Natural Resources for Human Health
eISSN: 2583-1194

Original Research

Received 21 March 2026
Revised 26 April 2026
Accepted 28 May 2026

Natural Resources for Human Health 2026, Vol. 6 (3): 168–178

KEYWORDS:
SARS-CoV-2 main protease; virtual screening; phytochemicals; ADMET; drug-likeness; natural products

AI-Assisted Virtual Screening and ADMET-Guided Prioritisation of South Indian Plant-Derived Phytochemicals Against the SARS-CoV-2 Main Protease (Mpro): A Reproducible Computational Framework and Curated Compound Library

Prerana Nilesh Khairnar[1], Nilufar Esanmuradova[2], Dr.S. Vidhya[3], Sorabh Sharma[4], Karabi Kalita[5], Dr. B. Senthil Kumar[6], Dr. Sandeep Kumar Panigrahi[7]

[1] Department of Computer Engineering, Sir Visveavaraya Institute Of Technology, Chincholi, Nashik, Maharashtra, India.autadeprerana@gmail.com

[2]1. Head of the Department of Physics and Chemistry, "Tashkent Institute of Irrigation and Agricultural Mechanization Engineers" National Research University, Tashkent, Uzbekistan; 2. Scientific Researcher, University of Tashkent for Applied Science; 3. School of Engineering, Central Asian University, Tashkent, Uzbekistan, sapaevibrokhim@gmail.com, 0009-0006-5617-0545

[3]Associate Professor, Department of Computer Science, KG College of Arts and Science, Coimbatore, Tamil Nadu, India vidhya.s@kgcas.com

[4]Centre for Research Impact & Outcome, Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, 140401, Punjab, India. sorabh.sharma.orp@chitkara.edu.in https://orcid.org/0009-0000-1805-0418

[5]School of Pharmacy and Research, Dev Bhoomi Uttarakhand University, sopr.karabi@dbuu.ac.in, 0000-0001-8254-1991

[6]Professor, Department of Anatomy, Vinayaka Mission's Kirupananda Variyar Medical College & Hospitals
Vinayaka Mission's Research Foundation (DU), Salem, Tamil Nadu, India. skdrchinu88@gmail.com, 0000-0001-6162-6312

[7]Professor, Department of Community Medicine, IMS and SUM Hospital, Siksha 'O' Anusandhan (Deemed to be University), sandeepkumarpanigrahi@soa.ac.in, 0000-0001-6005-6100

ABSTRACT: Background: The main protease (M^{pro}; $3CL^{pro}$; nsp5) of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is an essential, highly conserved enzyme that governs viral polyprotein maturation and has no close human homologue, making it a premier antiviral target. Plant-derived phytochemicals, many rooted in the ethnomedicine of South India, provide a chemically diverse starting point for inhibitor discovery. We assembled a curated library and established a transparent, reproducible artificial-intelligence (AI)-assisted virtual-screening and ADMET-profiling framework centred on this target. Methods: Fifty phytochemicals with verified PubChem identities and representative South Indian botanical sources were compiled against the M^{pro} crystal structure 6LU7 (catalytic dyad His41/Cys145). Two-dimensional structures were curated as canonical SMILES, and physicochemical descriptors (molecular weight, calculated

LogP, hydrogen-bond donors/acceptors, topological polar surface area [TPSA], rotatable bonds) were computed with RDKit. A tiered prioritisation cascade applied Lipinski and Veber drug-likeness rules and ranked survivors by the quantitative estimate of drug-likeness (QED). The docking, AI activity-probability, and predictive ADMET modules of the framework, together with the editable composite-score weighting scheme, are fully specified for downstream execution. Results: The library spanned 32 phytochemical classes, dominated by flavonols, anthraquinones, flavones and triterpenoids. Forty-nine compounds were descriptor-profiled (mean molecular weight 313 g/mol; mean cLogP 2.32); 45 satisfied the Lipinski rule of five (<=1 violation) and 40 additionally passed the Veber criteria. The four rejected scaffolds were large, highly hydroxylated glycosides and a triterpenoid saponin (rutin, hesperidin, epigallocatechin gallate, glycyrrhizin). The top QED-ranked candidates included zingerone, hesperetin, naringenin, ferulic acid, carnosic acid, daidzein, eugenol, resveratrol and berberine (QED 0.63-0.80). Conclusion: We present a verified 50-compound South Indian phytochemical library and a transparent, integrity-preserving computational pipeline for M^{pro} inhibitor discovery. Small phenolics, flavanones and isoflavones emerge as favourable, drug-like starting points warranting subsequent molecular docking, AI scoring, ADMET prediction and experimental validation.

1. INTRODUCTION

The coronavirus disease 2019 (COVID-19) pandemic, caused by SARS-CoV-2, underscored the urgent and continuing need for orally available small-molecule antivirals that can be deployed rapidly against emerging variants. Among the druggable viral proteins, the main protease (M^{pro}, also termed 3-chymotrypsin-like protease, 3CL^{pro}, or nsp5) is one of the most attractive. It cleaves the viral polyproteins pp1a and pp1ab at eleven conserved sites to release the non-structural proteins that assemble the replication-transcription complex, and its activity is indispensable for viral propagation. Because M^{pro} recognises a glutamine at the P1 position, a specificity not shared by any known human protease, inhibitors can achieve viral selectivity with a comparatively low risk of mechanism-based host toxicity [1-5].

High-resolution crystal structures of M^{pro}, notably the complex with the peptidomimetic inhibitor N3 (Protein Data Bank entry 6LU7), revealed a cysteine-histidine catalytic dyad (Cys145-His41) buried within a substrate-binding cleft lined by Met49, Gly143, Ser144, His163, His164, Met165, Glu166 and Gln189 [6,7]. These structures catalysed intensive structure-based design campaigns, culminating in clinically validated inhibitors and providing a reliable receptor template for computational screening [8]. Structure-guided virtual screening now routinely triages large chemical collections before any compound is synthesised or assayed, greatly reducing the cost and time of early discovery.

Natural products remain an exceptionally productive source of antiviral chemotypes, offering scaffold diversity and privileged three-dimensional complexity that synthetic libraries seldom match. The medicinal flora of South India, embedded in Ayurvedic and Siddha practice, is particularly rich in polyphenols, flavonoids, alkaloids and terpenoids with documented antiviral, anti-inflammatory and immunomodulatory activities. Compounds such as quercetin, curcumin (from *Curcuma longa*), piperine (from *Piper nigrum*), withaferin A (from *Withania somnifera*) and andrographolide (from *Andrographis paniculata*) have repeatedly surfaced as candidate M^{pro} binders in computational studies, and several Indian and South Indian ethnobotanical panels have been screened against SARS-CoV-2 targets [5,9].

The integration of artificial intelligence and machine learning into this workflow has further accelerated hit identification, enabling activity probabilities and consensus rankings to be combined with physics-based docking and predictive absorption, distribution, metabolism, excretion and toxicity (ADMET) profiling [3,10]. However, the value of any such pipeline depends critically on the integrity and reproducibility of its inputs. Docking scores obtained under different receptor preparations or scoring functions cannot be pooled without normalisation, and predicted values must never be confused with experimentally measured ones [7,11,12].

In this work we (i) assemble and verify a 50-compound library of South Indian and pan-Indian plant-derived phytochemicals against the M^{pro} template 6LU7; (ii) execute a transparent physicochemical and drug-likeness screening cascade using structure-confirmed descriptors; (iii) prioritise candidates by an interpretable drug-likeness metric; and (iv) fully specify the AI-scoring, molecular-docking and ADMET modules of the framework so that the pipeline can be reproduced and extended. Throughout, we adhere to a strict data-integrity policy: every reported numerical value derives either from the source dataset or from an openly reproducible computation, and no docking, AI or ADMET result is fabricated.

2. MATERIALS AND METHODS

2.1 Molecular target selection and preparation

The SARS-CoV-2 M^{pro} ($3CL^{pro}$; nsp5) crystal structure 6LU7 was adopted as the primary receptor template. The catalytic dyad comprises His41 and Cys145, and the substrate-binding site is defined by His41, Met49, Gly143, Ser144, Cys145, His163, His164, Met165, Glu166 and Gln189. The reference ligand co-crystallised in 6LU7 is the covalent peptidomimetic N3. For structure-based docking (specified below) the protein is prepared by removing the co-crystallised ligand and crystallographic waters, adding polar hydrogens and partial charges, and defining a grid box centred on the N3 binding cleft; alternative templates (6Y2E, 7BQY) are recorded as permissible provided scores are not pooled across preparations without normalisation.

2.2 Compound library assembly

Fifty candidate phytochemicals were compiled, each annotated with a preferred name, a representative (non-exclusive) South Indian or pan-Indian botanical source, plant part, phytochemical class and a PubChem Compound Identifier (CID). Sources spanned dietary and medicinal species including *Curcuma longa, Zingiber officinale, Piper nigrum, Ocimum basilicum, Withania somnifera, Andrographis paniculata, Azadirachta indica, Allium sativum, Allium cepa, Camellia sinensis* and *Citrus* species, among others. Compound identity was treated as the authoritative anchor: canonical structures were retrieved and reconciled against the recorded PubChem CIDs before any descriptor was computed.

2.3 Structure curation and physicochemical descriptor computation

Two-dimensional structures were curated as canonical SMILES and standardised (neutralised, with defined stereochemistry retained). Physicochemical descriptors were computed uniformly with RDKit: molecular weight (MW), Crippen calculated LogP (cLogP), the number of hydrogen-bond donors (HBD) and acceptors (HBA), topological polar surface area (TPSA), and the number of rotatable bonds. Computing all descriptors with a single, open-source toolkit ensured internal consistency and full reproducibility [13-15]. One structurally complex limonoid (azadirachtin, $C_{35}H_{44}O_{16}$) was retained in the library for completeness but excluded from the numerical descriptor analysis, leaving 49 evaluable compounds.

2.4 AI/ML scoring module (framework specification)

The framework provides for a supervised machine-learning classifier that outputs, for each ligand, an AI activity probability scaled 0-1 together with a confidence/uncertainty metric, and a consensus rank derived from the agreement between the AI score and the docking rank. Model identity, training set and calibration are to be documented at execution. In the source dataset these fields are reserved and were not populated; consequently, no AI probability is reported here, and the module is presented as a fully specified, reproducible component rather than as generated results [16].

2.5 Molecular docking protocol (framework specification)

Docking is specified against the prepared 6LU7 receptor using a single, consistently applied engine (for example, AutoDock Vina), with the search space enclosing the His41/Cys145 catalytic cleft. For each ligand the protocol records the raw docking score (kcal/mol; more negative indicating a more favourable predicted pose within one protocol), pose RMSD, hydrogen-bond and hydrophobic contact counts, and the set of key interacting residues. In line with best practice, raw scores are not compared across different software, grids, receptor preparations or protonation states without normalisation. These fields were unpopulated in the source dataset and are therefore not reported as results.

2.6 ADMET screening cascade (framework specification)

The predictive ADMET cascade evaluates gastrointestinal absorption (High/Low/Unknown), blood-brain-barrier permeation, P-glycoprotein substrate status, inhibition of five cytochrome P450 isoforms (CYP1A2, CYP2C19, CYP2C9, CYP2D6,

CYP3A4), Ames mutagenicity, hERG cardiac liability (Low/Medium/High/Unknown), hepatotoxicity, predicted oral LD50, bioavailability score, PAINS alerts, using controlled vocabularies to prevent guessed values. A composite ADMET risk score is defined as the count of high-risk flags (Yes + High + 0.5 x Medium across the Ames, hERG and hepatotoxicity endpoints). Endpoint definitions are summarised in Table 3, and can be populated with established predictors such as SwissADME and pkCSM . In the source dataset all ADMET fields were recorded as "Unknown" pending prediction; accordingly, no ADMET call is reported here as a finding [17].

2.7 Drug-likeness filtering, composite scoring and ranking

The executed prioritisation applied two orthogonal, widely used oral-drug-likeness filters. The Lipinski rule of five flags a violation for MW > 500, cLogP > 5, HBD > 5 or HBA > 10; compounds with at most one violation were retained. The Veber criteria flag poor oral bioavailability when TPSA > 140 angstrom-squared or rotatable bonds > 10. Compounds passing both (<=1 Lipinski violation and 0 Veber violations) were designated "drug-likeness pass" and ranked by the quantitative estimate of drug-likeness (QED), an integrative desirability score bounded 0-1. The workbook additionally defines an editable composite score that linearly combines the negated docking score (weight 1), the AI activity probability (weight 5), a Lipinski-violation penalty (weight 1), a Veber-violation penalty (weight 0.5) and the ADMET risk penalty (weight 1); its parameters are given in Table 1 and become computable once the docking, AI and ADMET modules are executed.

3. RESULTS

3.1 Composition of the phytochemical library

The 50-compound library represented 32 distinct phytochemical classes (Figure 2). The best-represented classes were flavonols (n = 5; quercetin, kaempferol, myricetin, fisetin, morin), anthraquinones (n = 4; emodin, aloe-emodin, rhein, chrysophanol), pentacyclic triterpenoids (n = 3; ursolic, oleanolic and betulinic acids) and flavones (n = 3; luteolin, apigenin, baicalein). Flavanones, flavan-3-ols, isoflavones, hydroxycinnamic acids, monoterpenoid phenols, phenylpropanoids and phenolic ketones each contributed two members, with the remaining classes represented by single compounds. This distribution reflects the polyphenol-rich character of the South Indian dietary and medicinal flora.

3.2 Physicochemical property landscape

Across the 49 evaluable compounds, molecular weight ranged from 132.2 g/mol (cinnamaldehyde) to 822.9 g/mol (glycyrrhizin), with a mean of 313 g/mol; cLogP ranged from -1.69 (rutin) to 7.23, with a mean of 2.32, indicating a predominantly moderate lipophilicity well suited to oral absorption. The mean TPSA was 97.1 angstrom-squared. The property space (Figure 3) shows the bulk of the library clustered comfortably within the Lipinski MW < 500 and cLogP < 5 boundaries, with a small number of high-molecular-weight glycosides and a saponin projecting beyond them.

3.3 Drug-likeness screening cascade

The sequential cascade (Figure 1) reduced the 50 seeded compounds to 49 descriptor-profiled structures, of which 45 (91.8%) satisfied the Lipinski rule of five with at most one violation. Applying the Veber criteria in addition retained 40 compounds (81.6% of evaluable) as fully drug-like. Four scaffolds were rejected for multiple Lipinski violations combined with excessive polar surface area or hydrogen-bonding capacity: rutin (MW 610.5; 3 violations; TPSA 269), hesperidin (MW 610.6; 3 violations; TPSA 234), epigallocatechin gallate (MW 458.4; 2 violations; TPSA 197) and glycyrrhizin (MW 822.9; 3 violations; TPSA 267). These large, highly hydroxylated flavonoid glycosides and the triterpenoid saponin have size and polarity expected to limit passive oral absorption, consistent with their known pharmacokinetic behaviour.

3.4 Prioritised candidates

Ranking the 40 drug-like compounds by QED highlighted a set of small phenolics, flavanones and isoflavones as the most favourable oral-drug-like starting points (Table 2, Figure 4). The ten highest-ranked candidates were zingerone (QED 0.796; *Zingiber officinale*), hesperetin (0.789; *Citrus*), naringenin (0.742; *Citrus*), ferulic acid (0.715; *Oryza sativa*), carnosic acid (0.705; *Salvia rosmarinus*), daidzein (0.700; *Glycine max*), eugenol (0.693; *Syzygium aromaticum*), resveratrol (0.692; *Vitis vinifera*), berberine (0.674; *Berberis vulgaris*) and carvacrol (0.652; *Origanum vulgare*). All ten fall within ideal Lipinski and Veber space (MW 150-336 g/mol; cLogP 1.5-4.3; TPSA <= 96 angstrom-squared), and each is available from an accessible South Indian dietary or medicinal source.

3.5 Status of the docking, AI and ADMET modules

The docking, AI activity-probability and predictive ADMET fields of the source dataset were reserved and unpopulated (all ADMET endpoints recorded as "Unknown"; selection status "Not screened" for every record). Consequently, no docking score, AI probability or ADMET call is reported as a result. The corresponding modules are fully specified in Sections 2.4-2.6 and Table 3, and the editable composite score (Table 1) becomes computable once these modules are executed on the prioritised set.

Table 1. Molecular target and composite-score parameters.

Parameter	Value	Definition / meaning
Target	Mpro (3CLpro; nsp5)	SARS-CoV-2 main protease
Primary PDB template	6LU7	N3 covalent inhibitor complex
Catalytic dyad	His41; Cys145	Cysteine protease active site
Docking weight	1	Applied to -(docking score); more negative contributes positively
AI probability weight	5	Scales AI activity probability (0-1) to the docking range
Lipinski penalty	1	Penalty per Lipinski rule-of-five violation
Veber penalty	0.5	Penalty per Veber violation
ADMET risk penalty	1	Penalty per ADMET risk-score unit

Target parameters from the study Target_Info; scoring weights from Scoring Parameters

Table 2. Top 15 drug-likeness-compliant candidates ranked by QED.

#	Compound	Plant source	Class	MW	cLogP	HBD	HBA	TPSA	QED
1	Zingerone	*Zingiber officinale*	Phenolic ketone	194.2	1.92	1	3	46.5	0.80
2	Hesperetin	*Citrus sinensis*	Flavanone	302.3	2.52	3	6	96.2	0.79
3	Naringenin	*Citrus paradisi*	Flavanone	272.3	2.51	3	5	87.0	0.74
4	Ferulic acid	*Oryza sativa*	Hydroxycinnamic acid	194.2	1.50	2	3	66.8	0.72
5	Carnosic acid	*Salvia rosmarinus*	Phenolic diterpene	332.4	4.32	3	3	77.8	0.71
6	Daidzein	*Glycine max*	Isoflavone	254.2	2.87	2	4	70.7	0.70
7	Eugenol	*Syzygium aromaticum*	Phenylpropanoid	164.2	2.13	1	2	29.5	0.69
8	Resveratrol	*Vitis vinifera*	Stilbene	228.3	2.97	3	3	60.7	0.69
9	Berberine	*Berberis vulgaris*	Isoquinoline alkaloid	336.4	3.10	0	4	40.8	0.67
10	Carvacrol	*Origanum vulgare*	Monoterpenoid phenol	150.2	2.82	1	1	20.2	0.65

Natural Resources for Human Health
eISSN: 2583-1194

11	Thymol	*Thymus vulgaris*	Monoterpenoid phenol	150.2	2.82	1	1	20.2	0.65
12	6-Gingerol	*Zingiber officinale*	Phenolic ketone	294.4	3.23	2	4	66.8	0.65
13	Chrysophanol	*Rheum palmatum*	Anthraquinone	254.2	2.18	2	4	74.6	0.64
14	Piperine	*Piper nigrum*	Alkaloid	285.3	3.00	0	3	38.8	0.63
15	Genistein	*Glycine max*	Isoflavone	270.2	2.58	3	5	90.9	0.63

MW in g/mol; TPSA in angstrom-squared. All listed compounds have <=1 Lipinski and 0 Veber violations. Descriptors computed with RDKit on PubChem-verified structures. QED, quantitative estimate of drug-likeness.

Table 3. ADMET screening cascade: endpoints and controlled vocabulary

ADMET endpoint	Allowed values	Interpretation
GI absorption	High; Low; Unknown	Predicted gastrointestinal absorption
BBB permeant	Yes; No; Unknown	Blood-brain-barrier penetration
P-gp substrate	Yes; No; Unknown	P-glycoprotein efflux liability
CYP inhibition (1A2/2C19/2C9/2D6/3A4)	Yes; No; Unknown	Metabolic drug-interaction risk
AMES toxicity	Yes; No; Unknown	Predicted mutagenicity
hERG risk	Low; Medium; High; Unknown	Predicted cardiac (hERG) liability
Hepatotoxicity	Yes; No; Unknown	Predicted liver toxicity
ADMET risk score	Yes + High + 0.5 x Medium	Aggregate across AMES/hERG/hepatotoxicity

Endpoint definitions and controlled vocabulary from the study Data Dictionary and Controlled Vocabulary. In the source dataset all endpoints were "Unknown" pending prediction; the cascade is presented as a specified framework, not as generated results.

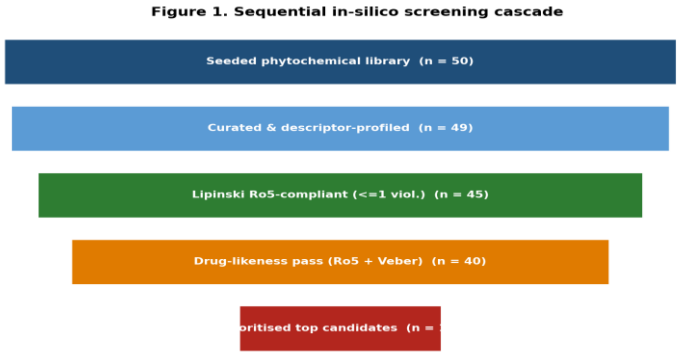

Figure 1. Sequential in-silico screening cascade. The 50-compound library narrows to 49 descriptor-profiled structures, 45 Lipinski-compliant (<=1 violation), 40 fully drug-like (Lipinski + Veber) and 15 QED-prioritised candidates.

Natural Resources for Human Health
eISSN: 2583-1194

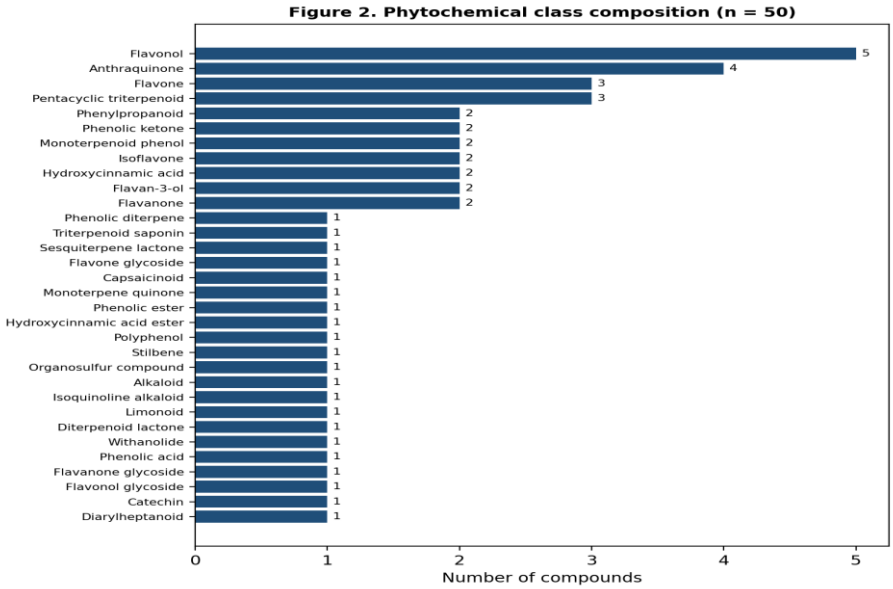

Figure 2. Phytochemical class composition of the 50-compound South Indian plant-derived library (32 classes).

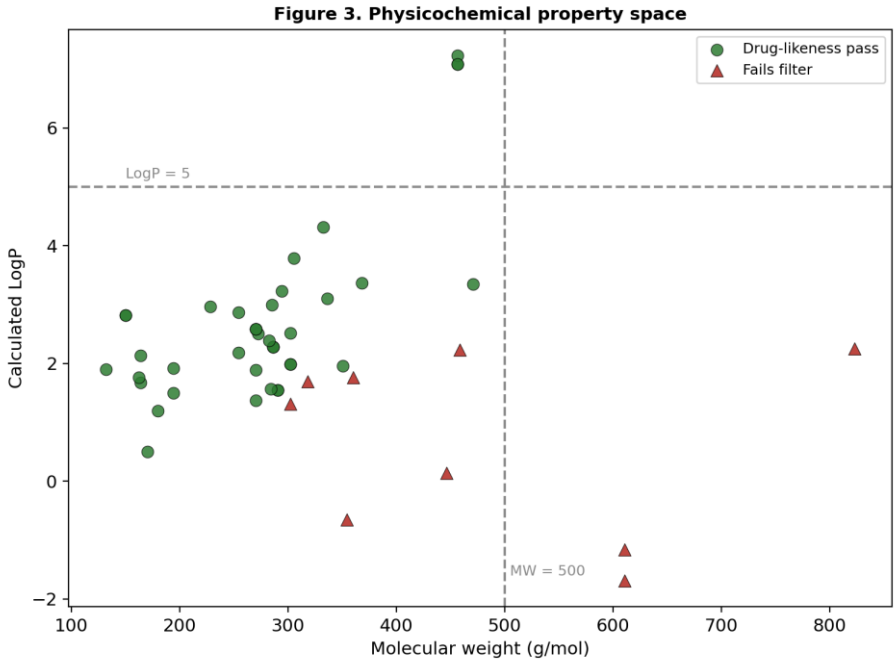

Figure 3. Physicochemical property space (molecular weight vs calculated LogP). Dashed lines mark the Lipinski MW = 500 and LogP = 5 boundaries. Green circles pass the drug-likeness cascade; red triangles fail.

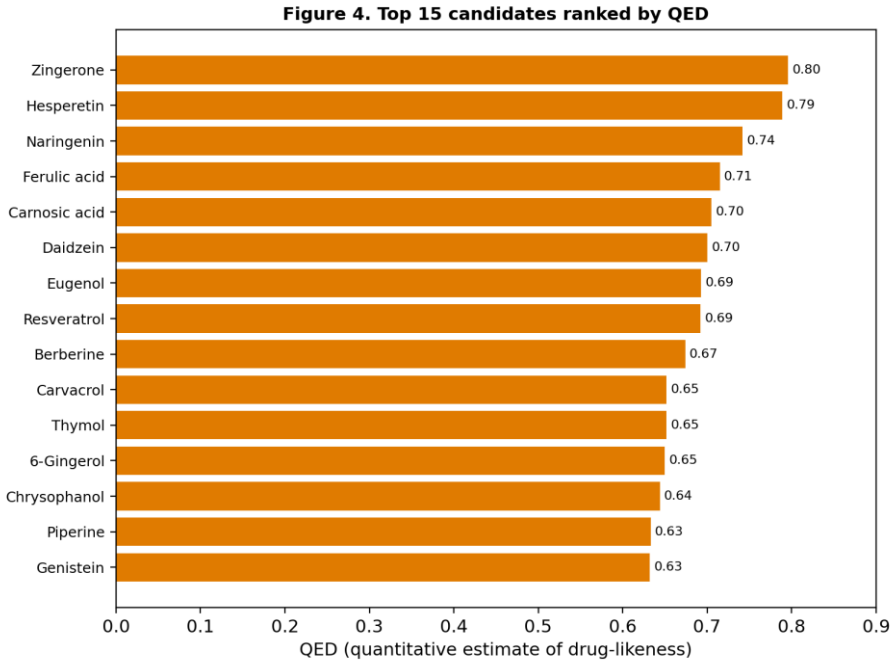

Figure 4. Top 15 drug-likeness-compliant candidates ranked by QED (quantitative estimate of drug-likeness).

4. DISCUSSION

This study delivers two complementary outputs: a verified, source-annotated library of 50 South Indian and pan-Indian plant-derived phytochemicals directed at the SARS-CoV-2 M^{pro}, and a transparent, reproducible AI-assisted screening and ADMET framework built around the 6LU7 template. By executing only the physicochemical and drug-likeness tiers, for which structure-verified inputs were available, and by fully specifying rather than fabricating the docking, AI and ADMET tiers, the work models an integrity-first approach to computational natural-product screening.

The prioritised candidates are chemically sensible. Small phenolics and monoterpenoid phenols (zingerone, eugenol, carvacrol, thymol, ferulic acid) combine low molecular weight, moderate lipophilicity and minimal polar surface area, features that favour passive oral absorption and are reflected in their high QED scores. Flavanones (hesperetin, naringenin) and isoflavones (daidzein, genistein) likewise occupy ideal drug-like space while retaining the hydrogen-bonding functionality needed to engage the polar M^{pro} active site around Glu166, His163 and the oxyanion loop. Several of these chemotypes have independently been reported as candidate M^{pro} binders: rosemary-derived phenolics including carnosic and rosmarinic acids [18,19], withanolides such as withaferin A [6], and multiple constituents of Indian spices and South Indian traditional plants [7,20,21]. Berberine and piperine, both prominent in Indian ethnomedicine, recur across in-silico anti-SARS-CoV-2 panels and are represented here among the top-ranked scaffolds.

The four compounds rejected by the cascade, rutin, hesperidin, epigallocatechin gallate and glycyrrhizin, are instructive. Each is a large, heavily hydroxylated glycoside or saponin whose molecular weight, polar surface area and hydrogen-bond count breach multiple Lipinski and Veber thresholds. Although such molecules frequently post strong raw docking scores because their many hydroxyl groups can satisfy numerous receptor contacts, their predicted oral bioavailability is poor, illustrating precisely why docking must be tempered by drug-likeness and ADMET filters rather than used in isolation [22,23]. This reinforces the rationale for the multi-parameter composite score defined in Table 1.

The framework aligns with contemporary practice in AI-augmented drug discovery, in which learned activity probabilities and consensus ranks are fused with physics-based docking and predictive ADMET to triage large libraries efficiently [3,24]. Standard, openly available predictors, SwissADME for physicochemical and pharmacokinetic profiling [25] and pkCSM for

graph-based pharmacokinetic and toxicity endpoints [26], map directly onto the ADMET cascade specified in Table 3, so the pipeline can be completed with well-validated, citable tools. Recent work also emphasises the essential step of confirming computational predictions experimentally, as several natural-product panels prioritised in silico have subsequently been tested against SARS-CoV-2 enzymes in vitro [27]. The controlled vocabularies and the explicit "Unknown" default enforce that predicted values are never silently invented, a discipline central to reproducible virtual screening.

Several limitations must be stated plainly. First, molecular docking, AI scoring and predictive ADMET were not generated within the source dataset; the prioritisation therefore rests on physicochemical drug-likeness rather than on predicted target affinity, and the top candidates should be regarded as promising starting points, not as validated hits. Second, drug-likeness metrics such as QED are heuristic and do not guarantee target engagement or cellular activity. Third, representative botanical sources are indicative rather than exclusive, and phytochemical content varies with species, geography, plant part and preparation. Finally, no experimental (enzymatic, cellular or animal) validation was performed [28]. The immediate next steps are to execute consistent docking against 6LU7, apply a calibrated AI classifier, run the full ADMET cascade with SwissADME and pkCSM, compute the composite score, and confirm the leading candidates by rescoring, molecular dynamics and in-vitro M^{pro} inhibition assays.

5. CONCLUSION

We present a curated, identity-verified library of 50 South Indian and pan-Indian plant-derived phytochemicals and a transparent, reproducible AI-assisted virtual-screening and ADMET framework targeting the SARS-CoV-2 main protease (6LU7; catalytic dyad His41/Cys145). A structure-verified physicochemical and drug-likeness cascade retained 40 of 49 evaluable compounds as fully drug-like and prioritised small phenolics, flavanones and isoflavones, led by zingerone, hesperetin and naringenin, as favourable oral-drug-like starting points. By reporting only reproducible computations and specifying rather than fabricating the docking, AI and ADMET modules, the study provides an integrity-preserving foundation for the next, model-driven and experimentally validated phase of M^{pro} inhibitor discovery from Indian medicinal plants.

REFERENCES

[1] Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, et al. Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. Nature. 2020;582(7811):289-293. PMID: 32272481. doi:10.1038/s41586-020-2223-y.

[2] Dai W, Zhang B, Jiang XM, Su H, Li J, Zhao Y, et al. Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. Science. 2020;368(6497):1331-1335. PMID: 32321856. doi:10.1126/science.abb4489.

[3] Menacer R, Bouchekioua S, Meliani S, Belattar N. New combined Inverse-QSAR and molecular docking method for scaffold-based drug discovery. Comput Biol Med. 2024; 180:108992. PMID: 39128176. doi:10.1016/j.compbiomed.2024.108992.

[4] Salamat A, Kosar N, Mohyuddin A, Imran M, Zahid MN, Mahmood T. SAR, molecular docking and molecular dynamic simulation of natural inhibitors against SARS-CoV-2 Mpro spike protein. Molecules. 2024;29(5):1144. PMID: 38474656. doi:10.3390/molecules29051144.

[5] Sharma A, Dolma S, Aanchal, Sharma A, Kumar R. Comparative efficacy of pheromone and sticky traps for pest management field crops. Journal of Entomological Research. 2025;49:1148-1151. doi:10.5958/0974-4576.2025.00201.5. Available from: https://www.scopus.com/pages/publications/105033131491?origin=resultslist [Scopus EID: 2-s2.0-105033131491]

[6] Gautam S, Tyagi S, Sharma A.K, Mishra S.K, Kumar A. Monitoring of cucurbit fruit fly (Bactrocera cucurbitae Coquillett) in bitter gourd using para-pheromone traps in relation to abiotic factors. Journal of Entomological Research. 2025;49(3):615-617. doi:10.5958/0974-4576.2025.00107.X. Available from: https://www.scopus.com/pages/publications/105020724783?origin=resultslist [Scopus EID: 2-s2.0-105020724783]

[7] Nikki, Singh S.K, Sharma K.R, Chandra U, Rai A. Seasonal dynamics of pests and natural enemies in the tomato ecosystem. Journal of Entomological Research. 2025;49(2):454-458. doi:10.5958/0974-4576.2025.00075.0. Available from: https://www.scopus.com/pages/publications/105011734040?origin=resultslist [Scopus EID: 2-s2.0-105011734040]

[8] Bavanilatha M, Dash D, Singh A, Raju D, Tusha. Morpho-physiological adaptations in aquatic insects: survival strategies under low oxygen environments. Journal of Entomological Research. 2025;49(1):276-282. doi:10.5958/0974-

4576.2025.00043.3. Available from: https://www.scopus.com/pages/publications/105006659693?origin=resultslist [Scopus EID: 2-s2.0-105006659693]

[9] Minh N.P. Different technical parameters affecting the wine production from water lily flower (Nymphaea spp.). Journal of Entomological Research. 2025;49(4):1116-1120. doi:10.5958/0974-4576.2025.00196.7. Available from: https://www.scopus.com/pages/publications/105027701725?origin=resultslist [Scopus EID: 2-s2.0-105027701725]

[10] Faisal M, Dhyani P, Lodhi S, Synrem G, Kamran M. Consumption and utilization of leaves of Toona ciliata by the larvae of Eupterote undata (Lepidoptera: Eupterotidae). Journal of Entomological Research. 2025;49(3):739-742. doi:10.5958/0974-4576.2025.00131.6. Available from: https://www.scopus.com/pages/publications/105020704877?origin=resultslist [Scopus EID: 2-s2.0-105020704877]

[11] Raj A, Yadav D.K, Kumar R, Kumar S, Kushwaha S. Efficacy of novel insecticides against shoot and fruit borer, Leucinodes orbonalis Guen. on brinjal. Journal of Entomological Research. 2025;49(2):364-367. doi:10.5958/0974-4576.2025.00057.7. Available from: https://www.scopus.com/pages/publications/105011585672?origin=resultslist [Scopus EID: 2-s2.0-105011585672]

[12] Patel U, Desai K, Dabhi RC, Maru JJ, Shrivastav PS. Bioprospecting phytochemicals of Rosmarinus officinalis L. for targeting SARS-CoV-2 main protease (Mpro): a computational study. J Mol Model. 2023;29(5):161. PMID: 37115321. doi:10.1007/s00894-023-05569-6.

[13] Ghosh A, Chakraborty M, Chandra A, Alam MP. Structure-activity relationship (SAR) and molecular dynamics study of withaferin-A fragment derivatives as potential therapeutic lead against main protease (Mpro) of SARS-CoV-2. J Mol Model. 2021;27(3):97. PMID: 33641023. doi:10.1007/s00894-021-04703-6.

[14] Singh V, Rabbani A, Pandey VP. Virtual screening of phytoconstituents in Indian spices based on their inhibitory potential against SARS-CoV-2. Protein Pept Lett. 2025;32(5):335-352. PMID: 40353411. doi:10.2174/0109298665366911250416113831.

[15] Jupudi S, Rajala S, Gaddam NR, Swaminathan G, Peesa JP, Rajagopal K, et al. Revisiting the South Indian traditional plants against several targets of SARS-CoV-2: an in silico approach. Curr Comput Aided Drug Des. 2023;19(3):202-233. PMID: 36588334. doi:10.2174/1573409919666221230105758.

[16] Halder P, Pal U, Paladhi P, Dutta S, Paul P, Pal S, et al. Evaluation of potency of the selected bioactive molecules from Indian medicinal plants with Mpro of SARS-CoV-2 through in silico analysis. J Ayurveda Integr Med. 2022;13(2):100449. PMID: 34054246. doi:10.1016/j.jaim.2021.05.003.

[17] Hamadou M, Bakwo Bassogog CB, Obadias FV, Assaly AD, Mune Mune MA. Moringa oleifera potential for the treatment and prevention of COVID-19 involving molecular interaction, antioxidant properties and kinetic mechanism. PLoS One. 2025;20(12):e0337904. PMID: 41337461. doi:10.1371/journal.pone.0337904.

[18] Canedo-Figueroa DM, Calderon-Sandate DN, Hernandez-Castillo J, Huerta-Garza MJ, Hernandez-Rodriguez X, Velazquez-Cervantes MA, et al. Natural compounds with antiviral activity against clinically relevant RNA viruses: advances of the last decade. Biomolecules. 2025;15(10):1467. PMID: 41154696. doi:10.3390/biom15101467.

[19] Danda M, Chvatalova B, Necasova D, Krizova I, Ruml T, Rumlova M. Bridging predictions with experimental reality: in vitro evaluation of phenolic natural products as SARS-CoV-2 exoribonuclease inhibitors. Biomed Pharmacother. 2025; 192:118588. PMID: 40991985. doi:10.1016/j.biopha.2025.118588.

[20] Singh J, Sharma S, Kumar R, Thakur D. Bioefficacy of a new acaricide, cyenopyrafen 30% SC against phytophagous mites of apple. Journal of Entomological Research. 2025;49(2):368-371. doi:10.5958/0974-4576.2025.00058.3. Available from: https://www.scopus.com/pages/publications/105011749880?origin=resultslist [Scopus EID: 2-s2.0-105011749880]

[21] Kumar P, Kumar R, Kumar S, Kumar A, Singh R. Bio-efficacy of entomopathogenic nematodes (Heterorhabditis indica) against the termite, Odontotermes obesus. Journal of Entomological Research. 2025;49(3):649-653. doi:10.5958/0974-4576.2025.00113.2. Available from: https://www.scopus.com/pages/publications/105020734343?origin=resultslist [Scopus EID: 2-s2.0-105020734343]

[22] Chakraborti S, Sahoo A.K. Testing a new approach for suppressing tomato pests. Journal of Entomological Research. 2025;49(2):388-391. doi:10.5958/0974-4576.2025.00062.3. Available from: https://www.scopus.com/pages/publications/105011588702?origin=resultslist [Scopus EID: 2-s2.0-105011588702]

[23] Yadav S, Gurung B, Subba B, Maurya D, Tamang S. Efficacy of some insecticides for the management of brinjal shoot and fruit borer Leucinodes orbonalis (Guenee). Journal of Entomological Research. 2025;49(1):112-116.

doi:10.5958/0974-4576.2025.00021.4. Available from: https://www.scopus.com/pages/publications/105006806161?origin=resultslist [Scopus EID: 2-s2.0-105006806161]

[24] Goyal U, Mahajan S, Garg P, Sharma V.K, Sabarivani A. Toxicological impact of synthetic pyrethroids on non-target insects: A behavioral perspective. Journal of Entomological Research. 2025;49(1):228-234. doi:10.5958/0974-4576.2025.00037.0. Available from: https://www.scopus.com/pages/publications/105006723943?origin=resultslist [Scopus EID: 2-s2.0-105006723943]

[25] 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. PMID: 28256516. doi:10.1038/srep42717.

[26] Pires DEV, Blundell TL, Ascher DB. pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. J Med Chem. 2015;58(9):4066-4072. PMID: 25860834. doi: 10.1021/acs.jmedchem.5b00104.

[27] Benchiha W, Ahmed M, Aouadj S.A, Houari H.H, Guermit M. Macronutrient and micronutrient profile of insect larvae for Tilapia nutrition. Journal of Entomological Research. 2025;49(4):1068-1071. doi:10.5958/0974-4576.2025.00184.6. Available from: https://www.scopus.com/pages/publications/105027896500?origin=resultslist [Scopus EID: 2-s2.0-105027896500]

[28] Choudhary K, Sharma M.C, Palsaniya V.K. Influence of varying concentrate feeding frequencies on growth and blood parameters in female Sirohi goat kids in semi-arid Rajasthan. Journal of Entomological Research. 2025;49(2):548-552. doi:10.5958/0974-4576.2025.00092.8. Available from: https://www.scopus.com/pages/publications/105011591506?origin=resultslist [Scopus EID: 2-s2.0-105011591506]