Clean” module of YASARA Structure (version 19.12.14) which included removing crystallographic waters, atom typing with the Amber03 force field, adding polar hydrogens, assigning charges to titratable amino acids before, and carried out geometry optimization with the steepest gradient approach (100 iterations) [37,41].

2.3. Virtual screening of natural compounds

YASARA Structure (version 19.12.14) was used to perform the virtual screening exercise. The ligand binding site in Mpro proteins was selected based on the location of crystallized ligands (N3 and K36) coordinates. Initially, the coordinates of the crystallized ligand were obtained from PDB, prepared using the “Clean” module of YASARA suite and atom-typed using AMBER03 force field and subsequently redocked to its native position. The prepared 7809 natural compounds were

C.N. Patel et al.

Computers in Biology and Medicine 151 (2022) 106318
Docked into the Msupero site (20 × 20 × 20 Å grid box size) using the AutoDock Vina technique for pose search, and the AMBER03 force field. Was used to enumerate protein-ligand interactions [42]. Finally, a short energy-minimization step was pursued as a measure of post-docking refinement of docking solutions. Conjugate gradient-based energy minimization was carried out with 100 iterations with no restraint on heavy atoms to ensure the generation of optimal bonds and orders in the docked pose. The YASARA empirical energy function engineered in the YASARA Structure suite was used to score the best poses. A large positive value suggests strong binding of the ligand to its target receptor, which is in direct contrast to conventional scoring functions with negative energy implying better ligand binding [43]. The AMBER3 force field enumerates the intermolecular interactions between the sum of potential energy and solvation energy terms in the free state (receptor and ligand in isolation) and the sum of the same set of energy terms in the complex state (ligand-bound receptor conformation). The top-ranked natural compounds were visualized using ACCELrys Discovery Studio (DS) visualizer (freeware for academic use) and analyzed various intermolecular bond formations such as hydrogen bonds and hydrophobic interactions [44]. The compounds with the highest binding energy with better receptor interactions were selected for further study using the molecular dynamics approach [45]. The following empirical equation was used to calculate the binding free energy ΔGbind:

\[
\Delta G = \Delta G_{\text{vdw}} + \Delta G_{\text{bond}} + \Delta G_{\text{elec}} + \Delta G_{\text{rot}} + \Delta G_{\text{desolv}}
\]  

(1)

Where \(\Delta G_{\text{vdw}}\) = van der Waals term for docking energy; \(\Delta G_{\text{bond}}\) = H bonding term for docking energy; \(\Delta G_{\text{elec}}\) = electrostatic term for docking energy; \(\Delta G_{\text{rot}}\) = torsional free energy term for the compound when the compound transits from unbounded to bounded state; \(\Delta G_{\text{desolv}}\) = desolvation term for docking energy.

2.4. Density functional theory calculation of top-scoring molecules

The Density Functional Theory (DFT) calculations for the optimization of N3, K36 and best natural hits were performed using Becke’s three-parameter hybrid exchange (B3LYP) in the Gaussian 09 (G09 B.01) software package (academic license) [46–49]. The 6–311g+ (d,p) basis set was employed for all compounds. Global electronic descriptors such as softness (S) [50], hardness (η) [51–53], and chemical potential (μ) for selectivity and reactivity of the DFT concept were studied. Further, Frontier molecular orbital (FMOs) difference i.e., the energy difference between Highest occupied molecular orbital (HOMO) and Lowest unoccupied molecular orbital (LUMO) was also calculated [50,54,55]. The B3LYP method with 6–311g+ (d,p) basis sets was used to visualize the HOMO and LUMO profiles. The optimized structure and HOMO-LUMO difference were visualized using GaussView (G09 B.01) software package (academic license). The stability [56], softness (S), hardness (η) and chemical potential (μ) were calculated using the bandgap energy (Eg) as below.

\[
E_g = E_{\text{HOMO}} - E_{\text{LUMO}}
\]

(2)

\[
\eta = \frac{E_{\text{LUMO}} - E_{\text{HOMO}}}{2}
\]

(3)

\[
S = \frac{1}{2\eta}
\]

(4)

\[
\mu = \frac{E_{\text{HOMO}} + E_{\text{LUMO}}}{2}
\]

(5)

The maximum hardness principle (MHP) states that the systems in their ground or valence states will tend to arrange themselves to be as hard as possible. The global index is the electrophilicity index (ω), an important term for energy reduction due to the maximum current of the electron between donor and acceptor.

The stabilization energy (ΔE) calculated as follow.

\[
\Delta E = -\frac{\mu^2}{2\eta}
\]

(7)

2.5. Molecular dynamics simulation of top-scoring molecules

Molecular dynamics (MD) simulations were performed to evaluate the stability of the Msupero-ligand complexes and the binding ability of ligands [57]. This computational technique, in addition to molecular docking, enabled the study of large conformational shifts as well as the stabilization of protein-ligand complexes [58,59]. The Desmond (Schrödinger, LLC, NY, USA) package was used to run MD simulations [60]. A total of 7 MD simulations were performed with the following molecular configurations, re-docked (N3)-6LU7, theaflavin-6LU7, ginkgetin-6LU7, re-docked (K36)-7TOB, theaflavin-7TOB, and ginkgetin-7TOB for 1 µs time interval each. These ligand-bound Msupero structures were analyzed using the Minimization panel of Desmond to identify the most accurate and energy-minimized structure for MD simulations. The AMBER05 force field with TIP3P (Transferable Inter-molecular Potential with 3 Points) solvent model was used with a hybrid approach of steepest descent (100 steps) and L-BFGS (Broyden–Fletcher–Goldfarb-Shanno) approaches. Structures that secured energy differences less than 0.1 kcal/mol were retained and subjected as the starting input to MD simulations. Further, the apo-Msupero protein was subjected to MD simulations to account for the comparative analysis of ligand-bound changes. Both receptor-ligand complexes were again prepared using Protein Preparation Wizard for compliance in the Schrodinger’s Maestro workspace with default settings and not altering the ligand pose conformation. The periodic simulation box was built with the System Builder module and dissolved with TIP3P water model [61]. The Optimized Potentials for Liquid Simulations (OPLS) all-atom force field 2005 [62] was employed to apply bonded and non-bonded constraints and neutralized them by adding counter ions and energy minimization with the steepest descent technique (1000 iterations). After reaching equilibrium, an unrestrained protocol was enabled in the NPT ensemble for 1 µs at 300 K and 1.01325 bar (atoms, pressure, and temperature were kept constant). The isotropic Martyna-Tobias-Klein barostat (relaxation time = 2 ps) [63] and the Nose-Hoover thermostat (relaxation time = 1 ps) were used. The smooth particle mesh Ewald (PME) [64] system with a smooth particle network with the RESPA integrator [65] was used to measure short-range interactions (cut-off = 9 Å) and long-range Coulomb interactions. At an interval of 5 ps frames were captured to build the simulation trajectory. The system’s stability was assessed using various plots for root mean square deviation (RMSD), root mean square fluctuation (RMSF), hydrogen bond analysis, radius of gyration (Rg), and torsional bonds [66].

2.6. Binding free energy calculations of top-scoring molecules

The binding free energy of the top-scoring molecules in complex with the target proteins was calculated using the MM/GBSA (molecular mechanics (MM) with generalized Born and surface area continuum solvation (GBSA)) with the single trajectory (a trajectory containing single protein-ligand complex) approach [67,68]. The combination of the MM technique with implicit solvation models provides a most accurate ranking of ligands than empirical scoring schemes used in the docking exercise above and is more computationally efficient than rigorous alchemical perturbation methods. The binding free energy (ΔGBind) is composed of three terms, i.e. MM energy term (ΔEMM) is a combination of three different interaction energy terms such as internal energy of the molecular system studied in terms of bond, angle and dihedral energies (ΔEinternal), electrostatic (ΔEelectrostatic) and van der Waals (ΔEvdw), ii.
Solvation energy ($\Delta G_{\text{solv}}$) composed of polar ($\Delta G_{\text{PB}}$) and non-polar terms ($\Delta G_{\text{SA}}$), and iii. Entropy term (-T$\Delta S$), a conformation entropy term accounting for the loss of entropy when a ligand in a free state binds to a free state of the target receptor [69]. Although Poisson-Boltzmann (PB) methods are more accurate than GB, the Prime module of the Schrödinger suite features a thorough analysis of the optimized implicit solvent energy terms and physics-based corrected terms for H-bonding, hydrophobic interactions, intramolecular interaction among others in the VSGB 2.0 model [69] [71]. The MM/GBSA calculations were performed upon the 1 \mu s long MD simulation trajectory obtained for the top-scoring molecules in complex with the target receptors at a frame step size of 10. The following equations 8–11 were used to quantify the ligand-binding free energy values:

\begin{align*}
\Delta G_{\text{bind}} & = G_{\text{complex}} - (G_{\text{receptor}} + \Delta G_{\text{ligand}}) \\
\Delta G_{\text{bind}} & = \Delta H - T\Delta S = \Delta E_{\text{MM}} + \Delta G_{\text{sol}} - T\Delta S \\
\Delta E_{\text{MM}} & = \Delta E_{\text{internal}} + \Delta E_{\text{electrostatic}} + \Delta E_{\text{vdw}} \\
\Delta G_{\text{sol}} & = \Delta G_{\text{PB/GB}} + \Delta G_{\text{SA}}
\end{align*}

3. Results

3.1. Virtual screening of SARS-CoV-2 M\textsuperscript{pro} target

The best natural compounds with the potential to inhibit the SARS-CoV-2 M\textsuperscript{pro} were identified through a virtual screening technique using the YASARA Structure suite. First, the docking accuracy was evaluated by re-docking the two co-crystal ligands N3 (PDB entry: 6LU7) and K36 (7TOB) into their ligand-binding site and measuring the RMSD between the native and docked conformations. The close placement of the redocked ligand to its native positions demonstrated that the dimensions (20 $\times$ 20 $\times$ 20 Å; 0.375 Å spacing) of the grid box built upon the M\textsuperscript{pro} targets were sufficient to dock native-like ligands from the compiled natural compound dataset. In addition, the constitution of the catalytic dyad (Cys145 and His41) along with crucial polar and hydrophobic pocket-lining residues within the docking grid confirmed that the intermolecular interactions secured by natural compounds with these M\textsuperscript{pro} residues will highlight inhibitory functions. After the generation of docking solutions, a post-docking refinement process with a conjugate gradient approach (100 steps) was adopted to generate physically relevant docking conformations. Figure. 1 depicts the dock pose of co-crystallized ligand N3 and 6LU7 protein. The best dock pose with 7.55 kcal/mol binding energy was identified which showed an RMSD of 1.70 Å. The re-docked pose of N3 comprised of 7 H-bonds (with Gly 143, Phe140, His163, His164, Glu166 and Thr190), 1 amide $\pi$-stacked (Leu141), 5 $\pi$-alkyl (His41, Met49, Leu167, Pro168, and Ala191) contacts, 2 carbon-hydrogen bonds (Met165 and His172), and 1 van der Waals (with Asn142) (Figure. 1). Most of the intermolecular contacts were noted as in crystal pose demonstrating the ability to generate similar interaction profiles with new molecules. The docking validation of Omicron variant M\textsuperscript{pro} showed that K36 obtained close to near-native pose with a binding energy of 7.80 kcal/mol with an RMSD of 0.32 Å (Table-1). The re-docked pose of K36 contained 4 hydrogen bonds with Phe140, His164 and Glu166 and a single covalent bond with Cys145.
Table 1: The calculated binding energy, hydrogen bonds and contacting receptor residues of co-crystal ligand N3 and top-five natural compounds with SARS-CoV-2 Main Protease (PDB ID: 6LU7) using YASARA structure.

| Compounds Name | Binding energy (kcal/mol) | Hydrogen bonds | Contacting receptor residues |
|----------------|---------------------------|----------------|----------------------------|
| N3             | 7.55                      | 8              | Thr24, Thr25, Thr26, Leu27, His41, Val42, Thr45, Ser46, Met49, Tyr54, Phe40, Leu41, Asn42, Gly43, Ser144, Cys145, His163, His164, Met165, Glu166, His172, Asp187, Arg188, Gln189, His41, Met49, Phe40, Leu41, Asn42, Gly43, Ser144, Cys145, His163, His164, Met165, Glu166, Pro168, Arg187, Asp188, Gln189, Thr190, Ala191, Gln192 |
| Theaflavin     | 10.04                     | 4              | Thr24, Thr25, Thr26, Leu27, His41, Met49, Phe40, Leu41, Asn42, Gly43, Ser144, Cys145, His163, His164, Met165, Glu166, Leu167, Pro168, His172, Asp187, Arg188, Gln189, His41, Met49, Phe40, Leu41, Asn142, Gly143, Ser144, Cys145, His163, His164, Met165, Glu166, Pro168, Asp187, Arg188, Gln189, Thr190, Ala191, Gln192 |
| Ginkgetin      | 9.64                      | 4              | Leu27, His41, Met49, Phe140, Leu141, Asn142, Gly143, Ser144, Cys145, His163, His164, Met165, Glu166, Leu167, Pro168, His172, Asp187, Arg188, Gln189, Thr190, Ala191, Gln192 |
| Hesperadin     | 8.47                      | 4              | Thr24, Thr25, Thr26, Leu27, His41, Met49, Asn142, Gly143, Cys145, His164, Met165, Glu166, Pro168, Asp187, Arg188, Gln189, Thr190, Ala191, Gln192 |
| Withanolide-D  | 8.37                      | 1              | Thr25, Thr26, Leu27, His41, Ser46, Met49, Asn142, Gly143, Ser144, Cys145, His164, Met165, Glu166, Pro168, Arg188, Gln189, Thr190, Ala191, Gln192 |
| Psoraladin     | 8.35                      | 4              | His41, Met49, Phe40, Leu41, Asn142, Gly143, Ser144, Cys145, His163, His164, Met165, Glu166, Leu167, His172, Asp187, Arg188, Gln189, Thr190, Ala191, Gln192 |

There were 5 carbon-hydrogen bonds with His164, Met165 and His72, and a single unfavourable donor-donor bond with Gln189 due to steric hindrance, 1 alkyl and 1 π-alkyl bond with His41 and Met49 (Figure 2). The 7809 compounds compiled from Dr. Duke’s Phytochemical and Ethnobotanical databases, IMPPAT, PhytoHub, AromaDb and Zinc databases were docked against both SARS-CoV-2 and Omicron Mpro protein targets. The natural compounds were selected based on two aspects, the better binding energy of the compound with the target site with preferential interactions within the S1 subsite and the ability of the compound to recruit the highest number of amino acid residues from the binding cleft of N3. The top five natural compounds that secured interactions with the N3 binding cleft were theaflavin, ginkgetin, hesperadin, withanolide D and psoraladin (Figure 3). According to the ranking based on binding energy, theaflavin has the highest binding energy of 10.04 kcal/mol which was better than N3 and K36 binding energy and found to be interacting with His41, Met49, Phe40, Leu41, Asn142, Gly143, Ser144, Cys145, His163, His164, Met165, Glu166, Pro168, Asp187, Arg188, Gln189, Thr190 and Ala191 (Figure S1). The intermolecular contacts of N3 dock poses were charted using Discovery Studio Visualizer which included one hydrogen bond, two carbon-hydrogen bonds, one π-sulfur bond (covalent bond), one π- π stacked bond, and one π-π T-shaped bond formation was examined. The second best compound in the ranking order based on binding energy was ginkgetin, which had binding energy of 9.643 kcal/mol and interacted with Leu27, His41, Met49, Phe140, Leu141, Asn142, Gly143, Ser144, Cys145, His163, His164, Met165, Glu166, Leu167, Pro168, His172, Asp187, Arg188, Gln189, Thr190, Ala191 and Gln192 residues (Figure S2). Table 1 and Table 2 list the top-ranked compounds, binding energies, rankings and their amino acid interactions with SARS-CoV-2 Mpro protein.

The virtual screening exercise of our natural library was also carried out targeting the Omicron variant of the SARS-CoV-2 Mpro protein to study whether the best-ranked natural compounds were able to obtain the best ranks with the Omicron Mpro target. Similar to the SARS-CoV-2 compound ranking, the Omicron Mpro gauged the same set of natural compounds. This includes theaflavin, ginkgetin, hesperadin, withanolide D and psoraladin (Figure 3) (Table 2). The highest binding energy among the natural compounds was secured by theaflavin with a binding energy of 8.51 kcal/mol with 4 hydrogen bonds (Phe140, Met165, Glu166 and Thr190, 2 π-alkyl bonds (Met49 and Pro168) and 2 π-anion bonds (Glu166) (Figure S3). Ginkgetin made 2 hydrogen bonds (His163 and Asp187), 1 π-anion bond (Glu166), 2 alkyl bonds (Met49 and Pro168), 2 π-alkyl bonds (Met49 and Met165) with Omicron variant Mpro (Figure S4). Notably, the interaction profiles of best natural compounds with both SARS-CoV-2 and Omicron Mpro protein revealed common amino acid residues for hydrogen bonds, hydrophobic contacts, π-stacks, π-alkyl and alkyl contacts including (Met49, Phe140, Leu141, Asn142, Gly143, Ser144, Cys145, His164, Met165, Glu166, Leu167, Pro168, Asp187, Arg188, Gln189, Thr190, Ala191, Gln192).

3.2. DFT calculation of crystal and natural compounds

The geometry optimization of N3, K36, theaflavin and ginkgetin was performed in the gas phase and is given in Table 3 and Figure 4. The 6–311g+(d,p) basis sets achieved significantly better energy of optimization indicating more stability of the compounds’ geometry. The HOMO and LUMO energy calculations were done using the 6–311g(d,p) basis sets (Table 4, Figure S7). Table 4 documents the results of the \( \lambda_{\text{HOMO}} \) and \( \lambda_{\text{LUMO}} \) of N3, K36, theaflavin and ginkgetin. Electronic chemical potential (\( \mu \)) was studied to determine the direction of electron transfer (Figure S7). The \( \mu \) of ginkgetin was −3.829 a.u., an electron-negative value which was comparatively higher than theaflavin (~4.168 a.u.) and N3 (~4.293 a.u.) implying that electron will certainly be transferred from an occupied orbital to an unoccupied orbital of the ginkgetin molecule to form a stable complex upon binding to Mpro targets in comparison to theaflavin, N3 and K36. In addition, the higher electronegativity value (\( \omega \)) of the ginkgetin compared to theaflavin, N3 and K36 highlighted that ginkgetin molecule can act as a strong donor during ligand binding as well as a contender for tight binding throughout Mpro dynamics. A closer look at the bandgap energy suggests that the theaflavin has lesser energy among others which may exhibit better binding capabilities.

3.3. Molecular dynamics simulations of Mpro-ligand complexes

The virtual screening experiment yielded the best-docked complexes which were used to evaluate the conformational stability and time-dependent binding ability of natural ligands in the Mpro catalytic pocket using molecular dynamics simulations [72]. Two pairs of simulations were run by selecting the Mpro target of SARS-CoV-2 and Omicron with best-scored natural compounds together with their respective crystal ligands in re-docked conformations. A total of 6 simulations for...
M\text{pro}\text{-ligand complexes (SARS-CoV-2: M\text{pro}\text{-N3, M\text{pro}\text{-theaflavin, M\text{pro}\text{-ginkgetin; Omicron: M\text{pro}\text{-K36, M\text{pro}\text{-theaflavin and M\text{pro}\text{-ginkgetin) were carried out for a time period of 1 \(\mu\)s using Schrodinger Desmond package. Furthermore, the apo-M\text{pro} protein (PDB ID: 6Y84) was also subjected to simulations to draw a comparison between ligand-bound and unbound conformational changes. The RMSD measure was computed upon the conformations obtained from the simulation trajectories to study the deviations of ligand binding from its initial reference (docked) pose. Figure. S8 and Figure. S9 illustrate the C\text{\alpha}-RMSD profile for all simulated complexes. The ligands, N3, theaflavin and ginkgetin, experienced little fluctuations in the RMSD profile after 300 ns. Further, the structural comparisons of structures captured at 300 ns with a regular interval of 50 ns with respect to final conformation (1 \(\mu\)s) from the simulation trajectory constituted less than 3 \(\AA\) RMSD when C\text{\alpha} atoms were aligned. This indicated that the M\text{pro}\text{-ligand complexes attained stabilization and were used to compute the MM/GBSA binding energy for structures belonging to the time interval between 300 and 1000 \(\mu\)s. The RMSD values of N3, theaflavin and ginkgetin for SARS-CoV-2 M\text{pro} were within the range of 9 \(\AA\). Contrastingly, Omicron M\text{pro} obtained RMSD in the range of 10 \(\AA\) for K36 crystal ligand, theaflavin and ginkgetin. The RMSD difference computed over residue index of SARS-CoV-2 M\text{pro} and Omicron M\text{pro} variant with apo protein was 0.47 \(\AA\) and –0.014 \(\AA\). This indicates apo and SARS-CoV-2 M\text{pro} have distinct dynamic motions compared to the Omicron variant.

The RMSF evaluative measure locates atoms and amino acid residues experiencing large fluctuations throughout the simulation time. Figure. S10 shows the RMSF profile of simulated SARS-CoV-2 M\text{pro} complexes bound to N3, theaflavin and ginkgetin. Whereas the RMSF profile of Omicron M\text{pro} variant with K36, theaflavin and ginkgetin complexes were presented in Figure. S11. The apo protein was also included to gauge the residue index comparison with both SARS-CoV-2 M\text{pro} and Omicron M\text{pro} variant. RMSF plot shows that gink and Theaflavin bound conformation experienced lesser fluctuation than apo and n3 bound form except for the residue window of 48–60 ns. In a comparison of apo and omicron bound conformation, k36, the and gink experienced less variation in relation to apo form. These residues correspond to loop elements and are located in the N terminal region of the M\text{pro} protein which is in good agreement with earlier reports. The radius of gyration (rGyr) measures the ligand’s ‘compactness’ and is equivalent to its critical depiction of stability during the simulation time. Solvent accessible surface area (SASA) is a measure to study the accessibility of solvent molecules (usually water) towards the protein-ligand complex; the less SASA values indicate less exposure to an aqueous environment and the ability of the complex to preserve hydrophobic core thereby enhancing the stability of complex as a whole [73]. Figure. S12 depicts rGyr and SASA properties of N3, theaflavin and ginkgetin for SARS-CoV-2 M\text{pro}. In Theaflavin demonstrates a lower variation than N3 and ginkgetin for rGyr and SASA. As a comparison, the Omicron M\text{pro} variant with K36, theaflavin and ginkgetin complexes were studied which showed similar results for rGyr and SASA (Figure. S13). The intermolecular interactions of N3, theaflavin and ginkgetin were listed in Figure. S14 and Figure. S15 for SARS-CoV-2 M\text{pro} and Omicron M\text{pro}, respectively. We observed more than 5 hydrophobic, 15 hydrogen bonds and water bridges in the three simulated complexes of SARS-CoV-2 M\text{pro}.

Fig. 2. Interaction of K36 in the binding cleft of SARS-CoV-2 Omicron variant Mpro (PDB ID: 7TOB) of COVID-19 shown in (a) 3 D representation and (b) 2 D representation (for better clarity) describing ligands interactions by formation of various H-bonds and hydrophobic interactions with protein at the active site of the protein.
We spotted 5 hydrophobic contacts and >15 hydrogen and water bridges in the simulation trajectory of all Omicron M\textsuperscript{pro} complexes.

3.4. Conservation of intermolecular contacts in molecular dynamics simulations

The SARS-CoV-2 M\textsuperscript{pro} in complex with N3 consisted of hydrophobic interactions in high frequency (Met49, Met165, Leu167, Pro168 and Ala191). It also formed hydrogen bonds with 11 amino acids such as Thr45, Ser46, Asn142, Ser144, Cys145, His164, Glu166, Gln189, Thr190 and Gln192. Residues in the sequence window of Thr24 to His163 and Gln 189 to Gln192 formed water bridges. Various types of intermolecular interactions are detected and are shown in Figure. S16. The 2D interaction plots of N3, theaflavin and ginkgetin illustrate the conservation of interactions along the simulation course. The co-crystal ligand N3 retained almost the entire series of crystal interactions including its alkyl moiety interactions with Glu166 (86%) and Gln189 (55%) residues, His164 (water bridges, 56%), and Gly143 (carboxylate group, 35%) (Figure. S16A). The interactions registered in the simulation trajectory are in line with those of dock poses which provides evidence of the tight binding of ligands. The top-ranked theaflavin established 2 hydrophobic interaction with Phe140 (57%), Met165, 2 negative charge with Glu166 (43%) and Asp187 (90%), 1 negatively charged with Arg188 (55%) and 3 polar contacts with Asn142 (64%), Gln189 (30%) and Thr190 (30%) (Figure. S16B). Ginkgetin lost most of the docking-based interactions in the simulations. One polar and hydrophobic interaction was identified with Glu189 and Met165 in 30% frames (Figure. S16C). The co-crystal ligand K36 of Omicron M\textsuperscript{pro} exhibited hydrophobic contacts with His41 and Met49 residues which were similar to those present in SARS-CoV-2 M\textsuperscript{pro} and additionally interacted with Leu27, Cys44 (46%), Thr45 (31%), Cys145 (77%) and Glu166 (3%) (Figure. S17A). Theaflavin developed hydrophobic interactions with Gly166 (73%) and Asp187 (59%), while 1 hydrogen bond was formed with His164 (54%) (Figure. S17B). However, ginkgetin maintained only one polar bond with His 41 (34%) (Figure. S17C).

All simulated complexes were analyzed to recognize the structural level integrity and conformational changes at every 100 ns time interval. Figure. S18 shows the interaction analysis of SARS-CoV-2 M\textsuperscript{pro}-bound to N3. A total of 7 hydrophobic and 8 hydrogen bonds were noted at the beginning of the simulation. However, a single covalent bond was developed which was present in the interaction profile of ligand N3. After the completion of 1 μs simulation time, residues such as His41, Met49, Pro168 and Ala191 switched to hydrophobic interactions. In comparison to simulation of SARS-CoV-2 M\textsuperscript{pro}-theaflavin complex, His41, Cys145, Met165, Thr190 and Ala191 built new hydrophobic contacts which were absent in the starting structure (Figure. S19). The SARS-CoV-2 M\textsuperscript{pro}-ginkgetin complex developed 9 hydrogen and 7 hydrophobic interactions in its initial conformation but tend to maintain only 4 hydrogen bonds at the end of the simulation (Figure. S20). The comparative study of the Omicron M\textsuperscript{pro}-K36 complex revealed residues such as Gln16, Thr26, Cys44, Ser46, Gly143, Ser144, Cys145 and Glu166 coordinated hydrogen bonds at the completion of simulation (Figure. S21). Theaflavin possessed a huge difference in the interaction list of hydrogen and hydrophobic interactions as shown for SARS-CoV-2 (Figure. S22) and instead created contacts with Ser46, Glu47, Leu50, Cys145, Met165, Pro168, Gly170 and Gln189. Ginkgetin ended with a high number of hydrogen bonds at the end of the simulation. Figure. S23 illustrates the ginkgetin contacts with Arg40, Glu166, Phe185, Val186, Arg188, Thr190 and Gln192 whereas hydrophobic interactions were generated with Cys44, Met49, Met165, Leu167 and Phe181.

3.5. MM/GBSA binding free energy calculation of M\textsuperscript{pro}-ligand complexes

SARS-CoV-2 M\textsuperscript{pro} in complex with N3, theaflavin and ginkgetin, and Omicron M\textsuperscript{pro} in complex with K36, theaflavin and ginkgetin, were used to compute the free energy of ligand binding using the MM/GBSA approach. This was accomplished by performing a post-simulation analysis.
Table-2
The calculated binding energy, hydrogen bonds and contacting receptor residues of co-crystal ligand K36 and top-five natural compounds with SARS-CoV-2 Omicron variant (PDB ID: 7TOB) using YASARA structure.

| Compounds Name | Binding energy [kcal/mol] | Hydrogen bonds | Contacting receptor residues |
|----------------|---------------------------|----------------|-------------------------------|
| K36            | 7.80                      | 8              | Thr25, Thr26, Leu27, His41,  |
|                |                           |                | Met49, Pro52, Tyr54, Phe140,  |
|                |                           |                | Leu141, Am142, Gly143, Ser144,  |
|                |                           |                | Cys145, His163, His164, Met165,  |
|                |                           |                | Glu166, His172, Asp187, Arg188,  |
|                |                           |                | Glu189, Thr190                |
| Theaflavin     | 8.51                      | 4              | Thr25, Leu27, His41, Met49,  |
|                |                           |                | Phe140, Leu141, Am142, Ser144,  |
|                |                           |                | Cys145, His163, His164, Met165,  |
|                |                           |                | Glu166, His172, Asp187, Arg188,  |
|                |                           |                | Glu189, Thr190                |
| Ginkgetin      | 7.80                      | 2              | Thr26, His41, Met49, Phe140,  |
|                |                           |                | Leu141, Am142, Ser144, His164,  |
|                |                           |                | Met165, Glu166, Leu167, Arg188,  |
|                |                           |                | Gln189, Thr190, Ala191         |
| Hesperadin     | 7.27                      | 3              | Thr24, Thr25, Thr26, Leu27,  |
|                |                           |                | His41, Met49, Am142, Gly143,  |
|                |                           |                | Cys145, His164, Met165, Glu166,  |
|                |                           |                | Leu167, Arg188, Gln189, Thr190,  |
|                |                           |                | Ala191, Gln192                 |
| Withanolide_D  | 6.89                      | 2              | Thr25, Thr26, Leu27, His41,  |
|                |                           |                | Met49, Ser144, Cys145, His164,  |
|                |                           |                | Met165, Glu166, Pro168, Arg188,  |
|                |                           |                | Gln189, Thr190, Ala191, Gln192 |
| Psoralidin     | 6.26                      | 3              | Thr25, His41, Met49, Phe140,  |
|                |                           |                | Leu141, Am142, Gly143, Ser144,  |
|                |                           |                | Cys145, His164, Met165, Glu166,  |
|                |                           |                | Leu167, His172, Asp187, Arg188,  |
|                |                           |                | Gln189                        |

The details of geometry optimization (b3lyp method) of ligands.

| Sr No | Name  | Basis set | Optimization Energy (Hartree) | Fig. No. |
|-------|-------|-----------|------------------------------|----------|
|       |       |           | Gas                           | Diff     |
| 1     | N3    | 6-31 g(d,p) | 2299.07566                     | 0        | S7a     |
|       |       | 6-311 g(d,p) | 2299.61296                     | 0.5737   |         |
| 2     | Ginkgetin | 6-31 g(d,p) | 1984.94374                     | 0        | S7b     |
|       |       | 6-311 g(d,p) | 1985.40896                     | 0.46522  |         |
| 3     | Theaflavin | 6-31 g(d,p) | 2022.17087                     | 0        | S7c     |
|       |       | 6-311 g(d,p) | 2022.66853                     | 0.497668 |         |
| 4     | K36   | 6-31 g(d,p) | 1984.4851                      | 0        | S1d     |
|       |       | 6-311 g(d,p) | 1983.6052                      | 0.413189 |         |

enhancement in the intermolecular contacts with pocket residues in addition to Glu166 and Asp187 in the Alpha variant (Figure. S16B) is responsible for tight binding whereas sole interactions with Glu166 and Asp187 residues as noted in Omicron variant (Figure. S17B) do not necessarily indicate better binding. This notion was also reflected in the free energy of binding computed using the MM/GBSA approach. Collectively, the increased number of contacts together with Glu166 and Asp187 residues may be the strategy for developing potent Mpro inhibitory molecules.

4. Discussion

This COVID-19 outbreak is the third devastating global Coronavirus outbreak that has resulted in severe illness and fatalities. CoVs are a sizable virus family that affects a variety of species, including humans. The newly discovered COVID-19 coronavirus, also known as SARS-CoV-2 because of its 89.1% structural similarity to the MERS-CoV and SARS-CoV, has very high transmissibility to spread throughout humans [74]. One of the twelve proteins targeted for CoV regulation, Mpro is one of the most crucial targets due to its significance in genetic transcription and replication. The structures of SARS-CoV2 Mpro co-crystallized with different ligands (N3 and K36) formed the basis for providing crucial intermolecular contacts and aided in developing new compound leads. In this present study, N3 and K36 were used as reference molecules to assess the efficacy of identified natural compounds using a rigorous computational approach with inhibition potential. Further, the native ligands as well as identified compounds were investigated for their potential to acquire similar molecular properties at the DFT level to validate independently the potential to inhibit Mpro. The HOMO-LUMO energy profile gives an overview of the bandgap energy responsible for energy transfer and indicates inhibitory potential matching those of co-crystallized ligands [75]. Some previously developed compounds for SARS-CoV and MERS-CoV include Hesperetin, Calmidazolium, Cinanserin, Aza-peptide epoxides, FL-166, 8c, 2a, and 4o [76, 77].

A potential solution for an accelerated clinical trial of a lead compound to suppress SARS-CoV-2 targets is to identify bioactive compounds with better ADMET profiles by conducting theoretical studies at a high-throughput scale. Arya et al., recently suggested Procainamide, Tetrahydrozoline, and Levamisole interact with SARS-CoV-2 papain-like protease [78]. Theaflavin is a polyphenol abundant in tea and known for its efficacy against lung cancer. Various computational and in vitro studies showed significant inactivation of the SARS-CoV-2 targets using.
natural compounds [79] [-] [85]. A flavone with anti-adipogenic, anti-inflammatory, anti-microbial, and neuroprotective properties is ginkgetin [86]. By inhibiting the cell cycle, inducing apoptosis, encouraging autophagy, and addressing various out-of-control signaling pathways like JAK/STAT and MAPKs, it reduces the growth of cancer. Additionally, it prevents neuron apoptosis, eliminates cerebral microhemorrhage, lowers neurologic impairments, and offers effective neuroprotection against oxidative stress-induced cell death. It also contains anti-inflammatory, anti-microbial, and neuroprotective properties.

Table 4
The HOMO-LUMO details of ligands.

| Sr No. | Ligands | Energy Value (eV) | Energy Gap (eV) | Hardness (\(\eta\)) | Softness (\(S\)) | Chemical potential (\(\mu\)) | Electrophilicity index (\(\omega\)) | Stabilization energy (\(\Delta F\)) |
|--------|---------|------------------|----------------|-------------------|----------------|--------------------------|-----------------------------|-----------------------------|
| 1      | N3      | -6.834           | -1.752         | -5.082            | 2.541          | 0.196773                 | 3.626495                    | -3.626495                   |
| 2      | Theaflavin | -5.267          | -2.391         | -2.876            | 1.438          | 0.347705                 | 5.907789                    | -5.097789                   |
| 3      | Ginkgetin | -6.132          | -2.205         | -3.927            | 1.9635         | 0.254647                 | 4.1685                      | 4.424852                    |
| 4      | K36     | -6.825           | -1.054         | -5.771            | 2.8855         | 0.17328                | 3.9395                      | 2.689255                    |

Table 5
MM/GBSA profiles of N3, Theaflavin and Ginkgetin in interaction with SARS-CoV-2 M\(^{\text{pro}}\).

| Ligands     | \(\Delta G_{\text{bind}}\) (kcal/mol) | \(\Delta G_{\text{bind}}\) (kcal/mol) | \(\Delta H_{\text{bind}}\) (kcal/mol) | \(\Delta S_{\text{bind}}\) (kcal/mol) | \(\Delta G_{\text{bind}}\) (kcal/mol) | \(\Delta G_{\text{bind}}\) (kcal/mol) | \(\Delta H_{\text{bind}}\) (kcal/mol) | \(\Delta S_{\text{bind}}\) (kcal/mol) |
|-------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|
| N3          | -80.75                               | -65.44                               | -29.66                               | 4.89                                 | -2.98                               | -16.55155505                       | 29.85                               |
| Theaflavin   | -71.33                               | -55.52                               | -37.35                               | 9.09                                 | -4.39                               | -13.29233041                       | 33.36                               |
| Ginkgetin    | -58.52                               | -47.86                               | -13.85                               | 2.22                                 | -1.15                               | -14.08843961                       | 19.85                               |

Table 6
MM/GBSA profiles of K36, Theaflavin and Ginkgetin in interaction with the Omicron variant of the SARS-CoV-2 M\(^{\text{pro}}\) protein.

| Ligands     | \(\Delta G_{\text{bind}}\) (kcal/mol) | \(\Delta G_{\text{bind}}\) (kcal/mol) | \(\Delta H_{\text{bind}}\) (kcal/mol) | \(\Delta S_{\text{bind}}\) (kcal/mol) | \(\Delta G_{\text{bind}}\) (kcal/mol) | \(\Delta G_{\text{bind}}\) (kcal/mol) | \(\Delta H_{\text{bind}}\) (kcal/mol) | \(\Delta S_{\text{bind}}\) (kcal/mol) |
|-------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|
| K36         | -19.39                               | -15.41                               | -7.35                                | 0.38                                 | -0.54                               | -3.72                               | 7.36                                |
| Theaflavin   | -36.14                               | -27.11                               | -21.79                               | 3.33                                 | -2.00                               | -8.43                               | 20.87                               |
| Ginkgetin    | -41.54                               | -32.73                               | -11.45                               | 2.47                                 | -0.93                               | -8.93                               | 13.20                               |

Availability of data and material

All data generated or analyzed during this study are included in this published article (and its supplementary Information files).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2022.106318.

References

[1] T. Acter, N. Uddin, J. Das, A. Akhter, T.R. Choudhury, S. Kim, Evolution of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) as coronavirus disease 2019 (COVID-19) pandemic: a global health emergency, Sci. Total Environ. 730 (2020), https://doi.org/10.1016/j.scitotenv.2020.138966.
[2] E. Sifuentes-Rodríguez, D. Palacios-Reyes, Covid-19: the outbreak caused by a new coronavirus, Bol. Med. Hosp. Infant. Mex. 77 (2020) 47–53, https://doi.org/10.12487/BMHIM.200000039.
[3] A.E. Gorbalenya, S.C. Baker, R.S. Baric, R.J. de Groot, C. Drosten, A.A. Gulyaeva, B. Miguel Miguel, J. Vezina, G. Tollefsen, M. Fauque, J. Vezina, 2020.02.07.937862.
[4] F. Garcés-Ayalá, A. Araiza-Rodríguez, Edgar Mendieta-Condado, Abril P. Rodríguez-Maldonado, C. Wong-Arzuaga, M. Landa-Flores, J. Carlos Del Mazo-Esteban-Valencia, L. Lloret-Sanchez, L. Dayanaira, S. Arellano-Suarez, Tatiana E. Nuñez-García, Nervain, B. Contreras-González, N. Cruz-Ortiz, A. Ruiz-López, Miguel A. Fierro-Valdez, D. Regalado-Santigo, N. Martinez-Velazquez, M. Mederos-Michel, J. Vázquez-Pérez, José A. Martínez-Gonzalez, Eduardo Becerril-Ayala.
C. Chakraborty, M. Bhattacharya, S.S. Nandi, R.K. Mohapatra, K. Dhama, J.C. Leao, T.P. de L. Gusmao, A.M. Zarzar, J.C. Leao Filho, A. Barkokebas Santos de C.N. Patel et al. — A Handbook of Phytochemical Constituents of GRAS Herbs and Other Edible Fruits in Combating Emerging Viral Infections, J. Food Biochem. 45 (2021), 10.3389/FPHYAR.2018.01038.

ZINC — A Free Database of Commericially Available Compounds for Virtual Screening | Journal of Chemical Information and Modeling, (n.d.). https://pubs.acs.org/doi/10.1021/acs.jcim.1c00140 acs.jcim.1c00140.

J. A. Duke, Handbook of Phytochemical Constituents of GRAS Herbs and Other Economic Plants, 2017, 10.1080/07391102.2020.1796811.

E. Krieger, T. Darden, S.B. Nabuurs, A. Finkelstein, G. Vriend, Making optimal use of empirical energy functions: force-field parameterization in crystal space, Proteins Struct. Funct. Genet. 57 (2004) 678–683, 10.1002/PROT.20251.

H. Land, M.S. Humble, YASARA, a tool to obtain structural guidance in biocatalytic investigations, Methods Mol. Biol. 1685 (2018) 43–67, 10.1007/978-1-4939-7366-8_4.

E. Krieger, G. Nihor Koraimann, G. Vriend, Increasing the precision of comparative protein modeling applications, J. Comput. Chem. 36 (2015) 996–1008, 10.1002/jcc.24081.

J. A. Duke, Handbook of Phytochemical Constituents of GRAS Herbs and Other Economic Plants, 2017, 10.1080/07391102.2020.1796811.

E. Krieger, T. Darden, S.B. Nabuurs, A. Finkelstein, G. Vriend, Making optimal use of empirical energy functions: force-field parameterization in crystal space, Proteins Struct. Funct. Genet. 57 (2004) 678–683, 10.1002/PROT.20251.

H. Land, M.S. Humble, YASARA, a tool to obtain structural guidance in biocatalytic investigations, Methods Mol. Biol. 1685 (2018) 43–67, 10.1007/978-1-4939-7366-8_4.

E. Krieger, G. Nihor Koraimann, G. Vriend, Increasing the precision of comparative protein modeling applications, J. Comput. Chem. 36 (2015) 996–1008, 10.1002/jcc.24081.

M. S. Ramakrishnan, A.P. Patel, R. Melles, R.A. Vora, Multiple evanescent white dot syndrome: findings from a large northern California cohort, Ophthalmol. Retin. 5 (2021) 850–854, 10.1016/j.orx.2020.11.005.

C. Chakraborty, M. Bhattacharya, S.S. Nandi, R.K. Mohapatra, K. Dhama, J.C. Leao, T.P. de L. Gusmao, A.M. Zarzar, J.C. Leao Filho, A. Barkokebas Santos de C.N. Patel et al. — A Handbook of Phytochemical Constituents of GRAS Herbs and Other Economic Plants, 2017, 10.1080/07391102.2020.1796811.

E. Krieger, G. Vriend, Making optimal use of empirical energy functions: force-field parameterization in crystal space, Proteins Struct. Funct. Genet. 57 (2004) 678–683, 10.1002/PROT.20251.

H. Land, M.S. Humble, YASARA, a tool to obtain structural guidance in biocatalytic investigations, Methods Mol. Biol. 1685 (2018) 43–67, 10.1007/978-1-4939-7366-8_4.

E. Krieger, G. Nihor Koraimann, G. Vriend, Increasing the precision of comparative protein modeling applications, J. Comput. Chem. 36 (2015) 996–1008, 10.1002/jcc.24081.

M. S. Ramakrishnan, A.P. Patel, R. Melles, R.A. Vora, Multiple evanescent white dot syndrome: findings from a large northern California cohort, Ophthalmol. Retin. 5 (2021) 850–854, 10.1016/j.orx.2020.11.005.

C. Chakraborty, M. Bhattacharya, S.S. Nandi, R.K. Mohapatra, K. Dhama, J.C. Leao, T.P. de L. Gusmao, A.M. Zarzar, J.C. Leao Filho, A. Barkokebas Santos de C.N. Patel et al. — A Handbook of Phytochemical Constituents of GRAS Herbs and Other Economic Plants, 2017, 10.1080/07391102.2020.1796811.

E. Krieger, G. Vriend, Making optimal use of empirical energy functions: force-field parameterization in crystal space, Proteins Struct. Funct. Genet. 57 (2004) 678–683, 10.1002/PROT.20251.

H. Land, M.S. Humble, YASARA, a tool to obtain structural guidance in biocatalytic investigations, Methods Mol. Biol. 1685 (2018) 43–67, 10.1007/978-1-4939-7366-8_4.

E. Krieger, G. Nihor Koraimann, G. Vriend, Increasing the precision of comparative protein modeling applications, J. Comput. Chem. 36 (2015) 996–1008, 10.1002/jcc.24081.
H. Sun, Y. Li, S. Tian, L. Xu, T. Hou, Assessing the performance of MM/PBSA and MM/GBSA methods. 1. Accuracies of MM/PBSA and MM/GBSA methodologies based on a combination of molecular mechanics and continuum models, Curr. Comput. Aided Drug Des. 2 (2006) 287–293.

J. Wang, T. Hou, X. Xu, Recent advances in free energy calculations with a combination of molecular mechanics and continuum models, Curr. Comput. Aided Drug Des. 2 (2006) 287–293.

T. Zhou, W. Lu, W. Wu, L. Guo, W. Yuan, W. Hu, Y. Li, Y. Liu, Q. Li, E. Fang, L. Jiang, W. Zhang, J. Zhang, Y. Zhang, L. Wang, Z. Yang, Y. Liu, T. Hou, A. Zhong, Q. Li, Y. Wang, C. Tian, H. Sun, Development and application of the MMFF14SB force field for protein-ligand and protein-protein interactions, J. Chem. Theory Comput. 18 (2022) 655–667.

J. Zhang, Z. Zhang, L. Zhang, W. Hu, W. Li, Y. Li, Q. Li, J. Wang, Y. Li, T. Hou, A. Zhong, Q. Li, Y. Wang, C. Tian, H. Sun, Development of the MMFF14SB force field for protein-ligand and protein-protein interactions, J. Chem. Theory Comput. 18 (2022) 655–667.

L. Zhang, Y. He, Y. Xiong, Z. Xu, Z. Liu, J. Liao, Y. Li, Y. Wang, X. Guo, Q. Li, J. Tang, Y. Li, J. Wang, T. Hou, A. Zhong, Q. Li, Y. Wang, C. Tian, H. Sun, Development and application of the MMFF14SB force field for protein-ligand and protein-protein interactions, J. Chem. Theory Comput. 18 (2022) 655–667.

Y. Zhang, S. Liu, J. Li, Y. Li, Y. Wang, X. Guo, Q. Li, J. Tang, Y. Li, J. Wang, T. Hou, A. Zhong, Q. Li, Y. Wang, C. Tian, H. Sun, Development and application of the MMFF14SB force field for protein-ligand and protein-protein interactions, J. Chem. Theory Comput. 18 (2022) 655–667.