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.
In silico screening of antiviral compounds from Moringa oleifera for inhibition of SARS-CoV-2 main protease

Bala Mohan Sivani, Priyanka Venkatesh, T.P.Krishna Murthy, S.Birendra Kumar

Department of Biotechnology, M S Ramaiah Institute of Technology, Bengaluru, 560054, Karnataka, India

ABSTRACT

COVID-19 as per early April 2021, has globally exceeded 129 million cases and is the cause of about 2.8 million deaths. Since the illness is known to have a wide array of symptoms, in addition to its ability to spread rapidly through contact and the complexity involved in developing drugs, there is an immediate need to look into alternative treatment regimes. This study was intended to identify potential antiviral compounds from Moringa oleifera against the selected target main protease (M\textsuperscript{pro}), which is vital for the survival of the virus. In silico molecular docking and dynamics was performed to determine a potential inhibitor against M\textsuperscript{pro}. Phytochemicals of Moringa oleifera reported in literature were retrieved from public databases and employed for molecular docking studies against M\textsuperscript{pro}. PyRx software was used to perform docking analysis. Visualization of amino acid interactions between the ligand and target was performed using Maestro and Discovery studio visualizer to analyze the type of interactions. Compounds displaying high binding affinities were subjected for analysis of pharmacokinetic studies, later molecular dynamics (MD) and MM-PBSA studies was conducted over selected compounds using GROMACS. Rutin and Isorhamnetin-3-O-rutinoside, both flavonoids thoroughly studied for their medicinal properties showcased strong interactions and the highest binding affinity of −8.9 kcal/mol with the M\textsuperscript{pro}. The binding energy calculated employing MM-PBSA for Rutin and Isorhamnetin-3-O-rutinoside were −86.832 kJ/mol and −72.984 kJ/mol, respectively. The overall studies revealed that Rutin and Isorhamnetin-3-O-rutinoside are potential in inhibiting the SARS-CoV-2 M\textsuperscript{pro} and can be validated through in-vivo and in-vivo studies.

1. Introduction

The coronavirus (CoV) pandemic also known as COVID-19, has been epidemiologically associated with the seafood and animal market in Wuhan, China which started in December 2019 [1]. In a matter of a few weeks, over 100,000 cases were confirmed globally, with numbers increasing rapidly. Within a month, WHO designated it as “public health emergency of international concern” [2]. This novel CoV is a 30,000 bp single, positive-stranded RNA virus that belongs to the Coronaviridae family, along with severe acute respiratory syndrome coronavirus (SARS-CoV) and the Middle East respiratory syndrome coronavirus (MERS-CoV) [2]. Symptoms include fever, cough, myalgia or fatigue and dyspnea [3]. The six species of Human coronaviruses (HCoVs) strains that were associated with respiratory tract infections are: HCoV-229E, HCoV-OC43, HCoV- Hong Kong University 1 (HCoV-HKU1), HCoV-NL63, SARS-CoV and MERS-CoV. The new strain associated with COVID-19 is the SARS-CoV-2, which taxonomically belongs to Betacoronaovirus genre [4].

The infection process starts with the interaction between the spike (S) glycoprotein and the host ACE2 receptor and cleavage of the S protein by the host transmembrane serine protease 2 (TMPRSS2) occurs prior to the fusion to the host cell membrane [5]. This spike glycoprotein contains (i) a large ecto-domain (ii) a single-pass transmembrane anchor, and (iii) a short C-terminal intracellular tail [6]. The entry of the virus depends on the cleavage of this spike protein with the help of the protease enzyme. Lysosomal cathepsins are responsible for endocytosis-mediated viral entry [7].

Accounting to the fact that novel drug discovery is not only expensive but also time-consuming, repurposing has been done using previously used antiviral drugs for the time being. Lopinavir/ritonavir (HIV protease inhibitor/CYP450 inhibitor), Remdesivir (RNA-dependent RNA polymerase inhibitor), Favipiravir (RNA-dependent RNA polymerase...
inhibitor), Darunavir/cobicistat(HIV protease inhibitor/CYP450 inhibitor) and Immuno-modulatory and anti-inflammatory drugs such as Tocilizumab, Chloroquine/Hydroxychloroquine, Baricitinib, Aviptadil and Eculizumab [8]. Chloroquine has been reported to be a potential antiviral drug for H5N1 in an animal model as it increases the endosomal pH which prevents virus fusion [9].

Plant sources have been largely used for its medicinal value in traditional medicine across the globe for several centuries [10]. Even today, it is observed that around 80% of the global population resort to natural remedies, primarily because of their affordability and accessibility [11]. Plant sources have a wide array of applications and are known to provide several classes of new chemical entities. The existence of traditional literature and ethnomedicine makes the selection process for a drug candidate more efficient. Adding on, it also provides scope for semi-synthesis of compounds for targeting dose-limiting. Though there are several challenges, mostly ethical, associated with natural sources, they are still widely used in modern drug discovery [10,12].

The Moringa genus, which belongs to the family Moringaceae, has shown extensive utility ranging from its function as a nutritional supplement to its antimicrobial activity. M. oleifera, commonly known as a drumstick is one of the most widely grown species in this family, which are normally found to be grown in tropical and subtropical regions across the world [13]. Since every part of the plant is reviewed to possess diverse medicinal applications, the plant reflects a substantial phytochemical profile [11]. Several studies concerning compounds from this plant are associated with antiviral activity against herpes simplex virus type 1 (HSV-1); causes fever blisters and cold sores around mouth, Hepatitis-B virus (HBV) which affects the liver, causes cirrhosis and liver cancer and towards antiretroviral therapy for Human Immunodeficiency Virus (HIV); the causative agent for AIDS infection. It has been used against common cold [13] Epstein Barr Virus [14] and Warts [15]. Studies have reported that M. oleifera shows an increase in production of antibodies against ND virus, which is an animal virus [16]. These factors and conclusive evidence of its antiviral property give compounds from M. oleifera the opportunity to be potential drug candidates for the treatment of SARS-CoV-2 [17].

Genomic evidence shows that the virus primarily relies on the functioning of the viral proteins: main protease (M(pro) or 3Cl(pro)) [18]. M(pro), which is a non-structural protein-5 (nsp5), is known to automatically cleave itself from poly-proteins pp1a and pp1ab, producing mature enzymes which further cleave nsps at 11 sites to release nsp4-nsp16. Structurally, it forms a dimer where an N-terminal catalytic region and a C-terminal region is present in each monomer. It is also observed that 3Cl(pro) from SARS-CoV and SARS-CoV-2 share 96% identity, where the minimal difference between the two is observed on the surface of the protein [19]. The mediation of maturation of nsps is an essential process in the life cycle of the virus and hence of importance for its survival [7, 20]. Since it not only plays a crucial role in the viral life cycle but also lacks closely related homologues in humans, M(pro) is described to be an attractive target for designing antiviral drugs [7].

The availability of data associated with the structures of protein targets has proportionately increased the demand for computational tools which could aid in the prediction of active sites as well as potential drug molecules. Adding on, plant sources yield a large number of pharmaceutically essential molecules, whose pharmacological properties can be validated all together, making this approach more efficient [21]. Application of in silico techniques in early pharmaceutical research reduces the cost and time involved in drug designing, hence providing a drug relatively earlier to the consumer at a lesser cost. Utilization of various computational tools has been discussed here to identify potential drug molecules from M. oleifera against targets in SARS-CoV-2 [22,23].

The present works describe the screening of potent inhibitors from M. oleifera against the target M(pro) from SARS-CoV-2, the molecular docking studies with the known inhibitors from M. oleifera, pharmacore development from the best-docked alignments of the compounds with high binding affinities, ADMET and the detailed investigation of promising mechanisms of inhibition through molecular dynamics simulations of the most promising molecules.

2. Methodology

2.1. Datasets

A total of 216 phytochemical compounds from M. oleifera were obtained from the literature (Table S1) [11,17,24,25]. The 3D structures of the compounds chosen from M. oleifera [26] were acquired from ChemSpider (www.chemspider.com) [27], PubChem (https://pubchem.ncbi.nlm.nih.gov/) [28] and ChEMBL (https://www.ebi.ac.uk/chembl/) databases [29]. The structures of the target main protease (M(pro)) (PDB ID: 6LU7) (Fig. 1) were retrieved from the RCSB PDB (http://www.rcsb.org/) [20,30]. The co-crystallised ligand and additional small molecules were removed along with energy minimization was done using Swiss PDB viewer [31]. Biovia Discovery Studio Client 2019 aided in the compilation of ligands into single .sdf file and also for the visualization of the molecules [32].

2.2. Molecular docking

The pdbqt file of M(pro) required for docking is prepared using AutoDock tools [33]. The binding site residues of M(pro) was obtained from literature and the following amino acids were found to be part of binding site: THR24, THR26, PHE140, ASN142, GLY143, CYS145, HIS163, HIS164, GLU166, HIS172. Molecular docking was performed using the Autodock vina feature found in PyRx with the suitable grid box that includes all the above mentioned binding site residues [34]. The protein-ligand interaction analysis was carried out using PyMOL molecular graphics viewer, Biovia Discovery Studio Client 2019 and Maestro 11.7 [32].

2.3. ADMET studies

pkCSM (http://biosig.unimelb.edu.au/pkcsmprediction), a web server that predicts pharmacokinetic (PK) properties has been used to extract rational compounds based on the desirable PK properties such as Absorption, Distribution, Metabolism, Excretion and Toxicity; which...
further have been subcategorized to VDAs (Volume of Distribution), CNS (Central Nervous System) permeability, CYP450 inhibitors, total clearance, hepatotoxicity, etc. to name a few [35]. The top compounds obtained from the docking analysis exhibiting higher binding energies were subjected to ADME studies. The SMILES notation required for the compounds were obtained from PubChem. The absorption of drugs depends on membrane permeability (indicated by colon cancer cell line (Caco-2)), intestinal absorption, skin permeability levels, P-glycoprotein substrate or inhibitor. The distribution of drugs depends on factors that include the blood–brain barrier (logBB), CNS permeability, and the volume of distribution (VDs). Metabolism is predicted based on the CYP models for substrate or inhibition (CYP2D6, CYP3A4, CYP1A2, CYP2C19, CYP2C9, CYP2D6, and CYP3A4). Excretion is predicted based on the total clearance model and renal OCT2 substrate. The toxicity of drugs is predicted based on Ames toxicity, hERG inhibition, hepatotoxicity, and skin sensitization. These parameters were calculated and checked for compliance with their standard ranges [36].

2.4. Molecular dynamics

Molecular dynamics simulations using GROMACS version 2019.4 was performed on the X-ray crystal of Main Protease (6LU7) in complex with selected ligands determined from docking studies. gromos54a7 force field was used to generate the protein parameters. PRODRG web server was used to develop the ligand topology using the steepest descent algorithm, processed setup was first vacuum minimized for 1500 steps. This was followed by energy minimization which was performed to remove steric clashes and to optimize the structure. Post energy minimization, equilibration of the system was performed in two steps, temperature stabilization of the system by heating up the system to 300 K followed by stabilizing the pressure and density of the system. Both were performed for 100 ps of NVT equilibration. The resultant structures obtained from the NPT equilibration phase was directed for the final production run in the NPT ensemble for a simulation time of 100 ns. The simulation time step was set as 0.002 ps/2 fs. The v-rescale algorithm was employed for temperature coupling while the Parinello-Rahman algorithm was used for temperature coupling. Particle Mesh Ewald (PME) method was selected to evaluate electrostatic interactions. The default simulation box was set to cubic and the dimensions were 2.0 nm [37,38]. Data on SPC water molecules and number of counter ions added is discussed under supplementary data (Table S2).

2.5. Trajectory analysis and free energy calculation (MMPBSA)

GROMACS analysis tool was employed to perform trajectory analysis. Tools such as gmx rms and gmx rmsf were used to compute the Root Mean Square Deviation (RMSD) and Root Mean Square Fluctuation (RMSF) of the protein, respectively [39]. In addition, analysis of radius of gyration, solvent accessible surface area (SASA) and a number of hydrogen bonds formed between protein and ligand were assessed using the tools gmx gyrate, gmx sasa and gmx hbond respectively. The binding affinity of an inhibitor with a protein over the simulation time was inferred by employing the Molecular Mechanics Poisson-Boltzmann surface area (MM-PBSA) approach [40]. The binding free energy was estimated using a GROMACS utility g_mmpbsa. ∆Glocal for the last 20 ns with dt 1000 frames was computed to obtain accurate results [37].

3. Results and discussion

3.1. Screening and identification of potential inhibitors

Docking of candidate compounds into binding sites of the biological target and ranking them based on predicted binding affinity is referred to as Structure-based virtual screening (SBVS) [40,41]. High throughput virtual screening (HTVS) programs through softwares like PyRx involves docking of one or more ligands against targets. The foundation of docking engines comprises a combination of pose identification and scoring algorithms. The evaluation is performed qualitatively based on inspection of ligand position, and quantitatively by implementing scoring algorithms [42]. Successful docking is observed with maximum binding energy as the indicator.

Table 1 illustrates the top 10 ligands from *Moringa oleifera*, which showed potential activity against Mₚｒｏ, along with the prominent interactions observed. From this study, Rutin [43], which is a natural flavone derivative, showed the highest binding score of −8.9 kcal/mol. The same binding score of −8.9 kcal/mol is also observed in the case of Isoharmnettin-3-O-rutinoside, a methylated flavone, structurally sharing close resemblance with Rutin. Quercetin-3-gentiobioside and Apigenin-7-O-rutinoside, both are observed to display a binding score of −8.7 kcal/mol. Mulberroforan Q, which is a benzo-furan derivative displays a binding affinity of −8.6 kcal/mol. A score of −8.4 kcal/mol was observed for the binding between the target and the glycosidic ligand Mudanpioside J. Ellagic acid, the ligand with the least molecular weight among the top 10 ligands which also is a phenolic group derivative shows a score of −8.3 kcal/mol. Apin, a recognized flavonoid displays a binding score of −7.8 kcal/mol. Isoqueretin and Procyandins both are observed to have a binding affinity of −7.4 kcal/mol [22].The docking scores of all the 216 ligands has been provided as additional information (Table S3).

When the ∆G of the protein-ligand complex and unbound state is directly proportional to the stability of the protein-ligand interaction, it is accepted that a protein-ligand interaction occurs [44]. Low ∆G value is known to assist the binding between the protein and ligand [45]. Hence, the stability of a protein-ligand complex, which is an essential characteristic for an effective drug is indicated by a negative value of ∆G [46]. Observations show that most compounds with favourable binding affinities described above are categorized under flavonoids. Our present work depicts protein-ligand interactions displaying appreciable changes in ∆G levels, therefore, concluding the high binding affinity of Rutin and the drug target Mₚᵣ₀. Hence, we suggest that Rutin and Isoharmnettin-3-O-rutinoside has shown the potentiality to establish stable and strong complexes with the target employed in this study.

Additionally, reported literature supports our claim with respect to our choice of compounds. Rutin is a very commonly studied flavonoid since it is known to be antibacterial, anti-inflammatory, an antioxidant, an antiviral agent and also displays anti-cancer activities. Originating from the vitamin C2 family, it is widely found in many fruits, vegetables and cereals. Rutin could be recommended as a potential candidate for the treatment of COVID-19, not only because of its high score observed during simulation but also because of its easy availability and inexpensive nature. Isoharmnettin-3-O-rutinoside, also known as Narcissin is a disaccharide derivative. Though general analysis of flavonoids discusses its role as an anti-tumour agent and antiviral agent, the specific role of this compound has not been widely studied and needs further validation [47,48].

Simultaneously, the mapping of interactions of the amino acid residues present in the target with the ligands were analyzed. Favourable conformations are indicated based on the different types of amino acid interactions such as electrostatic, hydrogen bonding, hydrophobic, non-polar and polar, that are observed. Protein stability is assessed based on the available hydrophobic and hydrogen bond interactions. Hydrophobic interactions primarily determine the folding configuration equilibria of many native proteins. Table 1 discusses the predominant interactions between the target and the selected ligands. Target Mₚᵣ₀ interacted with rutin (Fig. 2a), predominantly, through hydrophobic, polar and hydrogen bonds. Hydrophobic interactions between the target and ligand rutin involved Met (49), Tyr (54), Phe (140), Leu (141), Cys (145), Met (165), Leu (167) and Pro (168) residues. Amino acids Leu (141), Glu (166), Asp (187) and Gln (189) participated in hydrogen bonding. Bond lengths of 1.83 Å, 1.94 Å, 2.27 Å and 2.29 Å, respectively, were observed. Polar interactions involved Thr (26), His (41), Asn (142), Ser (144), His (163), His (164), His (172) and Gln (189). Mₚᵣ₀ also
showed two residues involved in electrostatic interactions. The M^P^p, isorhamnetin-3-O-rutinoside complex (Fig. 2b) exhibited hydrogen bonding with the ligand and Leu (141) amino acid residue, with a bond length of 1.82 Å, hydrophobic interactions from the residues Met (49), Tyr (54), Phe (140), Leu (141), Cys (145), and Met (165), and polar interactions from Thr (26), His (41), Asn (142), Ser (144), His (163), His (164), His (172), and Gln (189). The complex is also observed to show minimal electrostatic interactions. Interactions between Quercetin-3-gentiobioside and the target Mpro (Fig. 2c) show a similar trend of interactions from Thr (26), His (41), Asn (142), Ser (144), His (163), His (164), Gln (189) and hydrophobic interactions with Leu (27), Met (49), Tyr (54), Phe (140), Leu (141), Cys (145), Met (165), Leu (167), Pro (168) respectively. Hydrophobic interactions ((Met (49), Tyr (54), Leu (141), Cys (145), Met (165), Pro (168)) and polar interactions (His (41), Asn (142), Ser (144), His (163), His (164), Gln (189), and Thr (190)) and hydrogen bonding (Gly (143)) was observed between M^P^p and the ligand Apigenin-7-O-rutinoside (Fig. 2d). M^P^p-Mulberrofuran Q (Fig. 2e) has a minimal number of amino acids involved in interactions where Asn (95), Asn (119), Ser (121), Ser (123) show polar interactions and Met (17), Trp (31), Ala (70), Pro (122) show hydrophobic interactions. No hydrogen bonds were formed. ADMET studies

Table 1

| Sl No. | Compound             | Binding Affinity (in kcal/mol) | RMS value | Type of Interaction               | Residue Information                          |
|--------|----------------------|--------------------------------|-----------|-----------------------------------|----------------------------------------------|
| 1      | Rutin                | -8.9                           | 0         | H-bond                            | Leu (141), Gln (166), Asp (187), Gln (189) |
|        |                      |                                |           | Hydrophobic                       | Met (49), Tyr (54), Phe (140), Leu (141), Cys (145), Met (165), Leu (167), Pro (168) |
| 2      | Isoquercetin-3-O-rutinoside | -8.9                         | 0         | H-bond                            | Leu (141)                                    |
|        |                      |                                |           | Hydrophobic                       | Met (49), Tyr (54), Phe (140), Leu (141), Cys (145), Met (165) |
| 3      | Quercetin-3-gentiobioside | -8.7                         | 0         | H-bond                            | Leu (27), Leu (141), His (163)               |
|        |                      |                                |           | Hydrophobic                       | Met (49), Tyr (54), Phe (140), Leu (141), Cys (145), Met (165) |
| 4      | Apigenin-7-O-rutinoside | -8.7                         | 0         | H-bond                            | Gly (143)                                    |
|        |                      |                                |           | Hydrophobic                       | Met (49), Tyr (54), Phe (140), Leu (141), Cys (145), Met (165), Pro (168), Val (186), Ala (191) |
| 5      | Mulberrofuran Q       | -8.6                           | 0         | Polar                              | Met (17), Thr (31), Ala (70), Pro (122)     |
| 6      | Mulbaniapioside J     | -8.4                           | 0         | H-bond                            | Gly (143), His (163)                         |
| 7      | Ellagic Acid          | -8.3                           | 0         | H-bond                            | Gln (110), Thr (111)                         |
| 8      | Apin                 | -7.8                           | 0         | Hydrophobic                       | Val (104), Ile (106), Phe (112), Ile (152), Phe (294), and polar interactions with Gln (107), Gln (110), Thr (111), Asn (151), Ser (158), and Thr (292) respectively. Hydrophobic interactions (Ile (200), Tyr (237), Leu (271), Leu (272), and Leu (287)), polar interactions (Thr (196), Thr (198), Thr (199), and Leu (287)), and hydrogen bonding (Lys (137), Asp (197), Asn (238), and Asp (289)) were exhibited between M^P^p and Apin (Fig. 2h). The target M^P^p interacts with Isoquercetin through electrostatic interactions (Asp (197), Glu (288), Asp (289), and Gln (290)) and a hydrogen bond (Lys (137)) (Fig. 2i). Electrostatic interactions (Asp (197), Glu (288), Asp (289), and Gln (290)), hydrophobic interactions (Tyr (237), Tyr (239), Leu (272), Leu (286), and Leu (287)), and hydrogen bonds (Tyr (237)) were exhibited by M^P^p interacting with procyandin (Fig. 2j). Fig. 2 describes and summarises the 2-D interactions observed between the amino acids from the target M^P^p and the selected ligands, which showed promising results. 3-D interactions have been depicted in Fig. S1.

Additionally, for better understanding of the interactions, surface mapping of the complexes was performed and the data is described in Fig. S2.

3.2. ADMET studies

An ideal oral drug should be absorbed from the gastrointestinal tract, distributed to the target specifically, metabolized without eliminating its property and removed without any damage. The relation between chemical structures and physiological parameters helps us in understanding the pharmacokinetic properties. The skin-permeability coefficient (logKp), apparent Caco-2 and MDCK were computed and analyzed. Caco-2 permeability, intestinal absorption (human), skin permeability, and P-glycoprotein substrate or inhibitor were used to predict the absorption level of the compounds. A CaCo-2 permeability value greater than 0.9 indicates high permeability. Mudanpioside J has shown the highest permeability value of 0.687. With regard to intestinal absorption level, value less than 30% is considered to be poorly absorbed. Many compounds have shown very good absorption. Mulberrofuran Q (100), Mudanpioside J (78.029) and Ellagic acid (86.68) are among the top compounds with regard to absorption. With regard to skin permeability the log Kp > -2.5, shows low skin permeability, all our top ten compounds have shown a constant skin permeability value of 2.735. P-
Fig. 2. Interaction between the selected drug target (Mpro) and the top 10 lead molecules. (a) Rutin (b) Isorhamnetin-3-O-rutinoside (c) Quercetin-3-gentiobioside (d) Apigenin-7-O-rutinoside (e) Mulberrofuran Q (f) Mudanpioside J (g) Ellagic acid (h) Apiin (i) Isoquercetin (j) Isoquercetin.
glycoprotein is a member of the ATP-binding transmembrane glycoprotein family [ATP-binding cassette (ABC)], which can excrete drugs or other exogenous chemicals from cells. The results suggest that all the compounds are substrates of P-Glycoprotein. Mulberryferan Q, Munapioside J and Procyanidins were predicted to be a P-glycoprotein 1 inhibitor. Isorhamnetin-3-O-rutinoside, Mulberryferan Q and Procyanidins were predicted to be P-glycoprotein 2 inhibitor.

The distribution volume (VDs), Fraction unbound (human), CNS permeability and blood–brain barrier membrane permeability (logBB) were used to characterize the distribution of compounds. VDs lower than 0.71 L kg\(^{-1}\) (log VDs < –0.15), the distribution volume is considered to be relatively low. When VDs is higher than 2.81 L kg\(^{-1}\) (log VDs > 0.45), the distribution volume is considered to be relatively high. Apigenin-7-O-rutinoside, Rutin and Isorhamnetin-3-O-rutinoside, Rutin and Isorhamnetin-3-O-rutinoside have shown comparatively better values. For blood–brain barrier membrane permeability, logBB > 0.3, the compounds were thought to cross the blood–brain barrier easily. A logBB < –1 suggested that the compounds did not easily cross the blood–brain barrier. For CNS permeability all compounds were predicted to be able to penetrate the CNS (logPS is < –3).

Cytochrome P450s is an important enzyme system for drug metabolism in liver. The two main subtypes of cytochrome P450 are CYP2D6 and CYP3A4. The results showed that Mulberryferan Q and Munapioside J were substrates for CYP3A4. None of the compounds were predicted to be CYP2C19 inhibitor, CYP2C9 inhibitor, and CYP3A4 inhibitor. This suggested that these compounds may be metabolized in the liver.

Drug elimination is related to the molecular weight and hydrophilicity of compounds. Rutin and Isorhamnetin-3-O-rutinoside showed low clearance value. The results also suggest that none of compounds are toxic in Ames test nor hepatotoxic. None of the compounds have cardiotoxicity or skin sensitization. Table 2 describe the ADMET properties discussed.

### Table 2

The result of ADMET parameters for compounds with high binding affinities.

| Property          | Model Name          | Unit             | 1     | 2     | 3     | 4     | 5     | 6     | 7     | 8     | 9     | 10    |
|------------------|---------------------|------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Absorption       | Water solubility    | Numeric (log mol/L) | –2.892 | –2.899 | –2.882 | –2.942 | –2.917 | –3.088 | –3.181 | –2.851 | –2.925 | –2.892 |
| Absorption       | Caco2 permeability  | Numeric (log Papp in 10–6 cm/s) | –0.949 | 0.201 | –1.253 | 0.285 | 0.266 | 0.687 | 0.335 | –0.966 | 0.242 | 0.131 |
| Absorption       | Intestinal absorption (human) | Numeric (%) | Absorbed | 23.446 | 30.095 | 24.364 | 32.393 | 100 | 78.029 | 86.684 | 17.411 | 47.999 | 55.527 |
| Absorption       | Skin Permeability P-glycoprotein substrate | Numeric (log Kp) | Categorical (Yes/No) | –2.735 | –2.735 | –2.735 | –2.735 | –2.735 | –2.735 | –2.735 | –2.735 |
| Absorption       | P-glycoprotein I inhibitor | Categorical (Yes/No) | No | No | No | No | Yes | Yes | Yes | Yes | Yes | Yes |
| Absorption       | P-glycoprotein II inhibitor | Categorical (Yes/No) | No | Yes | No | No | Yes | No | No | No | No |
| Distribution     | VDs (human)         | Numeric (log L/kg) | 1.663 | 1.622 | 1.333 | 1.815 | –1.777 | 0.771 | 0.375 | 1.004 | 1.846 | 0.193 |
| Distribution     | Fraction unbound (human) | Numeric (Fu) | 0.187 | 0.15 | 0.225 | 0.116 | 0.264 | 0.108 | 0.083 | 0.171 | 0.228 | 0.284 |
| Distribution     | BBD permeability     | Numeric (log BB) | –1.899 | –1.754 | –2.132 | –1.72 | –1.608 | –1.727 | –1.272 | –1.793 | –1.688 | –1.783 |
| Distribution     | CNS permeability     | Numeric (log PS) | –5.178 | –5.136 | –4.932 | –4.727 | –3.629 | –4.609 | –3.533 | –4.972 | –4.093 | –4.106 |
| Metabolism       | CYP2D6 substrate    | Categorical (Yes/No) | No | No | No | No | No | No | No | No | No | No |
| Metabolism       | CYP3A4 substrate    | Categorical (Yes/No) | No | No | No | No | Yes | Yes | No | No | No | No |
| Metabolism       | CYP1A2 inhibitor    | Categorical (Yes/No) | No | No | No | No | No | No | Yes | No | No | No |
| Metabolism       | CYP2C19 inhibitor   | Categorical (Yes/No) | No | No | No | No | No | No | No | No | No | No |
| Metabolism       | CYP2C9 inhibitor    | Categorical (Yes/No) | No | No | No | No | No | No | No | No | No | No |
| Metabolism       | CYP2D6 inhibitor    | Categorical (Yes/No) | No | No | No | No | No | No | No | No | No | No |
| Metabolism       | CYP3A4 inhibitor    | Categorical (Yes/No) | No | No | No | No | No | No | No | No | No | No |
| Excretion        | Total Clearance     | Numeric (log ml/ min/kg) | –0.369 | –0.252 | –0.35 | –0.016 | 0.252 | 0.077 | 0.537 | –0.054 | 0.394 | –0.058 |
| Excretion        | Renal OCT2 substrate | Categorical (Yes/No) | No | No | No | No | No | No | No | No | No | No |
| Toxicity         | AMES toxicity       | Categorical (Yes/No) | No | No | No | No | No | No | No | No | No | No |
| Toxicity         | Max. tolerated dose (human) | Numeric (log mg/kg/day) | 0.452 | 0.485 | 0.48 | 0.533 | 0.418 | –0.266 | 0.476 | 0.446 | 0.569 | 0.438 |
| Toxicity         | hERG I inhibitor    | Categorical (Yes/No) | No | No | No | No | No | No | No | No | No | No |
| Toxicity         | hERG II inhibitor   | Categorical (Yes/No) | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Toxicity         | Oral Rat Acute Toxicity (LD50) | Numeric (mol/kg) | 2.491 | 2.507 | 2.483 | 2.52 | 2.521 | 2.86 | 2.399 | 2.49 | 2.541 | 2.482 |
| Toxicity         | Oral Rat Chronic Toxicity (LOAEL) | Numeric (log mg/kg bw/day) | 3.673 | 3.424 | 4.945 | 3.338 | 2.298 | 3.344 | 2.698 | 4.574 | 4.417 | 3.857 |
| Toxicity         | Hepatotoxicity      | Categorical (Yes/No) | No | No | No | No | No | No | No | No | No | No |
| Toxicity         | Skin Sensitization  | Categorical (Yes/No) | No | No | No | No | No | No | No | No | No | No |
| Toxicity         | T. Pyriformis toxicity | Numeric (log ug/L) | 0.285 | 0.285 | 0.285 | 0.285 | 0.285 | 0.285 | 0.285 | 0.285 | 0.285 | 0.285 |
| Toxicity         | Minnow toxicity     | Numeric (log mM) | 7.677 | 7.187 | 13.088 | 4.806 | 0.455 | 4.47 | 2.11 | 3.835 | 8.061 | 10.607 |
3.3. MD simulations

The MD simulations for the M^{Pro}-Rutin, M^{Pro}-Isorhamnetin-3-O-rutinoside complexes and the apo-protein were performed for 100 ns to scrutinize the behaviour of molecules in complex systems. This is done in order to get an accurate analysis of changes in conformations, energetics and the interactions between the ligand and the protein. A force field is employed to calculate the overall potential energy of the system based on the accountability of interactions between individual atom-atom pair \[49,50\].

MD trajectories were at first evaluated with the help of RMSD and RMSF in order to deduce the fluctuations and the stability of the complex structures. Least-squares superposition method is used to correspond a reference structure and a snapshot selected during the simulation, by aligning one structure with respect to the other and by also employing a rigid body translation. The deviation from standard values is termed as root-mean-square deviation (RMSD).

The molecular dynamics studies were performed in duplicates and the average value of the two trials has been calculated to plot and calculate the respective results. The RMSD plot for Rutin, Isorhamnetin-3-O-rutinoside and the apo-protein are represented in Fig. 3 as a merged plot. The average RMSD value observed for the apo-protein was calculated to be 0.312 nm. The value remains almost between 0.2 nm and 0.4 nm throughout the entire simulation period. Rutin has consistently shown stability throughout the entire simulation between 0.13 nm and 0.28 nm with an average RMSD value of 0.254 nm, a lower value than the RMSD of the apo-protein. In the case of Isorhamnetin-3-O-rutinoside, an increase from 0.15 nm to close to 0.4 nm is observed in the first 40 ns, after which it stabilizes and remains between 0.3 nm and 0.4 nm for the rest of the simulation. This complex shows an average RMSD value of 0.352 nm. Low values of RMSD indicate that the complexes are highly stable \[51\]. M^{Pro}-Rutin complex expresses a higher stability than the apo-protein and the M^{Pro}-Isorhamnetin-3-O-rutinoside complex, from which we can conclude that the ligand Rutin, additionally also increases the stability of the apo-protein, which could indicate effective binding of the ligand and also potential activity.

The RMSF values were estimated from the respective 100 ns MD trajectories. The fluctuations in the protein Ca atoms were assessed in order to quantify the flexibility of the protein in the presence of the ligand. Highest fluctuations were observed towards the beginning and end residues in both complexes as well as the apo-protein. The average RMSF value for Rutin, Isorhamnetin-3-O-rutinoside and apo-protein were around 0.160 nm when plotted together. The RMSF of the carbon atoms analyzed were represented in Fig. 4.

The equilibrium conformation of the total system is described by the parameter referred to as the Radius of Gyration (Rg). The compactness and unbending nature of a molecule can be determined using the Rg value [52]. If protein folding is seen to occur in favourable conditions, it is likely to maintain a stable value of radius of gyration [53]. From Fig. 5, we observe that during the simulation period of 100 ns, the average Rg value for the apo-protein is 2.175 nm, for M^{Pro}-Rutin complex it is found to be 2.325 nm and for M^{Pro}-Isorhamnetin-3-O-rutinoside it was observed to be 2.194 nm. We observe that all three complexes have undergone protein folding in favoured conditions and also display high levels of compactness and stability. Amongst the three, the M^{Pro}-Isorhamnetin-3-O-rutinoside complex displays a relatively higher magnitude of compactness.

Solvent accessibility of the complex is assessed by calculating the solvent-accessible surface area (SASA) of the protein on binding to the protein. Expansion of the protein is observed for a higher SASA value. Table 3 shows the solvent surface accessible on the protein [53,54].

Binding affinity and activity of the ligand can be assessed by the binding of key residues at the binding site through non-bonding interactions such as hydrogen bonding. Conformational stability can also be analyzed by estimating the total number of hydrogen bonds formed during the simulation period. The average number of hydrogen bonds found to have occurred in the entire simulation for i.e., M^{Pro}-Rutin complex is around 7 and for M^{Pro}-Isorhamnetin-3-O-rutinoside it was around 5 bonds. This is described in the graph shown in Fig. 7. More
number of hydrogen bonds between the ligand and protein in the binding site result in more binding affinity as well as conformational stability [50]. Results suggest that the complex formed with Isorhamnetin-3-O-rutinoside exhibits a relatively lower conformational stability in comparison to the complex formed with Rutin. Since, the binding energy data obtained from molecular docking is the same for both the ligands, dynamics studies highlight the differences. A comparison of results for MD analysis for the respective trials is provided as additional data (Fig. S3, S4 and S5).

3.4. MM/PBSA – binding free energy calculation

Revalidation of inhibitor affinity for the protein-ligand complexes predicted by the docking studies was performed by analyzing the binding free energies of the complexes during the molecular simulations. The binding energies are calculated by summing up non-polar interactions, polar and non-bonded interaction energies [55]. The average free energies obtained during MM/PBSA calculations are tabulated in Table 3. For the Mpro-Rutin complex, it is observed that Van der waal energy (−180.038 kJ/mol) and Electrostatic energy (−107.335 kJ/mol) have relatively distant values with the electrostatic energy being slightly towards the lower side. High polar solvation energy (220.596 kJ/mol) was recorded. The solvent-accessible surface area (SASA) energy was recorded to be around −20.204 kJ/mol. The net binding energy was observed to be −86.832 kJ/mol. Mpro-Isorhamnetin-3-O-rutinoside complex shows a comparable Van der waal energy (−166.122 kJ/mol) and similar electrostatic energy observed (−110.379 kJ/mol). This complex displayed a polar solvation energy of 221.149 kJ/mol and SASA energy of −17.641 kJ/mol. Overall, the net binding energy was observed to be −72.984 kJ/mol. Individual energy calculations for the respective trials is also reported (Table S4).

On comparing the two complexes, we observe that the all binding energy values are higher for the Mpro-Rutin complex in comparison to the Mpro-Isorhamnetin-3-O-rutinoside complex though the difference in energies are limited. The interactions are highly favoured by Van der Waal interactions and electrostatic interactions together. Polar solvation energy individually dominates the binding energy [56,57].

4. Conclusion

In conclusion, we have thoroughly studied SARS-CoV-2 and have identified the target Mpro as one of the targets essential for the survival of the virus. Analysis of phytochemicals from Moringa oleifera has led to investigating several antiviral compounds. We have identified the ligands Rutin and Isorhamnetin-3-O-rutinoside from Moringa oleifera to have inhibitory action against the target Main protease (Mpro) of SARS-CoV-2 through molecular docking, ADME and dynamics studies. Both complexes showed high levels of stability, favourable pharmacokinetic properties, minimum levels of fluctuations and high degree of compactness. They also exhibit good surfaces for solvent exposure. Significant set of interactions include, Hydrogen bonds, Hydrophobic and electrostatic interactions. Binding energy results have also been analyzed which has given additional information on inhibitor affinity. Altogether, the results reveal that compounds present in Moringa oleifera can be potent anti-COVID-19 drug candidates. Also, this study opens up potential testing (in-vitro and in-vivo) possibilities for these two hit compounds against COVID-19.

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.crgsc.2021.100202.

References

[1] P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li, C.-L. Huang, A pneumonia outbreak associated with a new coronavirus of probable bat origin, Nature 579 (2020) 270–273.
[2] J. Cui, F. Li, Z.-L. Shi, Origin and evolution of pathogenic coronaviruses, Nat. Rev. Microbiol. 17 (2019) 181–192.
[3] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xia, X. Gu, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, Lancet 395 (2020) 497–506.
[4] T. Pillaiyar, S. Meenakshisundaram, M. Manickam, Recent discovery and development of inhibitors targeting coronaviruses, Drug Discov. Today Off. 25 (2020) 668–688.
