



An investigation of the drug discovery potential of phytochemicals from *Monodora* species through *in silico* ADMET profiling

Loveline N. Amen^{a,b}, Jude Y. Betow^{b,c} , Mathieu J.M. Tjegbe^b,
Gotoum Nadjinangar^d, Jean Duplex Wansi^a, Fidele Ntie-Kang^{b,c,e,*} 

^a Department of Chemistry, University of Douala, Douala, Cameroon

^b Center for Drug Discovery, University of Buea, Buea, Cameroon

^c Department of Chemistry, University of Buea, Buea, Cameroon

^d Département de Physique, Université de N'Djamena, N'Djamena, Chad

^e Institute of Pharmacy, Martin-Luther University Halle-Wittenberg, Halle (Saale), Germany

ARTICLE INFO

Editor: DR B Gyampoh

Keywords:

DMPK

Drug discovery

Modeling

Natural products

ABSTRACT

Phytochemicals are natural products of plant origin and hold a promise for drug discovery because of the traditional uses of medicinal plants. The genus *Monodora* is known for diverse traditional uses in ethnomedicine, and the present phytochemicals show diversity in compound classes. The compounds have been identified from the plants' different parts (e.g., roots, leaves, and stem bark), but the drug discovery potential of their constituent phytochemicals has not been investigated. The investigation of the DMPK profiles of small molecules early enough in drug discovery helps reduce the attrition rates. In this study, we have identified 132 phytochemicals with diverse biological properties from seven plant species from this genus, and these belong to several compound classes, including alkaloids, terpenoids, coumarins, etc. >100 molecular properties, often related to drug metabolism and pharmacokinetics (DMPK) were then computed for this dataset using three tools (SwissADME, QikProp and pkCSM), and compared with a dataset of >600 natural products reported with activities against the SARS-CoV-2 virus or its viral target proteins. It was demonstrated that the *Monodora* compounds outperformed the SARS-CoV-2 compounds for several key parameters, including those related to drug absorption, distribution, and toxicity. This small, diverse *Monodora* dataset holds promises for drug discovery and has been made available for working groups interested in molecular docking techniques and putative target determination. This study demonstrates that phytochemicals from *Monodora* not only show drug-likeness and favorable DMPK profiles but also outperform benchmark datasets of natural products already studied for antiviral activity.

Introduction

Natural products (NPs) have emerged as an outstanding source of promising new lead compounds for drug discovery because they have been tailored by biosynthetic processes that make them unique when compared with synthetic drugs [1]. Many drugs currently in use as effective therapeutic agents have been derived from naturally occurring sources. In the last couple of decades, drug discovery

* Corresponding author.

E-mail address: fidele.ntie-kang@ubuea.cm (F. Ntie-Kang).

based on natural products has received limited coverage from pharmaceutical companies, mainly owing to their inherent difficulties in isolation and the limited quantities directly available from extracts [2]. However, recent progress in isolation techniques, including metabolomics-based methods coupled with advanced mass spectrometry, has rekindled scientific interest in the discovery of new drugs from natural sources. Apart from the fact that many drugs in the market are either NPs or their derivatives [1], those from plants often possess various therapeutic activities and are constantly being examined to develop new medicines because of the long use of these plants in traditional medicines [3].

Monodora is a plant genus that belongs to the Annonaceae. This genus is made up of sixteen identified species, often small trees, native to the tropical forests of East and West Africa [4,5]. The species *Monodora carolinae* grows in Coastal bushes on sandy soils that are deep and well-drained and are threatened due to the decline in its population, e.g. in Tanzania [6]. *Monodora crispate* and *M. brevipes* are referred to as possibly endangered species in Côte d'Ivoire (or Ivory Coast) [7,8]. *Monodora myristica* spices from seeds are used in food preparation as flavour boosters, preservatives, and for other purposes. They are distinguished from herbs by their more pronounced aroma and taste. The wood of *M. myristica* is suitable for use in home furnishings, carpentry, and joinery [8]. Presently, a trend to use the oil obtained from *M. myristica* for flavouring popcorn has supported the use of the spice as a flavoring with good acceptance and no side effects [9].

This study directly addresses one of Africa's pressing challenges: the gap between rich ethnomedicinal heritage and modern drug discovery pipelines. Despite Africa's immense biodiversity and reliance on medicinal plants for primary healthcare, systematic evaluation of their drug-likeness and pharmacokinetic potential is still lacking. By applying computational tools to phytochemicals from the genus *Monodora*, we provide evidence-based insights that can guide drug discovery efforts using African resources. In doing so, this work contributes to the African Union's Agenda 2063 aspirations for health innovation and scientific self-reliance, while also advancing the United Nations Sustainable Development Goals, particularly SDG 3 (Good Health and Well-Being), SDG 9 (Industry, Innovation and Infrastructure), and SDG 15 (Life on Land). The curated dataset generated here thus represents not only a research advance but also a strategic contribution toward leveraging Africa's biodiversity for global health solutions.

Globally, plants belonging to the genus *Monodora* have been reported to contain alkaloids, the main constituents of which are prenylated indoles, which are considered to be chemotaxonomic markers of the genus [10]. Studies have proven that the genus *Monodora* is a major source of bioactive compounds from sesquiterpenes, saponins, flavonoids, steroids, tannins, monoterpenes, diterpenes, indoles, and essential oils [11], which are present in different parts of the plant species [12,13].

Many plants, including *Monodora* species, are being investigated for their potential as sources of antimicrobial drug leads [14,15]. We hypothesize that plants from the genus *Monodora* contain NPs that could hold the potential for the discovery of next-generation antimicrobial (antiviral, antibacterial, and antifungal) agents. Owing to the need for pandemic preparedness, our group has been engaged in the investigation of plants whose extracts and constituent metabolites are known for antimicrobial activities, and which could eventually become lead compounds for the discovery of next-generation antimicrobials. An investigation of the genus *Monodora*, which happens to fall under this category, is warranted at this time for its potential in the treatment of infections resulting from the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), causing the coronavirus disease (COVID-19). The SARS-CoV-2 dataset was used for comparison for several reasons, amongst which it is a recently published (updated) dataset, containing antimicrobial agents with proven biological properties, and it is composed of NPs, demonstrating that they are likely to occupy similar chemical space and, hence, comparable predicted DMPK profiles as the NPs from the *Monodora* genus. The objective of this study was to identify all plant species from this genus, as well as their identified and well-characterized phytochemicals, followed by an investigation of their potential to be further investigated as lead compounds for the development of potential drug candidates. An *in silico* analysis was performed to compare the drug metabolism and pharmacokinetics (DMPK) profiles of these *Monodora* compounds in comparison with a recently published manually curated dataset of phytochemicals with reported activities against the SARS-CoV-2 virus and/or any of its protein targets. This research advances the field by moving beyond simple phytochemical cataloging, providing a detailed *in silico* assessment of drug-likeness and pharmacokinetic properties of *Monodora* compounds. While earlier studies on African medicinal plants have mostly described ethnomedicinal uses or reported isolated compounds, few have applied systematic computational approaches to evaluate their drug discovery potential. By using three independent computational platforms, we ensured the reliability of the predictions and uncovered favorable absorption, distribution, and toxicity profiles for *Monodora* phytochemicals compared to a benchmark dataset of SARS-CoV-2 active natural products. Furthermore, by curating and making available a dataset of 132 compounds from seven *Monodora* species, we provide a valuable resource for the community to pursue molecular docking, target prediction, and structure–activity relationship studies, thus bridging the gap between ethnomedicine and modern drug discovery.

Since the drug discovery process, which is often very time-consuming and requires heavy investments, one approach to avoid high attrition of failure to make it to the market is by incorporating the investigation of DMPK properties, often summarized as absorption, distribution, metabolism, elimination, and toxicity (ADMET) property predictions early enough in the drug discovery pipeline. If such calculations are carried out early enough, it would lead to discarding those molecules that do not hold a promise towards drug development, i.e., molecules that either do not fit the category of known drugs or have predicted toxicity profiles.

Materials and methods

Data collection of *Monodora* species from natural sources

Manually screened libraries of compounds were generated from the literature reported on the genus *Monodora*. The search was obtained from relevant sources like the African Natural Products Database [16,17] (ANPDB), freely available on (<http://african-compounds.org/anpdb/>) and Google Scholar (<https://scholar.google.com>) with several keywords related to the *Monodora*, e.g.

search terms included "Monodora", "Monodora genus", "Monodora species", "Monodora compounds", "Monodora phytochemicals", etc. The articles found were checked, and data found on source species, compounds, class, and bibliographic references of the articles were selected and compiled on a spreadsheet. All phytochemicals reported from the genus were included in the study, irrespective of their reported biological or devoid of biological activities. The chemical structures available in the PubChem database were downloaded. Then, the PubChem molecules were accessed individually with their proper PubChem compound identifier as 2D structures, followed by the compound names. The IUPAC names of the compounds in these articles were checked with the PubChem and ChemSpider databases, an approach outlined earlier by [18,19]. In addition, the mol files for molecules not found in PubChem were hand-drawn using ChemDraw (version 15.0) based on the 2D structures as given in the literature sources.

Data preparation

During this process, the source of the species and the information on the compound were carefully and handily obtained and checked on two occasions. Further, the process of obtaining SMILES strings for a set of molecules was streamlined. First, duplicates were removed and, for molecules identified in PubChem, the SMILES strings were generated. For those not found in PubChem, SMILES were extracted directly from their existing mol format structures. This automated workflow was implemented using a custom pipeline available on our GitHub repository (https://github.com/UB-CeDD/Merge_smiles). The pipeline includes scripts that efficiently extract SMILES information from both SDF and MOL files, consolidating everything into a single spreadsheet for further analysis. Data was compiled from relevant published literature and analyzed on a spreadsheet. This resulted in a collection of 132 compounds, which were further analyzed.

Generation of ADMET properties

For ADMET property calculations using QikProp (Schrödinger Release 2024-4: QikProp, Schrödinger, LLC, New York, NY, 2024), the SMILES were uploaded on the Maestro interface and prepared using LigPrep (Schrödinger Release 2024-4). The same SMILES strings generated from the prepared dataset were uploaded to two servers to calculate ADMET properties, i.e. DeepPK (an update of pkCSM: <https://biosig.lab.uq.edu.au/deepk/>) and SwissADME (<http://www.swissadme.ch>; [20]). The most important computed properties included molecular weight (MolWt), the number of H-bond donors (HBD) and acceptors (HBA), the number of non-hindered rotatable bonds (#RotB), The number of violations of Lipinski's "rule of five" (RuleOfFive), Jorgensen's "rule of three" violations (RuleOfThree), Predicted solubility (log S), HERG K⁺ channel blocking activity, blood/brain partition coefficient, and skin permeability, among others. The distribution of selected properties that relate to drug absorption, distribution, metabolism, elimination, and toxicity (ADMET) was analyzed and used in the evaluation of the drug metabolism and pharmacokinetic (DMPK) profiles of the molecules.

Scaffold analysis

Before conducting scaffold analysis, the Retrosynthetic Combinatorial Analysis Procedure (RECAP) [21] was used within the molecular operating environment (MOE, version 2016.08, Chemical Computing Group, 2016) software. RECAP disassembles each molecule by breaking bonds predicted to be producible during synthesis from building blocks using common reactions. Each resulting fragment was uniquely identified by an extended SMILES string and a name reflecting its chemical context around the broken bond, as outlined by Weininger [22]. This process was applied to both the *Monodora* NPs dataset and the manually curated anti-SARS-CoV-2 NPs dataset. Following this, the frequency of these fragments was analyzed, identifying the most common chemical scaffolds. Only scaffolds meeting fragment-likeness criteria [23] and containing at least 10 atoms were retained for further analysis.

Principal component analysis

To evaluate and appreciate the chemical space occupied by the compounds in our dataset, principal component analysis (PCA) was conducted using the Jupyter Notebook. A reference set of manually curated NPs with documented anti-SARS-CoV-2 activity was included for comparative purposes [24]. Using descriptors from the Molecular Operating Environment (MOE, version 2016.08, 2016), a 2D PCA plot was generated, and the most influential molecular features in the formation of the first two principal components were analyzed, providing insights into the underlying chemical space relationships. A comprehensive set of 37 well-established physicochemical descriptors were computed, encompassing a wide range of molecular properties. The descriptors cover a broad range of properties including from simple property descriptors related to drug-likeness, e.g. molecular weight (MolWt), lipophilicity (log P), as well as size-based descriptors, e.g. size (reflected by the number of atoms, a_atom), descriptors related to the electronic properties and reactivities of chemicals (e.g. AM1_LUMO, AM1_HOMO), topological features (rotatable bonds, chiral centers), drug-likeness scores (Lipinski violations, Oprea Lead), and various other physicochemical and electronic descriptors such as solubility (logS), formal charge (FCharge), polar surface area (PSA), and dipole moment.

Results and discussion

Distribution of *Monodora* compounds in the literature

Of the 16 *Monodora* species with published phytochemical studies in the literature, only 7 included a description of identified or isolated compounds, i.e., *M. angolensis*, *M. brevipes*, *M. carolinae*, *M. crispata*, *M. minor*, *M. myristica*, and *M. tenuifolia*. This formed the basis of the compound data being discussed in this work. The compounds isolated from these species encompassed a diverse range of classes, including alkaloids (predominantly indoles), coumarins, fatty acids, flavonoids, phenolic acids, phenylpropanoids, terpenoids (e.g., monoterpenes, sesquiterpenes, and diterpenes), and glycosides.

Fig. 1. shows a pie chart that represents the distribution of the compounds into their respective species of origin and into their respective compound classes.

Concerning the species of origin (Fig. 1A), the majority of the compounds are from *M. myristica* (~47 %), *M. tenuifolia* (~29 %) and *M. crispata* (~15 %). The rest of the species have very few compounds described in the literature (often ≤ 2 %). Meanwhile, in Fig. 1B, the majority of *Monodora* compounds are of the terpenoid class (~58 %), followed by the alkaloids (~19 %) and fatty acids (~10 %), while the rest of the classes each represent ≤ 5 %. Regarding the most abundant class (terpenoids), the most abundant subclass is made up of the sesquiterpenoids, which represent 30.6 % of the compound collection. Most of the bioactive compounds originated from *M. myristica*. In terms of the reported biological activities of the compounds, most of the compounds were reported to exhibit anti-inflammatory, anticancer, antioxidant, antihypertensive, cholesterol-lowering activities, antispasmodic, antiulcer, hepatoprotective, antibacterial, and antifungal activities [10], Table 1. Besides, compounds from *M. minor* were shown to be highly effective for their anti-trypanosome and antiplasmodial activities, while *M. carolinae* compounds showed predominantly antimycobacterial and cytotoxic properties [11]. Several literature reports have shown that compounds from these species are biologically very active.

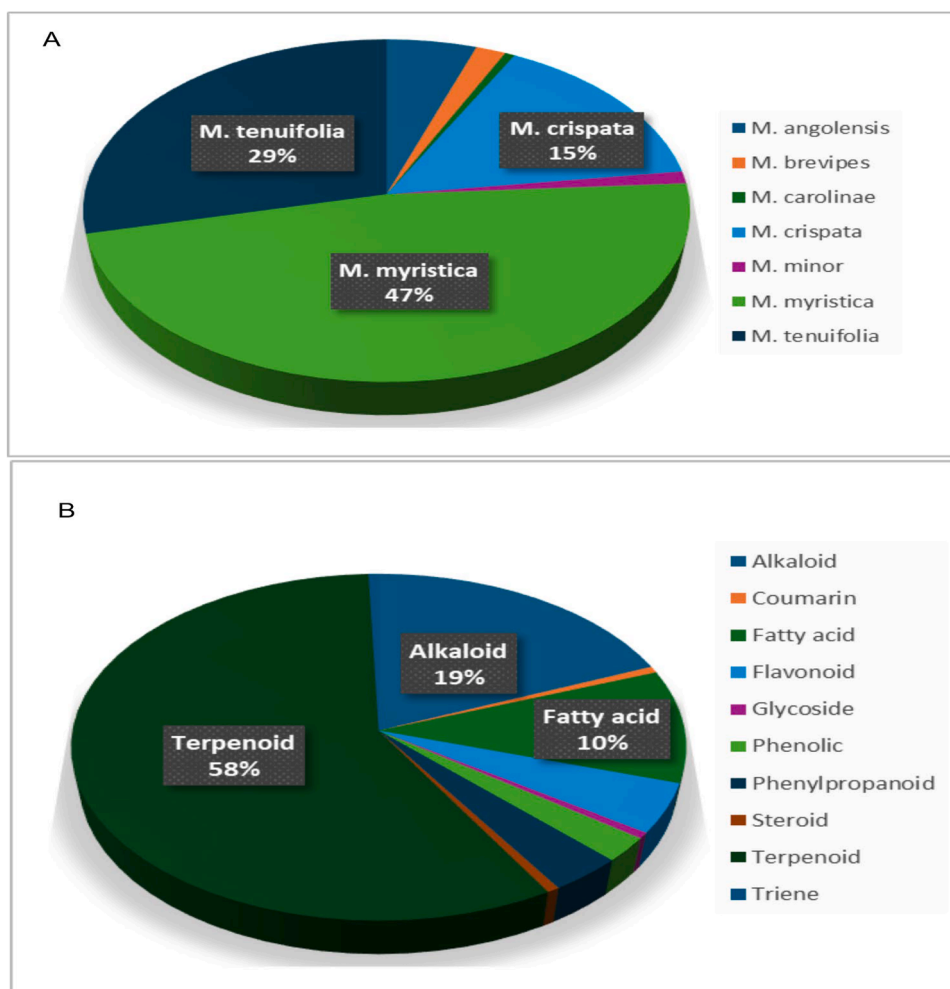
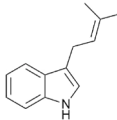
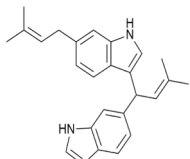
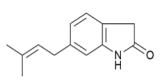
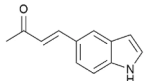
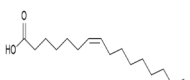
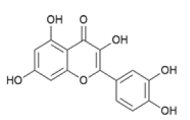
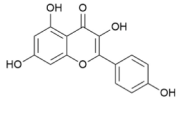
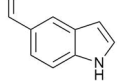


Fig. 1. Distribution of *Monodora* compounds into; (A) species of origin and (B) compound classes (classes which contribute ≤ 2 % have been left out for clarity of the figure).

Table 1

The most active compounds, their structures, classes, the source species, and their reported biological activities, showing the minimum inhibitory concentrations (MIC) and 50 % inhibitory concentrations (IC₅₀) values.

Origin of species	Activity	Compound	Class	Structure	Reference
<i>M. tenuifolia</i>	Antifungal; against dermatophyte, <i>Trichophyton mentagrophytes</i> and phytopathogen <i>Helminthosporium</i> sp. (MIC values of 1.56 and 0.78 µg/mL).	3-dimethylallylindole	Alkaloid		[25]
<i>M. angolensis</i>	Antimalarial; against <i>Plasmodium falciparum</i> (IC ₅₀ = 21 µg/mL).	Annonidine F	Alkaloid		[10]
	Antimalarial; against <i>P. falciparum</i> (IC ₅₀ = 21 µg/mL)	6-(3-methyl-but-2-enyl)-1,3-dihydro-indol-2-one	Alkaloid		[10]
<i>M. carolinae</i>	Antimicrobial (<i>Mycobacterium indicus pranii</i>) at an MIC value of 250 µg/mL.	(E)-4-(1H-indol-5-yl)-but-3-en-2-one	Alkaloid		[6]
	Antimicrobial (<i>Mycobacterium madagascariense</i> MIC value of 0.25 µg/mL)	(Z)-hexadec-7-oneic acid	Fatty acid		
<i>M. myristica</i>	Antioxidant; with 2, 2-diphenyl-1-picryl hydrazyl (DPPH)-radical scavenging activity (IC ₅₀ = 1.840 µg/mL)	Quercetin	Flavonoid		[26]
	Antioxidant; DPPH-radical scavenging activity with IC ₅₀ = 5.318 µg/mL	Kaempferol	Flavonoid		[26]
<i>M. minor</i>	Antimycobacterial; <i>Mycobacterium madagascariense</i> (IC ₅₀ = 1.25 µg/mL)	5-formyl-1H-indole	Alkaloid		[6]

The tested antimicrobial and antiparasitic properties of some of the compounds shown in Table 1, e.g., the indoles, as well as their promising drug-likeness and DMPK profiles, warrant further investigations in vivo, e.g., the antimalarial compound, Annonidine F, could be investigated in the mouse model against *Plasmodium berghei*.

Chemical space analysis and general properties related to drug-likeness prediction

The general parameters of "drug-likeness" used are important for a molecule to conform to Lipinski's rule of five (ro5) [27], although several FDA-approved drugs, including several kinase inhibitors, are known to violate Lipinski's ro5 [28]. This rule is a useful and popular checklist for determining whether chemicals have "drug-like" properties and are related to simple physico-chemical properties like MolWt, log P, HBD, and HBA, whose distributions have been shown in Fig. S1 (Supplementary Data), alongside those of our recently published manually curated dataset of 620 NPs with activities against the SARS-CoV-2 virus and its viral drug targets [24]. Only absolute numbers and the various ranges of values were used for comparison for each class of properties. The SARS-CoV-2 NPs (SNP) have been highlighted in green, while the *Monodora* NPs (MNP) have been highlighted in orange. For molecular weight (MolWt, Fig. S1A), both histograms showed a Gaussian distribution, with the difference being that the distribution for the *Monodora* compounds showed a peak at 201–300 Da, while the peak for SNP was at 301–400 Da. Within the Lipinski recommended range (≤ 500 Da), 97.5 % MNP versus 52.7 % of the SNP respected this criterion. HBA and HBD are descriptors that determine a molecule's ability to cross cell membranes (permeability) and its binding capacity to a drug target. A high capacity for hydrogen bonding results in a low permeability, which is a direct reflection of its low solubility. Histograms of HBA for MNP and SNP showed 98.3 % and 51.9 % compliance ($\log P < 5$), respectively (Fig. S1B), with a peak at $\log P = 1$ for MNP. The distribution of compounds in

HBD is higher for MNP than for SNP, with most compounds having 0 to 1 HBD (Fig. S1C). Besides, the majority of MNP and SNP compounds have the recommended number of HBD (< 5), i.e., 92.4 % and 48.7 %, respectively.

The number of rotatable bonds (#RotB) contributes to the flexibility of a molecule, which influences its ability to interact with biological targets, while also influencing bioavailability. The greater the number of rotatable bonds, the higher the flexibility of the compound, which has a negative impact on its permeability [29]. A good orally bioavailable drug usually has 10 or fewer rotatable bonds (#RotB), as well as following Lipinski's rule. Our results show that 87.1 % of MNP and 92.7 % of SNP fall into the recommended range. The log P value increases with lipophilicity, so the guiding range of a compound expected to attend as a lead candidate is as shown ($-2 < \log P < 5$) [30,31]. log P shows the greatest variance of compounds for the two datasets within the defined range, i.e., 87.4 % MNP fell into the defined range, compared to 72.9 % for SNP (Fig. S1E). For Lipinski's Violation, 98.4 % MNP showed no violation alongside 78.4 % SNP, indicating that MNP outperformed SNP. Both histograms showed slanted Gaussian distributions with peaks at 0 for both datasets (Fig. S1F), with most log P values lying between 0 and 1, an indication of good predicted oral bioavailability.

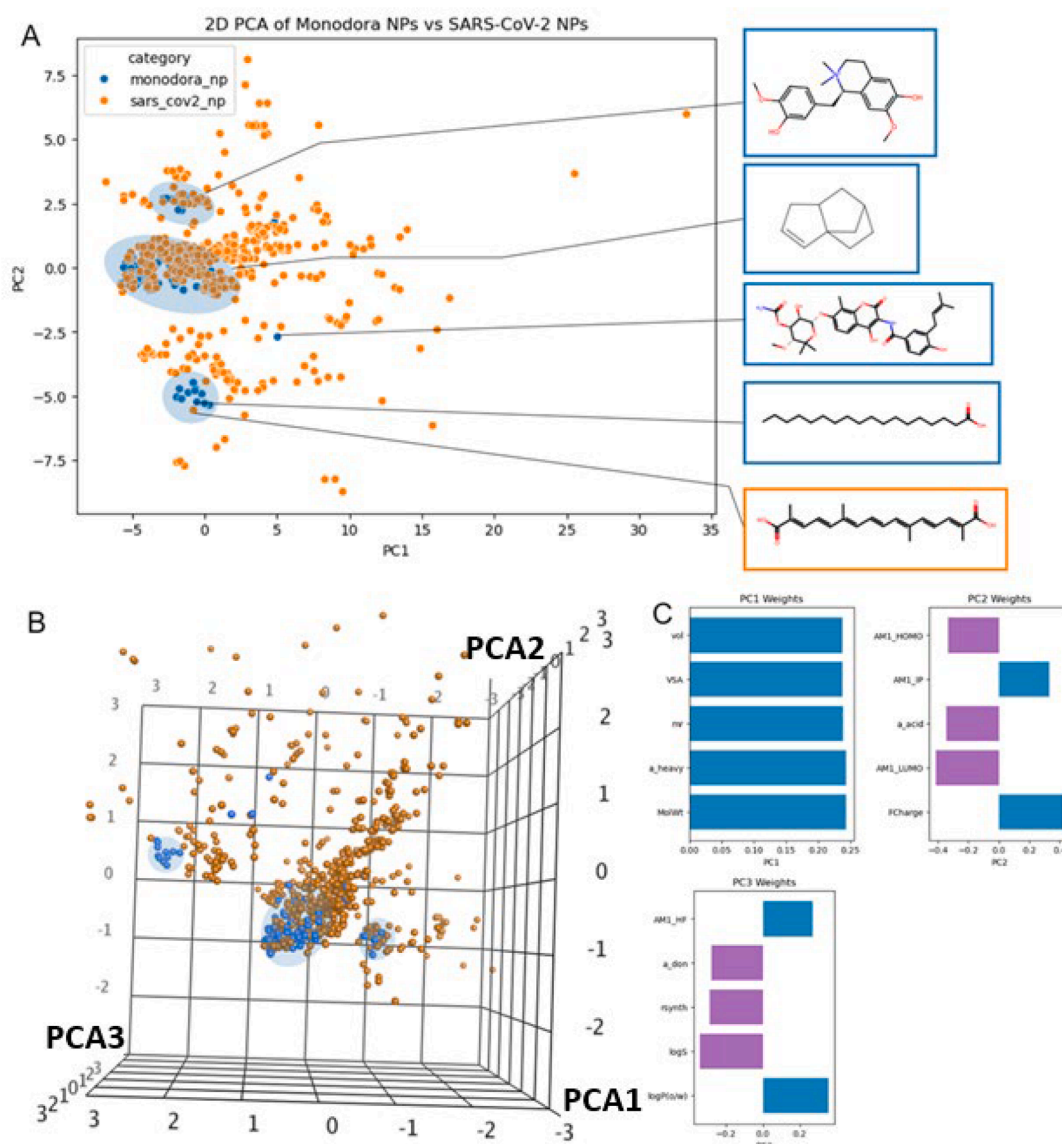


Fig. 2. (A) 2D PCA plot of the *Monodora* and SARS-CoV-2 natural products (B) 3D PCA distribution of the *Monodora* natural products (Explained variance; PC1: 46.00 %, PC2: 13.10 %, PC3: 10.59 %) (C) Top 5 descriptors contributing to the 3 principal components. The orange dots refer to SNPs, while the blue dots refer to the MNPs.

Chemical diversity of the compound scaffolds with respect to natural products having activities against SARS-CoV-2

The diversity of compounds in the MNP and SNP datasets have been summarized in the scatter PCA plots in Fig. 2. We observed that both datasets clustered around three areas, corresponding to lipids (fatty acids), sesquiterpenes and polyphenols (Fig. 2A). The chemical space of the SNP is a lot wider than the MNP and this was foreseeable as the SNP has about 4 to 5 times more compounds than MNP. The most important parameters captured by the three most important principal components (PC1, PC2, and PC3) are MolWt, formal charge, and log P, respectively.

Fig. S2 (Supplementary Data) shows the top-ten abundant scaffolds derived from the RECAP analysis of the *Monodora* compounds compared with SARS-CoV-2 compounds (*Monodora* compounds in blue, SARS-CoV-2 compounds in orange). It was observed that most of the SNP scaffolds are polyphenols, while indoles and N-containing scaffolds like indoles are more present in MNP.

Properties related to drug absorption

Absorption in drug pharmacokinetics refers to the process of how drugs go through the organs of the body to reach the systemic circulation. The absorption of drug formulation is a highly complicated procedure that can be affected by various parameters. For drug substances administered topically, the capacity of a substance to cross the outer layer of the skin is described by skin permeability (measured by log k_p). A compound is acceptable as an effective skin permeability agent when log k_p value ranges from -8.0 to -1.0 for 95 % known drugs. The distribution of compounds of MNP and SNP is as shown in Fig. 3A, with 95.4 % and 68.7 %, respectively, within the recommended range predicted to transit the corneum layer. This distribution showed a Gaussian-shaped graph with a peak value at log $k_p = -2$ unit for both datasets (Fig. 3A). For a drug to be orally administered, it must be able to dissolve in water. The data analysed for log S (water solubility) shows that most of the molecules in the two datasets, i.e. 62.1 % SNP and 70.4 % MNP, were found in the recommended range of -6.0 to 0.5 . Both distributions have the highest peak at -4 (Fig. 3B). The percentage of human oral absorption (%HOA), which is also known as oral bioavailability, is the percentage of a substance that is absorbed into the bloodstream after oral administration. The pie chart for the %HOA parameter is shown in (Fig. 3C and D). It was observed that %HOA is predicting a higher absorption with most of the molecules in the range (81–100) for MNP and SNP, that is 90 % and 52 % for MNP and SNP, respectively. Therefore, the absorption curves and pie charts in Fig. 3 conclude that most of the compounds in the defined range of 95 % of drugs were from MNP, which show a higher absorption prediction than those of SNP.

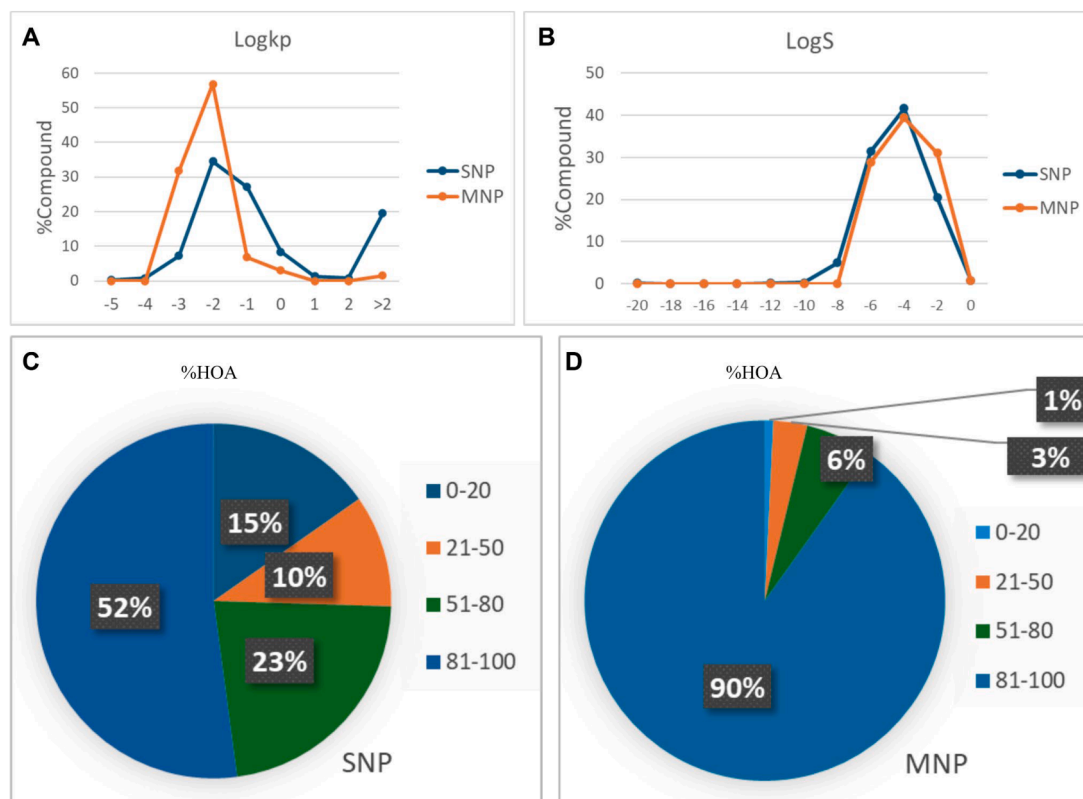


Fig. 3. Distribution curves and pie charts for (A) log k_p , (B) LogS and (C and D) % Human oral absorption of MNP and SNP compounds.

Properties related to drug distribution

The movement of the drug from the systemic circulation to the tissues (sites of action) in sufficient concentration for generating therapeutic action is known as drug distribution. To be effective, a drug must be distributed to the organs or tissues after absorption. Distribution is influenced by various parameters, including the blood-brain barrier (BBB). BBB permeability is among the important parameters to be taken into consideration when assessing a drug's ability to be distributed in the body. The BBB is a collection of endothelial cells that protects the brain by enabling the uptake of only those chemical substances required for brain metabolism. It regulates the permeation of active drugs through the cerebral vascular endothelium into the CNS. So chemical agents that are highly lipophilic can easily permeate the BBB by passive diffusion, unlike less lipophilic and polar agents. It was observed that 89 % of MNP and 51 % of SNP were predicted to penetrate the BBB, as indicated in Fig. 4A. Again, the distribution of drugs into the brain is an important focus in the discovery of drugs that target the central nervous system (CNS). Most polar drugs are prevented by the BBB from entry into the CNS. CNS activity can be computed on a scale of -2 (inactive) to $+2$ (active) [32]. With regards to the CNS defines the drug's ability to cross into the brain while improving the efficacy of drugs (fewer side effects). Fig. 4B shows that 9.1 % of MNP and 17 % of SNP fall within the given range that were predicted to penetrate the CNS. Another key parameter for drug distribution is the logarithm of predicted binding constant to human serum albumin ($\log K_{hsa}$). When a drug is absorbed into the intestinal wall, it enters the circulatory system and binds to plasma proteins. It is vital to evaluate the extent to which the drug is bound to plasma proteins, as it determines the amount of free drug that penetrates the membranes to get to the desired target. Therefore, $\log K_{hsa}$ descriptor is used to predict the binding capacity of substances to plasma proteins like the glycoproteins, human serum albumin, lipoproteins, and globulins [33]. While the recommended range for $\log K_{hsa}$ is -1.5 to 1.5 [34], it was observed that 99 % of MNP and 97 % of SNP compounds

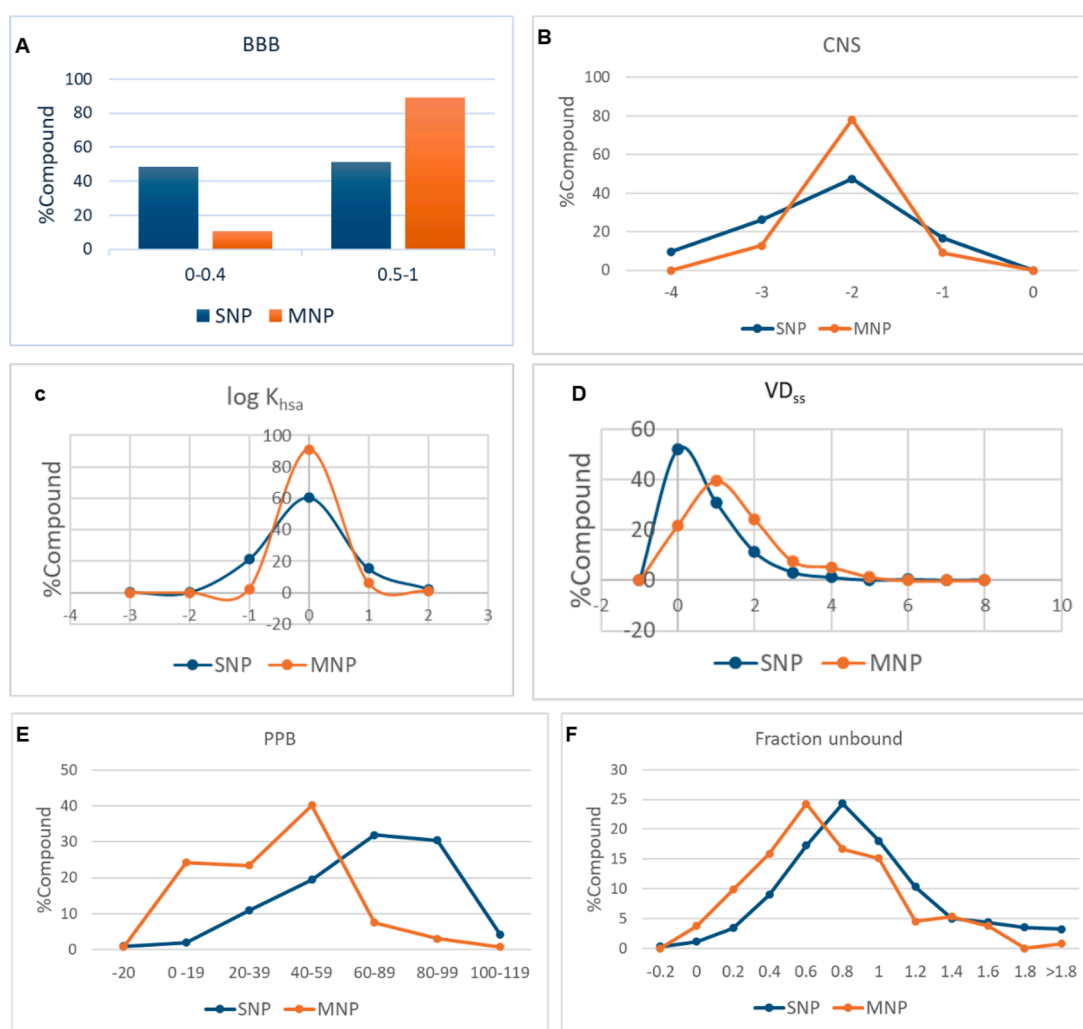


Fig. 4. Plots showing curves and histograms of properties related to drug distribution, for the MNP and SNP compounds: (A) Blood-brain barrier (BBB) penetration (B) Central Nervous System (CNS) activity, (C) $Q\log K_{hsa}$, (D) Steady-state volume of distribution, (E) Plasma protein binding (PPB), and (F) Fraction of unbound drug.

could not penetrate the membrane to get to the desired target. A smooth Gaussian-shaped curve was observed for the two datasets with a maximum value at 0 log K_{hsa} unit.

The steady state volume of distribution (VD_{ss}) is a parameter used to assess a drug's ability to be distributed throughout the body. When the VD_{ss} is high, more of the drug is distributed to the tissues than to the plasma [35,36]. A compound has good distribution if its VD_{ss} value is higher than 0.45, and our observations show that 78 % of MNP compounds and 47 % of SNP were predicted to be distributed within the range of > 0.45 predicted to be uniformly distributed to tissue and blood plasma (Fig. 4D). Plasma protein binding (PPB) is referred to the extent to which a drug or compound binds to protein in the blood stream such as albumin, lipoprotein, and globulins. So, when a drug binds to the plasma protein, its quantity in circulation is reduced, and the effectiveness of the drug may also be affected compared to when it is less bound [37]. The distribution (Fig. 4E) showed that about 25 % of the MNP compounds and 3 % of the SNP predicted to be less bound are compliant with this parameter, indicating that these compounds are likely to circulate freely within the bloodstream and hence may have access to the target site. The predicted fraction of unbound drug (F_u) is a pharmacokinetic parameter that represents the proportion of a drug that is not bound to plasma proteins. This is the proportion of the drug that exerts a pharmacological effect and can be readily distributed to tissues and organs as free drugs, and can also be eliminated. Fig. 4F showed that 56 % MNP and 60 % SNP were predicted to be unbound. Therefore, MNP compounds were predicted to have more potential drug candidates than the SNP.



Fig. 5. Histograms, curves, and pie charts of toxicity-related properties: (A and B) AMES mutagenesis, (C) liver injury I, (D) maximum tolerated dose (MTD), (E) Rat acute toxicity, (F) Rat chronic oral toxicity, (G and H) carcinogenesis, (I) hERG blockers, and (J) QPlogHERG showing the distribution of MNP and SNP compounds.

Properties related to drug metabolism and general pharmacokinetics

Metabolism is defined as the chemical modification or the biotransformation of foreign substances to enhance their hydrophilicity and water solubility to facilitate excretion. The metabolism of drugs is essentially hepatic, where the liver plays the key role (key organ) in the metabolism and detoxification of xenobiotic compounds [38]. Descriptors like number of metabolic reactions (#metab), affinity to organic anion-transporting polypeptide 1B1 (OATP1B1), and susceptibility to be biotransformed by the enzyme CYP1A2 were used for this analysis, particularly since CYP1A2 inhibition predictions are relevant for assessing drug-drug interaction potential, particularly with polyphenolic or aromatic phytochemicals. Other clinically important and relevant CYP isoforms include CYP3A4 and CYP2D6. Given that CYP3A4 metabolizes approximately half of all marketed drugs, it was observed that only about 9 % of the MNP compounds were predicted to inhibit CYP3A4, thereby posing a potential problem for drug-drug interactions. Besides, based on probabilities, only less than one-third of those predicted to inhibit CYP3A4 could be classified as "high confidence". The majority were "medium" and "low" confidence predictions. For CYP2D6, a highly significant enzyme in drug metabolism due to its role in the oxidative metabolism of a wide range of pharmaceutical compounds (e.g., it metabolizes ~25 % of all clinically used drugs, despite accounting for only ~2–5 % of total hepatic CYP enzymes), even fewer of the MNP compounds were predicted to inhibit this isoform. It must be mentioned that CYP2D6 is often evaluated early in drug development to predict metabolism, drug-drug interactions, and personalized dosing strategies. For a drug to be biologically transformed, it needs to be controlled by the number of metabolic reactions (from 1 to 8). Compounds metabolized by more than eight metabolic reactions are generally known to be complex secondary metabolites [39]. Therefore, our compounds that were predicted to fall in the recommended ranges of #metab had 95 % MNP and 75 % SNP as indicated in Fig. S3A (Supplementary Data). OATP1B1 is a protein that plays a crucial role in the transport of substances like drugs, metabolites across cell membranes. It facilitates the uptake of substances from the bloodstream into the liver cells, as is primarily expressed in the liver. In this study, it was observed that 98 % of compounds were predicted to be able to transit cell membranes, whereas 78 % of SNP were also predicted (Fig. S3B). CYP1A2 is a member of the cytochromes P450 (CYP) superfamily of enzymes. It plays a crucial role in the metabolism of substances, including drugs. It was observed that 57 % of SNP and 55 % of MNP compounds have non-inhibitory properties of the CYP1A2 isoform of CYP450, respectively (Figs. S3C and S3D).

Properties related to excretion

Excretion is the clearance of chemicals from the body. Excretion parameters used in this study include total clearance (Cl_{Tot}), the half-life, and the propensity to be an organic cation transporter 2 (OrCT2) substrate. Clearance is a pharmacokinetic parameter that measures the rate at which a drug is removed from the body. It also defines together the volume of distribution, the half-life, and the frequency of dosing of a drug. It assesses the clearing of a compound through the liver and kidneys. It helps in the determination of the dosage rate and dose concentration of the drug at steady state. It was observed that 89 % of MNP compounds and 93 % of SNP were predicted to be readily excreted from the human body after executing the therapeutic function within the range of 1 to greater than 100 ml/min/kg. The half-life of a drug is the time needed for the plasma concentration of a drug to reach 50 % due to elimination processes. Half-life is therefore affected by each ADME effect of a drug. This half-life is an essential parameter for dose schedules. If the half-life is greater, it will turn to reduced efficacy, i.e., if the dosing interval is too long, the drug's effect may wear off, requiring increased dosing frequency, and lead to fluctuation in concentration. Observations show that MNP has 78 % of compounds with a short half-life of < 3 hr, while the SNP has 85 % (Fig. S4, Supplementary Data). Therefore, these compounds may have a short duration of action, accompanied by less accumulation and faster elimination.

Toxicity-related properties

Selected toxicity-related parameters include predictions for the AMES mutagenesis test, liver injury, maximum tolerated dose (MTD), rat acute toxicity, chronic oral toxicity, carcinogenesis, and human ether-a-go-go (hERG) channel blocking, often related to cardiac toxicity of drugs. Fig. 5 shows the distribution of the toxicity-related predicted parameters for the MNP and SNP datasets. For the AMES mutagenesis test, it was predicted that 84 % of MNP were safe when compared with 77 % (Fig. 5A and 5B). The same trend is seen for the risk of liver injury, MTD, and hERG blockage, e.g., <20 % of the MNP showed risk of liver injury between 0.5 and 1, when compared with 40 % for SNP (Fig. 5C), a very similar trend to the hERG distribution (Fig. 5I). On the contrary, a higher percentage of SNP (86 %) were predicted to be safe in terms of carcinogenesis, when compared with only 62 % for MNP. Overall, the MNP outperformed the SNP regarding the toxicity-related parameters, even though this property is often quite difficult to predict using *in silico* models [40].

Conclusion

In this work, we have compiled a small chemically diverse dataset of chemical constituents identified from literature sources in the genus *Monodora*, which are widely used in the treatment of several ailments. The additional advantage of this work is that the supplementary data shows a summary of the literature information regarding the compounds, their plant species of origin, as well as the computed DMPK parameters. Unfortunately, several computer models are often not good at predicting DMPK parameters, e.g., in the prediction of toxicity parameters like hERG channel blockage [40]. To avoid this inconvenience, the evaluation of the computed DMPK-related parameters of the compound collection was carried out using three separate tools that employ machine learning predictions (e.g. pkCSM) and QSAR models (e.g. QikProp) or a combination of both (e.g. SwissADME). The generated data was compared

with those of a recently published dataset of phytochemicals having activities against the coronavirus SARS-CoV-2 or any of its drug targets, and the results reveal that a significant proportion of the phytochemicals from *Monodora* species have "druglike" properties and, hence, hold a promise for drug discovery based on compounds from medicinal plants. The 600 SARS-CoV-2 compounds already have proven activities against the virus and/or its vital enzyme targets, e.g., M^{pro}, spike/ACE2, RdRp, etc. [24]. These are all manually curated natural products with proven in vitro activities against SARS-CoV-2 in phenotypic assays of inhibiting any of the SARS-CoV-2 target enzymes in biochemical assays. Some of the compounds have shown early clinical trial efficacy in the treatment of SARS-CoV-2. A comparison of the chemical space, most common scaffolds, and DMPK profiles of the two datasets would provide clues as to whether the genus *Monodora* could provide phytochemicals that could potentially target this virus. Prior work mainly catalogued ethno-medicinal uses or chemical constituents, but no study had comprehensively evaluated their drug-likeness and pharmacokinetic properties.

This represents the most extensive curated dataset for this genus to date, covering multiple compound classes (alkaloids, terpenoids, coumarins, etc.). This is the first comparative benchmarking against >600 SARS-CoV-2 natural products, demonstrating that *Monodora* compounds show superior absorption, distribution, and toxicity profiles compared with natural products already reported for viral targets is both new and of broad scientific interest. Interestingly, this work provides a publicly available and well-curated dataset for downstream drug discovery applications. The new resource is the curated dataset itself, made available for community use as a ready-to-use resource for molecular docking, target identification, and cheminformatics studies, which is not common in studies of African medicinal plants. The molecular structures have been made available as additional files to facilitate their target elucidation through molecular docking. In the summary shown in Table 1, ten selected compounds have been highlighted because of their interesting biological activities. There are 171 entries in the Supplementary Data, showing the various plant species of origin and literature references. However, some of the compounds are present in several species, meaning that the number of unique compounds is 132, and these constitute the basis of this investigation and further investigations on compound prioritization for lead discovery from phytochemicals in this genus. Further investigation of these 132 compounds in vitro is possible in the future as a follow-up to this study, particularly against SARS-CoV-2 and some of its protein targets.

Declarations

Funding

We acknowledge financial support from the Bill & Melinda Gates Foundation through the Calestous Juma Science Leadership Fellowship awarded to Fidele Ntie-Kang (grant award number: INV-036848 to University of Buea). FNK also acknowledges joint funding from the Bill & Melinda Gates Foundation and LifeArc (award number: INV-055897 and Grant ID: 10646) under the African Drug Discovery Accelerator program. FNK acknowledges further funding from the Alexander von Humboldt Foundation for a Research Group Linkage project (award number Ref [3].4- 1156361-CMR-IP).

Author contributions

LNA: Conceptualization, Methodology, Data curation, Formal analysis, Investigation, Writing –original draft. **JYB:** Conceptualization, Methodology, Data curation, Formal analysis, Investigation, Writing – original draft. **MJMT:** Formal analysis, Investigation, **GN:** Conceptualization, Investigation, Methodology, **JDW:** Conceptualization, Formal analysis, Supervision, Writing – review & editing. **FN-K:** Funding acquisition, Investigation, Methodology, Supervision, Writing – original draft, Writing – review & editing. **All authors:** reviewed the manuscript.

Ethical approval

Not applicable

Consent to participate

Not applicable

Clinical trial number

Not applicable

Consent to publish

Not applicable

Data availability

All data have been made available in the manuscript and supplementary materials

Code availability

Not applicable

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.sciaf.2025.e03060](https://doi.org/10.1016/j.sciaf.2025.e03060).

References

- [1] D.J. Newman, Natural products and drug discovery, *Natl. Sci. Rev.* 9 (11) (2022) nwac206, <https://doi.org/10.1093/nsr/nwac206>.
- [2] M. Lahlou, The success of natural products in drug discovery, *Pharmacol. Pharm.* 4 (3A) (2013) 17–31, <https://doi.org/10.4236/pp.2013.43A003>.
- [3] R.J. Singh, A. Lebeda, O. Tucker, Chapter 2. Medicinal plants nature's pharmacy, in: R.J. Singh (Ed.), *Genetic resources, Chromosome Engineering, and Crop improvement. Medicinal Plants, Genetic resources, Chromosome Engineering, and Crop improvement. Medicinal Plants*, 6, CRC Press, Boca Raton, 2012, pp. 13–51.
- [4] L.W. Chatrou, M.D. Pirie, R.H.J. Erkens, T.L.P. Couvreur, K.M. Neubig, J.R. Abbott, J.B. Mols, W. Jan, J.W. Maas, R.M.K. Saunders, M.W. Chase, A new subfamilial and tribal classification of the pantropical flowering plant family Annonaceae informed by molecular phylogenetics, *Bot. J. Linn. Soc.* 169 (2012) 5, <https://doi.org/10.1111/j.1095-8339.2012.01235.x>.
- [5] M.N. Adiko, B.K.F.P. Kouamé, A.F. Kabran, A.O. Akoubet, J.R. Kablan, C. Siomenan, et al., Isoquinoline alkaloids from *Monodora crispata* Eng. and Diels, *Monodora myristica* Gaertn. and *Monodora tenuifolia* Benth. (Annonaceae), *Int. J. Pharmaceut. Sci. Invent.* 7 (2018) 11–16.
- [6] N. Magori, S.S. Nyandoro, E.J.J. Munissi, M. Heydenreich, Antimycobacterial and cytotoxicity evaluation of the constituents of *Monodora carolinae*, *Tanz. J. Sci.* 39 (1) (2013) 12–18.
- [7] L. Kablan, J. Dade, T. Okpekon, F. Roblot, L.A. Djakouré, P. Champy, Alkaloids from the leaves *Monodora crispata* Engl. And Diels and *M. brevipes* Benth. (Annonaceae), *Biochem. Syst. Ecol.* 46 (2013) 162–165, <https://doi.org/10.1016/j.bse.2012.09.024>.
- [8] P.O. Adomi, J.M. Nana, Antibacterial and phytochemical activities of *Monodora myristica* (African nutmeg) seeds, *Trop. J. Sci. Technol.* 4 (2) (2023) 17–22, <https://doi.org/10.47524/tjst.v4i2.18>.
- [9] A. Agiriga, M. Siwela, *Monodora myristica* (Gaertn.) Dunal; a plant with multiple food, health and medicinal applications, *Am. J. Food Technol.* 12 (4) (2017) 271–284, <https://doi.org/10.3923/ajft.2017.271.284>.
- [10] M.H.H. Nkunya, J.J. Makangara, S.A. Jonker, Prenyl indoles from Tanzanian *Monodora* and *Isolona* species, *Nat. Prod. Res.* 18 (2004) 253–258, <https://doi.org/10.1080/14786410310001620529>.
- [11] D.J.M. Eric, C. Siomenan, B.K.P. Kouamé, A.L.C. Kablan, C.G. Kodjo, Z. Doumade, A. Akoubet, M. Adiko, A.S. Aka, J.R. Kablan, A.L. Djakouré, A.K. Barthélemy, A new natural indole and three aporphine alkaloids from *Monodora bevipipes* Benth. (Annonaceae), *Int. Curr. Pharmaceut. J.* 6 (7) (2017) 40–43, <https://doi.org/10.3329/ICPJ.V6i7.34330>.
- [12] R.N. Nwaoguikpe, C.O. Ujowundu, A.A. Emejulu, The antioxidant and free radical scavenging effects of extracts of seeds of some neglected legumes of South-East Nigeria, *Sch. Acad. J. Biosci.* 2 (1) (2014) 51–59.
- [13] O. Miediegha, A.D.C. Owaba, L. Okori-west, Acute toxicity studies, physicochemical and GC/MS analyses of *Monodora myristica* (Gaertn.) Dunal oil, *Nig. J. Pharm. Res.* 18 (1) (2022) 91–99, <https://doi.org/10.4314/njpr.v18i2.1>.
- [14] O.O. Ogunrinola, R.I. Kanmodi, A. Oluwaseyi, A.O. Ogunrinola, Medicinal plants as immune booster in the palliative management of viral diseases: a perspective on coronavirus, *Food Front.* 3 (1) (2022) 83–95, <https://doi.org/10.1002/fft2.107>.
- [15] S.A. Onikanni, B. Lawal, A.O. Fadaka, O. Bakare, E. Adewole, M. Taher, J. Khotib, D. Susanti, B.E. Oyinloye, B.O. Ajiboye, O.A. Ojo, N.R.S. Sibuyi, Computational and Preclinical prediction of the antimicrobial properties of an agent isolated from *Monodora myristica*: a novel DNA gyrase inhibitor, *Molecules* 28 (4) (2023) 1593, <https://doi.org/10.3390/molecules28041593>.
- [16] F. Ntie-Kang, K.K. Telukunta, K. Döring, C.V. Simoben, A.F. A Moumbock, Y.I. Malange, L.E. Njume, J.N. Yong, W. Sippl, S. Günther, NNPDB: a resource for natural products from Northern African sources, *J. Nat. Prod.* 80 (7) (2017) 2067–2076, <https://doi.org/10.1021/acs.jnatprod.7b00283>.
- [17] C.V. Simoben, A. Qaseem, A.F.A. Moumbock, K.K. Telukunta, S. Günther, W. Sippl, F. Ntie-Kang, Pharmacoinformatic investigation of medicinal plants from East Africa, *Mol. Inform.* 39 (11) (2020) e2000163, <https://doi.org/10.1002/minf.202000163>.
- [18] H.E. Pence, A. Williams, ChemSpider: an online chemical information resource, *J. Chem. Educ.* 87 (2010) 1123–1124, <https://doi.org/10.1021/ed100697w>.
- [19] O.T. Eboh, S.B. Babiaka, F. Ntie-Kang, Natural products as potential lead compounds for drug discovery against SARS-CoV-2, *Nat. Prod. Bioprospect.* 11 (6) (2021) 611–628, <https://doi.org/10.1007/s13659-021-00317-w>.
- [20] A. Daina, O. Michielin, V. Zoete, SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules, *Sci. Rep.* 7 (2017) 42717, <https://doi.org/10.1038/srep42717>.
- [21] X.Q. Lewell, D.B. Judd, S.P. Watson, M.M. Hann, RECAP-retrosynthetic combinatorial analysis procedure: a powerful new technique for identifying privileged molecular fragments with useful applications in combinatorial chemistry, *J. Chem. Inf. Comput. Sci.* 38 (3) (1998) 511–522, <https://doi.org/10.1021/ci970429i>.
- [22] D. Weininger, SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules, *J. Chem. Inf. Comput. Sci.* 28 (1) (1988) 31–36, <https://doi.org/10.1021/ci00057a005>.
- [23] M. Congreve, R. Carr, C. Murray, H. Jhoti, A 'rule of three' for fragment-based lead discovery? *Drug Discov. Today* 8 (19) (2003) 876–877, [https://doi.org/10.1016/s1359-6446\(03\)02831-9](https://doi.org/10.1016/s1359-6446(03)02831-9).
- [24] J.Y. Betow, G. Turon, C.S. Metuge, S. Akame, V.A. Shu, O.T. Eboh, M. Duran-Frigola, F. Ntie-Kang, The chemical space spanned by manually curated datasets of natural and synthetic compounds with activities against SARS-CoV-2, *Mol. Inform.* 44 (2025) e202400293, <https://doi.org/10.1002/minf.202400293>.
- [25] A.O. Adeoye, B.O. Oguntim, 3-dimethylallylindole: an antibacterial and antifungal metabolite from *Monodora tenuifolia*, *J. Nat. Prod.* 49 (3) (1986) 534–537, <https://doi.org/10.1021/np50045a031>.
- [26] C. Tian, X. Liu, Y. Chang, R. Wang, L. Tianmeng, C. Cui, M. Liu, Investigation of the anti-inflammatory and antioxidant activities of luteolin, kaempferol, apigenin and quercetin, *S. Afr. J. Bot.* 137 (2021) 257–264, <https://doi.org/10.1016/j.sajb.2020.10.022>.
- [27] C.A. Lipinski, Lead- and drug-like compounds: the rule-of-five revolution, *Drug Discov. Today Technol.* 1 (2004) 337–341, <https://doi.org/10.1016/j.ddtec.2004.11.007>.

- [28] R. Roskoski Jr., Rule of five violations among the FDA-approved small molecule protein kinase inhibitors, *Pharmacol. Res.* 191 (2023) 106774, <https://doi.org/10.1016/j.phrs.2023.106774>.
- [29] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv. Drug Deliv. Rev.* 46 (1–3) (2001) 3–26, [https://doi.org/10.1016/s0169-409x\(00\)00129-0](https://doi.org/10.1016/s0169-409x(00)00129-0).
- [30] D.F. Veber, S.R. Johnson, H.Y. Cheng, B.R. Smith, K.W. Ward, K.D. Kopple, Molecular properties that influence the oral bioavailability of drug candidates, *J. Med. Chem.* 45 (12) (2002) 2615–2623, <https://doi.org/10.1021/jm020017n>.
- [31] D. Lagorce, D. Douguet, M.A. Miteva, B.O. Villoutreix, Computational analysis of calculated physicochemical and ADMET properties of protein-protein interaction inhibitors, *Sci. Rep.* 7 (2017) 46277, <https://doi.org/10.1038/srep46277>.
- [32] Schrödinger Press, *QikProp 3.4 User Manual*, LLC, New York, NY, 2011.
- [33] P. Wils, A. Warnery, V. Phung-Ba, S. Legrain, D. Scherman, High lipophilicity decreases drug transport across intestinal epithelial cells, *J. Pharmacol. Exp. Ther.* 269 (2) (1994) 654–658, [https://doi.org/10.1016/s0022-3565\(25\)38749-5](https://doi.org/10.1016/s0022-3565(25)38749-5).
- [34] C. Bertucci, E. Domenici, Reversible and covalent binding of drugs to human serum albumin: methodological approaches and physiological relevance, *Curr. Med. Chem.* 9 (15) (2002) 1463–1481, <https://doi.org/10.2174/0929867023369673>.
- [35] G. Colmenarejo, A. Alvarez-Pedraglio, J.-L. Lavandera, Cheminformatic models to predict binding affinities to human serum albumin, *J. Med. Chem.* 44 (2001) 4370–4378, <https://doi.org/10.1021/jm010960b>.
- [36] G.J.M. Mvondo, A. Matondo, D.T. Mawete, S.N. Bambi, B.M. Mbala, O.P. Lohohola, *In silico* ADME/T properties of quinine derivatives using SwissADME and pkCSM web servers, *Int. J. Trop. Dis. Health* 42 (11) (2021) 1–12, <https://doi.org/10.9734/IJTDH/2021/v42i1130492>.
- [37] D.E. Pires, T.L. Blundell, D.B. Ascher, pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using Graph- Based Signatures, *J. Med. Chem.* 58 (2015) 4066–4072, <https://doi.org/10.1021/acs.jmedchem.5b00104>.
- [38] F. Ntie-Kang, An in-silico evaluation of the ADMET profile of the StreptomeDB database, *Springerplus* 2 (2013) 353, <https://doi.org/10.1186/2193-1801-2-353>.
- [39] B.Z. Gbolo, K.N. Ngbolua, D.S.T. Tshibangu, P.B. Memvanga, D.D. Tshilanda, A. Matondo, J.T. Kilembe, B.M. Lebwaze, A. Nachtergaele, P.T. Mpiana, P. Duez, In vivo evaluation and *in silico* prediction of the toxicity of Drepanoalpha hard capsules, *bioRxiv*. (2019), <https://doi.org/10.1101/2020.12.03.411124> preprint.
- [40] G. Falcón-Cano, A. Morales-Helguera, H. Lambert, M.A. Cabrera-Pérez, C. Molina, hERG toxicity prediction in early drug discovery using extreme gradient boosting and isometric stratified ensemble mapping, *Sci. Rep.* 15 (1) (2025) 15585, <https://doi.org/10.1038/s41598-025-99766-3>.