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.
Identification of bioactive molecules from Triphala (Ayurvedic herbal formulation) as potential inhibitors of SARS-CoV-2 main protease (Mpro) through computational investigations

Mithun Rudrapal a,⇑, Ismail Celik b, Johra Khan c,d,⇑, Mohammad Azam Ansari e, Mohammad N. Alomary f, Fuad Abdullah Alatawi p, Rohitash Yadav g, Tripti Sharma h, Trina Ekawati Tallei i,j, Praveen Kumar Pasalak, Sanjay G. Walode a

⇑ Corresponding authors at: Department of Pharmaceutical Chemistry, Rasiklal M. Dhariwal Institute of Pharmaceutical Education & Research, Pune 411019, Maharashtra, India (M. Rudrapal); Department of Medical Laboratory Sciences, College of Applied Medical Sciences, Majmaah University, Al Majmaah 11952, Saudi Arabia (J. Khan).

E-mail addresses: rsmrpal@gmail.com (M. Rudrapal), j.khan@mu.edu.sa (J. Khan), falatawi@ut.edu.sa (F.A. Alatawi).

Peer review under responsibility of King Saud University.

Article history:
Received 4 October 2021
Revised 20 December 2021
Accepted 5 January 2022
Available online 10 January 2022

Keywords:
SARS-CoV-2
COVID-19
Triphala
Bioactive molecules
Molecular docking
Molecular dynamics simulation

ABSTRACT

Severe acute respiratory syndrome coronavirus disease (SARS-CoV-2) induced coronavirus disease 2019 (COVID-19) pandemic is the present worldwide health emergency. The global scientific community faces a significant challenge in developing targeted therapies to combat the SARS-CoV-2 infection. Computational approaches have been critical for identifying potential SARS-CoV-2 inhibitors in the face of limited resources and in this time of crisis. Main protease (Mpro) is an intriguing drug target because it processes the polyproteins required for SARS-CoV-2 replication. The application of Ayurvedic knowledge from traditional Indian systems of medicine may be a promising strategy to develop potential inhibitor for different target proteins of SARS-CoV-2. With this endeavor, we docked bioactive molecules from Triphala, an Ayurvedic formulation, against Mpro followed by molecular dynamics (MD) simulation (100 ns) to investigate their inhibitory potential against SARS-CoV-2. The top four best docked molecules (terflavin A, chebulagic acid, chebulinic acid, and corilagin) were selected for MD simulation study and the results obtained were compared to native ligand X77. From docking and MD simulation studies, the selected molecules showed promising binding affinity with the formation of stable complexes at the active binding pocket of Mpro and exhibited negative binding energy during MM-PBSA calculations, indicating their strong binding affinity with the target protein. The identified bioactive molecules were
1. Introduction

Coronavirus disease 2019 (COVID-19) is a highly infectious disease with a high rate of reproducibility. Severe acute respiratory syndrome coronavirus disease-2 (SARS-CoV-2) is the causative agent for COVID-19 (Rudrapal et al., 2020). This deadly disease has infected nearly every country on the planet, with 266,946,588 confirmed cases and 5,282,046 fatalities, according to the latest reports of World Health Organization (https://covid19.who). With the majority of the world under lockdown and the alarming threat of thousands of deaths every day, rapid drug discovery for the identification of anti-COVID drugs is an urgent need of the hour. Conventional drug discovery processes could take years. To aid in the rapid discovery of new drugs, computational approaches have been effective in identifying potential drugs/medications for COVID-19 therapy (Rudrapal et al., 2021a). The use of computational algorithms such as molecular docking, molecular dynamics (MD) simulation and density functional theory (DFT) method has made the search for drugs of natural or synthetic origin that target the viral system, particularly SARS-CoV-2 main protease more feasible while causing fewer side effects with improved therapeutic performance in human populations. In this context, repurposing the traditional herbal remedies such as Ayurvedic formulations can play a significant role in the development of effective therapeutic entities. Moreover, the use of phytochemicals over synthetic drugs are advantageous since they are safer and a higher dose is well tolerated.

Ayurvedic formulation *Triphala*, which literary means *Tri* (three) and *Phala* (fruit) in Sanskrit), is a ‘tridosic rasayan’ that balances and rejuvenates the “tridosha”, i.e., vata, pitta, and kapha, as well as promotes immunity and health. It is a polyherbal formulation made up of a proportion al amount of dried fruit powder from three plants: *Phyllanthus emblica* (Indian gooseberry or Amalaki), *Terminalia bellirica* (Bibhitaki), and *Terminalia chebula* (Haritaki) (Peterson et al., 2017). In the clinical practice of Ayurveda, the formulation (Triphala) is consumed as a ‘health tonic’. It is also used to treat digestive problems such as gastric acidity, constipation and poor food assimilation. Additionally, several diseases such as fever, cough, jaundice, anemia, asthma, chronic ulcers, pyorrhea and leucorrhoea are also treated with *Triphala*. Additionally, obesity, inflammation, fatigue, cardiovascular disorders, liver dysfunction and ophthalmic problems are some other health ailments that *Triphala* is being recommended for. Moreover, it demonstrates cardiotoxic properties by lowering serum cholesterol levels, reducing myocardial necrosis, enhancing blood circulation, and strengthening capillaries. Flavonoids, tannins, gallic acid, pectin, vitamin C, corilagin, and phyllaemblinc compounds are present in *P. emblica* fruits (Habit-ur-Rehman et al., 2007). *Terminalia bellirica* extract contains β-sitosterol, tannins and phenolic acids such as elagic acid, gallic acid, galloyl glucose, ethyl gallate, and chebulaginic acid (Kumar and Khurana, 2018). The main bioactive active constituents present in *T. chebula* are chebulic, chebulinic acid, 2,4-chebulyl-D-glucopyranose, terflavin A, terchelin, elagic acid, gallic acid, tannic acid, ethyl gallate, arjunigenin, arjunglucoside I, punicalagin and luteloin.

Recently, computational techniques are being extensively used to investigate the inhibitory potential of phytochemicals against SARS-CoV-2 (Singh et al., 2021; Sharma et al., 2021; Bhardwaj et al., 2021). Numerous studies have been conducted to determine the efficacy of traditional medicine as anti-SARS-CoV-2 agents (Verma et al., 2020; Chakraborty et al., 2021; Dutta et al., 2021). Furthermore, effective design of SARS-CoV-2 inhibitors can be accomplished through the selection of druggable targets. Main protease (M<sub>pro</sub>) or 3-chymotrypsin-like protease 3 (CLpro) cleaves non-structural proteins (Nsp) at eleven different sites and leads to the formation of Nsp4-Nsp16, which is essential for viral replication (Yang et al., 2005). The S protein and Mpro are important targets for designing SARS-CoV-2 inhibitors (Idroes et al., 2021). No adverse toxic effect is likely to be observed by inhibiting Mpro as there are no known similar enzymes in humans with similar cleavage specificity (Khater and Nassar, 2021). Papain-like protease (PL<sub>pro</sub>) is another enzyme similar to M<sub>pro</sub> (Khan et al., 2021a); it generates Nsp1-Nsp3 by breaking at various sites, thereby generating essential proteins that suppress the immune system by speeding up the infection process in the host cell (Harcourt et al., 2004). One of the important targets is RNA-dependent RNA polymerase (RdRp) as the viral genome depends on it for its replication and transcription (Borgio et al., 2020; Rudrapal et al., 2021b). In literature (Ahmad et al., 2021), the possible potential of Indian medicinal plants and formulations against SARS-CoV-2/COVID-19 has been mentioned based upon the recommendation (https://www.ayush.gov.in/docs/ayush-Protocol-covid-19.pdf; https://www.ayush.gov.in/docs/ayurved-guidelines.pdf) of AYUSH (Ayurveda, Yoga, Naturopathy, Unani, Siddha, Sowa-Rigpa and Homoeopathy). The usefulness of *Triphala* formulation and other Ayurvedic herbal remedies as therapeutic supplements in COVID-19 has also been reported in some recent literature (Ozah, 2020). Khan et al. (2021b) reported the inhibitory potential of *Triphala* against spike glycoprotein of SARS-CoV-2 by molecular docking. In view of these information, in the urge to design new SARS-CoV-2 inhibitors by blockage of its main protease (M<sub>pro</sub>), we performed structure-based molecular docking studies of bioactive molecules from *Triphala* formulation. The selected candidates were subjected to MD simulations for confirmation of the SARS-CoV-2 inhibitory potential of bioactive molecules. Further, *in silico* screening of the selected compounds was conducted with the application of Lipinski’s filters to determine their drug-likeness and ADMET properties.

2. Materials and methods

2.1. Protein preparation

The crystal structure of the Mpro of SARS-CoV-2 bound with non-covalent inhibitor X77 (N-(4-tert-butylyphenyl)-N-[(1R)-2-(cy clohexylamino)-2-oxo-1-(pyridin-3-yl) ethyl]-1H-imidazole-4-car boxamide) at a resolution of 2.1 Å (PDB id: 6W63) (Rudrapal et al., 2021b) was retrieved from RCSB Protein data bank (https://www.rcsb.org/). The maestro’s protein preparation wizard (version 2021-2 Schrodinger suite) was used to prepare proteins by removing all water within 5 Å distance from the ligand, filling amino acid side chains, adding missing hydrogen atoms, creating disulfide bonds, and optimizing hydrogen bonds. For restrained minimization, OPLS4 force field was used until the average RMSD (root-mean-square deviation) of the non-hydrogen atom converged to
0.30 Å, allowing heavy atoms to move sufficiently to release strained bonds, angles, and clashes (Roos et al., 2019; Kumar et al., 2021).

2.2. Ligand preparation

The Triphala formulation’s bioactive molecules were chosen from the literature (Prasad and Srivastava, 2020), and their chemical structures were obtained from a popular online database, PubChem (https://pubchem.ncbi.nlm.nih.gov/). LigPrep was used to generate the ligands’ three-dimensional (3D) coordinates (Chen and Fopolpe, 2010). Ionization state (pH 7.4), tautomer, and chirality were predicted by Epik module. The OPLS4 force field was employed for energy minimization until an energetically stable confirmation was obtained (Roos et al., 2019).

2.3. Molecular docking study

The selected set of phytoconstituents was subjected for molecular docking against Mpro with the Extra Precision (XP) protocol of the glide docking module of Maestro (Friesner et al., 2006). The glide-receptor grid generation module was used with the default settings to identify the binding site, and the centroid of the bound ligand was taken into account for grid generation (Halgren et al., 2004). The single best pose for each molecule was saved and the compounds were then classified based on their docking scores, which represent free binding energies (Othman et al., 2021).

2.4. Molecular dynamics simulations

Following that, the top four compounds with the best binding free energies, namely terflavin A, corilagin, chebulagic acid, and chebulinic acid, were selected to further ensure the flexibility and stability at the binding site of M^{pro}. A 100 ns MD simulation was performed using GROdin-gen MMachine for Chemical Simulation (GROMACS) 2019.2 software (Abraham et al., 2015). The topologies of the protein and ligand were determined using the ‘pdb2gmx’ script and the GlycoBioChem PRODRG2 server (http://davapc1.bioch.dundee.ac.uk/cgi-bin/prodrg), respectively. Energy minimization was accomplished using the GROMOS96 54a7 force field. The complex was solvated using the SCP water model, electrically neutralized by adding sodium ions and chloride ions using ‘gmxgenion’ script, equilibrated under an 0.3 ns isothermal-isochoric (NVT) and 0.3 ns isothermal-isobaric (NPT) ensembles and subjected to MD simulations (100 ns) (Ghosh et al., 2021). The pressure and temperature of the system were maintained at 1 atm and 300 K, respectively. The leap frog algorithm was used to integrate the motion equation in triclinic at a time step of 2 fs. Finally, the protein–ligand complexes were put to generate trajectories of RMSD, RMSF, Rg and H bond by GROMACS scripts. The trajectories were analyzed and visualized by VMD 1.9.3 (Humphrey et al., 1996).

2.5. MM-PBSA calculations

The binding free energies of protein–ligand complexes were quantified by the Molecular Mechanics Poisson Boltzmann Surface Area (MM-PBSA). This technique allows us to study the biomolecular interactions between the target and bioactive molecules. The g_mmpbsa script was utilized for MM-PBSA binding free energy calculations (Kumari et al., 2014). Van der Waals, polar solvation, and electrostatic contributions to energy were calculated. Solvent accessible surface area (SASA) was used to calculate non-polar energy contribution.

2.6. Drug-likeness, ADMET and toxicity prediction

The selected bioactive molecules were further checked for drug-likeness properties using Lipinski’s rule of five (RO5) (Hussain et al., 2021; Rudrapal and Sowmya, 2019). The ADMET properties including absorption, distribution, metabolism, excretion and toxicity of the selected were determined by the admetSAR server (Cheng et al., 2012; Ghost et al., 2021). Oral in silico toxicity prediction was performed using web-based server ProTox-II (Banerjee et al., 2018).

3. Results and discussion

3.1. Molecular docking

Triphala (Ayurvedic herbal formulation) is made up of dried fruits from three different herbal plants in equal parts (Jantraprom et al., 2021; Tarasiuk et al., 2018). Following a thorough review of the literature on the three plants T. chebula, T. bellica, and P. emblica, 19 bioactive molecules were chosen for molecular docking against the M^{pro} of SARS-CoV-2. The binding pocket around the inhibitor X77 of the target protein M^{pro} was used as a reference for determining the grid box. Docking analysis generated many configurations, which were scored to determine the most favorable binding modes. The bioactive molecules docked against M^{pro} of SARS-CoV-2 exhibited excellent binding, as evidenced by high docking scores that exceeded X77. The validation of the docking method was carried out by redocking of co-crystal ligand, X77 in the original ligand binding site of Mpro. The redocking was successful. The measured RMSD value of redocking was 0.4493 Å between the experimental docking pose of X77 and the natural binding mode of the native ligand. The superimposed conformations were obtained from a popular online database, PubChem from the literature (Prasad and Srivastava, 2020), and their chemistry was performed using web-based server ProTox-II (Banerjee et al., 2018). Following that, the top four compounds with the best binding free energies are given in Table 1.

| Compounds               | XPG Score (kcal/mol) | Glide emodel (kcal/mol) | MMGBSA ΔG Bind (kcal/mol) |
|-------------------------|----------------------|-------------------------|---------------------------|
| Terflavin A             | −14.521              | −123.949                | −86.36                    |
| Corilagin               | −12.845              | −87.203                 | −68.70                    |
| Chebulagic acid         | −11.359              | −58.079                 | −55.83                    |
| Chebulinic acid         | −11.217              | −137.224                | −64.05                    |
| Phyllaemblicin B        | −10.237              | −85.229                 | −52.22                    |
| Punicalagin             | −9.388               | −75.623                 | −70.60                    |
| Bellericoside           | −9.290               | −65.127                 | −54.02                    |
| Emblicanin B            | −8.416               | −65.211                 | −54.04                    |
| Quercetin               | −7.641               | −52.406                 | −43.32                    |
| Gallolyd glucose        | −7.625               | −56.523                 | −49.99                    |
| Epicatechin             | −7.548               | −48.274                 | −46.39                    |
| Phyllaemblicin acid     | −7.510               | −48.331                 | −37.68                    |
| Arjungluicoside I       | −7.220               | −53.970                 | −35.40                    |
| Luteolin                | −7.186               | −45.854                 | −41.32                    |
| Ellagic acid            | −7.104               | −46.060                 | −43.04                    |
| Terchebin               | −6.850               | −92.637                 | −61.45                    |
| Belleric acid           | −6.595               | −45.484                 | −46.38                    |
| Emblicanin A            | −6.502               | −109.354                | −55.77                    |
| β-Sitosterol            | −6.277               | −46.520                 | −51.84                    |
| Apigenin                | −6.072               | −44.327                 | −40.62                    |
| Gallic acid             | −5.930               | −32.277                 | −28.29                    |
| Methyl gallate          | −5.916               | −37.714                 | −35.54                    |
| Arjunglenin             | −5.862               | −47.716                 | −40.94                    |
| Ascorbic acid           | −5.445               | −27.964                 | −21.89                    |
| Syringic acid           | −5.380               | −34.749                 | −31.81                    |
| Ethyl gallate           | −4.332               | −38.651                 | −35.41                    |
| Phyllaemblicin B        | −10.237              | −85.229                 | −52.22                    |
| X77                     | −5.966               | −5.966                  | −85.045                   |
Terflavin A achieved the highest docking score, with a value of $-14,521 \text{ kcal/mol}$. Corilagin, chebulagic acid, chebulinic acid and X77 had XP Gscore values of $-12,845 \text{ kcal/mol}$, $-11,359 \text{ kcal/mol}$, and $-5,966 \text{ kcal/mol}$, respectively. Chemical structures of the four best-docked compounds and the native ligand X77 are depicted in Fig. 1. Chebulagic acid has a lower XP Gscore value than chebulinic acid, but the binding energy value of chebulic acid is better ($-64.05 \text{ kcal/mol}$). Furthermore, despite having a higher XP Gscore, the dG bind value of the X77 ligand is $-85,045 \text{ kcal/mol}$, lower than corilagin, chebulagic acid, and chebulinic acid. The interactions with Mpro of the top four docked compounds terflavin A, chebulagic acid, chebulinic acid, corilagin, including the native ligand X77 are visualized in Fig. 3.

Terflavin A, chebulagic acid, chebulinic acid, corilagin are types of hydrolysable tannins, which are typically found on *Terminalia* species leaves (L.). Terflavin A and corilagin are ellagitannin, while chebulagic acid and chebulinic acid are benzopyran tannin. Tannins can inhibit SARS-CoV-2 through binding with catalytic dyad residues of Mpro (Chang et al., 2019; Das et al., 2020). Tannic acid even inhibits both Mpro and TMPRSS2 protease, effectively suppressing SARS-CoV-2 (Wang et al., 2020). The catalytic residues of 3CLpro (or Mpro in the case of SARS-CoV-2) are Cys145 and His41, which are surrounded by other residues that confer substrate specificity (Rudrapal et al., 2021b; Singh et al., 2020).

Terflavin A interacts with Mpro on its active site by H-bonds at residues such as Asn142 (2.10 Å), Gly143 (2.13 Å), Ser144 (2.50 Å), Cys145 (2.25 Å), Glu166 (1.84 Å), Arg188 (2.13 Å), Gln189 (2.13 Å), and Thr190 (2.50 Å) and then H-carbon bonds at Met165 (2.06 Å) and pi-sulphur interactions at Cys44 (2.14 Å) and Met49 (2.43 Å) residues. Terflavin A exhibits H-bond with Cys145 (1.96 Å), but has no interaction with His41. Chebulagic acid forms H-bonds with residues Asn142 (1.67), Gly143 (2.18 Å), Glu166 (2.01 Å), Pro168 (2.97 Å), Arg188 (2.16 Å), and Gln189 (1.67 Å), H-carbon bonds in Met49 (1.90 Å) and Asp187 (1.61 Å) and pi-alkyl interactions in Met165 (1.72 Å). While the binding of free energy value is high, chebulagic acid does not interact with Cys145 or His41. Chebulinic acid forms H-bonds on the active site, namely residues Thr26 (1.89 Å), Phe140 (2.07 Å), Asn142 (2.18 Å), Cys145 (2.16 Å), Glu166 (2.26 Å), Gln189 (2.27 Å) and Thr190 (1.52 Å) as well as pi-alkyl, pi-pi stacked and pi-sigma interactions on Pro168 (2.04 Å), His41 (2.19 Å), and Gln189 (2.67 Å), respectively. Chebulinic acid, like Terflavin A, has only one H-bond with Cys145 (2.08 Å) and no interaction with His41. Corilagin forms H-bonds at residues Thr25 (1.87 Å), His41 (2.11 Å), Asn142 (1.03 Å), Cys145 (1.66 Å), and Glu166 (1.78 Å) then H-carbon bonds at Met165 (2.06 Å) and pi-alkyl and pi-sulphur interactions at residues Met49 (2.13 Å) and Cys145 (2.12 Å), respectively. As evidenced, H-bond interactions exist between corilagin and Cys145 (1.72 Å) and His41 (1.08 Å). As a native ligand, the X77 forms H-bonds at residues Asn142 (2.16 Å), Gly143 (1.89 Å), His163 (1.67 Å), and Glu166 (2.05 Å), then H-carbon bonds at Thr26 (2.10 Å) and Met165 (2.90) and pi-pi stacked and pi-sulphur interactions at residues His41 (2.17 Å), Met49 (2.04 Å), and Cys145 (1.14 Å), respectively. These findings support previous findings that the selected compounds bind to the catalytic residues of Mpro (Silva et al., 2021). The substrate-binding cleft contains a catalytic dyad composed of His41 and Cys145, which is located between domain I and II at
the N-terminus, and a hydrogen-bonded water molecule bound to His41 acts as a third component of the catalytic triad (Tan et al., 2005).

3.2. Molecular dynamics simulation

MD simulations and other related techniques are on the verge to act as routine techniques for the discovery of drugs (De Vivo et al., 2016). In this study, the stability and dynamic behavior of the four ligand–protein (bound) complexes were verified by MD simulations for a 100 ns time. The RMSD, root mean square fluctuation (RMSF), radius of gyration (Rg) and solvent accessible surface area (SASA) graphs were plotted of the apoprotein, co-crystal ligand complex and four ligand–protein complexes (Junejo et al., 2021). The RMSD analysis provides information insight into the structural deviation and conformational stability and convergence during the time of the simulation. Thereby, in an equilibrated, stable system, the RMSD towards the end of MD simulations should be within 1–3 Å for a globular protein. The time-dependent RMSD graph was plotted for the backbone of the apoprotein, co-crystal ligand–protein complex and ligand (terflavin A, corilagin, chebulagic acid, chebulinic acid) protein (M\textsuperscript{pro}) complex. Fig. 4 (A) shows the RMSD values of apoprotein and co-crystal ligand–protein complex fluctuate in the beginning from 0.2 nm and at the end at around 0.3 nm, confirming the stability of the protein during 100 ns of simulation. At first, the RMSD values of the apoprotein and the protein–ligand complex are different, but after 50 nm, the RMSD values become stable and remain stable until 100 ns. Fig. 4 (B) shows the RMSD values of protein (M\textsuperscript{pro}) and ligand (terflavin A, corilagin, chebulagic acid, chebulinic acid) complex, indicated with colour code including M\textsuperscript{pro}-TerA (green), M\textsuperscript{pro}-Cor (pink), M\textsuperscript{pro}-Cgic (cyan), and M\textsuperscript{pro}-Cnic (indigo). The RMSD of M\textsuperscript{pro}-TerA fluctuates at the beginning until 70 ns (0.2 nm to ~ 0.6 nm) and at 100 ns, it fluctuates at around 0.3 nm. The RMSD values of both the Protein-Corilagin (M\textsuperscript{pro}-Cor) and Protein-Chebulagic acid (M\textsuperscript{pro}-Cgic) complex are stable from the beginning (0.2 nm), and stay stable until 100 nm (0.25 nm), confirming the stability of both protein–ligand complexes during 100 ns of simulation. Whereas, Protein-Chebulinic acid complex fluctuates (between 0.2 nm and 0.3 nm) up to 70 ns, but after 70 nm, it becomes stable and stays stable (0.2 nm) until 100 nm. The overall RMSD results reveal that the binding of the ligands (corilagin, chebulagic acid, chebulinic acid) in the active site of M\textsuperscript{pro} is stable.

The RMSF study gives data regarding the amino acid residues fluctuation from the reference structure. Fig. 4 (C) shows that the co-crystal ligand fluctuation is almost similar to apoprotein, whereas Fig. 4 (D) shows the fluctuation of the ligand's binding to active site residues Thr24, Thr25, His41, Cys44, Met49, Tyr54, Phe140, Asn142, Gly143, Cys145, His163, His164, Met165, Glu166, Leu167, Pro168, Asp187, Arg188, Gln189, and Thr190 in each of M\textsuperscript{pro}-TerA, M\textsuperscript{pro}-Cor, M\textsuperscript{pro}-Cgic, and M\textsuperscript{pro}-Cnic indicates that there is nearly similar RMSF fluctuation compared to apoprotein and co-crystal ligand–protein residues. Thus, analysis of the RMSF suggested a stable protein–ligand complex since there are no significant changes at the active site’s reference location after successful binding by the apoprotein.

Further, the ‘sasa’ function was used to calculate the SASA to examine the accessibility of solvent molecules on the protein surface. The SASA value of apoprotein, co-crystal ligand–protein complex, and protein–ligand complexes M\textsuperscript{pro}-TerA, M\textsuperscript{pro}-Cor, M\textsuperscript{pro}-Cgic, and Mpro-Cnic were stable during 100 ns MD. Thus, it indicates that ligand–protein complexes have comparable stability to apoprotein and co-crystal ligand–protein complexes [Fig. 4 (E)].

During MD simulations, a study on the radius of gyration (Rg) indicates the protein compactness. Rg of the C\alpha atoms of the Mpro SARS-CoV-2 protein was computed to ascertain conformational changes in it. The Rg graph was plotted for protein using 100 ns trajectories as represented in Fig. 4 (F). The Rg plot indicates that the Rg value apoprotein (mean: 2.20 nm) and protein–ligand complex (mean: 1.23 nm) are nearly identical. Finally, Rg results suggest that all secondary structure of protein complexes is closely packed during 100 ns MD simulation.
Previous literature has highlighted the potential impact of medicinal plants and their phytoconstituents against SARS-CoV-2 (Rudrapal et al., 2021b). Studies have reported that thioflavonoids and epicatechin-3-O-gallate interacted with (binding sites) Mpro and PLpro, and were considered as SARS-CoV-2 inhibitors (Gogoi et al., 2021; Ismail et al., 2021). Thus, bioactive molecules bind to or interact with the active residues which are present at the binding cavity of Mpro and PLpro.

3.3. MM-PBSA binding free energy (BFE)

Python script MmPbSaStat.py provided in the g_mmpbsa was utilized to calculate average BFE and its standard deviation error. The energy released during the interaction of protein–ligand is represented in the form of BFE. It is a cumulative sum of Van der Waals, polar salvation, electrostatic, and SASA energy. Precisely, the MM-PBSA method estimates the binding free energy of a ligand–protein complex as the difference between the free energy of the complex and the free energies of the unbound components (receptor and ligand). More negative the BFE, the more effectively the ligand binds to the protein (Table 2). The results indicate that all the forms of energy contributed fairly to the interactions between Mpro and the selected molecules. Mpro-TerA complex showed the least binding BFE (−233.782 kJ/mol) among the selected bioactive molecules followed by Mpro-Cgic (−221.916 kJ/mol). The binding free energy (BFE) of the X77 (−160.625 kJ/mol) is less negative compared to the selected bioactive molecules. A comparison of the BFE of all the complexes reveals that, except for corilagin, all of the other selected bioactive molecules outperform X77 in inhibiting the protein Mpro of SARS-CoV-2.

3.4. Drug-likeness, ADMET and toxicity properties

Lipinski’s rule of five (Ro5) is a set of rules used to assess the drug-likeness of a chemical compound that has a pharmacological activity that could lead to it being an orally active drug for human use (Lipinski et al., 2001; Kalita et al., 2020). The Ro5 specifies that the molecular weight must be less than 500 Da, the number of H-bond donors must be less than 5, the number of H-bond acceptor must be less than 10, the xlogP must be less than 5, and the molar

Fig. 3. Interaction between the ligands and main protease (Mpro) (PDB: 6W63). (a) Terflavin, (b) Chebulagic acid, (c) Chebulinic acid, (d) Corilagin, (e) X77 or N-(4-tert-butylphenyl)-N-[(1R)-2-(cyclohexylamino)-2-oxo-1-(pyridin-3-yl)ethyl]-1H-imidazole-4-carboxamide.

M. Rudrapal, I. Celik, J. Khan et al. Journal of King Saud University – Science 34 (2022) 101826
refractivity must be between 40 and 130 (Fatimawali et al., 2021; Kashyap et al., 2016; Sefren Geiner et al., 2020). Regarding this statement, Table 3 indicates that only compound X77 appears to meet the Ro5 criteria. However, strict adherence to Ro5 is not required for naturally occurring compounds or phytocomponents. Lipinski specifically states that the natural products and substrates of biological transporters are not covered by Ro5 (Lipinski et al., 2001; Rudrapal et al., 2017; Rudrapal et al., 2018). It is, therefore

Table 2

| Complexes | ΔE_{binding} (kJ/mol) | SASA (kJ/mol) | ΔE_{polar solvation} (kJ/mol) | ΔE_{electrostatic} (kJ/mol) | ΔE_{van der Waals} (kJ/mol) |
|-----------|-----------------------|---------------|-------------------------------|-----------------------------|-----------------------------|
| Terflavin A | -233.782 ± 19.600 | -28.158 ± 1.819 | 201.507 ± 22.760 | -41.154 ± 8.852 | -356.977 ± 16.074 |
| Corilagin | -134.539 ± 13.270 | -18.924 ± 1.200 | 81.964 ± 16.865 | -5.503 ± 7.878 | -192.076 ± 12.655 |
| Chebulagic acid | -221.916 ± 17.867 | 27.490 ± 1.279 | 171.064 ± 16.916 | -34.604 ± 11.949 | -330.887 ± 15.585 |
| Chebulinic acid | -215.150 ± 20.128 | 29.860 ± 1.612 | 223.952 ± 19.289 | -76.837 ± 13.932 | 332.405 ± 19.549 |
| X77 | -160.625 ± 15.840 | -20.961 ± 1.249 | 123.945 ± 13.577 | -29.776 ± 6.919 | -233.832 ± 13.802 |

Fig. 4. (a) Root-mean-square deviation (RMSD) analysis of apoprotein, and co-crystal ligand–protein complex; (b) Root-mean-square deviation (RMSD) analysis of ligand–protein complexes (M^{TerA}, M^{Cor}, M^{Cgic}, and M^{Cnic}); (c) Root-mean-square fluctuation (RMSF) analysis of apoprotein, and co-crystal ligand–protein complex; (d) Root-mean-square fluctuation (RMSF) analysis of ligand–protein complexes (M^{TerA}, M^{Cor}, M^{Cgic}, and M^{Cnic}); (e) SASA analysis and (f) Radius of gyration (Rg) of the apoprotein, co-crystal ligand–protein complex and protein–ligand complexes (M^{TerA}, M^{Cor}, M^{Cgic}, and M^{Cnic}).
important to note here that the desired pharmacological efficacy of Ayurvedic/herbal formulations is possibility due to the multi-component nature (different components) of such formulations.

On its way to the target site, the drug candidate undergoes pharmacokinetic events such as absorption, distribution, metabolism, and excretion (ADME). ADME parameters, in conjunction with toxicity, are critical for the discovery and development of potential drug candidates (Guan et al., 2019). Not only should a high-quality drug candidate be efficacious against the therapeutic target, but it should also exhibit appropriate ADME properties at the dosage for therapeutic purposes. Based on the results of the admetSAR and ProTox analyses (Table 4), all of the compounds were from any signs of toxicity such as hepatotoxic, carcinogenic, or mutagenic. The acute oral toxicity analysis reveals that all compounds are classified as category III (500 mg/kg LD50 5000 mg/kg), indicating a moderate level of toxicity (Herkenne et al., 2008; Li et al., 2014). Typically, animal experiments are used to determine a compound’s toxicology, carcinogenicity, and mutagenicity. However, this takes a long time and necessitates the sacrifice of animals. Thus, prediction of toxicity within in silico technique is a faster and less expensive alternative, as it is based on existing data on a compound’s toxicity. As a result, it is used to develop models capable of predicting the toxicity, carcinogenicity, and mutagenicity of novel compounds.

Bioavailability is a pharmacological term that refers to the level and amount to which a bioactive compound can be absorbed and made available (Kim et al., 2014). A compound’s low oral bioavailability is a frequent issue during the drug design process. As a result, a drug may fail in clinical trials despite its high efficacy in vitro and in vivo experiments (Fukunishi et al., 2014). A compound is predicted to have low oral bioavailability if its bioavailability score is lower than 0.5. If the score is greater than 0.5, the compound is expected to have a high oral bioavailability (Rath et al., 2021). According to the findings, all compounds, except for X77, have a high oral bioavailability.

If, on the other hand, the LD50 value of a compound is between 1,000 and 5,000 mg/kg, it is considered to be less or slightly toxic. Thus, chebulagic acid, chebulinic acid, and X77 are classified as less toxic. Whereas corilagin and terflavin A are classified as moderately toxic.

4. Conclusions

To develop drugs from ancient Indian knowledge against SARS-CoV-2, we made an attempt to analyze the binding modes of the phytoconstituents of Triphala formulation by docking against the active site of Mpro. Terflavin A, corilagin, chebulagic acid and chebulinic acid demonstrated promising binding affinity with the target protein, Mpro. Molecular dynamics simulation was carried out for these selected bioactive molecules to further elucidate their binding affinity towards the Mpro and explored the obtained trajectories for RMSD, RMSF, Rg, and SASA. The trajectories were found to have increased flexibility and higher stability. Further, strong binding affinity between the bioactive molecules and Mpro was confirmed by low binding free energies, dynamic behaviors of complexes as well as residue contribution energies obtained by MM-GBSA. This study identifies four bioactive molecules from Triphala formulation as promising SARS-CoV-2 Mpro inhibitors; hence these bioactive molecules can be further examined using experimental studies for confirmation of their therapeutic efficacies.

Funding

The author would like to acknowledge the Deanship of Scientific Research, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia, for providing Grant through project number Covid19-2020-002-IRMC.

Conflicts of interest
The authors declare no conflicting interests.

7. Availability of data and material

Data and material will be made available after proper request to the corresponding author.
Mithun Rudrapal: Software, Investigation. Ismail Celik: Software, Investigation. Johra Khan: \textit{Innovating for increased productivity, Investigation.}.

Hannah J. Khairnar: Funding acquisition. Mohammad N. Alomary: Funding acquisition. Rohtish Yadav: Software, Investigation, Writing – original draft. Tripti Sharma: Software, Investigation, Writing – original draft. Trina Ekawati Tallei: Software, Investigation. Praveen Kumar Pasala: Ranjan Kumar Sahoo: Funding acquisition. Shubham J. Khairnar: Funding acquisition. Atul R. Bendale: Funding acquisition. James H. Zothanluanga: Software, Investigation. Dipak Chetia: Software, Investigation. Sanjay G. Walode: Software, Investigation.

Acknowledgements

All MD simulations presented in the study were carried out with resources provided by TÜBİTAK (Turkish Scientific and Technological Research Council), ULAKBİM (Turkish Academic Network and Information Centre), and TRÜBA (High Performance and Grid Computing Centre). The author would like to acknowledge the Deanship of Scientific Research, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia, for providing Grant through project number Covid-19-2020-002-IRMC.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jskus.2022.101826.

References

Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J.C., Hess, B., Lindahl, E., 2015. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 1–2, 19–25. https://doi.org/10.1016/j.softx.2015.06.001.

Ahmad, S., Zahiruddin, S., Parveen, A., Parveen, A., Rahman, A., Ahmad, M., 2021. Indian medicinal plants and formulations and their potential against COVID-19–preclinical and clinical research. Front. Pharmacol. 11, 2470. https://doi.org/10.3389/fphar.2020.01211-9.

Guan, L., Yang, H., Cai, Y., Sun, D., Li, P., Li, W., Liu, G., Tang, Y., 2019. ADMET-score – a comprehensive scoring function for evaluation of chemical drug-likeness. MedchemComm 10, 317–327. https://doi.org/10.1039/c8md00586e.

Chosh, S., Chetia, D., Gogoi, N., Rudrapal, M., 2021. Design, molecular docking, drug-likeness and molecular dynamics studies of 1,2,4-trioxane derivatives as novel F. palorus Falcipain-2 (FP-2) inhibitors. Biotechnology 102, 257–275. https://doi.org/10.1186/s40128-021-00882-3.

Gogoi, B., Chowdhury, P., Goswami, N., Gogoi, N., Niazy, T., Chetia, P., Mahanta, S., Chetia, D., Tanti, B., Borah, P., Handique, P.J., 2021. Identification of potential plant-based inhibitor against viral proteases of SARS-CoV-2 through molecular docking. MM-PBSA binding energy calculations and molecular dynamics simulation. Mol. Divers. 25 (3), 1963–1977. https://doi.org/10.1007/s11030-021-02111-9.

Hussain, N., Kakoti, B.B., Banerjee, A.K., Choudhary, M.K., Khaliq, N., Atta-ur-Rahman, Choudhary, M.I., Malik, C.H., 2007. Studies on the chemical composition of Phyllanthus emblica. Nat. Prod. Res. 21, 775–781. https://doi.org/10.1080/17486690601124664.

Hargena, T.H., Murphy, R.B., Friesner, R.A., Beard, H.S., Frye, L.L., Pollard, W.T., Banks, J.L., 2004. Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. J. Med. Chem. 47, 1750–1759. https://doi.org/10.1021/jm030644x.

Harcourt, B.H., Jukneliene, D., Kanjanahaluethai, A., Bechill, J., Severson, K.M., Smith, C.M., Rota, P.A., Baker, S.C., 2004. Identification of severe acute respiratory syndrome coronavirus replicase products and characterization of papain-like protease activity. J. Virol. 78, 13600–13612. https://doi.org/10.1128/JVI.78.24.13600-13612.2004.

Kher, I., Nasser, A., 2021. In silico molecular docking analysis for repurposing approved antiviral drugs against SARS-CoV-2 main protease. Biochem. Biophys. Reports 27, 101032. https://doi.org/10.1016/j.bbrep.2021.101032.
