

A COMBINED VIRTUAL SCREENING OF COMPOUNDS FROM *CARICA PAPAYA* LEAVES AGAINST THE SARS-CORONAVIRUS-2 PROTEIN TARGET

SHANNON IAN FERNANDES*

Department of Biotechnology, NMAM Institute of Technology, Nitte, Karnataka, India. Email: shannonfernandes25@gmail.com

Received: 15 October 2022, Revised and Accepted: 25 October 2022

ABSTRACT

Objective: The coronavirus known as severe acute respiratory syndrome (SARS)-Cov-2 was a new strain that had generated a pandemic that was spreading due to an increase in cases and fatalities. Coronavirus causes respiratory and intestinal infections in animals and humans and accounted for the highest number of fatalities during the pandemic research. Numerous research focused on developing supplementary therapy that alleviates the symptoms and addresses comorbidities.

Methods: In the present study, we examined the outcomes of an *in silico* investigation to develop a novel oral medication against the SARS coronavirus. In this investigation, chemicals from *Carica papaya* leaves, which have been suggested as an herbal remedy for main protease (Mpro) inhibitor of SARS-Coronavirus-2, were subjected to an integrated virtual screening. The chosen substances were then subjected to pharmacological and toxicological screening. Finally, the screening was carried out by docking the chosen compound against the protein target of SARS-Coronavirus-2 using Biovia and PyRx.

Results: Docking studies identified myristoleic acid, palmitic acid, and thiamine as effective therapeutic agents against Mpro of SARS-Coronavirus-2.

Conclusion: It is worthwhile to suggest *C. papaya* leaves for additional *in vitro* research against SARS-Coronavirus-2 at both the molecular and cellular levels.

Keywords: *Carica papaya*, Coronavirus, Severe acute syndrome-Cov-2, Main protease, Auto-docking, Screening, ADMET analysis, Drug, Phytocompounds.

© 2022 The Authors. Published by Innovare Academic Sciences Pvt Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>) DOI: <http://dx.doi.org/10.22159/ijms.2022v10i6.46619>. Journal homepage: <https://innovareacademics.in/journals/index.php/ijms>

INTRODUCTION

Severe acute respiratory syndrome (SARS)-COVID is a highly transmissible virus that emerged in late 2019. Coronavirus causes respiratory and intestinal infections in animals and humans. The coronavirus that circulated in humans before the pandemic of SARS in 2002 and 2003 was not thought to be very harmful to humans since it often produced moderate infections in immunocompromised persons [1]. The 2019 coronavirus illness (COVID-19) began as an outbreak in Wuhan and has since spread to other parts of the world. The virus that caused COVID-19, known as SARS-coronavirus-2 (SARS-CoV-2), was discovered and sequenced by Chinese researchers. They ended up finding that it was a novel coronavirus that got to share a high degree of sequence identity with SARS-like coronaviruses derived from bats and pangolins, indicating a zoonotic origin.

SARS-CoV-2 is an enveloped, positive-sense, and single-stranded RNA virus that targets a wide range of animals. It belongs to the *Coronaviridae* family [2]. The main means of transmission are respiratory droplets a healthy individual that may contract SARS-CoV-2 if he comes into contact with an infected person or any of his possessions, such as clothes and doorknobs. Aerosol transmission (Airborne transmission) of SARS-CoV-2 has been observed in studies; however, there is no conclusive research on neonatal infections (mother to child). An affected individual may show the following symptoms such as fever, cough, tiredness, and loss of taste or smell. However, the transmission can be stopped by maintaining a 2 m space between people, using properly fitted masks outside, keeping oneself clean, sanitizing hands, and isolating affected individuals [3].

The replica gene of SARS-CoV-2 encodes two overlapping polyproteins, pp1a and pp1ab, which are essential for viral replication and transcription. The SARS-CoV-2 genome has roughly 30,000 nucleotides [4]. The protein chain is broken down by Main protease

(Mpro), also known as 3C-like protease, in at least 11 conserved places, commencing with the autolytic cleavage of this enzyme from pp1a and pp1ab [5]. Three domains make up the Mpro for SARS-CoV-2: domain I (residues 8–101), domain II (residues 102–184), and domain III (residues 201–303). The third domain, which has five α -helices, forms an antiparallel conglomeration and is bound to domain II by a prolonged loop region. The first two domains have an antiparallel-barrel structure (residues 185–200). The substrate-binding site of the Cys-His catalytic dyad observed in the Mpro of the SARS-CoV-2 viruses is positioned between domains I and II [6]. All CoV Mpros have quite a substrate-recognition pocket that is structurally quite conservative. Ways to further investigate this have been made accessible by the recent discovery of novel CoVs and the information about the structure of CoV Mpros from various strains. All CoV Mpros possess the same substrate-binding area between domains I and II as a result of structural preservation, according to the superposition of 12 crystal structures of Mpros (SARS-CoV-2, SARS-CoV, MERS-CoV, HCoV-HKU1, BtCoV-HKU4, MHV-A59, PEDV, FIPV, 312 TGEV, HCoV-NL63, and IBV) [7].

Tropical regions are the place of residence of the fruit known as *Carica papaya*, and it belongs to the family Caricaceae [8,9]. Protein, vitamins, and carbohydrates are all present. Flavonoid inhibits SARS-CoV by suppressing it. Due to its high vitamin A, B, and C content, *C. papaya* possesses antiviral, antifungal, and antibacterial and proteolytic enzymes such as papain and chymopapain. Burns and wounds, fever, intestinal nematode infection, asthma, and gastric ulcers can all be treated with *C. papaya*. The study focuses on various papaya features, including antioxidant and free radical scavenging action, anticancer activity, anti-inflammatory activity, treatment for dengue fever, anti-diabetic activity, wound healing activity, and influence on infertility [10].

Ten phytochemicals from *C. papaya* were used in molecular docking against the protein Mpro of the SARS Cov. It was an *in silico* study. Using ADMET and toxicity screening, a pharmaceutical evaluation of

ten phytocompounds was conducted. The top three binding ligands for proteins were obtained using molecular docking after ten hits of phytocompounds, and their characteristics were examined.

METHODS

Retrieval of macromolecule structure

SARS-Cov-2 crystal structure of Mpro, (PDB ID: 7LTN) which had a resolution of 1.79, a macromolecule residue count of 306, and just an A chain, was the protein model selected (Fig. 1). It was retrieved in pdb format from the RCSB Protein Data Bank (<https://www.rcsb.org/structure/7LTN>) [11]. Its molecular visuals were edited with Biovia Discovery Studio Visualizer (<https://discover.3ds.com/discovery-studio-visualizer-download>). The protein was purified with the following protocol; the crystallographic structure does not match the free energy of the water molecule. Water molecules were completely removed before docking, because they can have an impact on docking scores. The prebound ligands are taken out of the crystal structures to speed up binding with the ligands selected for the investigation. To improve purified structures, polar hydrogen atoms are introduced (Fig. 2). The purified structure was subjected to the prediction of protein statistics.

Finally, using Pepstat (https://www.ebi.ac.uk/Tools/seqstats/emboss_pepstats/) [12] and Pepwindow (https://www.ebi.ac.uk/Tools/seqstats/emboss_pepwindow/), the hydropathy plot and characteristics of the simulated protein were produced. Pepwindow computes hydropathy over the input sequence in windows of a given size. The secondary structure of the protein and its motifs was acquired when the modeled

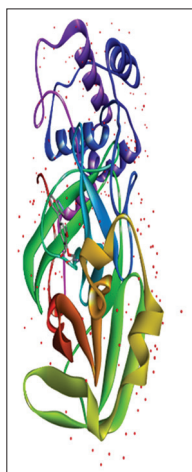


Fig. 1: Unpurified Mpro protein

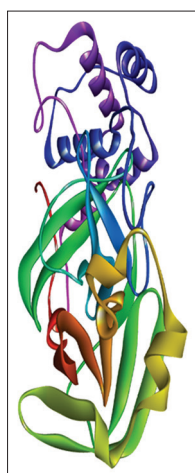


Fig. 2: Purified model protein

protein was submitted to PDBsum generate (<http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=index.html>) [13] to produce a Ramachandran plot and its statistics.

Retrieval of ligands

Based on their therapeutic potential, specific phytochemicals from *C. papaya* were retrieved from IMPPAT (<https://cb.imsc.res.in/imppat/>) server which is a repository for the phytochemicals of Indian medicinal plants [14]. The 3D structure data format structure of these ten phytocompounds was then acquired from PubChem (<https://pubchem.ncbi.nlm.nih.gov/>) [15]. Using an online smiles translator (<https://www2.chemie.uni-erlangen.de/services/translate/>) [16], the 2D structures were rendered into 3D structures. To assess the physiochemical characteristics of phytochemicals, canonical SMILES was collected.

Drug likeness studies

Swiss absorption, distribution, metabolism, and excretion (ADME) (<http://www.swissadme.ch/>) [17] performs ADME analysis. It establishes parameters for drug similarity, such as blood-brain barrier (BBB), gastrointestinal (GI) absorption, synthetic accessibility (SA) score, MLogP, molar refractivity, hydrogen acceptors, and donors. The BOILED-Egg is also obtained from SwissADME. The determination of surface area, water solubility, and rotatable bonds was also made. These crucial variables were selected and created following the Lipinski Rule of 5. For ADME analysis, the Admet lab 2.0 (<https://admetmesh.scbdd.com/>) [18] was also used. It provided crucial information on physiochemical characteristics. ChemAGG (<https://admet.scbdd.com/ChemAGG/index/>) [19] was used to find the aggregator class of the chemicals and the toxicity of phytocompounds was tested by utilizing Protox2 (https://tox-new.charite.de/protox_II/) [20].

Molecular docking

In structural molecular biology and computer-assisted drug creation, molecular docking is a crucial tool. It assists in identifying the protein-ligand interaction, locating the active location where interaction may occur, and providing binding affinity scores. Through the removal of water molecules and additional chains which are not required followed by the addition of hydrogen polar, the protein was purified in Biovia Discovery Studio. The purified protein was then uploaded in PyRx (<https://sourceforge.net/projects/pyrx/>) [21], which includes embedded versions of auto dock wizard, open babel, and vina wizard. By minimizing energy and optimizing the grid dimensions (X:37.4781 Y:66.7479 Z:61.7533), the ligands were docked to the protein in Vina Wizard, and the binding energy was obtained in a CSV file. Ligands with the most negative binding affinity score were chosen and submitted to Biovia, where the interactions between the ligand and protein were predicted, and hydrophobic and 2D structures were examined.

RESULTS

Retrieval of protein

The characterization of the function, structure, and interactions of the target protein relies on its purification. Before separating the target

Table 1: Statistical analysis of different types of amino acids present in Mpro

Property	Residues	Number	Mole%
Tiny Small	(A+C+G+S+T)	95	31.046
Small	(A+B+C+D+G+N+P+S+T+V)	173	56.536
Aliphatic	(A+I+L+V)	84	27.451
Aromatic	(F+H+W+Y)	38	12.418
Non-polar	(A+C+F+G+I+L+M+P+V+W+Y)	176	57.516
Polar	(D+E+H+K+N+Q+R+S+T+Z)	130	42.484
Charged	(B+D+E+H+K+R+Z)	55	17.974
Basic	(H+K+R)	29	9.477
Acidic	(B+D+E+Z)	26	8.497

protein from all other proteins, the purification process may first separate the protein and non-protein components. Water molecules and the additional chains are not involved in the binding process during docking and can alter the docking score and cause complexity in the protein structure. Hence, before docking purification is done in Biovia, in which water molecules and B chains are removed from the protein to make it pure and polar charges are added (Figs. 1 and 2).

Pepstat and pepwindow analysis

The results of the Pepstat analysis for the modeled protein showed that it had a molecular weight (MW) of 33796.64, 306 residues, and a charge of -0.5. The average residue weight was found to be 110.447, and the protein's isoelectric point was 6.3628. The pH level at which a specific molecule has no net electrical charge is known as the isoelectric point (pI). The amino acid profile of the purified Mpro protein structure is enlisted in Table 1.

Protein hydrophobicity is determined by their hydropathy plots along the complete length of the protein sequence. The hydropathy plot combines both the hydrophobic and hydrophilic characteristics of the peptide's amino acid sequences. These diagrams aid in identifying potential protein architectures, domains, or motifs. These diagrams are essential for analyzing how well the proteins operate. The hydropathy plot has amino acid number on the X-axis and degree of hydrophobicity and hydrophilicity on the Y-axis. There are several ways to measure hydropathy plots, including the Kyte Doolittle and Hopp Woods scales. In Fig. 3, positive results suggest that amino acids are hydrophobic and may form an alpha-helix that spans a lipid bilayer. On the other hand, negative values mean that the amino acids are in contact with a solvent or water, which means that they are more likely to be found on the protein's outer surface and that they are hydrophilic.

Analysis of the macromolecule structure

The statistical distribution of the phi and psi angles for each of the amino acids in the structure is displayed using a Ramachandran plot. In Fig. 4, the Red region represents favored regions in the Ramachandran plot. The most favoring regions in the Ramachandran plot contained 243 residues and covered 91.4% of the amino acids, while the additional allowed regions had 19 residues and made up 7.2%, the generously allowed regions had one residue and made up 0.4%, and the disallowed regions had two residues and made up 0.8%. There were found to be 26 and 13 glycine and proline residues, respectively.

From PDBSum generate, the secondary structure of pure Mpro protein was acquired. Eleven helices, two sheets, seven beta hairpins, eight beta bulges, 13 strands, 11 helices, 13 helix-helix interactions, 23 beta turns, and 4 gamma turns to make up the secondary structure of the pure protein, as shown in Fig. 5.

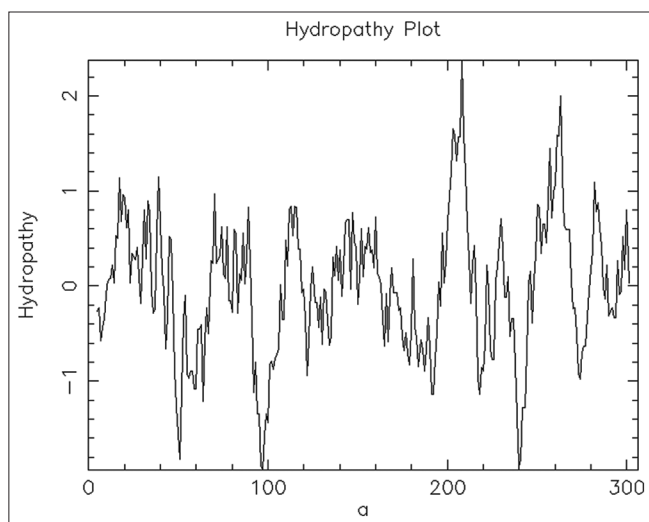


Fig. 3: Hydropathy plot for the protein 7LTN

Retrieval of ligands

Drug likeness and ADMET analysis

Medication-likeness evaluates a molecule's potential to develop into an oral drug in light of the Lipinski rule of 5. To determine whether phytocompounds have structural or physicochemical features advanced enough to be regarded as oral drug candidates, ADMET analysis is used to determine drug-likeness (Tables 2-6).

The fraction of Csp3 that is made up of sp3 hybridized carbons should be at least 0.25 and preferably greater than that. The MW should range from 150 to 500 g/mol in terms of size. The topological polar surface area for polarity should range from 20 to 130. The affinity of a medication for a lipid environment is known as lipophilicity. The ideal

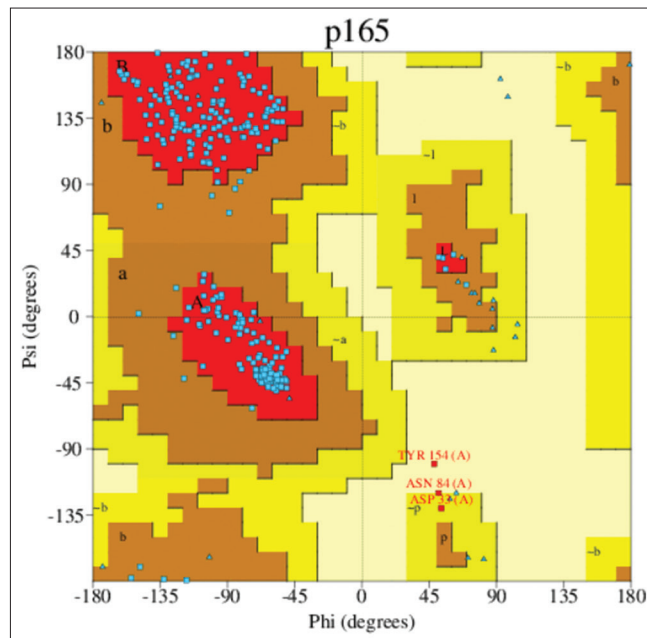


Fig. 4: Ramachandran plot of the protein 7LTN

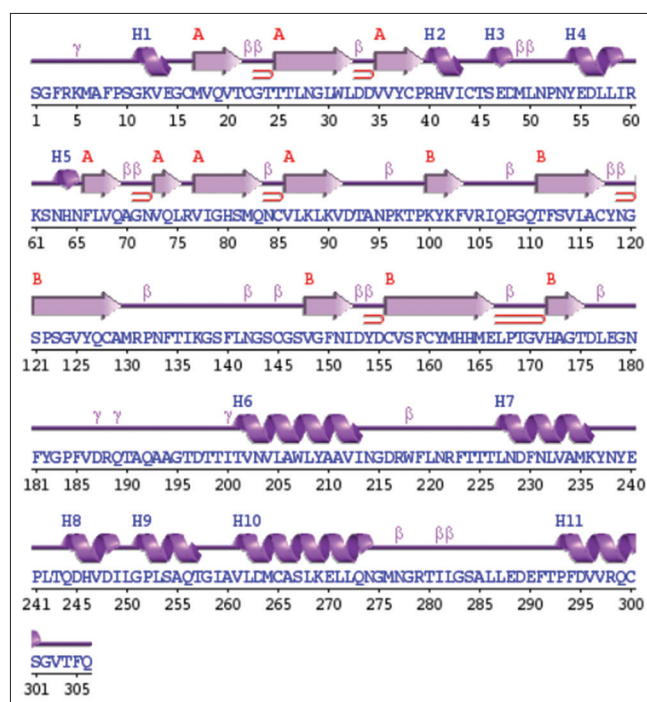


Fig. 5: Secondary structure of protein 7LTN

Table 2: Physicochemical properties of modeled protein

Ligand	MW	Fraction Csp3	RotaTable bonds	TPSA	Lipophilicity
Myristic acid	228.37	0.93	12	37.3	6.11
Lauric acid	200.32	0.92	10	37.3	4.2
Stearic acid	284.48	0.94	16	37.3	8.23
Myristoleic acid	226.36	0.79	11	37.3	5.5
Palmitic acid	256.42	0.94	14	37.3	7.17
Arachidic acid	312.53	0.95	18	37.3	9.29
Palmitoleic acid	254.41	0.81	13	37.3	6.58
Oleic acid	282.46	0.83	15	37.3	7.64
Linoleic acid	280.45	0.72	14	37.3	6.98
Thiamine	265.35	0.42	4	104.15	1.02

Table 3: Lipinski filter analysis

Ligand	MW	MLogP	H donors	H acceptors	MR
Myristic acid	228.37	3.69	1	2	71.18
Lauric acid	200.32	3.15	1	2	61.57
Stearic acid	284.48	4.67	1	2	90.41
Myristoleic acid	226.36	3.58	1	2	70.71
Palmitic acid	256.42	4.19	1	2	80.8
Arachidic acid	312.53	5.13	1	2	100.03
Palmitoleic acid	254.41	4.09	1	2	80.32
Oleic acid	282.46	4.57	1	2	89.94
Linoleic acid	280.45	4.47	1	2	89.46
Thiamine	265.35	0.05	2	3	73.26

MR: Molar refractivity

lipophilicity (LogP) range is between -0.7 and +5.0. From Table 2, it can be seen that all the characteristics are followed by ligands except for lipophilicity. Only thiamine and lauric acid have the appropriate lipophilicity range.

MW must be <500 Daltons, MlogP must be <4.15, hydrogen atom donors and acceptors must be <5 and 10, respectively, and the ideal range for molar refractivity must be between 40 and 130. These are the criteria for the Lipinski rule. From Table 3 it is evident that the ligands Stearic acid, arachidic acid, palmitic acid, oleic acid, and linoleic acid have MlogP values greater than 4.15 and therefore, these ligands do not obey the Lipinski Rule of 5. While the remaining ligands (myristic acid, lauric acid, myristoleic acid, palmitoleic acid and thiamine) fulfil the Lipinski parameters.

BBB enables the blood vessels to precisely control how ions, chemicals, and cells travel between the blood and the brain. Drug development must consider critical factors such as GI absorption. Permeability glycoprotein (PGP) substrate aids in pumping substances out of the cell through the efflux system. Its presence signals that substances can be pumped out of the cell. For drug-likeness, the solubility of the ligand should be <6. If the SA score is <6, the compounds can be made easily. A substance's bioavailability, which depends on absorption and secretion, is a measurement of how much of it can access circulation and reach the target area. bioavailability score of 0.55 indicates that the compound has fulfilled the Lipinski rule of 5. PAINS are molecules giving false-positive results for interaction. Pan assay interference (PAIN) molecules interact with the target protein in a non-specific manner. Here, no ligands exhibit PAIN properties, as shown in Table 4.

Aggregator class ligands fall under category 1, while non-aggregator class ligands belong under category 0. The probability score defines and supports the category class as well. From Table 5, it is seen that all the ligands are non-aggregators.

Following are the several categories for toxicity:

- Class I: if swallowed, deadly ($LD_{50} \leq 5$)
- Class II: If ingested, lethal ($5 < LD_{50} \leq 50$)
- Class III: poisonous if ingested ($50 < LD_{50} \leq 300$)

- Class IV: dangerous if ingested ($300 < LD_{50} \leq 2000$)
- Class V: May be dangerous if ingested ($2000 < LD_{50} \leq 5000$)
- Class VI: Non-toxic ($LD_{50} > 5000$).

According to Table 6, linoleic acid is not harmful if taken as a medication. Based on the tox prediction, swallowing myristoleic, palmitoleic, and oleic acids can be lethal. However, these acids support a healthy heart, cholesterol mechanism, skin, and insulin sensitivity. These could be utilized in the future to make medications. The remaining ligands belonged to class 4.

Boiled-egg analysis

The GI absorption capabilities of the therapeutic molecule are a crucial characteristic in the design and development of drugs and are predicted by the Brain or Intestinal Estimated Permeation Predictive Model (BOILED-Egg). The BOILED-Egg's white section forecasts that ligand molecules falling, there will have higher rates of GI absorption, while those dropping in the yolk and yellow regions will have a better likelihood of crossing the BBB. Fig. 6 depicts that the boiled-egg's white region displays higher GI absorption rates for thiamine, stearic acid, and oleic acid, whereas the yellow region displays higher rates for BBB crossing for the remaining ligands. A blue circle point indicates the presence of PGP in the ligands, whereas a red circle point indicates the absence of PGP. Only thiamine exhibits PGP+, demonstrating its ability to be effluxed from cells efficiently.

Molecular docking and visualization

From Table 7, the ligands which had the most negative binding were selected for visualization (Highlighted in yellow). Three ligands were selected, namely, thiamine, myristoleic acid, and palmitic acid with binding affinities -5.9, -5.0, and -5.0, respectively.

Visualization of the docked complexes

- Docking of Myristoleic acid with modeled protein Mpro.
- Docking of Palmitic acid with modeled protein Mpro.
- Docking of Thiamine with modeled protein Mpro.

DISCUSSION

Human coronaviruses are members of coronaviridae that causes a variety of respiratory illnesses, ranging in severity from pneumonia to bronchiolitis and the common cold [22]. Patients with SARS-CoV-2 infection may not show any symptoms or have signs that are similar to those of other acute viral and bacterial illnesses [23]. The SARS-CoV-2 virus has a spherical form and is enveloped but not segmented. A lipid bilayer makes up the envelope, and spike proteins protrude from the surface of the viral particles [24]. SARS-CoV-2 Mpro is a homodimer protease; it is composed of the two protomers (A and B) that make up its structure. On dimerization and activation, these two protomers arrange themselves in the proper conformation to carry out the catalytic action [25]. Given its very unique structure, SARS-CoV-2 Mpro is unquestionably the most intriguing molecular target for the pharmaceutical therapy of COVID-19. All coronaviruses' replication cycles are dominated by the SARS-CoV-2 Mpro [26,27].

Table 4: Admesar analysis

Ligands	BBB barrier	GI absorption	PGP substrate	Solubility (LOGSw-SILICOS IT)	PAINS	Bioavailability	SA SCORE
Myristic acid	Yes	High	No	-4.51 (Moderately soluble)	0	0.85	2.09
Lauric acid	Yes	High	No	-3.69 b (Soluble)	0	0.85	1.87
Stearic acid	No	High	No	-6.11 (Poorly soluble)	0	0.85	2.54
Myristoleic acid	Yes	High	No	-3.79 (Soluble)	0	0.85	2.63
Palmitic acid	Yes	High	No	-5.31 (Moderately soluble)	0	0.85	2.31
Arachidic acid	No	Low	No	-6.91 (Poorly soluble)	0	0.85	2.77
Palmitoleic acid	Yes	High	No	-4.59 (Moderately soluble)	0	0.85	2.84
Oleic acid	No	High	No	-5.39 (Moderately soluble)	0	0.85	3.07
Linoleic acid	Yes	High	No	-4.67 (Moderately soluble)	0	0.85	3.1
Thiamine	No	High	Yes	-3.3 (Soluble)	0	0.55	2.99

BBB: Blood-brain barrier; PGP: Permeability glycoprotein, GI: Gastrointestinal, SA: Synthetic accessibility

Table 5: Aggregation parameters of the ligands

Compound	Canonical smiles	Probability score	Aggregator class
Myristic acid	CCCCCCCCCCCC(=O)O	0.009	0
Lauric acid	CCCCCCCCCCC(=O)O	0.009	0
Stearic acid	CCCCCCCCCCCCCCCC(=O)O	0.009	0
Myristoleic acid	CCCC=CCCCCCCC(=O)O	0.011	0
Palmitic acid	CCCCCCCCCCCCCCCC(=O)O	0.009	0
Arachidic acid	CCCCCCCCCCCCCCCCCCC(=O)O	0.009	0
Palmitoleic acid	CCCCC=CCCCCCCC(=O)O	0.011	0
Oleic acid	CCCCCCCC=CCCCCCCC(=O)O	0.011	0
Linoleic acid	CCCCC=CCC=CCCCCCCC(=O)O	0.01	0
Thiamine	OCCc1sc[n+](c1C) Cc1cnc (nc1N)C	0.005	0

Table 6: Toxicity analysis

Compound	Canonical smiles	Predicted LD 50 (mg/kg)	Predicted toxicity class
Myristic acid	CCCCCCCCCCCC(=O)O	900	4
Lauric acid	CCCCCCCCCCC(=O)O	900	4
Stearic acid	CCCCCCCCCCCCCCCC(=O)O	900	4
Myristoleic acid	CCCC=CCCCCCCC(=O)O	48	2
Palmitic acid	CCCCCCCCCCCCCCCC(=O)O	900	4
Arachidic acid	CCCCCCCCCCCCCCCCCCC(=O)O	900	4
Palmitoleic acid	CCCCC=CCCCCCCC(=O)O	48	2
Oleic acid	CCCCCCCC=CCCCCCCC(=O)O	48	2
Linoleic acid	CCCCC=CCC=CCCCCCCC(=O)O	10000	6
Thiamine	OCCc1sc[n+](c1C) Cc1cnc (nc1N)C	1000	4

Table 7: Binding affinity of ligands

Ligand	Binding affinity
Arachidic acid	-4
Myristic acid	-4.5
Thiamine	-5.9
Lauric acid	-4.5
Palmitoleic acid	-4.9
Oleic acid	-4.5
Linoleic acid	-4.8
Myristoleic acid	-5
Stearic acid	-4.7
Palmitic acid	-5

C. papaya belongs to the family Caricaceae and is well known for its therapeutic and nutritional properties all over the world. The different parts of the papaya plant have been used since ancient times for its therapeutic applications [28]. The papaya plant, comprising the fruit, leaves, seeds, bark, latex, and other parts of it, is crucial in controlling the spread of disease. Alkaloids, glycosides, tannins, saponins, and flavonoids are some of the active ingredients in *C. papaya* leaf that give it its medicinal properties. Papaya leaf juice also raises platelet counts in dengue fever patients which are an additional benefit [29]. Flavonoid-rich plants aid in reducing SARS-Cov-2's Mpro activity [30]. Therefore, phytochemicals from *C. papaya* were chosen to study their efficiency against Mpro.

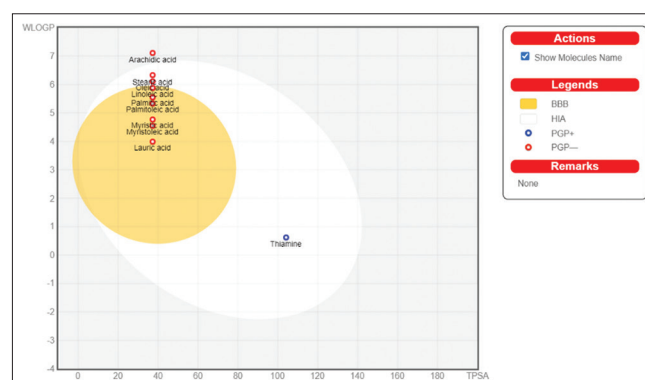


Fig. 6: Brain or Intestinal Estimated permeation method (BOILED-EGG)

Myristic acid, lauric acid, stearic acid, myristoleic acid, palmitic acid, arachidic acid, palmitoleic acid, oleic acid, linoleic acid, and thiamine were the ten phytochemicals of *C. papaya* (Fig. 7) chosen for the *in silico* experiment. Tannins, flavonoids, alkaloids, saponins, phytosterol, flavanol, phenylpropanoid, glycosides, and substances of the flavin class were revealed by these phytochemicals. These compounds' physicochemical, drug-like, ADMET, toxicological, and pharmacodynamic

characteristics made them useful in the development of therapeutic medicines (Table 2-6).

To comply with the Lipinski rule of five (Table 3), an ADME study was performed on the 10 phytochemicals that were chosen. The chemicals' solubility and permeability were examined (Table 4). Bio tools were also used to test the compound's absorption and binding to the target. To assess drug excretion, the ligands were examined to see if they could be readily effluxed from the system (Table 4 and Fig. 6). To ensure safe drug consumption, the ligands were examined to see if they aggregated at the protein site (Table 5) and all the compounds had their toxicities predicted (Table 6). The ten ligands were then screened to determine the molecular interactions between the ligand-protein complex after the pharmacological investigations. Three compounds, myristoleic acid, palmitic acid, and thiamine, were shown to have the most effective binding (Table 7), *in silico* experiments were conducted to demonstrate their molecular interactions by docking these three chemicals into the Mpro protein targets of the SARS-Cov-2 (Figs. 8-10).

The noteworthy chemicals chosen might contain flavonoids. According to studies [30], flavonoids have a pleiotropic effect on the

Mpro protein target of the SARS-Cov-2 virus. The Biovia program was used to dock the three chosen ligands with protein Mpro. Both a 2D and 3D structural diagram of them showed how they interacted (Figs. 8-10), as well as the characteristics of their amino acids were observed (Tables 8-10).

Numerous amino acids, including glutamine, threonine, serine, valine, isoleucine, asparagine, phenylalanine, aspartic acid, and arginine, were present in the 2D structure of the Myristoleic acid-Mpro protein complex (Fig. 8). Conventional hydrogen bonds were used to attach threonine and aspartic acid, while alkyl and pi-alkyl groups were used to bind valine and phenylalanine to myristoleic acid showing hydrophobic characters (Table 8).

Amino acids such as phenylalanine, asparagine, leucine, glutamic acid, glycine, tryptophan, methionine, and arginine were seen in the 2D structure of the palmitic acid-Mpro protein complex (Fig. 9). A typical hydrogen bond held glycine to palmitic acid. Hydrophobic interactions were seen among other amino acids. Tryptophan demonstrated pi-alkyl binding to the palmitic acid, whereas leucine demonstrated alkyl binding to the ligand (Table 9).

The structure of thiamine-Mpro protein complex consisted of valine, aspartic acid, asparagine, glutamine, threonine, arginine, phenylalanine, isoleucine, serine, and cysteine (Fig. 10). Only the amino acid phenylalanine exhibited hydrophobic properties, whereas those of the other amino acids were hydrogen-binding. While phenylalanine was pi-stacked to thiamine, aspartic acid, serine, and glutamine were bound to thiamine by conventional hydrogen bonds (Table 10).

The production of hormones and proteins depends on amino acids, which also support healthy immunological responses to illnesses. In the SARS-Cov-2, phenylalanine is used as a severity marker. Increased inflammation is associated with higher phenylalanine concentration levels. In addition, phenylalanine is linked to greater mortality [31]. Leucine and isoleucine are examples of branched-chain amino acids that support anabolic pathways, promote healing, and combat sepsis. Due to the drop in glycine, immune cell phagocytosis and antibody synthesis become reduced [32]. Glutamine promotes immune system development, and lymphocyte production by producing antibodies, and cell growth while preventing apoptosis [33,34]. To boost the immune system, aspartic acid aids in the creation of antibodies. Threonine influences the immune system and plays a significant role in immunoglobulins and intestinal mucins [35]. Tryptophan contributes to the health of the immune system and is involved in neuronal function [36]. The synthesis of tryptophan requires serine. Serine facilitates protein synthesis, neurotransmission, and the manufacture of purines and pyrimidines [37]. The phytochemicals' amino acids may or may not contribute to the Mpro protein's inhibition. However, these findings need to be supported by larger-scale clinical and

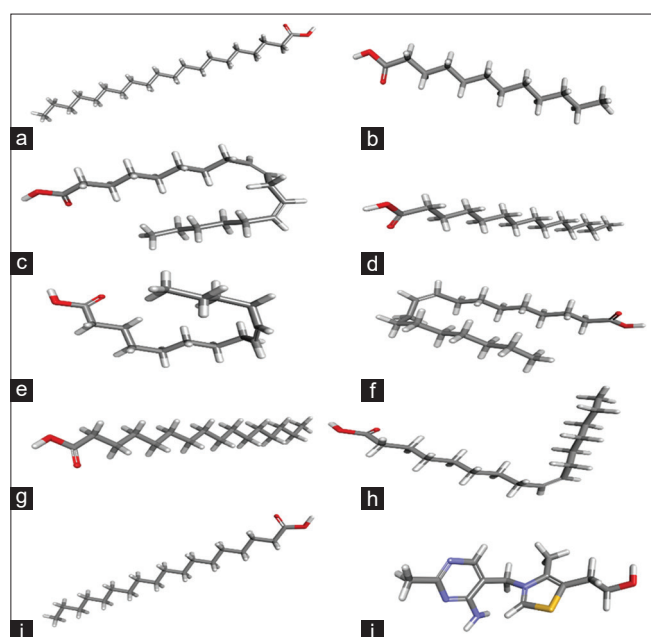


Fig. 7: 3D structures of phytochemicals from *Carica papaya*. (a) Arachidic acid, (b) lauric acid, (c) linoleic acid, (d) myristic acid, (e) myristoleic acid, (f) oleic acid, (g) palmitic acid, (h) palmitoleic acid, (i) stearic acid, and (j) thiamine

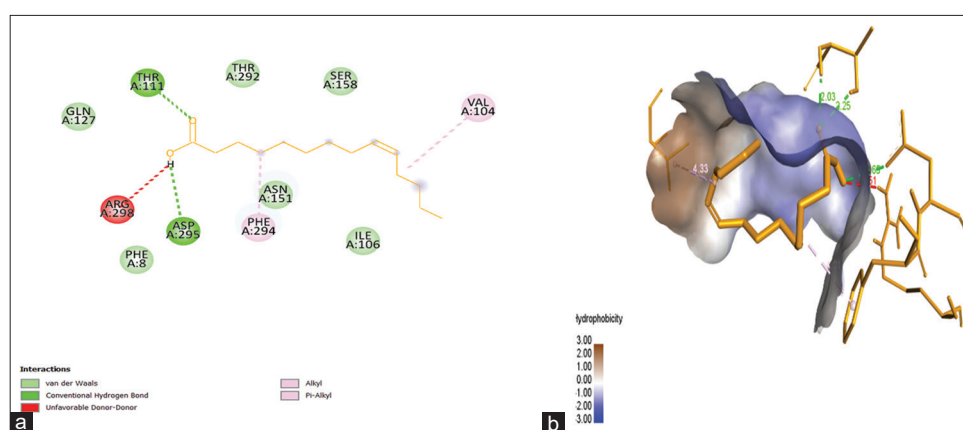


Fig. 8: Visualization of myristoleic acid with modeled protein Mpro. (a) 2D interaction diagram and (b) 3D interaction structure

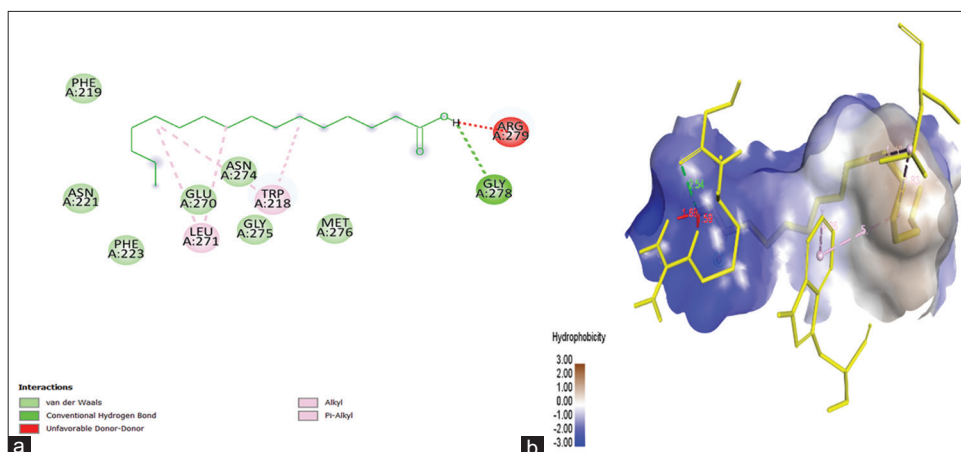


Fig. 9: Visualization of palmitic acid with modeled protein Mpro. (a) 2D interaction diagram and (b) 3D interaction structure

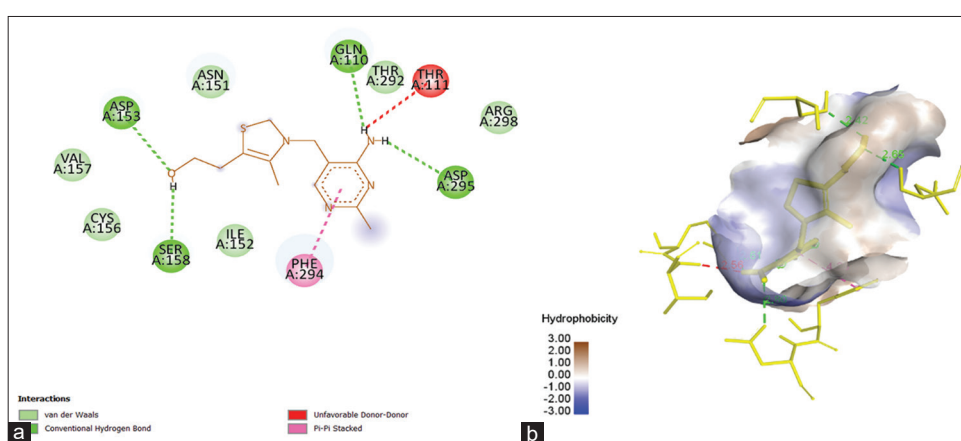


Fig. 10: Visualization of thiamine with modeled protein Mpro. (a) 2D interaction diagram and (b) 3D interaction structure

Table 8: Docking interactions of Myristoleic acid-protein complex

Name	Distance	Category	Types
A: THR111:HN-N:	2.03261	Hydrogen bond	Conventional hydrogen bond
UNK1:O			hydrogen bond
A: THR111:HG1-N:	2.25096	Hydrogen bond	Conventional hydrogen bond
UNK1:O			hydrogen bond
N: UNK1:H-A:	2.6455	Hydrogen bond	Conventional hydrogen bond
ASP295:OD1			hydrogen bond
A: VAL104-N: UNK1	4.33151	Hydrophobic	Alkyl
A: PHE294-N: UNK1	4.01363	Hydrophobic	Pi-Alkyl

Table 9: Docking interactions of Palmitic acid-protein complex

Name	Distance	Category	Types
N: UNK1:H-A:	2.54428	Hydrogen bond	Conventional hydrogen bond
GLY278:O			hydrogen bond
N: UNK1-A: LEU271	5.43381	Hydrophobic	Alkyl
N: UNK1-A: LEU271	4.83213	Hydrophobic	Alkyl
A: TRP218-N: UNK1	4.85982	Hydrophobic	Pi-Alkyl
A: TRP218-N: UNK1	5.17118	Hydrophobic	Pi-Alkyl

Table 10: Docking interactions of thiamine-protein complex

Name	Distance	Category	Types
A: ASP153:HN-N:	2.6478	Hydrogen bond	Conventional hydrogen bond
UNK1:O			hydrogen bond
N: UNK1:H-A:	2.41533	Hydrogen bond	Conventional hydrogen bond
SER158:OG			hydrogen bond
N: UNK1:H-A:	2.61048	Hydrogen bond	Conventional hydrogen bond
GLN110:OE1			hydrogen bond
N: UNK1:H-A:	2.79663	Hydrogen bond	Conventional hydrogen bond
ASP295:OD1			hydrogen bond
A: PHE294-N: UNK1	4.25707	Hydrophobic	Pi-Pi stacked

that the flavonoids in *C. papaya* may aid in reducing activity [30]. The three lastly selected phytochemicals were used in *in silico* experiments using molecular docking and visualization, and it was discovered that they can be exploited in medication development. However, these ligands are amenable to *in vitro* testing and can be used to create powdered medications.

CONCLUSION

C. papaya possesses a wide range of therapeutic benefits, including antibacterial, antiviral, anticancer, hypoglycemic, and anti-inflammatory activities. Studies have shown that *C. papaya* leaves may be an effective SARS-Coronavirus-2 antiviral. Ten phytochemicals were used for the *in silico* tests and later filtered to three ligands with the strongest binding affinities that are myristoleic acid, palmitic acid, and thiamine.

experimental research.

The Mpro protein of SARS-Cov-2 may be inhibited by the *C. papaya*'s immune actions, which may stop its proliferation by reducing the activity of the Mpro protein of SARS coronavirus. Studies have shown

These can be used to power pharmaceutical drugs for the creation of SARS-Coronavirus-2 medications. The therapeutic value of similar plants with flavonoid chemicals to prevent Mpro of SARS-Cov-2 in the future can also be assessed.

ACKNOWLEDGMENT

I hereby acknowledge the Department of Bioinformatics, BioNome, Bengaluru, India for providing computational facilities and support in the scientific research services. I thank Ms. Susha Dinesh and Mr. Sameer Sharma for their assistance throughout the project.

REFERENCES

- Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol* 2019;17:181-92.
- Asselah T, Durantel D, Pasmant E, Lau G, Schinazi RF. COVID-19: Discovery, diagnostics and drug development. *J Hepatol* 2021;74:168-84.
- Yesudhas D, Srivastava A, Gromiha MM. COVID-19 outbreak: History, mechanism, transmission, structural studies and therapeutics. *Infection* 2021;49:199-213.
- Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* 2020;579:265-9.
- Hegyi A, Ziebuhr J. Conservation of substrate specificities among coronavirus main proteases. *J Gen Virol* 2002;83:595-9.
- Yang H, Xie W, Xue X, Yang K, Ma J, Liang W, *et al.* Design of wide-spectrum inhibitors targeting coronavirus main proteases. *PLoS Biol* 2005;3:e324.
- Ji Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, *et al.* Structure of M^{pro} from SARS-CoV-2 and discovery of its inhibitors. *Nature* 2020;582:289-93.
- Hariyono P, Patramurti C, Candrasari DS, Hariono M. An integrated virtual screening of compounds from *Carica papaya* leaves against multiple protein targets of SARS-Coronavirus-2. *Results Chem* 2021;3:100113.
- Da Silva JA, Rashid Z, Nhut DT, Sivakumar D, Gera A, Souza MT, *et al.* Papaya (*Carica papaya* L.) biology and biotechnology. *Tree Forest Sci Biotechnol* 2007;1:47-73.
- Yogiraj V, Goyal PK, Chauhan CS, Goyal A, Vyas B. *Carica papaya* Linn: An overview. *Int J Herbal Med* 2014;2:1-8.
- Behzadi P, Gajdacs M. Worldwide protein data bank (wwPDB): A virtual treasure for research in biotechnology. *Eur J Microbiol Immunol (Bp)* 2021;11:77-86.
- Zappa M, Verdecchia P, Spanevello A, Angeli F. Structural evolution of severe acute respiratory syndrome coronavirus 2: Implications for adhesivity to angiotensin-converting enzyme 2 receptors and vaccines. *Eur J Intern Med* 2022;104:33-6.
- Laskowski RA. PDBsum new things. *Nucleic Acids Res* 2009;37:D355-9.
- Mohanraj K, Karthikeyan BS, Vivek-Ananth RP, Chand R, Aparna SR, Mangalapandi P, *et al.* IMPPAT: A curated database of indian medicinal plants, phytochemistry and therapeutics. *Sci Rep* 2018;8:4329.
- Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, *et al.* PubChem in 2021: New data content and improved web interfaces. *Nucleic Acids Res* 2021;49:D1388-95.
- Sakshuwong S, Weir H, Raucchi U, Martínez TJ. Bringing chemical structures to life with augmented reality, machine learning, and quantum chemistry. *J Chem Phys* 2022;156:204801.
- 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* 2017;7:42717.
- Xiong G, Wu Z, Yi J, Fu L, Yang Z, Hsieh C, *et al.* ADMETlab 2.0: An integrated online platform for accurate and comprehensive predictions of ADMET properties. *Nucleic Acids Res* 2021;49:W5-14.
- Stork C, Mathai N, Kirchmair J. Computational prediction of frequent hitters in target-based and cell-based assays. *Artif Intell Life Sci* 2021;1:100007.
- Fareed MM, El-Esawi MA, El-Ballat EM, Batiha GE, Rauf A, El-Demerdash FM, *et al.* *In silico* drug screening analysis against the overexpression of PGAM1 gene in different cancer treatments. *BioMed Res Int* 2021;2021:5515692.
- Dallakyan S, Olson AJ. Small-molecule library screening by docking with PyRx. *Methods Mol Biol* 2015;1263:243-50.
- Kirtipal N, Bharadwaj S, Kang SG. From SARS to SARS-CoV-2, insights on structure, pathogenicity and immunity aspects of pandemic human coronaviruses. *Infect Genet Evol* 2020;85:104502.
- Cohen PA, Hall LE, John JN, Rapoport AB. The early natural history of SARS-CoV-2 infection: Clinical observations from an urban, ambulatory COVID-19 clinic. *Mayo Clin Proc* 2020;95:1124-6.
- Fiorino S, Tateo F, De Biase D, Gallo CG, Orlandi PE, Corazza I, *et al.* SARS-CoV-2: Lessons from both the history of medicine and from the biological behavior of other well-known viruses. *Future Microbiol* 2021;16:1105-33.
- Zhong N, Zhang S, Zou P, Chen J, Kang X, Li Z, *et al.* Without its N-finger, the main protease of severe acute respiratory syndrome coronavirus can form a novel dimer through its C-terminal domain. *J Virol* 2008;82:4227-34.
- Zhang L, Lin D, Sun X, Curth U, Drosten C, Sauerhering L, *et al.* Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α -ketoamide inhibitors. *Science* 2020;368:409-12.
- Hilgenfeld R. From SARS to MERS: Crystallographic studies on coronavirus proteases enable antiviral drug design. *FEBS J* 2014;281:4085-96.
- Singh SP, Kumar S, Mathan SV, Tomar MS, Singh RK, Verma PK, *et al.* Therapeutic application of *Carica papaya* leaf extract in the management of human diseases. *Daru* 2020;28:735-44.
- Vij T, Prashar Y. A review on medicinal properties of *Carica papaya* Linn. *Asian Pac J Trop Dis* 2015;5:1-6.
- Jo S, Kim S, Shin DH, Kim MS. Inhibition of SARS-CoV 3CL protease by flavonoids. *J Enzyme Inhib Med Chem* 2020;35:145-51.
- Luporini RL, Pott-Junior H, Leal MC, Castro A, Ferreira AG, Cominetti MR, *et al.* Phenylalanine and COVID-19: Tracking disease severity markers. *Int Immunopharmacol* 2021;101:108313.
- Holeček M. Branched-chain amino acids in health and disease: Metabolism, alterations in blood plasma, and as supplements. *Nutr Metab (Lond)* 2018;15:33.
- Wu GY, Field CJ, Marliss EB. Elevated glutamine metabolism in splenocytes from spontaneously diabetic BB rats. *Biochem J* 1991;274:49-54.
- Newsholme P, Curi R, Curi TC, Murphy CJ, Garcia C, de Melo MP. Glutamine metabolism by lymphocytes, macrophages, and neutrophils: Its importance in health and disease. *J Nutr Biochem* 1999;10:316-24.
- Zhang H, Chen Y, Li Y, Zhang T, Ying Z, Su W, *et al.* l-Threonine improves intestinal mucin synthesis and immune function of intrauterine growth-retarded weanling piglets. *Nutrition*. 2019;59:182-7.
- Comai S, Bertazzo A, Brughera M, Crotti S. Tryptophan in health and disease. *Adv Clin Chem* 2020;95:165-218.
- Holeček M. Serine metabolism in health and disease and as a conditionally essential amino acid. *Nutrients* 2022;14:1987.