Identification of Possible Inhibitors of SARS-CoV-2 Main Protease from some Bioactive Compounds of *Artemisia annua*: An in silico Approach

Titilayo Omolara Johnson (✉ titijohnson2004@yahoo.com)

University of Jos

Abayomi Emmanuel Adegboyega

University of Jos

Oluwafemi Adeleke Ojo

Landmark University

Amina Jega Yusuf

Usmanu Danfodiyo University

Opeyemi Iwaloye

Federal University of Technology, Akure

Chinenye Jane Ugwah-Oguejiofor

Usmanu Danfodiyo University

Rita Onyekachukwu Asomadu

University of Nigeria

Ifeoma Felicia Chukwuma

University of Nigeria

Stephen Adakole Ejembi

University of Jos

Emmanuel Ike Ugwuja

Ebonyi State University

Research Article

Keywords: Coronavirus, SARS-CoV-2, main protease, *Artemisia annua*, Kaemferol, Quercimeritrin

Posted Date: August 23rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-612899/v2>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

The inhibitory potential of *Artemisia annua*, a well-known antimalarial herb, against several viruses including the coronavirus is increasingly gaining recognition. The plant extract has shown significant activity against both the Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) and the novel SARS-CoV-2, that is currently ravaging the world. It is therefore necessary to identify the bioactive compounds of the plants that are responsible for this activity, for the purpose of designing drugs against SARS-CoV-2. In this study, we employed *in silico* techniques comprising of molecular docking, binding free energy calculations, pharmacophore modelling as well as druglikeness, pharmacokinetics and toxicity predictions, to identify potential inhibitors of the SARS-CoV-2 main protease (Mpro) from 168 bioactive compounds of *Artemisia annua*. Rhamnocitrin, Isokaempferide, Kaempferol, Quercimeritrin, Apigenin, Penduletin, Isoquercitrin, Astragalin, Luteolin-7-glucoside and Isorhamnetin were ranked highest; with docking scores ranging from -7.84 to -7.15 kcal/mol compared with -6.59 kcal/mol demonstrated by the standard ligand. Rhamnocitrin, Isokaempferide and Kaempferol, like the standard ligand interacted with important active site amino acid residues like HIS 41, CYS 145, ASN 142, and GLU 166, among others. These compounds also possess acceptable druglike properties and safety profile. Hence, they could be considered for experimental studies and further development into drugs against SARS-CoV-2.

Introduction

The outbreak of COVID-19, also known as severe acute respiratory syndrome - coronavirus 2 (SARS-CoV-2) has placed the entire world in a stop for some time now. The disease began from Wuhan (Hubei, China) in December 2019 ^{1,2}, spread to the rest of the world, and has been announced a pandemic by the World Health Organization (WHO) ³. They are the biggest known RNA infections comprising of positive single-stranded RNA and are connected with serious respiratory illness ^{4,5}. SARS-CoV-2 has constrained mankind to be on lockdown as a preventive measure to maintain a strategic distance from disease transmission (contraction and spreading of the illness). The side effects of the sickness incorporate but not restricted to fever, dry hack, sore throat and trouble breathing. As of 3:09pm CEST, 2 May 2021, a total of 151,803,822 established cases of COVID-19 and 3,186,538 deaths have been reported to WHO ⁶. Shockingly, there is neither a particular medication endorsed for the treatment of the disease, yet the contamination and death rates are consistently expanding.

Medications like Lopinavir/Ritonavir ⁷ and Remdesivir ⁸ have been under examination and clinical preliminaries towards managing COVID-19. As many countries of the world are continuously relaxing their lockdown limitations, there are worries about the chances of the disease expanding dramatically. Subsequently there is a critical need to distinguish and create novel drugs for the treatment and regulation of SARS-CoV-2 infections. Creating conventional medications and improvement pipelines could however be tedious, expensive and sometimes connected with high clinical disappointments ^{9,10}. Nevertheless, medicinal plants have been explored as a rich source of effective bioactive agents against numerous viral infections ^{11,12} and several antiviral plant products are suggested to be active against the SARS-CoV-2 Mpro protease ¹³⁻¹⁶.

Artemisia annua is a renowned medicinal plant belonging to the Asteraceae family (Compositae) ^{17,18} and it is generally known as sweet wormwood or Qinghao ^{19,20}. The identification of the plant and its artemisinin component constituted a major breakthrough in the fight against malaria. Other pharmacological activities have since then been associated with the plant. These include analgesic, anti-inflammatory, antioxidant, immunomodulatory, antibacterial, anticancer and antiviral activities ^{21,22}, among others.

Artemisia annua tea is reported to be effective against several viral infections including herpes simplex virus, dengue virus, human immunodeficiency virus (HIV) and of course, coronavirus. The plant extract showed significant activity against the SARS coronavirus that occurred in 2002 and it was selected for its significant inhibitory effect among 200 Chinese medicinal herbs screened for antiviral activities against SARS-CoV ²³. The plant was also recently reported to possess *in vitro* inhibitory activity against SARS-CoV-2 replication, but the active compound was not identified. It was however suggested that some non-artemisinin bioactive phytochemicals in the plant extracts could be responsible for the SARS-CoV-2 inhibitory activity, since the observed anti-viral effect did not correlate to the artemisinin component(s) in the extracts.

Over the years, many bioactive phytochemicals such as terpenoids, tannins, coumarins, essential oils, biflavonoids and polyphenols have been identified from different parts of *Artemisia annua* ²² and some have been extracted, characterized and analyzed for

antiviral activity ^{24–27}. Identification of anti-viral compounds of *Artemisia annua* with SARS-CoV-2 inhibitory activity could therefore help provide a promising therapeutic alternative against COVID-19.

The discovery of novel therapeutics has advanced in the last two decades towards the use of ground-breaking complementary approaches such as computational procedures ²⁸. In this manner, bioinformatic devices have been used widely in many intriguing investigations; and as of late for the SARS-CoV-2 drug research ^{29–32}. Molecular docking, pharmacophore modeling and ADMET (Absorption, Distribution, Metabolism, Excretion and Toxicity) studies are some of the major computational methods that offer great applicability in a short duration of time ³³. The molecular docking technique assists to predict the binding affinity and inhibitory potential of a test compound (ligand) against a target protein whose modification can produce a therapeutic benefit ³⁴. A well-recognized molecular target for the development of anti-SARS-CoV-2 drugs is the coronavirus main protease (Mpro), which plays a vital role in viral replication by processing the polyproteins that are translated from the viral RNA. Therefore, the present work carried out detailed *in silico* analysis, comprising of molecular docking, pharmacophore modeling and ADMET studies of the phytocompounds from *Artemisia annua*, for the purpose of identifying promising inhibitors of SARS-CoV-2 main protease, thereby controlling COVID-19.

Results

Molecular docking analysis

The molecular docking analysis showed that the compounds of *Artemisia annua* possess varying levels of binding affinities for the SARS-CoV-2 main protease, the top ten being Rhamnocitrin (7-Methylkaempferol), Isokaempferide (3-Methylkaempferol), Kaempferol, Quercimeritrin, Apigenin, Penduletin, Isoquercitrin, Astragalin, Luteolin-7-glucoside and Isorhamnetin (figure 1). The binding affinities of these compounds are higher than that of the standard ligand which is -6.59 kcal/mol. Rhamnocitrin scored highest with a docking score of -7.83 kcal/mol followed by Isokaempferide with -7.81 kcal/mol and Kaempferol with -7.65 kcal/mol (table 1). The entire list of the 168 compounds with their docking scores (kcal/mol) and binding free energy ($D G_{bind}$) MM-GBSA against the SARS-CoV-2 main protease can be found as Supplementary Table S1 online.

Table 1: The docking scores (kcal/mol) of the ten top-scoring compounds of *Artemisia annua* against SARS-CoV-2 main protease

Compound	Docking score (kcal/mol)
K36 (Standard ligand)	-6.59
Rhamnocitrin	-7.83
Isokaempferide	-7.81
Kaempferol	-7.65
Quercimeritrin	-7.55
Apigenin	-7.49
Penduletin	-7.38
Isoquercitrin	-7.33
Astragalin	-7.23
Luteolin-7-glucoside	-7.16
Isorhamnetin	-7.15

K36 - (1S,2S)-2-({N-[(benzyloxy)carbonyl]-L-leucyl}amino)-1-hydroxy-3-[(3S)-2-oxopyrrolidin-3-yl]propane-1-sulfonic acid.

The binding free energy of the top ten compounds are shown in table 2 and figure 2 shows the scatter plot of the binding free energy (DG_{Bind}) versus docking score (kcal/mol) of the 168 phytochemical compounds of *Artemisia annua* against SARS-CoV-2 Mpro. This post docking analysis showed that the ten top-scoring compounds of *Artemisia annua* had binding free energy values ranging between -39.67 kcal/mol and -50.61 kcal/mol. Astragalin gave the highest value followed by 7-Methylkaempferol and Quercimeritrin. Most of the points are however close to the regression line of the scatter plot.

Table 2: The binding free energy (DG_{bind}) MM-GBSA of ten top-scoring compounds of *Artemisia annua* against SARS-CoV-2 main protease

Compound	DG _{Bind}	DG _{Bind} Coulomb	DG _{Bind} Covalent	DG _{Bind} Hbond	DG _{Bind} Lipo	DG _{Bind} Packing	DG _{Bind} Solv GB	DG _{Bind} vdW
Rhamnocitrin	-49.53	-14.04	4.05	-1.48	-7.8	-4.39	12.69	-38.57
Isokaempferide	-45.34	-11.95	4.72	-1.35	-6.27	-4.66	14.89	-40.72
Kaempferol	-42.09	-12.94	3.53	-1.36	-5.68	-4.83	16.46	-37.28
Quercimeritrin	-47.93	-35.28	5.34	-4.07	-9.88	-3.31	35.39	-36.12
Apigenin	-39.67	-9.98	4.48	-1.35	-5.09	-4.88	13.62	-36.47
Penduletin	-46.7	-24.39	3.1	-1.73	-11.02	-3.18	25.17	-34.66
Isoquercitrin	-44.77	-9.6	2.05	-2.87	-10.85	-2.67	28.01	-48.84
Astragalin	-50.61	-11.26	2.98	-2.26	-11.52	-2.54	23.26	-49.26
Luteolin-7-glucoside	-44.92	-36.17	5.52	-2.87	-9.78	-2.85	37.88	-36.64
Isorhamnetin	-44.74	-13.02	3.87	-1.38	-6.31	-4.74	17.57	-40.73

Molecular modelling of Biological Interactions

Analysis of the 3D and 2D structures of the complexes formed by the top three compounds (Rhamnocitrin, Isokaempferide and Kaempferol) with SARS-CoV-2 Mpro showed that the compounds occupied the active site of the enzyme (figures 3 and 4). The three compounds, like the standard ligand interacted with active site amino acid residues like HIS 41, TRY 54, ASN 142, CYS 145, HIS 164, MET 165, GLU 166 and LEU 167, among others.

Pharmacophore modelling

The pharmacophore models of the standard ligand, Rhamnocitrin, Isokaempferide and Kaempferol on Mpro are shown in Figure 4. Two hydrogen bond donors contributed to the binding of the standard ligand to the enzyme while two aromatic rings and one hydrogen bond donor are involved in the molecular interactions of the three test compounds with the enzyme.

2.3 ADMET profile

As displayed on table 3, the molecular weights of the top ten *Artemisia annua* compounds are between 270.24 and 464.38. Log S values of the compounds ranged from -3.04 (soluble) to -4.16 (moderately soluble) and Log P values from -0.25 to 2.43. Rhamnocitrin, Isokaempferide, Kaempferol, Apigenin, Penduletin and Isorhamnetin have zero Lipinski violations and bioavailability score of 0.55. The remaining four compounds have 2 Lipinski violations and bioavailability score of 0.17. Gastrointestinal (GI) absorption potential is high for Rhamnocitrin, Isokaempferide, Kaempferol, Apigenin, Penduletin and Isorhamnetin, but low for the remaining compounds. None of the compounds show blood brain barrier (BBB) permeability and only Isorhamnetin is a

permeability glycoprotein (Pgp) substrate. For the cytochrome P450 inhibitory potencies, Rhamnocitrin, Isokaempferide, Kaempferol, Apigenin, Penduletin and Isorhamnetin are potential inhibitors of CYP1A2, CYP2D6 and CYP3A4. Penduletin is also a potential inhibitor of CYP2C9.

According to the ProTox II toxicity prediction, all the compounds fall under the oral toxicity class 5, with LD₅₀ ranging from 2,500 to 5000 mg/kg and none of the then show the tendency for hepatotoxicity, carcinogenicity, mutagenicity and cytotoxicity. Quercimeritrin, Penduletin, Isoquercitrin, and Isorhamnetin are however shown to be immunotoxic (table 4).

Table 3: Druglikeness and Pharmacokinetics Prediction of the ten top-scoring *Artemisia annua* compounds

	A	B	C	D	E	F	G	H	I	J
Molecular Weight	300.26	300.26	286.24	464.38	270.24	344.32	464.38	448.38	448.38	316.26
ESOL Log S	-3.51	-3.51	-3.31	-3.04	-3.94	-4.16	-3.04	-3.18	-3.65	-3.36
Solubility Class	Soluble	Soluble	Soluble	Soluble	Soluble	Moderately soluble	Soluble	Soluble	Soluble	Soluble
Mean Log P	1.98	1.94	1.58	-0.37	2.11	2.43	-0.25	-0.25	0.16	1.65
Lipinski violations	0	0	0	2	0	0	2	2	2	0
Bioavailability Score	0.55	0.55	0.55	0.17	0.55	0.55	0.17	0.17	0.17	0.55
GI absorption	High	High	High	Low	High	High	Low	Low	Low	High
BBB permeant	No	No	No	No	No	No	No	No	No	No
Pgp substrate	No	No	No	No	No	No	No	No	Yes	No
CYP1A2 inhibitor	Yes	Yes	Yes	No	Yes	Yes	No	No	No	Yes
CYP2C19 inhibitor	No	No	No	No	No	No	No	No	No	No
CYP2C9 inhibitor	No	No	No	No	No	Yes	No	No	No	No
CYP2D6 inhibitor	Yes	Yes	Yes	No	Yes	Yes	No	No	No	Yes
CYP3A4 inhibitor	Yes	Yes	Yes	No	Yes	Yes	No	No	No	Yes

A = 7-Methylkaempferol, B = 3-Methylkaempferol, C = Kaempferol, D = Quercimeritrin, E = Apigenin, F = Penduletin, G = Isoquercitrin, H = Astragalin, I = Luteolin-7-glucoside, J = Isorhamnetin

Table 4: Toxicity profile of the ten top-scoring *Artemisia annua* compounds

	A	B	C	D	E	F	G	H	I	J
LD₅₀ (mg/kg)	4000	3919	3919	5000	2500	5000	5000	5000	5000	5000
Toxicity Class	5	5	5	5	5	5	5	5	5	5
Hepatotoxicity	-	-	-	-	-	-	-	-	-	-
Carcinogenicity	-	-	-	-	-	-	-	-	-	-
Immunotoxicity	-	-	-	+	-	+	+	-	-	-
Mutagenicity	-	-	-	-	-	-	-	-	-	-
Cytotoxicity	-	-	-	-	-	-	-	-	-	-

Active = (+); Inactive = (-)

A = 7-Methylkaempferol, B = 3-Methylkaempferol, C = Kaempferol, D = Quercimeritrin, E = Apigenin

F =Penduletin, G = Isoquercitrin, H = Astragalin, I = Luteolin-7-glucoside, J = Isorhamnetin

Discussion

The main protease (Mpro) is an indispensable enzyme responsible for viral replication and it is among the top characterized drug targets of coronaviruses ³⁵. Upon host infection by SARS-CoV-2, replicase polyproteins 1a and 1ab are synthesized to produce various functional subunits essential for viral replication. This is accomplished by site specific hydrolysis of the polyproteins by two viral proteases, one of which is the Mpro ³⁶. Inhibition of this enzyme has been recognized to be an important strategy for therapeutic intervention against the SARS-CoV-2. In this study, in silico techniques were employed to identify possible SARS-CoV-2 inhibitors among the compounds of the *Artemisia annua*, an antimalarial plant with reported activity against the SARS-CoVs.

The compounds of *Artemisia annua* showed varying levels of binding affinities for the SARS-CoV-2 main protease in the molecular docking analysis, the highest being -7.83 kcal/mol by 7-Methylkaempferol (Rhamnocitrin) followed by -7.81 kcal/mol by 3-Methylkaempferol (Isokaempferide) and -7.65 by Kaempferol. It is worth noting that all the ten top-scoring compounds are flavonoids, which mainly include Kaempferol and its derivatives and Quercetin derivatives. Kaempferol, a tetrahydroxyflavone with its four hydroxy groups at positions 3, 5, 7 and 4' is a plant-derived aglycone flavonoid commonly found in medicinal herbs, fruits, seeds and vegetables. This phytochemical has been demonstrated to possess several pharmacological actions, among which are, antioxidant, anti-inflammatory, anticancer and antiviral activities ³⁷. Kaempferol has been found to display antiviral activity against several viruses including coronavirus ^{38,39}. A recent study by Xia *et al.* identified Kaempferol and Quercetin along with Luteolin as the main active ingredients of the Chinese herbal combination (Amygdalus Communis Vas and Ephedra sinica Stapf) used in the treatment of COVID-19 ⁴⁰. They also reported that these active ingredients demonstrated good binding affinity for SARS-CoV-2 Mpro ⁴⁰ as substantiated in this present study.

The reliability of the molecular docking result was determined by estimating the binding free energy through the Prime MM-GBSA module of Maestro. This approach is one of the most reliable methods of validating docking results as it helps to determine the stability of the receptor-ligand complex ^{41,42}. This denotes that a favorable binding free energy correlate to a reliable output from molecular docking study. This post docking analysis showed that the ten top-scoring phytochemical constituents of *Artemisia annua* possess rich binding free energy values (-39.67 kcal/mol and -50.61 kcal/mol) towards SARS-CoV-2 main protease. Studies have shown that the higher the binding free energy, the more favorable and stable the ligand-bound protein ^{43,44}. With this in mind, Astragalin showed the highest stability with the SARS-CoV-2 main protease, followed by Rhamnocitrin and Quercimeritrin. The ten selected compounds could however be said to form favorable complexes with the protein crystal structure of SARS-CoV-2 main protease. This can be further explained from the scatter plot on figure 2, as most of the points are close to the regression line, which is also an indication that the molecular docking result is comparable to a large extent with the binding free energy of the compounds.

The inhibitory potentials of *Artemisia annua* compounds against SARS-CoV-2 Mpro was further revealed by the interactions of Rhamnocitrin, Isokaempferide and Kaempferol with the active site amino acid residues of the enzyme (figures 3 and 4). HIS 41 and CYS 145 to which these compounds bind are known to play very crucial roles at the active site of the SARS coronavirus Mpro. Previous studies on the crystal structure of the SARS-CoV-2 main protease showed that the protein is composed of three domains: domain I comprising of residues 10–99, domain II comprising of residues 100–182 and domain III comprising of residues 198–303. The active site holds a HIS 41-CYS 145 catalytic dyad in a cleft between domains I and II, with CYS 145 acting as a nucleophile during the first step of the enzymatic process and HIS 41 acting as a base catalyst^{45,46}. Also of interest is the interaction of the compounds with GLU 166 which is essential for Mpro substrate-induced dimerization required for catalysis and ASN-142 which forms hydrogen bonding with GLU 166 to block the substrate-binding subsite entrance in the monomer. Mutation of GLU 166 is reported to significantly block the substrate-induced dimerization process, thereby preventing enzyme activation⁴⁷. Therefore, interaction of *Artemisia annua* Kaempferol and its 3 and 7 methyl derivatives with these amino acid residues, which are targets of most SARS-CoV-2 main protease inhibitors, further validates them as potential therapeutic agents against the virus.

The in silico pharmacophore modeling was employed to identify the structural features of the test compounds responsible for their affinity for the target protein. This is to further authenticate the inhibitory potentials of the compounds against SARS-CoV-2 main protease. Two aromatic rings and one hydrogen bond donor were identified as the structural features of Rhamnocitrin, Isokaempferide and Kaempferol responsible for their interactions with the enzyme. The involvement of aromatic rings, in addition to hydrogen bond formation, in the interaction of the compounds with the enzyme might have contributed to the higher binding affinity of the compounds compared with the standard ligand which depended on two hydrogen bond donors only. Aromatic rings are vital residues for molecular interactions and frequently exist in several protein–ligand and protein–protein interactions. Owing to their natural existence in amino acids residues like histidine, tryptophan, phenylalanine and tyrosine, they are considered to be very important for protein stability and molecular recognition processes. Furthermore, aromatic rings are frequently used in drug design due to their role in the improvement of binding affinity and specificity of drug-like molecules⁴⁸.

In spite of the Mpro inhibitory potential exhibited by the *Artemisia annua* compounds, the ADMET properties of the compounds is an essential factor that will determine their pharmacological activity against SARS-CoV-2. In silico ADMET prediction is a fast and low-cost approach to determine whether the compounds will be easily absorbable, well distributed to their target site of action, favorably metabolized and easily eliminated from the body without leaving toxic side effects⁴⁹. The Lipinski filter has been an effective method for screening potential drug candidates for oral drug-likeness based their molecular weights, hydrogen bond acceptors and donors and lipophilicity⁵⁰. Therefore, compounds like Rhamnocitrin, Isokaempferide, Kaempferol, Apigenin, Penduletin and Isorhamnetin with zero Lipinski violations are likely to be orally active and this is further validated by their oral bioavailability scores. The 0.55% bioavailability score signify that these compounds have about 55% probability of a minimum of 10% oral absorption in rat or human colon carcinoma absorptivity⁵¹. The remaining four compounds with 2 Lipinski violations and bioavailability score of 0.17 are not likely to be orally active. This is also reflected in the GI absorption potential of the compounds which is high for Rhamnocitrin, Isokaempferide, Kaempferol, Apigenin, Penduletin and Isorhamnetin, but low for the remaining compounds. However, none of the compounds show BBB permeability.

Despite its oral bioavailability potential as explained above, Isorhamnetin as a Pgp substrate is likely to be restricted from entering its target site of action. Pgp is a member of the ATP-binding cassette transporters that plays a role of active efflux of foreign chemicals through cell membranes in order to protect the body from these chemicals, and it is a major factor contributing to drug resistance. Hence, the bioavailability and consequent therapeutic effectiveness of Isorhamnetin might be limited by this factor. Additionally, the cytochrome P450 inhibitory potentials shown by Rhamnocitrin, Isokaempferide, Kaempferol, Apigenin, Penduletin and Isorhamnetin is an indication that the compounds could cause drug-drug interactions. This is because more than 50% of drugs are metabolized by these CYP isoforms and their inhibition, which could hinder the metabolism of these drugs, is a major cause of pharmacokinetics-related drug-drug interactions⁵².

It is important to note that the toxicity prediction of the compounds appear to be favorable (table 4). All the compounds belong to the oral toxicity class 5, and their LD₅₀ is between 2,500 and 5000 mg/kg, which means they could be safely used within these dosage limits. Furthermore, none of the compounds is likely to be hepatotoxic, carcinogenic, mutagenic and cytotoxic, making them relatively safe as potential therapeutic agents against SARS-CoV-2.

In conclusion, out of 168 bioactive compounds of *Artemisia annua* screened for possible inhibitory activity against SARS-CoV-2 main protease, Rhamnocitrin exhibited the highest binding affinity followed by Isokaempferide and Kaempferol. The three compounds, like the standard ligand occupied the active site of Mpro, where they interacted with important amino acid residues like HIS 41, ASN 142, CYS 145 and GLU 166, among others. Two aromatic rings and one hydrogen bond donor are involved in the molecular interactions of the three test compounds with the enzyme. The compounds also possess favorable ADMET profile and none of these three compounds showed the tendency for hepatotoxicity, carcinogenicity, mutagenicity, cytotoxicity and immunotoxicity. Therefore, these *Artemisia annua* compounds could be considered for experimental studies and further development into drugs for the treatment of SARS-CoV-2.

Methods

Protein preparation

The crystal structure of SARS-CoV-2 main protease (Mpro), also known as 3-Chymotrypsinlike protease (PDB ID: 7C6U) was gotten from Protein Data Bank (PDB) repository. Using Glide (Schrödinger Suite 2020-3), the protein structure obtained was prepared via the protein preparation wizard panel. During the preparation processes, hydrogen was added, bond orders were allocated, disulfide bonds were generated, the side chains and loops that are missing were replaced using prime. Water molecules located outside 3.0 Å of the heteroatoms were removed and the protein structure was minimized using OPLS3e and optimized using PROPKA^{43,53}.

Generation of receptor grid

The receptor grid was created to outline the position and size of the protein's active site for ligand docking. This was done using the receptor grid generation tool of Schrödinger Maestro 12.5. The position of the co-crystallized ligand (K36) in the active site of the protein was used as the scoring grid.

Ligand preparation

One hundred and sixty-eight (168) bioactive compounds of *Artemisia annua* obtained from Dr. Duke's Phytochemical and Ethnobotanical Databases and the standard ligand ((1S,2S)-2-((N-((benzyloxy)carbonyl)-L-leucyl)amino)-1-hydroxy-3-[(3S)-2-oxopyrrolidin-3-yl]propane-1-sulfonic acid or K36) were prepared using Lipprep panel of Maestro 12.5, Schrödinger Suite 2020-3, for the molecular docking analysis. The ligands were prepared to obtain low-energy 3D structures with appropriate chiralities. The ionization state for each ligand structure was generated at a physiological pH of 7.2 ± 0.2. Stereoisomers of each ligand were computed by retaining specified chiralities while others were varied.

Protein-Ligand Docking

The molecular docking analysis was carried out using Glide-Ligand Docking panel of Maestro 12.5 on Schrödinger Suite 2020-3. The prepared ligands and the receptor grid file were imported into the work space of Maestro, using standard precision (SP) docking, the ligands were docked into the binding pocket of the target protein. The vdW radius scaling factor was scaled at 0.80 with a partial charge cut-off of 0.15 for ligand atoms and the ligand sampling method was set to be flexible.

Binding free energy calculation

MM-GBSA Prime panel of the Schrödinger Suite 2020-3 was used to estimate the binding free energy of the receptor-ligand complex. Following energy minimization, MM-GBSA measures the change in energy amid the free and the complex states of both the test compound and the protein.

Receptor-ligand complex pharmacophore modelling

The PHASE module of the Schrödinger Suite 2020-3 was used to generate an auto/e-pharmacophore model as previously described⁵⁴. The process was set to generate a maximum of 7 features at a minimum feature-feature distance of 2.00 and the minimum distance between features of the same type was set at 4.00 and donors were set as vectors.

3.8 ADMET Profiling

The ADMET properties of the test compounds were determined using in silico predictive models. The SwissADME server was used to determine the ADME properties of the compounds which include: lipophilicity (Log P), water solubility (ESOL Log S), drug-likeness based on Lipinski rule and bioavailability score; and pharmacokinetics based on gastrointestinal (GI) absorption, blood brain barrier (BBB) permeant, permeability glycoprotein (Pgp) binding and Cytochrome P450 (CYP) inhibition. The ProTox-II online server was used to predict the acute toxicity class, LD₅₀, hepatotoxicity, carcinogenicity, mutagenicity, cytotoxicity and immunotoxicity of the compounds.

Declarations

Author contributions

TOJ and AEA initiated the project and all authors contributed to the design of the work. AEA generated the data; TOJ, OAO, AYJ and OI interpreted the results; TOJ, OAO, OI, CJU, ROA and IFC prepared the first draft of the manuscript and all authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Data availability

All data generated and analyzed during this study are included in this published article and its Supplementary Information files.

References

1. Liu, Y., Gayle, A. A., Wilder-Smith, A. & Rocklöv, J. The reproductive number of COVID-19 is higher compared to SARS coronavirus. *J. Travel Med.* (2020) doi:10.1093/jtm/taaa021.
2. Wang, D. *et al.* Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus–Infected Pneumonia in Wuhan, China. *JAMA* **323**, 1061–1069 (2020).
3. Cucinotta, D. & Vanelli, M. WHO declares COVID-19 a pandemic. *Acta Biomed.* **91**, 157–160 (2020).
4. Shereen, M. A., Khan, S., Kazmi, A., Bashir, N. & Siddique, R. COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses. *J. Adv. Res.* **24**, 91–98 (2020).
5. Yang, D. & Leibowitz, J. L. 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. *Virus Res.* **206**, 120–133 (2015).
6. WHO. WHO Coronavirus (COVID-19) Dashboard. WHO Coronavirus (COVID-19) Dashboard With Vaccination Data. *WHO* 1–5 (2021).
7. Lim, J. *et al.* Case of the index patient who caused tertiary transmission of coronavirus disease 2019 in Korea: The application of lopinavir/ritonavir for the treatment of COVID-19 pneumonia monitored by quantitative RT-PCR. *J. Korean Med. Sci.* **35**, 1–6 (2020).

8. Wang, M. *et al.* Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res.* **30**, 269–271 (2020).
9. Aruleba, R. T., Adekiya, T. A., Oyinloye, B. E. & Kappo, A. P. Structural Studies of Predicted Ligand Binding Sites and Molecular Docking Analysis of Slc2a4 as a Therapeutic Target for the Treatment of Cancer. *Int. J. Mol. Sci.* **19**, (2018).
10. Ojo, O. A. *et al.* Deciphering the interaction of puerarin with cancer macromolecules: An in silico investigation. *J. Biomol. Struct. Dyn.* **0**, 1–12 (2020).
11. Cazarolli, L. *et al.* Flavonoids: Prospective Drug Candidates. *Mini-Reviews Med. Chem.* **8**, 1429–1440 (2008).
12. Harborne, J. B. & Williams, C. A. Advances in flavonoid research since 1992. *Phytochemistry* **22**, 481–505 (2000).
13. Costa, A. N., de Sá, É. R. A., Bezerra, R. D. S., Souza, J. L. & Lima, F. das C. A. Constituents of buriti oil (*Mauritia flexuosa* L.) like inhibitors of the SARS-Coronavirus main peptidase: an investigation by docking and molecular dynamics. *J. Biomol. Struct. Dyn.* (2020) doi:10.1080/07391102.2020.1778538.
14. Das, S., Sarmah, S., Lyndem, S. & Singha Roy, A. An investigation into the identification of potential inhibitors of SARS-CoV-2 main protease using molecular docking study. *J. Biomol. Struct. Dyn.* **0**, 1–11 (2020).
15. Ghosh, R., Chakraborty, A., Biswas, A. & Chowdhuri, S. Evaluation of green tea polyphenols as novel corona virus (SARS CoV-2) main protease (Mpro) inhibitors—an in silico docking and molecular dynamics simulation study. *J. Biomol. Struct. Dyn.* **0**, 1–13 (2020).
16. Umesh, U., Kundu, D., Selvaraj, C., Singh, S. K. & Dubey, V. K. Identification of new anti-nCoV drug chemical compounds from Indian spices exploiting SARS-CoV-2 main protease as target. *J. Biomol. Struct. Dyn.* **0**, 1–7 (2020).
17. Soleimani, T., Keyhanfar, M., Piri, K. & Hasanloo, T. Morphological evaluation of hairy roots induced in *Artemisia annua* L. and investigating elicitation effects on the hairy roots biomass production. *Int. J. Agric. Res. Rev.* **2**, 1005–1013 (2012).
18. Zanjani, K. E. *et al.* Physiological Response of Sweet Wormwood to Salt Stress under Salicylic Acid Application and Non Application Condition. *Life Sci. J.* **9**, 4190–4195 (2012).
19. Huang, L., Xie, C., Duan, B. & Chen, S. Mapping the potential distribution of high artemisinin-yielding *Artemisia annua* L. (Qinghao) in China with a geographic information system. *Chin. Med.* **5**, 1–8 (2010).
20. Emadi. *Phytochemistry of Artemisia annua*. (2013).
21. Lutgen, P. La tisane d'artemisia annua, une puissante polythérapie! *Mal. Trop. Asp. Humanit. Sci.* **25**, 6–7 (2009).
22. Nkuitchou-Chougouo, R. D. K. *et al.* Comparative study of chemical composition of *Artemisia annua* Essential oil growing wild in Western Cameroon and Luxembourg by μ -CTE/TD/GC/MS. *North Asian Int. Res. J. Res. J. Consort.* **2**, (2016).
23. Li, S. Y. *et al.* Identification of natural compounds with antiviral activities against SARS-associated coronavirus. *Antiviral Res.* **67**, 18–23 (2005).
24. Brown, G. D. Two new compounds from *artemisia annua*. *J. Nat. Prod.* **55**, 1756–1760 (1992).
25. Brisibe, E. A. *et al.* Nutritional characterisation and antioxidant capacity of different tissues of *Artemisia annua* L. *Food Chem.* **115**, 1240–1246 (2009).
26. Lai, J. P., Lim, Y. H., Su, J., Shen, H. M. & Ong, C. N. Identification and characterization of major flavonoids and caffeoylquinic acids in three Compositae plants by LC/DAD-APCI/MS. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **848**, 215–225 (2007).
27. Bhakuni, R. S., Jain, D. C., Sharma, R. P. & Kumar, S. Secondary metabolites of *Artemisia annua* and their biological activity. *Curr. Sci.* **80**, 35–48 (2001).

28. Meng, X.-Y., Zhang, H.-X., Mezei, M. & Cui, M. Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Curr. Comput. Aided-Drug Des.* **7**, 146–157 (2012).
29. Gupta, M. K. *et al.* In-silico approaches to detect inhibitors of the human severe acute respiratory syndrome coronavirus envelope protein ion channel. *J. Biomol. Struct. Dyn.* **39**, 2617–2627 (2020).
30. Ojo, Oluwafemi Adeleke *et al.* Deciphering the Interactions of Bioactive Compounds in Selected Traditional Medicinal Plants against Alzheimer's Diseases via Pharmacophore Modeling, Auto-QSAR, and Molecular Docking Approaches. *Molecules* **26**, 1996 (2021).
31. Bhattacharya, M. *et al.* Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-CoV-2): Immunoinformatics approach. *J. Med. Virol.* **92**, 618–631 (2020).
32. Adegboyega, A. E., Johnson, T. O. & Omale, S. Computational Modelling of the Pharmacological actions of some anti-viral agents against SARS-CoV-2. in *Data Science for COVID-19* (eds. Kose, U., Gupta, D., Albuquerque, V. de & Khanna, A.) 467–482 (Elsevier, 2021).
33. Alonso, H., Bliznyuk, A. A. & Gready, J. E. Combining docking and molecular dynamic simulations in drug design. *Med. Res. Rev.* **26**, 531–568 (2006).
34. Wang, G. & Zhu, W. Molecular docking for drug discovery and development: A widely used approach but far from perfect. *Future Med. Chem.* **8**, (2016).
35. Suárez, D. & Díaz, N. SARS-CoV-2 Main Protease: A Molecular Dynamics Study. *J. Chem. Inf. Model.* **60**, 5815–5831 (2020).
36. Ziebuhr, J. Molecular biology of severe acute respiratory syndrome coronavirus. *Curr. Opin. Microbiol.* **7**, 412–419 (2004).
37. Imran, M. *et al.* Kaempferol: A key emphasis to its anticancer potential. *Molecules* **24**, 1–16 (2019).
38. Schwarz, S. *et al.* Kaempferol derivatives as antiviral drugs against the 3a channel protein of coronavirus. *Planta Med.* **80**, 177–182 (2014).
39. Ahmadian, R., Rahimi, R. & Bahramsoltani, R. Kaempferol: An encouraging flavonoid for COVID-19. *Bol. Latinoam. y del Caribe Plantas Med. y Aromat.* **19**, 492–494 (2020).
40. Xia, S. *et al.* The important herbal pair for the treatment of COVID-19 and its possible mechanisms. *Chinese Med. (United Kingdom)* **16**, 1–16 (2021).
41. Kikiowo, B., Ogunleye, J. A., Iwaloye, O. & Ijatuyi, T. T. Therapeutic potential of Chromolaena odorata phyto-constituents against human pancreatic α -amylase. *J. Biomol. Struct. Dyn.* **0**, 1–12 (2020).
42. Elekofehinti, O. O. *et al.* Molecular docking studies, molecular dynamics and ADME/tox reveal therapeutic potentials of STOCK1N-69160 against papain-like protease of SARS-CoV-2. *Mol. Divers.* (2020) doi:10.1007/s11030-020-10151-w.
43. Iwaloye, O., Elekofehinti, O. O., Momoh, A. I., Babatomiwa, K. & Ariyo, E. O. In silico molecular studies of natural compounds as possible anti-Alzheimer's agents: ligand-based design. *Netw. Model. Anal. Heal. Informatics Bioinforma.* **9**, 54 (2020).
44. Iwaloye, O., Elekofehinti, O. O., Oluwarotimi, E. A., Kikiowo, B. iwa & Fadipe, T. M. Insight into glycogen synthase kinase-3 β inhibitory activity of phyto-constituents from Melissa officinalis: in silico studies. *Silico Pharmacol.* **8**, 1–13 (2020).
45. Zhang, L. *et al.* Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α -ketoamide inhibitors. *Science (80-.)*. **368**, 409–412 (2020).
46. Jin, Z. *et al.* Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature* **582**, 289–293 (2020).

47. Cheng, S. C., Chang, G. G. & Chou, C. Y. Mutation of glu-166 blocks the substrate-induced dimerization of SARS coronavirus main protease. *Biophys. J.* **98**, 1327–1336 (2010).
48. Lanzarotti, E., Defelipe, L. A., Marti, M. A. & Turjanski, A. G. Aromatic clusters in protein-protein and protein-drug complexes. *J. Cheminform.* **12**, 1–9 (2020).
49. Ntie-Kang, F. An in silico evaluation of the ADMET profile of the StreptomeDB database. *Springerplus* **2**, 1–11 (2013).
50. Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **64**, 4–17 (2012).
51. Testa, B. & Krämer, S. D. The biochemistry of drug metabolism - An introduction part 5. Metabolism and bioactivity. *Chem. Biodivers.* **6**, 591–684 (2009).
52. Daina, A., Michielin, O. & Zoete, V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* **7**, 1–13 (2017).
53. Olsson, M. H. M., SØndergaard, C. R., Rostkowski, M. & Jensen, J. H. PROPKA3: Consistent treatment of internal and surface residues in empirical p K a predictions. *J. Chem. Theory Comput.* **7**, 525–537 (2011).
54. Johnson, T. O. *et al.* Computational study of the therapeutic potentials of a new series of imidazole derivatives against SARS-CoV-2. *J. Pharmacol. Sci.* 1–10 (2021) doi:10.1016/j.jphs.2021.05.004.

Figures

Figure 1

The ten top-scoring compounds of *Artemisia annua* against SARS-CoV-2 main protease

Figure 3

The 3D view of the molecular interaction of a standard inhibitor A = K36, A = Rhamnocitrin, C = Isokaempferide, D = Kaempferol with SARS-CoV-2 main protease

Figure 4

The 2D view of the molecular interaction of A = K36, B = Rhamnocitrin, C = Isokaempferide, D = Kaempferol with SARS-CoV-2 main protease

Figure 5

The receptor-ligand complex pharmacophore models of A = K36, A = 7-Methylkaempferol, C = 3-Methylkaempferol, D = Kaempferol on SARS-CoV-2 main protease

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryinformationSR.docx](#)