Article

Not peer-reviewed version

In Silico and In Vitro Studies of Terpenes from the Fabaceae Family Using the Phenotypic Screening Model against the SARS-CoV-2 Virus

Natália Ferreira de Sousa , Gabrielly Diniz Duarte , Carolina Borsoi Moraes , Cecília Gomes Barbosa , Holli-Joi Martin , Nail N. Muratov * , Yuri Mangueira do Nascimento , Luciana Scotti , Lúcio Holanda Gondim de Freitas-Júnior , José Maria Barbosa Filho , Marcus Tullius Scotti *

Posted Date: 11 June 2024

doi: 10.20944/preprints202406.0716.v1

Keywords: Virtual Screening; SARS-CoV-2; Natural products; Phenotypic screening; Fabaceae family

 doi:10.20944/preprints202406.0716.v1

Article

In silico and In Vitro Studies of Terpenes from the Fabaceae Family Using the Phenotypic Screening Model against the SARS-CoV-2 Virus

Natália Ferreira de Sousa [1], Gabrielly Diniz Duarte [2], Carolina Borsoi Moraes [3], Cecília Gomes Barbosa [3], Holli-Joi Martin [4], Nail N. Muratov [5,6], Yuri Mangueira do Nascimento [1], Luciana Scotti [1], Lúcio Holanda Gondim de Freitas-Júnior [3], José Maria Barbosa Filho [1] and Marcus Tullius Scotti [1,*]

[1] Postgraduate Program in Natural and Synthetic Bioactive Products, Federal University of Paraíba, João Pessoa 58051-900, Brazil; nataliafsousa@ltf.ufpb.br (N.F.S.); yurimangueira@ltf.ufpb.br (Y.M.N.); luciana.scotti@gmail.com (L.S); barbosa.ufpb@gmail.com (J.M.B.F.).
[2] Postgraduate Program in Development and Innovation of Drugs and Medicines, Federal University of Paraíba, João Pessoa 58051-900, Brazil; gabriellydduarte@gmail.com (G.D.D.).
[3] Institute of Biomedical Sciences of the University of São Paulo (ICB-USP), São Paulo-SP 05508-000, Brazil; cbmoraes@unifesp.br (C.B.M.); cecigomes.barbosa@gmail.com (C.G.B); luciofreitasjunior@gmail.com (L.H.G.F.J).
[4] Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, USA. holli27@unc.edu (H.J.M).
[5] Odessa National Polytechnic University, Odessa, 65000, Ukraine. nail_muratov@ukr.net (N.N.M)
[6] A.V. Bogatsky Physical-Chemical Institute NASU, Odessa, 65047, Ukraine
* Correspondence: mtscotti@gmail.com* (M.T.S.); Tel.: +55-83-99869-0415 (M.T.S).

Abstract: In 2019, the emergence of the seventh known coronavirus to cause severe illness in humans triggered a global effort towards the development of new drugs and vaccines for the SARS-CoV-2 virus. These efforts are still ongoing in 2024, including the present work where we conducted a ligand-based virtual screening of terpenes with potential anti-SARS-CoV-2 activity. We constructed a Quantitative Structure Activity Relationship (QSAR) model from compounds with known activity against SARS-CoV-2 with a model accuracy of 0.71. We utilized this model to predict the activity of a series of 217 terpenes isolated from the Fabaceae Family. Four compounds, predominantly triterpenoids from the lupane series, were subjected to an *in vitro* phenotypic screening in Vero CCL-81 cells to assess their inhibitory activity against SARS-CoV-2. The compounds which showed high rates of SARS-CoV-2 inhibition along with substantial cell viability underwent Molecular Docking at the SARS-CoV-2 Main-protease, Papain-like protease, Spike protein and RNA-dependent RNA polymerase. Overall, virtual screening through our QSAR model successfully identified compounds with the highest probability of activity, as validated by the *in vitro* study. This confirms the potential of the identified triterpenoids as promising candidates for anti-SARS-CoV-2 therapeutics.

Keywords: virtual screening; SARS-CoV-2; natural products; phenotypic screening; Fabaceae family

1. Introduction

The Coronavirus Disease 2019 (COVID-19) pandemic has had a profound global impact on global human health, with over 774 million confirmed cases and over 7 million deaths worldwide as of March 2024 [1]. This viral disease demonstrates a wide range of symptoms, from mild to severe pneumonia and acute respiratory distress, to a multi-organ disorder affecting various systems, including the pulmonary, cardiovascular, neurologic, renal, endocrine, dermatologic, and gastrointestinal systems [2]. The virus has undergone continuous evolutionary adaptations since its emergence which impact its interactions with our biological systems, enabling the virus to bypass immune defenses while diminishing its virulence. Such evolution has been observed in variants as Omicron BA.1 and others, with extensive sequencing and data analysis conducted over the past four years [3–5].

 doi:10.20944/preprints202406.0716.v1

Since its initial emergence in Wuhan, China, in 2019, the management of COVID-19 has significantly changed, transitioning from early strategies such as social distancing and lockdowns to more advanced approaches, including intensive care, vaccination campaigns, and the development of antiviral drugs and monoclonal antibodies [6].The SARS-CoV-2 virus, responsible for COVID-19, comprises a structural spike protein which is pivotal in its binding with host cell receptors to initiate the viral life cycle [7]. Remarkably, the virus can mutate its spike proteins to evade host defenses, presenting challenges for vaccinated individuals and contributing to recurrent infections [8]. Conversely, in anti-SARS-CoV-2 drug discovery, many studies have concentrated on attacking SARS-CoV-2 with small molecule inhibitors that block viral proteases and polymerases, including RNA-dependent RNA polymerase (RdRp) [9], the main protease (Mpro or 3CLpro) and the papain-like protease (PLpro) [10] which are instrumental in the development of novel compounds [11]. One popular method for testing in vitro activity of these compounds is phenotypic screening based on the activation of caspase 3/7 in Vero cells [12,13].

In addition to the development of pharmaceutical drugs and monoclonal antibodies, reports have surfaced from Zimbabwe [14], Nigeria [15], and India [16], suggesting the efficacy of medicinal plants in managing COVID-19 symptoms. These studies often involve the utilization of crude plant extracts or purified compounds from the plant families Fabaceae and Lamiaceae [17]. The Fabaceae family is considered the second most diverse and economically important plant family, and includes several medicinally significant plants known for their antimicrobial, anticancer, antibacterial, diuretic, and anti-inflammatory properties such as *Melilotus officinalis* (MO), *Coronilla varia* (CV), *Ononis spinosa* (OS), and *Robinia pseudoacacia* (RP) [18–23]. These are particularly intriguing because they contain an abundance of secondary metabolites, some of which have demonstrated pharmacological activity *in vitro* [24]. Among these compounds are terpenes, a class of natural volatile compounds with more than 80,000 screened for potential therapeutic applications [25], including antiviral activity against various Human Coronaviruses (HCoVs) [26].

This study aims to explore the anti-SARS-CoV-2 potential of terpenes isolated from the Fabaceae family. To do this we created an *in silico* Quantitative Structure Activity Relationship (QSAR) model from compounds with known activity against SARS-CoV-2. We utilized this model to predict the activity of a series of 217 terpenes isolated from the Fabaceae Family and subjected promising compounds to an *in vitro* phenotypic screening in Vero CCL-81 cells to assess their inhibitory activity against SARS-CoV-2. The compounds which showed high rates of SARS-CoV-2 inhibition along with substantial cell viability underwent Molecular Docking at the SARS-CoV-2 Main-protease, Papain-like protease, Spike protein and RNA-dependent RNA polymerase.

2. Materials and Methods

2.1. Extraction of Compounds the Study

Betulinic acid was extracted from the bark of *Zizhyphus joazeiro* Mart. (Rhamnaceae) using a method previously described by Barbosa and collaborators in 1985 [27,28]. Column chromatography was employed for the extraction, and acid hydrolysis was performed. The isolation process involved comparing physical properties obtained through various spectrometric methods (Infrared Spectrometry (IV), Ultraviolet Spectrometry (UV), Mass Spectrometry (MS), Hydrogen Nuclear Resonance Spectrometry (H-NMR) and Carbon Nuclear Resonance Spectrometry (C-NMR)).

Lupeol was extracted from the bark of *Lonchocarpus araripensis* Benth. (Fabaceae) using a method previously outlined [29,30] by Barbosa and collaborators (2013). The fractions corresponding to the crude ethanolic extract were monitored by Analytical Thin Layer Chromatography, with Lupeol being identified using nuclear magnetic resonance spectroscopic data and comparison with values reported in the literature. The substances Betulinic Acid Acetate and Betulinic Acid Methyl Ester were purchased from the company Sigma Aldrich (https://www.sigmaaldrich.com/BR/pt).

2.2. Data Collection and Curation

The CHEMBL database (EMBL-EBI, Wellcome Genome Campus, Cambridgeshire, England; https://www.ebi.ac.uk/chembl/) was used to extract 412 compounds with reported *in vitro* anti-SARS-

 doi:10.20944/preprints202406.0716.v1

CoV-2 activity (pIC_{50}) (ChEMBL ID: 4303835 - organism) in Vero E6 cells, Vero C1008 Cells, A549-ACE2, Caco-2 , Huh-7 and Calu-3 in various assays. We used a binary classification system where compounds with reported pIC50 ≥ 5.9355 were considered active (105 compounds), while compounds with pIC50 ≤ 5.8894 were classified as inactive (305 compounds). We used a standard 80-20 split for the training to test set, where 330 compounds constituted the training series, and 82 compounds formed the test series. We used a 10-fold cross-validation for our training set, meaning that 10% of the training series (33 compounds) was left out for internal validation and this was repeated 10 times, with a new set of 33 compounds used for training each iteration. The cross validation was employed using a stratified approach, ensuring that the proportion of active and inactive compounds was maintained during the removal process. We used the external test set of 82 compounds (21 active compounds and 61 inactive compounds) to validate our model and generate our models statistics.

To obtain our prediction set, we queried the Web of Science database (https://www.webofscience.com/wos/woscc/basic-search) using the keywords "Fabaceae", "terpene" and "*Leguminosae*". We collected data from 77 articles published between 1991 and 2023 and identified 217 compounds derived from natural products belonging to the class of terpenes found in the Fabaceae Family (*Leguminosae*) (Table S1 – Supplementary Material). We cross referenced this list of compounds with those used to develop our QSAR model to ensure there was no overlapping compounds.

The compounds chemical structures were designed using the Marvin Sketch 18.14 software program 2017 by ChemAxon (https://chemaxon.com/), and subsequently converted into SMILES. We used Chemaxon Standardizer v.18.17.0, (ChemAxon, Boston, USA, www.chemaxon.org) to transform the chemical structures into a uniform representation to avoid inconsistencies. This process included the addition of explicit hydrogen atoms, neutralizing charged fragments or functional groups, recognizing and converting legacy representations of functional groups (like aliases), removing water and salt counterions, expanding abbreviated groups, and conversion to 3D representation. The tool also ensures a unified representation of aromatic rings, tautomers and mesomers [31,32]. We compiled these compounds in a database and integrated it into the Sistemat X Web platform (https://sistematx.ufpb.br/) (Table S1 - Supplementary material).

2.3. QSAR Modelling

The Knime 3.6.2 software (Knime 3.6.2, Copyright Miner, de Konstanz Information, Zurich, Suíça, www.knime.org) was employed to build and evaluate the QSAR models. Given the success of previous studies conducted by our group [33,34], we utilized 3D QSAR analysis. To accomplish this, all compounds were converted into 3D structures, saved in SDF format, and then imported into AlvaDesc descriptors (https://www.alvascience.com/alvadesc/) [35,36] to obtain the necessary descriptors.

The Random Forest (RF) algorithm was selected to build the predictive model. The applicability domain was calculated based on the Euclidean distances present in the surveyed chemical space [34]. External cross-validation was conducted to estimate the predictive power of the developed model. The performance of the external predictions were assessed through Area Over the Curve analysis (ROC). Additionally, the models underwent analysis using the Matthews Correlation Coefficient (MCC) confusion matrix to provide a comprehensive evaluation of the model's effectiveness.

2.4. Biological Assays

2.4.1. Cell Line

Vero cells (CCIAL 057) were obtained from the "Núcleo de Cultura de Células - Instituto Adolfo Lutz, São Paulo, Brazil". The cells were cultured in high glucose DMEM medium (Sigma-Aldrich), supplemented with 10% healt-inactivated fetal bovine serum (FBS) (Thermo Scientific, USA) and 100 U mL^{-1} Streptomycin (Thermo Scientific) at 37 °C with 5% CO_2.

2.4.2. Virus Strain

 doi:10.20944/preprints202406.0716.v1

All procedures involving the SARS-CoV-2 virus were performed in the level 3 biosafety laboratory of the Institute of Biomedical Sciences of the University of São Paulo. The SARS-CoV-2 virus used in this study (HIAE-02: SARS-CoV-2/SP02/human/2020/ARB, GenBank Accession No.MTI26808.1) was isolated from a nosopharyngeal sample of a confirmed COVID-19 patient at Hospital Israelita Albert Einstein, São Paulo (SP) Brazil.

2.4.3. Phenotypic Screening with SARS-CoV-2

Test compounds were initially diluted to a concentration of 2 mg/mL in Dimethyl sulfoxide (DMSO). Subsequently, these compounds were tested at a single concentration of 10 µg/mL. Before treating the cells, the compounds underwent a 33.33x dillution in Phosphate-buffered saline (PBS), and 10 µL of each dilution were transferred to the assay plates, resulting in a final dilution factor of 200x. The tests were conducted in duplicate, with chloroquine utilized as a control [13].

For the phenotypic screening of compounds, 2000 Vero cells were seeded per well in 384-well plates in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% heat-inativacted fetal bovine serum (Thermo Scientific). The cells were incubated at 37 °C with 5% CO_2. After 24 h, the cells were treated with the compounds as described above, followed by the addition of the virus at a Multiplicity of Infection (MOI) of 0.1 viral particles per cell. The final concentration of DMSO in the assay plates was 0.5% (v/v). After 33 hours, the plates were fixed, immunofluorescence was performed using serum from Covid-19 patients, and the images were acquired and analyzed by the Operetta High-Content Analysis System (HCS) equipment [13].

Images were subjected to automated detection of infected and non-infected cells. The parameters measured in each well included the total number of cells and total number of infected cells. Infection was quantified as the percentage of infected cells relative to the total number of cells. The reduction in the number of infected cells reflected the percentage of antiviral activity exhibited by the samples. The activity of each compound was normalized against infected and uninfected controls, as was the cell survival rate. The cell survival rate is expressed as the percentage of the number of cells in the test well in relation to the average number of cells in the infected control wells [13].

The EC_{50} was defined as the compound concentration causing a 50% reduction in viral infection compared to infected controls. The CC_{50} value was defined as the compound concentration causing a 50% reduction in cell survival compared to infected controls. The Selectivity Index was calculated as the ratio between CC_{50} and EC_{50} (CC_{50}/EC_{50}). Maximum Activity represented the maximum inhibition of infection observed compared to controls.

2.5. Molecular Docking Studies

Proteins were downloaded from the Protein Data Bank (PDB) library (https://www.rcsb.org/) [37]. The targets were selected through bibliographical research concerning the mechanism of action involved in the inhibition of the SARS-CoV-2, taking into account their structural similarity. The obtained structures were as follows: Main-protease (M-pro) in complex with NCL-00024905 (PDB: 5RG1), resolution: 1.65 Å and method: X-Ray Diffraction [38]; Papain-like Protease (PL-pro) in complex with inhibitor 3k (PDB: 7TZJ), resolution: 2.66 Å and method: X-Ray Diffraction [39] Spike glycoprotein in complex with the 10D12 heavy-chain-only antibody (local refinement) (PDB: 8C8P), resolution: 4.10 Å and method: Electron microscopy and SARS-CoV-2 RNA dependent RNA Polymerase in complex with cofactors (PDB: 6M71) [40], resolution: 2.9 Å.

The active binding sites of the proteins were determined based on a literature search and were included in the docking study [41]. The active site was defined based on the active site information available in the referenced articles. For proteins: M-protease (PDB: 5RG1) and PL-protease (PDB: 7TZJ), which had a co-crystallized ligand, the active site was defined through the template established by the coordinates of the ligand in contact with the protein. Regarding the Spike glycoprotein (PDB: 8C8P) and the RNA dependent RNA Polymerase (PDB: 6M71), which did not have a co-crystallized ligand, the active site was determined using molecular pocket predictions from the platform Bite Net - Skolteck I Molecule, 2023 (https://sites.skoltech.ru/imolecule/tools/bitenet). These predictions indicated that the active site comprises the region equivalent to the terminal end of the A subunit.

 doi:10.20944/preprints202406.0716.v1

Furthermore, for the Spike protein, the drug Nirmatrelvir was employed as a positive control [42] and for the RNA dependent RNA Polymerase, the drug Rendesivir was employed as a positive control [40].

Redocking was conducted as a preliminary step to validate the docking simulation. Both procedures were performed using Molegro Virtual Docker (MVD) v.6.0.1 software [43]. Enzymes and compounds were prepared according to predefined parameters within the software.

In the coupling procedure (linker-enzyme) a Grid of 15 Å radius and a resolution of 0.30 was utilized. This grid emcompassed the binding site, as defined by a known ligand for each enzyme. A model was generated to perform and evaluate the fit with expected characteristics between the ligand and the enzyme, using the MOLDOCK Score (GRID) algorithm with the scoring function and search algorithm, corresponding to Moldock. The MolDock scoring function enhances these scoring functions with a new hydrogen bonding term and new charge schemes. The docking scoring function, Escore, is defined by the following energy terms:

Escore = Einter + Eintra

The visualization of the established interactions was performed in the Discovery Studio Visualizer program, Biovia, 2020 (https://www.3dsbiovia.com/) [44].

3. Results

3.1. Compounds in Study

The study comprises 217 natural products classified as terpenes, including sesquiterpenes, diterpenes, monoterpenes and triterpenes, occurring in the Fabaceae family (*Leguminosae*). These compounds were identified through an exhaustive literature search conducted on the Web of Sciences database (https://www.webofscience.com/wos/woscc/basic-search). Upon compiling the database, it was deposited on the Sistemat X web platform (https://sistematx.ufpb.br/). For a detailed description of the compounds under investigation, obtained through the literature review, please refer to Table S1 of the supplementary material (TS1 – Supplementary material).

3.2. Quantitative Structure-Activity Relationship (QSAR) Modeling

A classification model was developed for ligand-based virtual screening, employing the Random Forest (RF) algorithm. The physicochemical properties of the compounds were determined by AlvaDesc descriptors (https://www.alvascience.com/alvadesc/) [35,36]. The developed model underwent validation, and its predictive capacity was assessed using parameters such as specificity, sensitivity, accuracy and precision. The performance and robustness of the models were appraised through the Receiver Operating Characteristic Curve (ROC). Table 1 provides detailed information on the parameters of the model created with the AlvaDesc descriptors.

Table 1. Summary of parameters corresponding to the results obtained for all models with VolSurf descriptors.

Specie	Validation	Specificity	Sensitivity	Accuracy	Precision	Recall	ROC	MCC
SARS-CoV-2	Test	0.75	0.759	0.75	0.75	0.75	0.855	0.5
	Cross	0.702	0.729	0.716	0.716	0.729	0.802	0.431

Our model underwent cross-validation that affirmed excellent performance results with accuracy values exceeding 70%. The ROC Curve values for the developed model were greater than 0.80, signifying a robust and predictive model. The model demonstrated high specificity, with values of 0.75 (Test) and 0.702 (Cross). Similarly, high and satisfactory values for sensitivity were observed, as these corresponded to 0.759 (Test) and 0.716 (Cross). In general, it was observed that the model presents a good prediction as it presented Matthews Correlation Coefficient (MCC) values corresponding to 0.5 (Test) and 0.431 (Cross).

The model built on VolSurf descriptors for SARS-CoV-2 (Table S2 – Supplementary material) successfully predicted most compounds within the applicability domain, with the exception of 119

 doi:10.20944/preprints202406.0716.v1

compounds (1-17, 19, 24-25, 27, 30-34, 36-51, 53-59, 62, 64, 66-70, 72-75, 80, 82-102, 105-118, 130, 133-134, 139-142, 144-146, 148-151, 163-165, 175-176, 186, 188-189, 213 and 214) (Table S2 and Figure S1 – Supplementary Material). Furthermore, the model classified 85 compounds with a probability of activity above 50% and the applicability domain reliable, with probability values corresponding to 0.5 to 0589 (Table S3 and Figure S2 - Supplementary Material). These selected compounds corresponded to 168, 204 and 2015 (p = 0.589); 123, 172, 194, 198, 206, 210 and 211 (p = 0.579), 76, 77, 120, 138, 157, 162, 170, 171, 178, 195, 196 and 212 (p = 0.569); 28, 52, 60, 78, 119, 135, 136, 158, 160, 169, 192, 193 and 197 (p = 0.560); 104, 121, 128, 132, 153, 154, 173, 180, 181, 182, 199 and 209 (p = 0.550); 29, 63, 122, 129, 131, 147, 155, 161, 167, 205, 207 and 208 (p = 0.540); 126, 127, 137, 156, 174, 183, 184, 187, 202, 216 and 217 (p = 0.529); 124, 125, 166, 179, 185, 200 and 203 (p = 0.519) and 103, 143, 159, 190, 191, 65, 152 and 201 (p = 0.509); 65, 152 and 201 (p = 0.5). The chosen compounds represented the classes of diterpenes, rare monoterpenes substituted with osidic units and cumaric group and triterpenes. Figure S2 (Supplementary material) illustrates the chemical structure of the compounds that exhibited activity probability values above 0.5 and a reliable applicability domain in the prediction model created with AlvaDesc descriptors.

6.3. Selection of Molecules for Biological Test

After conducting in silico screening, 86 compounds were chosen for evaluation of in vitro biological activity, namely: 168, 204 and 2015 (p = 0.589); 123, 172, 194, 198, 206, 210 and 211 (p = 0.579), 76, 77, 120, 138, 157, 162, 170, 171, 178, 195, 196 and 212 (p = 0.569); 28, 52, 60, 78, 119, 135, 136, 158, 160, 169, 192, 193 and 197 (p = 0.560); 104, 121, 128, 132, 153, 154, 173, 180, 181, 182, 199 and 209 (p = 0.550); 29, 63, 122, 129, 131, 147, 155, 161, 167, 205, 207 and 208 (p = 0.540); 126, 127, 137, 156, 174, 183, 184, 187, 202, 216 and 217 (p = 0.529); 124, 125, 166, 179, 185, 200 and 203 (p = 0.519) and 103, 143, 159, 190, 191, 65, 152 and 201 (p = 0.509); 65, 152 and 201 (p = 0.5). The primary selection criterion was a probability of activity above 0.50 in the developed classification model, and the second criterion considered the availability of the substance and ease of acquisition. Consequently, of the 86 compounds selected by the model, it was only feasible to obtain a sufficient quantity for in vitro testing of the compounds (60) Betulinic acid (p = 0.560) and (136) Lupeol (p = 0.560). To broaden the test series for experimental validation of the model, two synthetic derivatives of the selected compounds were introduced. This aimed to evaluate whether modifications made through the insertion of acetate and methyl ester groups contributed to greater activity and enhanced cell viability. Therefore, the compounds (219) Betulinic acid acetate and (220) Betulinic acid methyl ester were included in the test series (Figure 1). We ran both compounds through our QSAR model to ensure they were predicted to be active before subjecting them to the in vitro assay.

 doi:10.20944/preprints202406.0716.v1

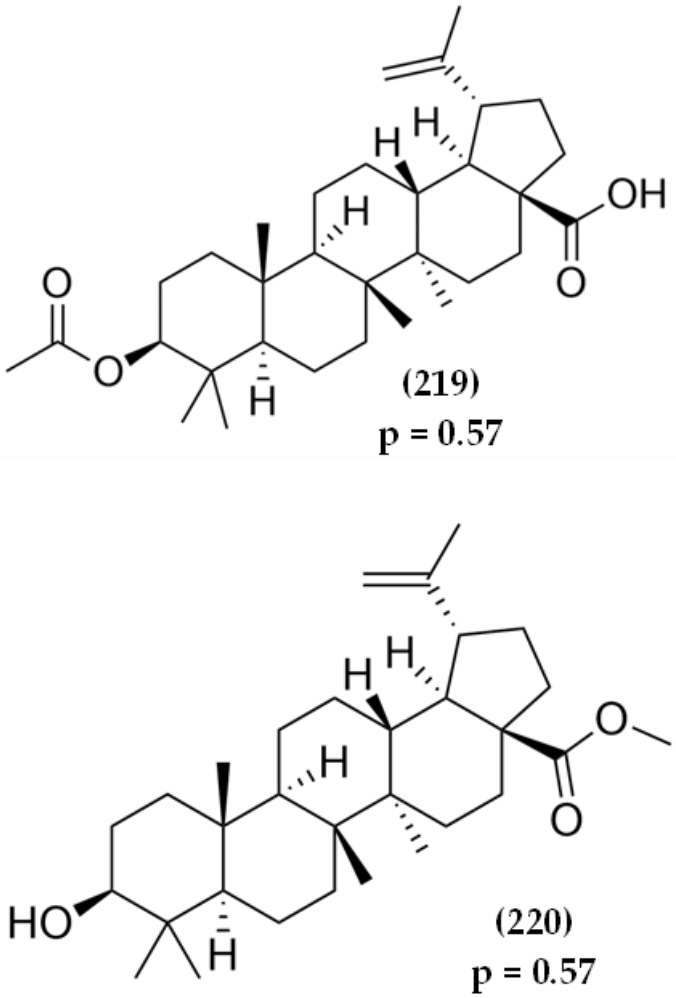

Figure 1. Structure of the synthetic compounds derived from the triterpenes under study.

According to our QSAR model, Betulinic acid acetate (219) presented activity probability values of 0.57, while the compound Betulinic acid methyl ester (220) exhibited probability values corresponding to 0.57. Hence, as these compounds showed activity probability values above 0.5 (random), they were subjected to in vitro testing, bringing the total to four compounds in the experimental validation of the study, namely: (60) Betulinic acid (p = 0.560), (136) Lupeol (p = 0.560), (219) Betulinic acid acetate (p = 0.57) and (220) Betulinic acid methyl ester (p = 0.57).

6.4. In Vitro Activity Assessment

A High Content Screening (HCS) Assay was devised to evaluate compounds inhibiting infection and cytotoxicity in Vero cells infected with a SARS-CoV-2 isolate [13]. The potential antiviral activity of four terpene-class compounds against SARS-CoV-2 in Vero CCL-81 cells was evaluated. An initial screening was conducted and the compounds were tested at a single concentration of 10μM, as indicated in Table 2.

Table 2. Anti-SARS-CoV-2 activity by Vero CCL-81 cells-high content screening-at 10 μM.

Compound	CS/%	AA/%
Lupeol	149.31	47.29
Betulinic acid	76.42	59.20
Betulinic acid methyl ester	30.99	77.41
Betulinic Acid Acetate	39.78	91.11

 doi:10.20944/preprints202406.0716.v1

(CS: cell survival; AA: antiviral activity).

Our results show that Betulinic Acid and its derivatives exhibited the highest percentages of inhibition. However, the derivatives Betulinic acid methyl ester and Betulinic Acid Acetate did not show high rates of cell viability. In contrast, Betulinic acid demonstrated inhibition of SARS-CoV-2 above 50% and mantained cell viability percentages above 70%. Lupeol showed inhibition percentages of 47.29%, indicating moderate activity, and a cell viability rate of 149.31% [13]. These findings demonstrate that the compounds under study showed high rates of SARS-CoV-2 inhibition, particularly betulinic acid, which also demonstrated substantial cell viability.

3.5. Molecular Docking

Molecular Docking simulations were conducted on our top compounds to further substantiate our predictions for these compounds to act as potential inhibitors of SARS-CoV-2. The molecular docking simulation aimed to show a proposed binding pose and estimate a binding affinity of the compounds (60) Betulinic acid and (136) Lupeol to targets related to this effect, including Main-protease (M-pro) enzymes (PDB : 5RG1), Papain-like protease (PL-pro) (PDB: 7TZJ), Spike protein (PDB: 8C8P) and RNA-dependent RNA polymerase (PDB: 6M71). Prior to the molecular docking simulation, a redocking procedure was carried out between the ligands and the co-crystallized proteins to validate our docking procedures (Figure S2). The redocking, including the RMSD (Root Mean Square Deviation) values which measure the deviation between the experimentally determined crystallographic structure and the coupled pose, are shown in Table S3 and Figure S3–Supplementary Material. [45,46]. Table 3 presents the affinity results for the compounds under study, according to the energetic values obtained from the Moldock Score algorithm (KJ.mol^{-1}).

Table 3. Binding energy (KJ.mol-1) and affinity probability (p) values of the compounds under study with the SARs-CoV-2 enzymes.

Compounds	M-pro (PDB: 5RG1)	PL-pro (PDB: 7TZJ)	Spike (PDB: 8C8P)	RNA-dependent RNA polymerase (PDB: 6M71)
Lupeol	-88.88 (p = 0.86)	-80.46 (p = 0.65)	-96.91 (p = 0.64)	-81.97 (p = 0.46)
Betulinic acid	-91.66 (p = 0.88)	-97.57 (p = 0.79)	-102.35 (p = 0.68)	-79.97 (p = 0.45)
PDB Ligand/ Positive control	-103.03 (p = 1)	**-123.07 (p = 1)**	**-149.45 (p = 1)**	**176.59 (p = 1)**

[1] The compound with the highest affinity is highlighted in bold.

The compounds under investigation exhibit negative binding energy values for all targets, indicating their favorable interaction with the proteins. While the terpenes under study did not show higher affinity than the PDB ligand in any of the enzymes, it is noteworthy that for the Main-protease (M-pro) (PDB: 5RG1) target, the compounds demonstrated binding energy values close to those of the controls, with affinity probability values exceeding 0.8. Specifically, Lupeol exhibited -88.88 KJ.mol-1 (p = 0.84), Betulinic Acid showed -91.66 KJ.mol-1 (p = 0.87), the PDB Ligand presented -103.03 KJ.mol-1 (p = 0.98), being the most stable compound, displayed -104.71 KJ.mol-1 (p = 1). Interestingly, Betulinic acid demonstrated a favorable binding free energy and high probablity of binding to Papain-like protease (PDB: 7TZJ) at -97.57 KJ.mol-1 (p = 0.79). The M-pro enzyme was the target with the highest affinity demonstrated by the compounds under study. The description of the affinity of the compounds with the other targets is described in section 01 of the supplementary material. Figure 2 illustrates the interaction of the compounds: Lupeol (A), Betulinic acid (B), the control drug and the PDB Ligand (C) with the target Main-protease (M-pro) (PDB: 5RG1).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 June 2024 doi:10.20944/preprints202406.0716.v1

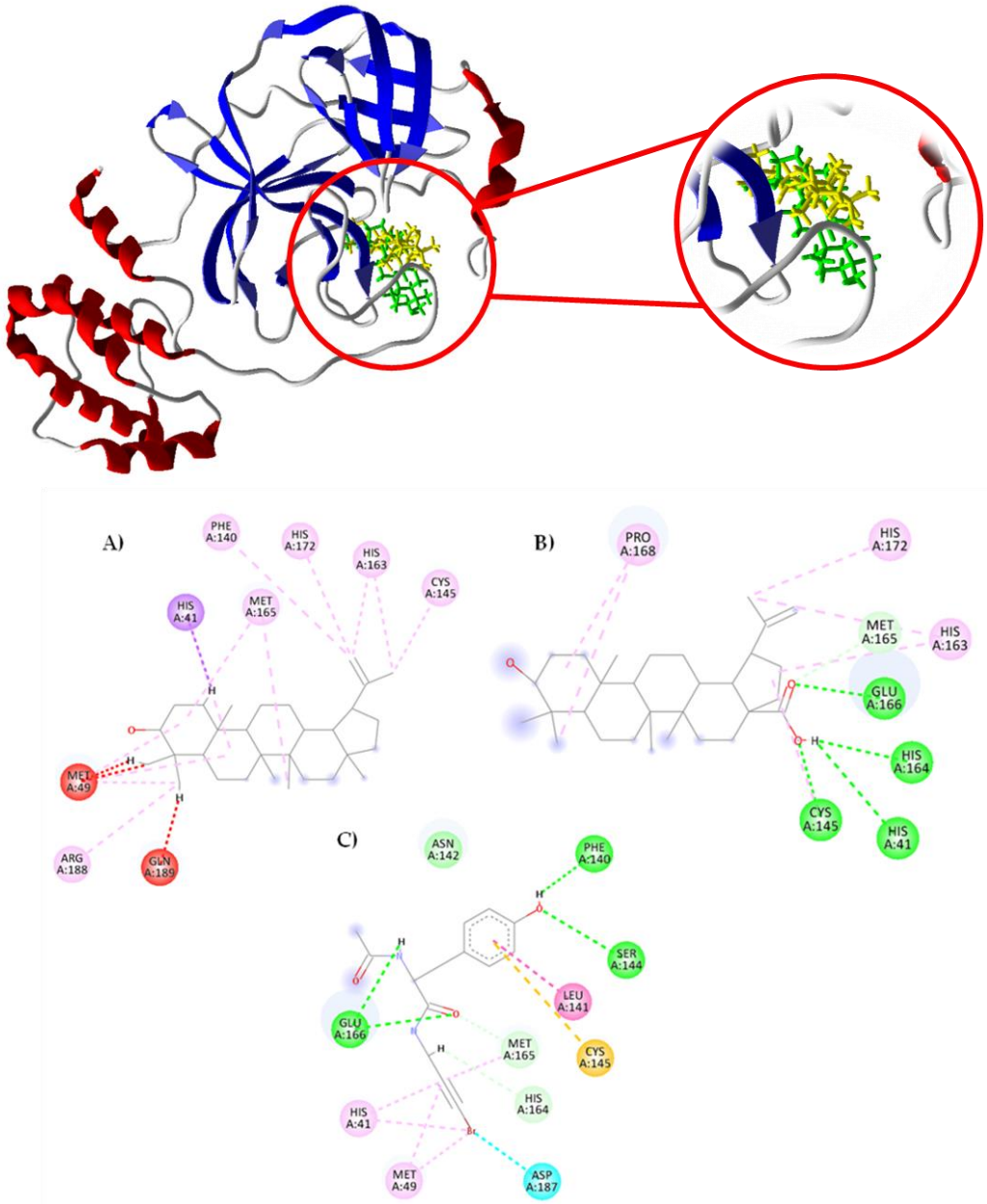

Figure 2. 2D molecular interactions established by the compounds Lupeol (A), Betulinic acid (B), and PDB Ligand (C) with the target Main-protease (M-pro) (PDB: 5RG1). Residues: His (Histidine), Phe (Phenylalanine), Cys (Cysteine), Met (Methionine), Arg (Arginine), Gln (Glutamine), Pro (Proline), Glu (Glutamic Acid), Asn (Asparagine), Leu (Leucine) and Asp (Aspartic acid). Interactions: Red (unfavorable bump), Purple (pi-sigma), Pink (alkyl, pi-alkyl, amide-pi stacked), Green (Conventional hydrogen bond and carbon hydrogen bond), Orange (pi-anion and pi-sulfur) and Blue (halogen – Cl, Br, I).

In Figure 2, the molecular interactions with the Main-protease (M-pro) (PDB: 5RG1) involved hydrogen bonds (green dashed lines), hydrophobic interactions (pink, blue and orange dashed lines) and steric interactions (red dashed lines).

In the interaction of the Lupeol compound, the involvement of crucial residues for maintaining the enzyme's activity was observed. For instance, residues His 41 and Cys 145 play a key role in forming the enzyme's catalytic dyad and ensuring complete dimerization of the active site. His residue 163 is important for the formation of the side chain. Met residue 165 contributes to the formation of the central monomer of the protein, representing a large hydrophobic cavity. The residues Met 49 and Gln 189 suggest a degree of plasticity in the enzyme's side chains [38]. The PDB

 doi:10.20944/preprints202406.0716.v1

ligand also demonstrated interactions similar to those observed in the molecular interaction of Lupeol, involving residues His 41, Met 49 and Met 165 in hydrophobic interactions of alkyl and pi-alkyl type.

Similar to Lupeol, the triterpene betulinic acid displayed interactions with crucial residues for maintaining the activity of the Main—protease enzyme. These included residues Met 165, His 163, and Pro 168, related to the plasticity of the enzyme's side chain. The residue Glu 166 is involved in dimerization of the chain and the formation of the catalytic dyad, while His164 is crucial for the formation of the central monomer of the hydrophobic enzyme [38]. Additionally, betulinic acid shared hydrogen bond interactions with residues Met 165, His 164, and Glu 166 with the PDB ligand. Both compounds under study also exhibited similar interactions with residue His 163 through alkyl and pi-alkyl type interactions.

4. Discussion

The Fabaceae family, or Leguminosae, is the third largest family of angiosperm plants [47,48]. This family to presented 770 genera and around 19,500 species, distributed across several subfamilies [49,50]. It is the largest family of plants in Brazil, with approximately 2,834 species found in different ecosystems [51]. The Fabaceae Family exhibits a rich diversity of chemical compounds, with a particular emphasis on phenolic compounds and alkaloids [52].

Terpenes, are another significant group associated with numerous biological activities, including antimicrobial and antiviral activity [53]. Compounds within the terpene class are promising natural compounds for the creation of new antiviral agents [54–58]. Natural terpenoids and their synthetic analogues are considered valuable sources for novel medicines for the treatment of various diseases, due to their diverse molecular structures, low toxicity and the ability to impact several specific cellular targets, resulting in a wide range of biological activities [58–63].

In silico studies indicate that triterpenes, such as lupeol and betulinic acid, have potential anti-SARS-CoV-2 activity [64]. Lupeol, found in several plants, is known for its anti-inflammatory, antioxidant and anticancer properties, and is being investigated as an inhibitor of the main protease (Mpro) of the virus, essential for its replication [65–67]. Molecular modeling suggests that lupeol can bind to the active site of Mpro, inhibiting its activity by interacting with amino acids such as cysteine and histidine [68]. In summary, lupeol shows potential as an inhibitor of Mpro of SARS-CoV-2, but to date its efficacy has not yet been experimentally validated. Similarly to Lupeol, it was observed that betulinic acid in molecular docking studies and molecular dynamics simulations showed inhibition of the activity of the enzyme M-Protease and papain-like protease with high stability, an enzyme that belongs to the family of cysteine proteases, a type of enzyme which cleaves proteins through the hydrolysis of peptide bonds, as they have a cysteine residue in the active center that acts as a nucleophile in the protein cleavage process [69]. Like Lupeol, the use of betulinic acid, the anti-SARS-CoV-2 activity requires further evidence, thus affirming the importance of the present work, as it provides experimental evidence on the anti-SARS-CoV-2 activity 2 and also addresses the machine learning methodology and studies focused on the ligand and structure approaches.

In this study, a classification model was developed to identify with a probability of inhibiting SARS-CoV-2. The created model successfully identified 86 compounds that presented a probability of activity exceeding 50% (p=0.5) and assessed their applicability domain. The selected compounds belong to classes of diterpenes, which have shown promise in treating or preventing viral infections caused by enveloped viruses that undergo hemagglutinin-mediated fusion with a host cell and/or the resulting symptoms. Previous research by Tret' yakowa and collaborators (2022) [58] reported the synthesis of Mannich diterpenic bases as potential therapeutic agents for Influenza A and SARS-CoV-2. Another class identified was triterpenes, widely used in traditional herbal medicine and represent an interesting case of natural compounds that play an important role in plant defence. Triterpenes are known for their antiviral activity against various diseases, including human immunodeficiency virus 1 (HIV-1), hepatitis B virus (HBV), hepatitis C virus (HCV), influenza A virus (IAV) , Ebola virus (EBOV) and SARS-CoV [56,64,70–73]. The third class worth mentioning consists of rare monoterpenes substituted with osidic units, which have demonstrated the ability to bind and

 doi:10.20944/preprints202406.0716.v1

interfere with the functions of different proteins in the SARS-Cov-2 virus, including the main protease, endoribonuclease, ADP ribose phosphatase, RNA-dependent polymerase, and spike protein [74–77]. These compounds also impact human cell proteins crucial for viral internalization and replication, including angiotensin-converting enzyme and cellular proteases, transmembrane serine protease 2, cathepsin B and cathepsin L [74,77–80].

We selected compounds lupeol and betulinic acid based on their probability of activity and availability and subjected them to biological testing. Additionally, two synthetic derivatives of betulinic acid were included: betulinic acid acetate and betulinic acid methyl ester, in order to investigate whether the activity would be enhanced or reduced. The results revealed that betulinic acid exhibited the highest percentage of inhibition and demonstrated a high rate of cell survival. Lupeol showed moderate activity but exhibited good cell viability. However, the synthetic derivatives displayed a higher percentage of inhibition, but did not show favorable percentages of cell viability.

The Main-protease was identified as the target with the highest affinity, and the compounds exhibited greater stability concerning the evaluation of free energy parameters. Lupeol and Betulinic acid, both terpenes, have substantial evidence in the literature supporting their anti-SARS-CoV-2 potential. Elkousy and collaborators (2022) [81] identified lupeol as a promising candidate for a therapeutic agent against SARS-CoV-2 through an in silico study involving Castor Oil Plant (*Ricinus communis*). Betulinic acid is addressed by Patel and collaborators (2023) in an in silico study on bioprospecting of *Rosmarinus officinalis* for M-protease [82].

5. Conclusions

The bibliographic study documented and standardized 217 compounds from the Fabaceae family with potential therapeutic activities for several emerging diseases. A predictive model developed successfully classified 83 compounds, including betulinum and lupeol, that demonstrated inhibitory activity against SARS-CoV-2, validating the model. The Main-Protease enzyme emerged as the most likely target. From this perspective, it can be concluded that the virtual screening identified compounds with a high probability of activity and stability, corroborated by *in vitro* tests.

6. Acronyms

Area Over the Curve analysis (ROC);
Carbon Nuclear Resonance Spectrometry (C-NMR);
Coronilla varia (CV);
Dulbecco's Modified Eagle's Medium (DMEM);
Human Coronaviruses (HCoVs);
Hydrogen Nuclear Resonance Spectrometry (H-NMR);
Infrared Spectrometry (IV);
Main protease (Mpro or 3CLpro);
Mass Spectrometry (MS);
Matthews Correlation Coefficient (MCC);
Melilotus officinalis (MO);
Multiplicity of Infection (MOI);
Ononis spinosa (OS)
Operetta High-Content Analysis System equipment (HCS);
Papain-like protease (PLpro);
Phosphate-buffered saline (PBS);
Protein Data Bank (PDB);
Quantitative Structure Activity Relationship (QSAR);
Random Forest (RF);
RNA-dependent RNA polymerase (RdRp);
Robinia pseudoacacia (RP);
Ultraviolet Spectrometry (UV).

 doi:10.20944/preprints202406.0716.v1

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, Table S1. Terpene derivatives found in the Fabaceae family; Table S2. Prediction of the anti-SARS-CoV-2 activity of terpenes isolated in the Fabaceae Family according to VolSurf descriptors; Figure S01: Compounds that did not show reliability in the applicability domain in the model created with AlvaDesc descriptors; Figure S2: Compounds that presented activity probability values above 0.50 and reliable applicability domain in the prediction model created with AlvaDesc descriptors; *S.01. Molecular Docking;* Table S3. RMSD values for the protein selected in the study; Figure S3. Redocking of the co-crystallized ligands and their respective poses. A) M-protease target (PDB: 5RG1) and Papain-like protease target (PDB: 7TZJ).

Author Contributions: Conceptualization, M.T.S., L.S., N.F.S., Y.M.N., N.N.M. and J.M.B.F.; methodology, N.F.S., M.T.S., N.N.M., L.S., J.M.B.F., C.G.B., C.B.M., Y.M.N. and L.H.G.F.J.; validation, N.F.S., M.T.S., L.S., J.M.B.F.J., C.G.B., C.B.M and L.H.G.F.J.; formal analysis, N.F.S., H.J.M, M.T.S., L.S., J.M.B.F., C.G.B., C.B.M and L.H.G.F.J.; investigation, M.T.S., L.S., N.N.M, N.F.S. and J.M.B.F.; resources, L.S., M.T.S., L.H.G.F.J. and J.M.B.F.; data curation, N.F.S. and H.J.M writing—Original draft preparation, N.F.S., H.J.M., G.D.D. and M.T.S.; writing—Review and editing, N.F.S., G.D.D.; H.J.M., N.N.M., L.S., J.M.B.F., M.T.S. and Y.M.N.; visualization, N.F.S.; supervision, L.S., N.N.M., J.M.B.F., Y.M.N. and M.T.S.; project administration, N.F.S.; funding acquisition, M.T.S., J.M.B.F., L.S., and Y.M.N. All authors have read and agreed to the published version of the manuscript.

Funding: Coordination of Improvement of Higher Education Personnel—Brazil (CAPES) GrantNo. 88887.505029/2020-00 and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Universal Call 18/2021. Process: 402976/2021-5. Title: Biomonitored chemical study of some plants from the semi-arid region of the Northeast, especially the State of Paraíba, semi-synthesis of secondary metabolites and in silico studies of natural products.

Acknowledgments: Coordination of Improvement of Higher Education Personnel—Brazil (CAPES) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Organization, W.H. World Health Organization (WHO) COVID-19 Dashboard [Internet]. *World Heal. Organ.* **2023**.
2. Nalbandian, A.; Sehgal, K.; Gupta, A.; Madhavan, M. V; McGroder, C.; Stevens, J.S.; Cook, J.R.; Nordvig, A.S.; Shalev, D.; Sehrawat, T.S. Post-Acute COVID-19 Syndrome. *Nat. Med.* **2021**, *27*, 601–615.
3. Markov, P. V; Ghafari, M.; Beer, M.; Lythgoe, K.; Simmonds, P.; Stilianakis, N.I.; Katzourakis, A. The Evolution of SARS-CoV-2. *Nat. Rev. Microbiol.* **2023**, *21*, 361–379.
4. Minkoff, J.M.; tenOever, B. Innate Immune Evasion Strategies of SARS-CoV-2. *Nat. Rev. Microbiol.* **2023**, *21*, 178–194.
5. Sinha, A.; Sangeet, S.; Roy, S. Evolution of Sequence and Structure of SARS-CoV-2 Spike Protein: A Dynamic Perspective. *ACS omega* **2023**, *8*, 23283–23304.
6. Uddin, M.; Mustafa, F.; Rizvi, T.A.; Loney, T.; Al Suwaidi, H.; Al-Marzouqi, A.H.H.; Kamal Eldin, A.; Alsabeeha, N.; Adrian, T.E.; Stefanini, C. SARS-CoV-2/COVID-19: Viral Genomics, Epidemiology, Vaccines, and Therapeutic Interventions. *Viruses* **2020**, *12*, 526.
7. Baggen, J.; Vanstreels, E.; Jansen, S.; Daelemans, D. Cellular Host Factors for SARS-CoV-2 Infection. *Nat. Microbiol.* **2021**, *6*, 1219–1232.
8. Guruprasad, K. Mutations in Human SARS-CoV-2 Spike Proteins, Potential Drug Binding and Epitope Sites for COVID-19 Therapeutics Development. *Curr. Res. Struct. Biol.* **2022**, *4*, 41–50.
9. Pozzi, C.; Vanet, A.; Francesconi, V.; Tagliazucchi, L.; Tassone, G.; Venturelli, A.; Spyrakis, F.; Mazzorana, M.; Costi, M.P.; Tonelli, M. Antitarget, Anti-SARS-CoV-2 Leads, Drugs, and the Drug Discovery–Genetics Alliance Perspective. *J. Med. Chem.* **2023**, *66*, 3664–3702.
10. Puhl, A.C.; Godoy, A.S.; Noske, G.D.; Nakamura, A.M.; Gawriljuk, V.O.; Fernandes, R.S.; Oliva, G.; Ekins, S. Discovery of PLpro and Mpro Inhibitors for SARS-CoV-2. *ACS Omega* **2023**.
11. Hersi, F.; Sebastian, A.; Tarazi, H.; Srinivasulu, V.; Mostafa, A.; Allayeh, A.K.; Zeng, C.; Hachim, I.Y.; Liu, S.-L.; Abu-Yousef, I.A. Discovery of Novel Papain-like Protease Inhibitors for Potential Treatment of COVID-19. *Eur. J. Med. Chem.* **2023**, *254*, 115380.
12. Wasilewicz, A.; Zwirchmayr, J.; Kirchweger, B.; Bojkova, D.; Cinatl Jr, J.; Rabenau, H.F.; Rollinger, J.M.; Beniddir, M.A.; Grienke, U. Discovery of Anti-SARS-CoV-2 Secondary Metabolites from the Heartwood of Pterocarpus Santalinus Using Multi-Informative Molecular Networking. *Front. Mol. Biosci.* **2023**, *10*, 1202394.

 doi:10.20944/preprints202406.0716.v1

13. Lima, C.M.C.F.; Freitas Junior, L.H.; Moraes, C.B.; Barbosa, C.G.; Opatz, T.; Victor, M.M. Synthesis of Isatins and Oxindoles Derivatives as SARS-CoV-2 Inhibitors Evaluated through Phenotypic Screening with Vero Cells. *J. Braz. Chem. Soc.* **2023**, *34*, 745–753.
14. Nyagumbo, E.; Pote, W.; Shopo, B.; Nyirenda, T.; Chagonda, I.; Mapaya, R.J.; Maunganidze, F.; Mavengere, W.N.; Mawere, C.; Mutasa, I. Medicinal Plants Used for the Management of Respiratory Diseases in Zimbabwe: Review and Perspectives Potential Management of COVID-19. *Phys. Chem. Earth, Parts A/B/C* **2022**, 103232.
15. Sharaibi, O.J.; Oluwa, O.K.; Omolokun, K.T. Traditional Plant Based Medicines Used for the Treatment of COVID-19 Symptoms by AWORI Tribe in OJO Local Community of Lagos State, Nigeria. *J. Med. Plants* **2022**, *10*, 57–62.
16. Devi, R.S.; Dimri, R.; Ramadevi Devarakonda, V.O.; Kumar, S. Medicinally Important Species of Fabaceae Family of Loktak Lake, India.
17. Agbor, G.A.; Ndjib, R. Systematic Review of Plants Used against Respiratory Diseases Related to COVID-19 in Africa. *J. Drug Deliv. Ther.* **2021**, *11*, 141–153.
18. Shahrousvand, M.; Haddadi-Asl, V.; Shahrousvand, M. Step-by-Step Design of Poly (ε-Caprolactone)/Chitosan/Melilotus Officinalis Extract Electrospun Nanofibers for Wound Dressing Applications. *Int. J. Biol. Macromol.* **2021**, *180*, 36–50.
19. Adhikari, B.; Marasini, B.P.; Rayamajhee, B.; Bhattarai, B.R.; Lamichhane, G.; Khadayat, K.; Adhikari, A.; Khanal, S.; Parajuli, N. Potential Roles of Medicinal Plants for the Treatment of Viral Diseases Focusing on COVID-19: A Review. *Phyther. Res.* **2021**, *35*, 1298–1312.
20. Usman, M.; Khan, W.R.; Yousaf, N.; Akram, S.; Murtaza, G.; Kudus, K.A.; Ditta, A.; Rosli, Z.; Rajpar, M.N.; Nazre, M. Exploring the Phytochemicals and Anti-Cancer Potential of the Members of Fabaceae Family: A Comprehensive Review. *Molecules* **2022**, *27*, 3863.
21. Obistioiu, D.; Cocan, I.; Tîrziu, E.; Herman, V.; Negrea, M.; Cucerzan, A.; Neacsu, A.-G.; Cozma, A.L.; Nichita, I.; Hulea, A. Phytochemical Profile and Microbiological Activity of Some Plants Belonging to the Fabaceae Family. *Antibiotics* **2021**, *10*, 662.
22. Stojković, D.; Dias, M.I.; Drakulić, D.; Barros, L.; Stevanović, M.; CFR Ferreira, I.; D. Soković, M. Methanolic Extract of the Herb Ononis Spinosa L. Is an Antifungal Agent with No Cytotoxicity to Primary Human Cells. *Pharmaceuticals* **2020**, *13*, 78.
23. Yerlikaya, S.; Baloglu, M.C.; Altunoglu, Y.C.; Diuzheva, A.; Jekő, J.; Cziáky, Z.; Zengin, G. Exploring of Coronilla Varia L. Extracts as a Source of High-Value Natural Agents: Chemical Profiles and Biological Connections. *South African J. Bot.* **2021**, *143*, 382–392.
24. David, J.; Kolawole, J.A.; Alemika, T.E.; Agwom, F.M.; Ajima, U. Analgesic and Anti-Inflammatory Activities of the Leaf Extracts of Detarium Microcarpum Guill & Perr (Fabaceae). *Asian J. Pharm. Pharmacol.* **2020**, *6*, 348–355.
25. Diniz, L.R.L.; Perez-Castillo, Y.; Elshabrawy, H.A.; Filho, C. da S.M.B.; de Sousa, D.P. Bioactive Terpenes and Their Derivatives as Potential SARS-CoV-2 Proteases Inhibitors from Molecular Modeling Studies. *Biomolecules* **2021**, *11*, 74.
26. Chatow, L.; Nudel, A.; Nesher, I.; Hayo Hemo, D.; Rozenberg, P.; Voropaev, H.; Winkler, I.; Levy, R.; Kerem, Z.; Yaniv, Z. In Vitro Evaluation of the Activity of Terpenes and Cannabidiol against Human Coronavirus E229. *Life* **2021**, *11*, 290.
27. Filho, J.M.B.; Trigueiro, J.A.; Cheriyan, U.O.; Bhattacharyya, J. Constituents of the Stem-Bark of Zizyphus Joazeiro. *J. Nat. Prod.* **1985**, *48*, 152–153.
28. Rodrigues, G.C.S.; dos Santos Maia, M.; de Souza, T.A.; de Oliveira Lima, E.; Dos Santos, L.E.C.G.; Silva, S.L.; da Silva, M.S.; Filho, J.M.B.; da Silva Rodrigues Junior, V.; Scotti, L. Antimicrobial Potential of Betulinic Acid and Investigation of the Mechanism of Action against Nuclear and Metabolic Enzymes with Molecular Modeling. *Pathogens* **2023**, *12*, 449.
29. Almeida, J.; Da-Cunha, E.V.L.; Silva, M.S.; Athayde-Filho, P.F.; Braz-Filho, R.; Barbosa-Filho, J.M. Outros Constituintes Químicos de Diplotropis Ferruginea Benth.(Fabaceae). *Rev. Bras. Farmacogn.* **2003**, *13*, 44–46.
30. de Lima, F.O.; Alves, V.; Filho, J.M.B.; da Silva Almeida, J.R.G.; Rodrigues, L.C.; Soares, M.B.P.; Villarreal, C.F. Antinociceptive Effect of Lupeol: Evidence for a Role of Cytokines Inhibition. *Phyther. Res.* **2013**, *27*, 1557–1563.
31. Dong, J.; Yao, Z.-J.; Zhu, M.-F.; Wang, N.-N.; Lu, B.; Chen, A.F.; Lu, A.-P.; Miao, H.; Zeng, W.-B.; Cao, D.-S. ChemSAR: An Online Pipelining Platform for Molecular SAR Modeling. *J. Cheminform.* **2017**, *9*, 1–13.
32. ChemAxon Standardizer.
33. Dos Santos Maia, M.; Rodrigues, G.C.S.; De Sousa, N.F.; Scotti, M.T.; Scotti, L.; Mendonça-Junior, F.J.B. Identification of New Targets and the Virtual Screening of Lignans against Alzheimer's Disease. *Oxid. Med. Cell. Longev.* **2020**, *2020*, doi:10.1155/2020/3098673.
34. dos Santos Maia, M.; de Sousa, N.F.; Rodrigues, G.C.S.; Monteiro, A.F.M.; Scotti, M.T.; Scotti, L. Lignans and Neolignans Anti-Tuberculosis Identified by QSAR and Molecular Modeling. *Comb. Chem. High Throughput Screen.* **2020**, *23*, 1–13, doi:10.2174/1386207323666200226094940.

 doi:10.20944/preprints202406.0716.v1

35. Mauri, A.; Bertola, M. Alvascience: A New Software Suite for the QSAR Workflow Applied to the Blood–Brain Barrier Permeability. *Int. J. Mol. Sci.* **2022**, *23*, 12882.
36. Mauri, A. AlvaDesc: A Tool to Calculate and Analyze Molecular Descriptors and Fingerprints. *Ecotoxicological QSARs* **2020**, 801–820.
37. RCSB Protein Data Bank.
38. Douangamath, A.; Fearon, D.; Gehrtz, P.; Krojer, T.; Lukacik, P.; Owen, C.D.; Resnick, E.; Strain-Damerell, C.; Aimon, A.; Ábrányi-Balogh, P. Crystallographic and Electrophilic Fragment Screening of the SARS-CoV-2 Main Protease. *Nat. Commun.* **2020**, *11*, 1–11.
39. Calleja, D.J.; Kuchel, N.; Lu, B.G.C.; Birkinshaw, R.W.; Klemm, T.; Doerflinger, M.; Cooney, J.P.; Mackiewicz, L.; Au, A.E.; Yap, Y.Q. Insights Into Drug Repurposing, as Well as Specificity and Compound Properties of Piperidine-Based SARS-CoV-2 PLpro Inhibitors. *Front. Chem.* **2022**, *10*, 861209.
40. Gao, Y.; Yan, L.; Huang, Y.; Liu, F.; Zhao, Y.; Cao, L.; Wang, T.; Sun, Q.; Ming, Z.; Zhang, L. Structure of the RNA-Dependent RNA Polymerase from COVID-19 Virus. *Science (80-.).* **2020**, *368*, 779–782.
41. Rajiv Gandhi, G.; Sharanya, C.S.; Jayanandan, A.; Haridas, M.; Edwin Hillary, V.; Rajiv Gandhi, S.; Sridharan, G.; Sivasubramanian, R.; Silva Vasconcelos, A.B.; Montalvão, M.M. Multitargeted Molecular Docking and Dynamics Simulation Studies of Flavonoids and Volatile Components from the Peel of Citrus Sinensis L.(Osbeck) against Specific Tumor Protein Markers. *J. Biomol. Struct. Dyn.* **2023**, 1–30.
42. Ullrich, S.; Ekanayake, K.B.; Otting, G.; Nitsche, C. Main Protease Mutants of SARS-CoV-2 Variants Remain Susceptible to Nirmatrelvir. *Bioorg. Med. Chem. Lett.* **2022**, *62*, 128629.
43. CLC Bio Company Mollegro Virtual Docker 6.0.
44. Biovia Accelrys Discovery Studio 3.5.
45. Hung, L.-H.; Guerquin, M.; Samudrala, R. GPU-QJ, a Fast Method for Calculating Root Mean Square Deviation (RMSD) after Optimal Superposition. *BMC Res. Notes* **2011**, *4*, 97.
46. Yusuf, D.; Davis, A.M.; Kleywegt, G.J.; Schmitt, S. An Alternative Method for the Evaluation of Docking Performance: RSR vs RMSD. *J. Chem. Inf. Model.* **2008**, *48*, 1411–1422.
47. Xu, Z.; Deng, M.; Xu, Z.; Deng, M. Fabaceae or Leguminosae. *Identif. Control Common Weeds Vol. 2* **2017**, 547–615.
48. Gudavalli, D.; Pandey, K.; EDE, V.G.; Sable, D.; Ghagare, A.S.; Kate, A.S. Phytochemistry and Pharmacological Activities of Five Species of Bauhinia Genus: A Review. *Fitoterapia* **2024**, *174*, 105830.
49. Borges, L.; Bruneau, A.; Cardoso, D.; Crisp, M.; Delgado-Salinas, A.; Doyle, J.J.; Egan, A.; Herendeen, P.S.; Hughes, C.; Kenicer, G. Towards a New Classification System for Legumes: Progress Report from the 6th International Legume Conference. *South African J. Bot.* **2013**, *89*, 3–9.
50. Azani, N.; Babineau, M.; Bailey, C.D.; Banks, H.; Barbosa, A.R.; Pinto, R.B.; Boatwright, J.S.; Borges, L.M.; Brown, G.K.; Bruneau, A. A New Subfamily Classification of the Leguminosae Based on a Taxonomically Comprehensive Phylogeny: The Legume Phylogeny Working Group (LPWG). *Taxon* **2017**, *66*, 44–77.
51. Santos-Silva, J.; Araujo, T.J. Are Fabaceae the Principal Super-Hosts of Galls in Brazil? *An. Acad. Bras. Cienc.* **2020**, *92*.
52. Macêdo, M.J.F.; Ribeiro, D.A.; Santos, M. de O.; Macêdo, D.G. de; Macedo, J.G.F.; Almeida, B.V. de; Saraiva, M.E.; Lacerda, M.N.S. de; Souza, M.M. de A. Fabaceae Medicinal Flora with Therapeutic Potential in Savanna Areas in the Chapada Do Araripe, Northeastern Brazil. *Rev. Bras. Farmacogn.* **2018**, *28*, 738–750.
53. Stephane, F.F.Y.; Jules, B.K.J. Terpenoids as Important Bioactive Constituents of Essential Oils. *Essent. oils-bioactive Compd. new Perspect. Appl.* **2020**, 1–15.
54. Pompei, R.; Laconi, S.; Ingianni, A. Antiviral Properties of Glycyrrhizic Acid and Its Semisynthetic Derivatives. *Mini Rev. Med. Chem.* **2009**, *9*, 996–1001.
55. González, M.A. Synthetic Derivatives of Aromatic Abietane Diterpenoids and Their Biological Activities. *Eur. J. Med. Chem.* **2014**, *87*, 834–842.
56. Xiao, S.; Tian, Z.; Wang, Y.; Si, L.; Zhang, L.; Zhou, D. Recent Progress in the Antiviral Activity and Mechanism Study of Pentacyclic Triterpenoids and Their Derivatives. *Med. Res. Rev.* **2018**, *38*, 951–976.
57. Hodon, J.; Borkova, L.; Pokorny, J.; Kazakova, A.; Urban, M. Design and Synthesis of Pentacyclic Triterpene Conjugates and Their Use in Medicinal Research. *Eur. J. Med. Chem.* **2019**, *182*, 111653.
58. Tret'yakova, E. V; Ma, X.; Kazakova, O.B.; Shtro, A.A.; Petukhova, G.D.; Klabukov, A.M.; Dyatlov, D.S.; Smirnova, A.A.; Xu, H.; Xiao, S. Synthesis and Evaluation of Diterpenic Mannich Bases as Antiviral Agents against Influenza A and SARS-CoV-2. *Phytochem. Lett.* **2022**, *51*, 91–96.
59. Paduch, R.; Kandefer-Szerszen, M. Antitumor and Antiviral Activity of Pentacyclic Triterpenes. *Mini. Rev. Org. Chem.* **2014**, *11*, 262–268.
60. Smee, D.F.; Hurst, B.L.; Evans, W.J.; Clyde, N.; Wright, S.; Peterson, C.; Jung, K.-H.; Day, C.W. Evaluation of Cell Viability Dyes in Antiviral Assays with RNA Viruses That Exhibit Different Cytopathogenic Properties. *J. Virol. Methods* **2017**, *246*, 51–57.
61. Rodriguez-Rodriguez, R. Oleanolic Acid and Related Triterpenoids from Olives on Vascular Function: Molecular Mechanisms and Therapeutic Perspectives. *Curr. Med. Chem.* **2015**, *22*, 1414–1425.

62. Ma, X.; Chen, R.; Huang, M.; Wang, W.; Luo, L.; Kim, D.K.; Jiang, W.; Xu, T. DAMGO-Induced μ Opioid Receptor Internalization and Recycling Restore Morphine Sensitivity in Tolerant Rat. *Eur. J. Pharmacol.* **2020**, *878*, 173118.
63. Chen, Y.; Wang, X.; Ma, X.; Liang, S.; Gao, Q.; Tretyakova, E. V; Zhang, Y.; Zhou, D.; Xiao, S. Facial Synthesis and Bioevaluation of Well-Defined OEGylated Betulinic Acid-Cyclodextrin Conjugates for Inhibition of Influenza Infection. *Molecules* **2022**, *27*, 1163.
64. Hisham Shady, N.; Youssif, K.A.; Sayed, A.M.; Belbahri, L.; Oszako, T.; Hassan, H.M.; Abdelmohsen, U.R. Sterols and Triterpenes: Antiviral Potential Supported by in-Silico Analysis. *Plants* **2020**, *10*, 41.
65. Lalthanpuii, P.B.; Lalrinmawia, C.; Lalruatfela, B.; Ramliana, L.; Lalchhandama, K. Molecular Modeling of Lupeol for Antiviral Activity and Cellular Effects. *J. Appl. Pharm. Sci.* **2023**, *13*, 131–143.
66. Park, J.S.; Rehman, I.U.; Choe, K.; Ahmad, R.; Lee, H.J.; Kim, M.O. A Triterpenoid Lupeol as an Antioxidant and Anti-Neuroinflammatory Agent: Impacts on Oxidative Stress in Alzheimer's Disease. *Nutrients* **2023**, *15*, 3059.
67. Liu, K.; Zhang, X.; Xie, L.; Deng, M.; Chen, H.; Song, J.; Long, J.; Li, X.; Luo, J. Lupeol and Its Derivatives as Anticancer and Anti-Inflammatory Agents: Molecular Mechanisms and Therapeutic Efficacy. *Pharmacol. Res.* **2021**, *164*, 105373.
68. Antonopoulou, I.; Sapountzaki, E.; Rova, U.; Christakopoulos, P. Inhibition of the Main Protease of SARS-CoV-2 (Mpro) by Repurposing/Designing Drug-like Substances and Utilizing Nature's Toolbox of Bioactive Compounds. *Comput. Struct. Biotechnol. J.* **2022**, *20*, 1306–1344.
69. Liu, Y.; Nie, T.; Hou, J.; Long, H.; Zhang, Z.; Lei, M.; Xu, Y.; Wu, W. Design, Synthesis and Biological Evaluation of Betulinic Acid Derivatives as Potential Inhibitors of 3CL-Protease of SARS-CoV-2. *Steroids* **2024**, *202*, 109351.
70. Darshani, P.; Sen Sarma, S.; Srivastava, A.K.; Baishya, R.; Kumar, D. Anti-Viral Triterpenes: A Review. *Phytochem. Rev.* **2022**, *21*, 1761–1842.
71. Li, H.; Sun, J.; Xiao, S.; Zhang, L.; Zhou, D. Triterpenoid-Mediated Inhibition of Virus–Host Interaction: Is Now the Time for Discovering Viral Entry/Release Inhibitors from Nature? *J. Med. Chem.* **2020**, *63*, 15371–15388.
72. Elshabrawy, H.A. SARS-CoV-2: An Update on Potential Antivirals in Light of SARS-CoV Antiviral Drug Discoveries. *Vaccines* **2020**, *8*, 335.
73. Avelar, M.; Pedraza-González, L.; Sinicropi, A.; Flores-Morales, V. Triterpene Derivatives as Potential Inhibitors of the RBD Spike Protein from SARS-CoV-2: An In Silico Approach. *Molecules* **2023**, *28*, 2333.
74. Da Silva, J.K.R.; Figueiredo, P.L.B.; Byler, K.G.; Setzer, W.N. Essential Oils as Antiviral Agents, Potential of Essential Oils to Treat SARS-CoV-2 Infection: An in-Silico Investigation. *Int. J. Mol. Sci.* **2020**, *21*, 3426.
75. Das, A.; Pandita, D.; Jain, G.K.; Agarwal, P.; Grewal, A.S.; Khar, R.K.; Lather, V. Role of Phytoconstituents in the Management of COVID-19. *Chem. Biol. Interact.* **2021**, *341*, 109449.
76. Kim, C.-H. Anti–SARS-CoV-2 Natural Products as Potentially Therapeutic Agents. *Front. Pharmacol.* **2021**, *12*, 590509.
77. Ćavar Zeljković, S.; Schadich, E.; Džubák, P.; Hajdúch, M.; Tarkowski, P. Antiviral Activity of Selected Lamiaceae Essential Oils and Their Monoterpenes against SARS-Cov-2. *Front. Pharmacol.* **2022**, *13*, 893634.
78. Panikar, S.; Shoba, G.; Arun, M.; Sahayarayan, J.J.; Nanthini, A.U.R.; Chinnathambi, A.; Alharbi, S.A.; Nasif, O.; Kim, H.-J. Essential Oils as an Effective Alternative for the Treatment of COVID-19: Molecular Interaction Analysis of Protease (Mpro) with Pharmacokinetics and Toxicological Properties. *J. Infect. Public Health* **2021**, *14*, 601–610.
79. Salem, M.A.; Ezzat, S.M. The Use of Aromatic Plants and Their Therapeutic Potential as Antiviral Agents: A Hope for Finding Anti-COVID 19 Essential Oils. *J. Essent. Oil Res.* **2021**, *33*, 105–113.
80. Torres Neto, L.; Monteiro, M.L.G.; Galvan, D.; Conte-Junior, C.A. An Evaluation of the Potential of Essential Oils against SARS-CoV-2 from in Silico Studies through the Systematic Review Using a Chemometric Approach. *Pharmaceuticals* **2021**, *14*, 1138.
81. Elkousy, R.H.; Said, Z.N.A.; Ali, M.A.; Kutkat, O. Anti-SARS-CoV-2 in Vitro Potential of Castor Oil Plant (Ricinus Communis) Leaf Extract: In-Silico Virtual Evidence. **2022**.
82. Patel, U.; Desai, K.; Dabhi, R.C.; Maru, J.J.; Shrivastav, P.S. Bioprospecting Phytochemicals of Rosmarinus Officinalis L. for Targeting SARS-CoV-2 Main Protease (Mpro): A Computational Study. *J. Mol. Model.* **2023**, *29*, 161.