Trop J Nat Prod Res, January 2025; 9(1): 295 - 305 ISSN 2616-0684 (Print)
ISSN 2616-0692 (Electronic)

Tropical Journal of Natural Product Research

Available online at https://www.tjnpr.org

Original Research Article

Anti-Fungal Activity Prediction of *Curcuma Longa* and *Piper belte* through Molecular Docking

Marisca Evalina Gondokesumo [1], Muhammad Rezki Rasyak [2], Mansur Ibrahim[3]

[1]*Faculty of Pharmacy, University of Surabaya, Indonesia.*
[2]*Eijkman Research Centre for Molecular Biology, National Research and Innovation Agency, Indonesia.*
[3]*Faculty of Pharmacy, Megarezky University, Indonesia.*

ARTICLE INFO

Article history:
Received 08 July 2024
Revised 28 July 2024
Accepted 17 December 2024
Published online 01 February 2025

ABSTRACT

Traditional medicine is deeply rooted in Indonesian culture, yet concerns persist regarding the safety and efficacy of medicinal plants. This study focuses on *Curcuma longa* and *Piper betle*, two prevalent medicinal plants in Southeast Asia, known for their potential antifungal properties, though the specific bioactive compounds remain unidentified. A machine learning model was developed using a dataset from DUD-E (Directory of Useful Decoys) docking to predict antifungal compounds, successfully identifying substances capable of inhibiting Cytochrome P450 EryK protein. Notably, compounds such as curcumin, beta-caryophyllene, and eugenol were highlighted for their significant antifungal activity. From the analysis, 21 compounds from *Curcuma longa* and 15 from *Piper betle* were selected and screened for their inhibitory potential. These compounds were further filtered based on Lipinski's rule of five, resulting in 7 compounds from *Curcuma longa* and 5 from *Piper betle* exhibiting binding affinities comparable to the reference ligand. Tanimoto similarity analysis revealed that the identified compounds shared less than 10% similarity in chemical structure to the reference ligand. However, the binding amino acids in several compounds demonstrated over 50% similarity to those of the reference ligand. The findings underscore the therapeutic potential of these compounds, contributing to the development of natural antifungal agents.

Keywords: Cytochrome P450 EryK, Machine learning, Molecular docking, *Curcuma longa*, *Piper betle.*

Introduction

Candidiasis, particularly vaginal candidiasis, is a common fungal infection affecting women, especially in tropical regions like Indonesia, where warm and humid conditions promote fungal growth.[1–3] Despite the availability of antifungal treatments, resistance to these therapies and their associated side effects have become pressing concerns, underscoring the need for alternative treatment options.[4] In this context, traditional medicinal plants offer a promising avenue for discovering new antifungal agents, with Indonesia's rich biodiversity presenting numerous candidates for study.[5] Among these, *Curcuma longa* (turmeric) and *Piper betle* (betel leaf) are well-known for their medicinal properties. turmeric has been extensively studied for its antioxidant, anti-inflammatory, and even anticancer activities, while its potential antifungal activity is an emerging area of interest.[6-7] Similarly, betel leaf has demonstrated antifungal effects, with caryophyllene identified as a key active compound in preliminary studies.[8-9] These plants, widely used in traditional medicine, contain a variety of bioactive compounds, many of which remain uncharacterized in terms of their potential to treat fungal infections. [10-11]

*Corresponding author. E mail: marisca@staff.ubaya.ac.id
Tel: 6287851367988

Citation: Gondokesumo ME, Rasyak MR, Ibrahim M Anti-Fungal Activity Prediction of *Curcuma Longa* and *Piper betle* through Molecular Docking. Trop J Nat Prod Res. 2025; 9(1): 295 – 305 https://doi.org/10.26538/tjnpr/v9i1.38

Official Journal of Natural Product Research Group, Faculty of Pharmacy, University of Benin, Benin City, Nigeria

In particular, compounds such as curcumin from turmeric and caryophyllene from betel leaf have shown potential in inhibiting fungal pathogens like *Candida albicans* (C. albicans). However, these plants likely contain other, yet-to-be-identified compounds that could possess antifungal properties.[12] Identifying these novel compounds is critical, as they may offer new therapeutic strategies for overcoming the limitations of current antifungal treatments. To aid in the discovery of these novel compounds, this study applies machine learning (ML), a branch of artificial intelligence (AI), to predict antifungal activity from chemical compound datasets.[13] ML has proven valuable in drug discovery by analyzing large datasets to identify promising leads based on molecular descriptors.[14] Datasets such as Directory of Useful Decoys (DUD-E) docking (http://dude.docking.org/) provide crucial information on protein inhibitors, including decoy compounds that assist in developing robust predictive models for drug discovery.[15] This study focuses on the cytochrome P450 EryK enzyme in C. albicans, which plays a key role in the biosynthesis of ergosterol, a critical component of the fungal cell membrane.[16-17] Inhibiting EryK disrupts ergosterol synthesis, leading to membrane instability and fungal cell death, making it a promising target for antifungal therapies.[17–20] By leveraging machine learning to explore compounds from turmeric and betel leaf, this research aims to identify novel inhibitors of the EryK enzyme, potentially leading to the discovery of effective antifungal agents for the treatment of candidiasis.

Materials and Method

Data mining and fingerprint extraction

In this study, machine learning was employed to predict inhibitors of the EryK cytochrome P450 enzyme from compounds within turmeric and betel leaf. Compound data for these plants were retrieved from the Knapsack Core Database (http://www.knapsackfamily.com/) using the

relevant keywords.[21] RDKit, an open-source Python software, was used to extract SMILES structures from all compounds, which were then utilized to generate PubChem fingerprints.[22] These fingerprints, consisting of 881 substructures, were assigned binary values of either one or zero.

Subsequently, the DUD-E docking platform was used to generate a decoy dataset, including known antifungal small-molecule protein inhibitors currently available on the market.[15] The platform also provided decoy compounds to serve as non-active controls. Machine learning models were then constructed, incorporating both active and inactive substances. Before model development, fingerprint extraction for each substructure was performed using the RDKit fingerprint extractor.

Machine learning model development
The Python library Scikit-learn provided the framework for constructing various machine learning models.[23] Jupyter notebooks served as the development environment for writing and implementing the model code.[24] Lazypredict (https://lazypredict.readthedocs.io/_/downloads/en/latest/pdf/) was used to compare 27 classification models, with 10-fold cross-validation applied to ensure robustness. The model with the highest AUC/ROC score was selected for further analysis.

In addition to AUC/ROC, the following metrics were calculated:

Accuracy: The proportion of correct predictions (both active and inactive compounds).

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Samples}}$$

Sensitivity: The ability to correctly identify active compounds

$$\text{Sensitivity} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

Specificity: The ability to correctly identify inactive compounds.

$$\text{Specificity} = \frac{\text{True Negatives}}{\text{True Negatives} + \text{False Positives}}$$

These metrics offered a comprehensive assessment of the models, with the trade-offs between sensitivity and specificity carefully considered. The selected model was then used to predict antifungal activity in compounds from turmeric and betel leaf.

Molecular docking from predicted active compound and interaction analysis
Following the machine learning prediction of potentially active antifungal substances within turmeric and betel leaf, a molecular docking investigation was conducted to assess their binding affinities. PLANTS 1.1 software was utilized for this purpose.[25] This program leverages an innovative approach called an "artificial ant colony" to simulate how these ants find optimal paths. In the context of docking, the "ants" explore different conformations of the predicted active compounds within the target protein's binding pocket, ultimately identifying the conformation with the lowest energy state.[26] PLANTS employs two scoring functions (PLANTSCHEMPLP and PLANTSPLP) to evaluate these docked poses, balancing accuracy with computational efficiency.[25]

After the docking simulations, PRODIGY software was used to calculate the binding affinity (ΔG) of the docked poses.[27] This program analyzes protein-ligand interactions at the atomic level, providing a more detailed assessment compared to traditional methods that rely solely on residue contacts (https://bianca.science.uu.nl/prodigy/cryst). The key advantages of PRODIGY-LIG include its user-friendliness, broad applicability to various protein-ligand complexes, and its demonstrated effectiveness.[28]

Druglikeness and ADMET (Absorption, Distribution, Metabolism, Excretion and toxicity) analysis
A commonly used metric to assess drug-likeness is Lipinski's rule of five. This rule considers four key properties: Maximum of five hydrogen bond donors (bonds between hydrogen and oxygen or nitrogen atoms), ten or fewer hydrogen bond acceptors (all nitrogen or oxygen atoms), molecular weight less than 500 Daltons, LogP value no greater than 4.15.[29] By evaluating these properties, researchers can gain insights into how well a potential drug candidate might be absorbed, distributed, metabolized, and excreted within the body. Toxtree software is employed to evaluate the toxicity of each compound.[30] The threshold for toxicological concern is often referred to as the Threshold of Toxicological Concern (TTC). TTC determines a safe exposure level for all substances, below which there is no significant risk to human health.[30] Blood-brain barrier (BBB) and human intestinal absorption (HIA) analyses are often integrated with toxicity assessments to predict the number of compounds that can be absorbed by the gastrointestinal tract (GI). Alongside toxicity analysis, BBB and HIA analyses help to estimate how many compounds can be absorbed from the gut and whether they can cross the blood-brain barrier. This combined approach aids in identifying potential drug candidates.[31]

Tanimoto similarity for chemical structure and binding site similarity calculation
Following successful molecular docking simulations, the potential antifungal compounds identified from turmeric and betel leaf were subjected to further analysis. PyPLIF, a Python library, was employed to convert the interaction information obtained from the docking simulations into a format suitable for similarity analysis.[32] This conversion process involves generating a unique "fingerprint" for each compound using Python interaction fingerprinting (IFP). This fingerprint essentially captures the key interaction patterns between the compound and the target protein.

Next, Pyplif-HIPPOS, another Python library, was utilized to compare the fingerprints of the potential antifungal compounds with those of known reference ligands that are already bound to the protein.[33] Pyplif-HIPPOS calculates a similarity score for each comparison, allowing researchers to identify potential antifungal compounds that exhibit interaction patterns similar to established antifungal agents. This approach helps prioritize promising candidates for further investigation.

Results and Discussion

Data mining and fingerprint extraction
A total of 180 compounds were collected from the Knapsack database using the keyword turmeric. Additionally, 37 compounds were collected using the keyword betel leaf. For training and testing purposes, active small molecules were collecting from 25 publications that mentioned drugs for candidiasis and their similar subtstructure in PubChem (https://pubchem.ncbi.nlm.nih.gov/) while decoy dataset was compiled from DUD-E. This dataset included 1387 active compounds and 1579 decoy compounds. These active and decoy compounds were then employed to develop a model after fingerprint extracting using pubchem fingerprint. The entire dataset comprised 2936 compounds, with approximately 75% designated for training and 25% reserved for testing.

Machine learning model development
Following evaluation by Lazypredict, the Support Vector Classifier (SVC) model emerged as the top performer compared to other models considered. This model was then subjected to parameter tuning, a process that optimizes its performance. Consequently, the SVC model achieved impressive metrics, including sensitivity (0.97), specificity (0.87), accuracy (0.92), and AUC/ROC score (0.96) (Table 1). To further validate these findings, a k-fold cross-validation with five splits was employed. In this technique, the data is divided into multiple subsets (folds). The model is then iteratively trained on one fold and tested on the remaining folds. This process is repeated for all folds. The k-fold cross-validation yielded consistently strong results for the SVC model, with an average training score of 0.914 and an average testing score of 0.906. This robust performance across various data

partitions solidified the SVC model as the optimal choice for predicting active compounds within turmeric and betel leaf.

Table 1: The Score of each of models in Lazypredit for Cytochrome P450 EryK Protein

No	Model	Accuracy	Balanced Accuracy	ROC AUC	F1 Score
1	SVC	0.92	0.92	0.92	0.92
2	Ridge Classifier CV	0.9	0.9	0.9	0.9
3	Ridge Classifier	0.89	0.9	0.9	0.89
4	Linear Discriminant Analysis	0.89	0.89	0.89	0.89
5	LGBM Classifier	0.89	0.89	0.89	0.89
6	NuSVC	0.89	0.89	0.89	0.89
7	XGB Classifier	0.89	0.89	0.89	0.89
8	AdaBoost Classifier	0.89	0.89	0.89	0.89
9	Passive Aggressive Classifier	0.88	0.88	0.88	0.88
10	Perceptron	0.88	0.88	0.88	0.88
11	KNeighbors Classifier	0.88	0.88	0.88	0.88
12	Random Forest Classifier	0.88	0.88	0.88	0.88
13	Extra Trees Classifier	0.88	0.88	0.88	0.88
14	Bagging Classifier	0.88	0.88	0.88	0.88
15	Logistic Regression	0.86	0.86	0.86	0.86
16	Bernoulli NB	0.85	0.85	0.85	0.85
17	Nearest Centroid	0.85	0.85	0.85	0.85
18	Decision Tree Classifier	0.85	0.85	0.85	0.85
19	Calibrated Classifier CV	0.85	0.85	0.85	0.85
20	SGD Classifier	0.84	0.84	0.84	0.84
21	Quadratic Discriminant Analysis	0.83	0.83	0.83	0.83
22	Linear SVC	0.83	0.83	0.83	0.83
23	Extra Tree Classifier	0.83	0.83	0.83	0.83
24	Gaussian NB	0.78	0.78	0.78	0.77
25	Label Spreading	0.68	0.68	0.68	0.67
26	Label Propagation	0.68	0.68	0.68	0.67
27	Dummy Classifier	0.52	0.5	0.5	0.35

The machine learning model identified 21 compounds from turmeric and 15 compounds from betel leaf with potential Cytochrome P450 EryK Protein inhibitory activity.

Molecular docking from predicted active compound

All compounds predicted to inhibit the Cytochrome P450 EryK protein were analyzed using molecular docking software to assess their potential binding within the protein's binding pocket. The protein structure of Cytochrome P450 EryK (PDB: 2XFH) was used, which includes a reference ligand bound to the protein with the PubChem code 2812.

Before analyzing compounds predicted by the machine learning model, a re-docking process is employed to validate the molecular docking software, PLANTS 1.1. This involves re-docking a known reference ligand (PubChem code: 2812) already bound to the Cytochrome P450 The EryK protein (PDB: 2XFH) can be used to evaluate docking software accuracy. This is done by performing re-docking simulations of a known ligand 1000 times. The Root Mean Square Deviation (RMSD) of the ligand's position in each result is then calculated. A low RMSD (ideally below 2 Å) indicates that the software can accurately reproduce known binding interactions, thus providing confidence in its ability to assess predicted active compounds. In other words, the software's accuracy is assessed by comparing the re-docked ligand's position with a reference ligand using RMSD. A low RMSD value suggests the software's ability to reliably predict binding interactions for new compounds.

The successful re-docking of the reference ligand for the Cytochrome P450 EryK protein (performed 1000 times) validates the ability of PLANTS 1.1 software to accurately reproduce known binding interactions. In this process, all re-docked ligand poses showed RMSD values consistently below 2 Ångstroms compared to the reference pose (**Suppl Figure 1**). This paves the way for the confident use of PLANTS 1.1 to analyze the potential binding of compounds predicted to be active by the machine learning model.

Following validation, the predicted active compounds from the machine learning model were subjected to molecular docking simulations using the validated software. The docking score and binding affinity score of selected compounds from turmeric and betel leaf are shown in **Table 2** and **Table 3** below.

The results of molecular docking analysis revealed interesting trends in the binding affinities of compounds from turmeric and betel leaf. Arachidic acid, gitoxigenin, tauroursodeoxycholic acid, glycocholic acid, beta-stigmasterol, (-)-beta-sitosterol, and 7-hydroxy-1,7-bis(4-hydroxy-3-methoxyphenyl)-1-heptene-3,5-dione from turmeric displayed lower binding affinity scores compared to the reference ligand, indicating potentially strong interactions with the protein.

Table 2: Docking and binding affinity score (ΔG) for the predicted compound as an inhibitor for Cytochrome P450 EryK Protein from *Curcuma longa*

Compound	Docking_score	Binding affinity (ΔG) Kcal/mol
Hexadecanal	-96.951	-8.21
Palmitic acid	-99.9729	-8.34
Octadecanamide	-109.021	-8.49
Dihydrosphingosine	-116.904	-8.57
alpha-Linolenic acid	-102.835	-8.67
Nerolidyl propionate	-105.276	-8.76
Octadecanoic acid	-103.365	-8.78
1,5-Dihydroxy-1,7-bis(4-hydroxy-3-methoxyphenyl)-4,6-heptadien-3-one	-97.0613	-8.81
Dihydrocurcumin	-95.1483	-8.82
Linoleic acid	-103.17	-8.87
1,7-Bis(4-hydroxy-3-methoxyphenyl)-3,5-heptanedione	-95.1344	-8.88
8,11-Octadecadienoic acid	-103.445	-8.88
Oleic acid	-101.255	-8.92
Curcumin	-92.6657	-8.92
7-Hydroxy-1,7-bis(4-hydroxy-3-methoxyphenyl)-1-heptene-3,5-dione	-90.6087	-9.17
Arachidic acid	-106.846	-9.18
Gitoxigenin	-86.536	-9.39
Tauroursodeoxycholic acid	-95.2822	-9.74
Glycocholic acid	-93.4788	-9.92
beta-Stigmasterol	-104.578	-10.13
(-)-beta-Sitosterol	-100.409	-10.9
Reference_ligand (Pubchem CID: 2812)	-77.1437	-8.97

Table 3: Docking and binding affinity score (ΔG) for the predicted compound as an inhibitor for Cytochrome P450 EryK Protein from *Piper betle*

Compound	Docking_score	Binding affinity (ΔG) Kcal/mol
Riboflavin	-86.8635	-6.98
Cineole	-63.8568	-7.13
Thiamin	-88.0711	-7.31
eugenol-methyl-ether	-72.5855	-7.59
Piperine	-86.1873	-8.07
Caryophyllene	-79.8837	-8.17
Piperlonguminine	-88.6462	-8.21
cepharadione a	-78.1528	-8.23
Cadinene	-83.9878	-8.24
octadecanoic acid	-103.095	-8.91
dotriacontanoic acid	-119.888	-10.74
Triacontane	-115.492	-10.81
(-)-beta-sitosterol	-100.057	-10.86

ISSN 2616-0684 (Print)
ISSN 2616-0692 (Electronic)

beta-sitosterol palmitate	-117.959	-12.5
beta-carotene	-113.959	-14.29
Reference_ligand (Pubchem CID: 2812)	-77.1437	-8.97

Similarly, dotriacontanoic acid, triacontane, (-)-beta-sitosterol, beta-sitosterol palmitate, and beta-carotene from betel leaf exhibited lower binding affinity scores than the reference ligand. Curcumin and caryophyllene, often cited as antifungal agents, demonstrated binding affinities close to the reference ligand, with curcumin at -8.92 kcal/mol and caryophyllene at -8.17 kcal/mol, differing by less than 1 kcal/mol. Binding affinity, measured through computational tools, refers to the strength of interactions between small molecules and target proteins. Lower scores indicate stronger and more favorable interactions within the protein's binding pocket.

Druglikeness and ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) Analysis

Compounds passing Lipinski's Rule of Five included dihydrosphingosine, curcumin, dihydrocurcumin, tauroursodeoxycholic acid, 1,7-bis(4-hydroxy-3-methoxyphenyl)-3,5-heptanedione, 1,5-dihydroxy-1,7-bis(4-hydroxy-3-methoxyphenyl)-4,6-heptadien-3-one, 7-hydroxy-1,7-bis(4-hydroxy-3-methoxyphenyl)-1-heptene-3,5-dione, gitoxigenin, and glycocholic acid from turmeric (Figure 1). From betel leaf, compounds included caryophyllene, cadinene, piperine, cepharadione A, piperlonguminine, cineole, and eugenol-methyl-ether (Figure 2). Tauroursodeoxycholic acid, 7-hydroxy-1,7-bis(4-hydroxy-3-methoxyphenyl)-1-heptene-3,5-dione, and glycocholic acid exhibited higher binding affinities than the reference ligand in turmeric, while all compounds from betel leaf that passed Lipinski's rule showed higher affinities compared to the reference ligand. Cadinene displayed the highest binding affinity of -8.24 kcal/mol compared to the reference ligand's -8.97 kcal/mol.

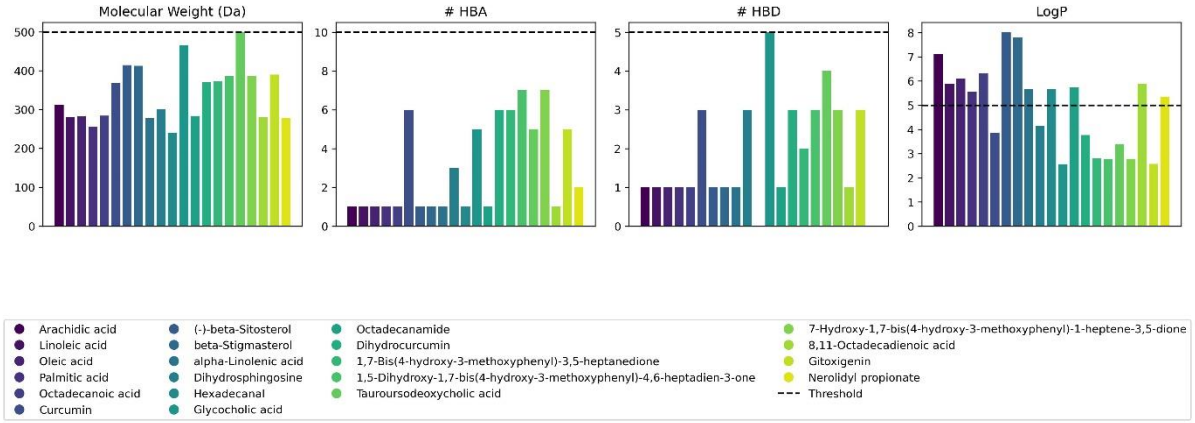


Figure 1: Lipinski rule of 5 for the Selected compound from *Curcuma longa*

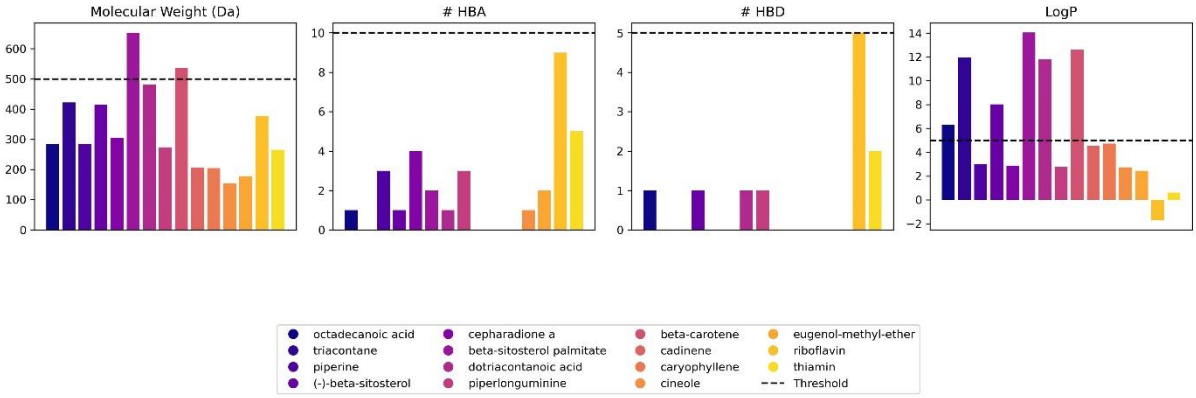


Figure 2: Lipinski rule of 5 for the Selected compound from *Piper betle*

In turmeric, gitoxigenin, linoleic acid, palmitic acid, alpha-linolenic acid, dihydrosphingosine, octadecanamide, 1,7-bis(4-hydroxy-3-methoxyphenyl)-3,5-heptanedione, and nerolidyl propionate were analyzed for GI absorption and BBB permeability. High GI absorption indicates good oral bioavailability, while good BBB permeability signifies the ability to cross the blood-brain barrier. Structural analysis using Toxtree Cramer rules identified dihydrosphingosine and 1,7-bis(4-hydroxy-3-methoxyphenyl)-3,5-heptanedione as exhibiting intermediate (Class II) toxicity, whereas octadecanamide was classified as high (Class III) toxicity. Linoleic acid, palmitic acid, alpha-linolenic acid, 8,11-octadecadienoic acid, and nerolidyl propionate were identified as low (Class I) toxicity compounds with better GI absorption and BBB permeation, all having near-similar scores to the reference ligand, differing by less than -1 kcal/mol.

Carcinogenicity and mutagenicity analysis with Toxtree indicated no carcinogenicity for linoleic acid, palmitic acid, alpha-linolenic acid, 8,11-octadecadienoic acid, and nerolidyl propionate (Table 4). Cramer rule analysis classified octadecanoic acid, triacontane, dotriacontanoic acid, cadinene, caryophyllene, thiamin, and eugenol-methyl-ether as low (Class I) toxicity from betel leaf (Table 5), although thiamin and eugenol-methyl-ether exhibited structural alerts for carcinogenicity, while the others showed negative results.

Table 4. ADME and Toxicity prediction of predicted ligand inhibitor for Cytochrome P450 EryK protein from *Curcuma longa*

Pubchem_id	Compound Name	GI Absorption	BBB	Creamer	Carcinogenicity
10467	Arachidic acid	Low	No	Low (Class I)	Negative
5280450	Linoleic acid	High	Yes	Low (Class I)	Negative
445639	Oleic acid	High	No	Low (Class I)	Negative
985	Palmitic acid	High	Yes	Low (Class I)	Negative
5281	Octadecanoic acid	High	No	Low (Class I)	Negative
969516	Curcumin	High	No	High (Class III)	Negative
222284	(-)-beta-Sitosterol	Low	No	High (Class III)	Negative
5280794	beta-Stigmasterol	Low	No	High (Class III)	Negative
5280934	alpha-Linolenic acid	High	Yes	Low (Class I)	Negative
91486	Dihydrosphingosine	High	Yes	Intermediate (Class II)	Negative
984	Hexadecanal	High	No	Low (Class I)	Structural Alert for genotoxic carcinogenicity
10140	Glycocholic acid	High	No	High (Class III)	Negative
31292	Octadecanamide	High	Yes	High (Class III)	Negative
10429233	Dihydrocurcumin	High	No	High (Class III)	Negative
124072	1,7-Bis(4-hydroxy-3-methoxyphenyl)-3,5-heptanedione	High	Yes	Intermediate (Class II)	Negative
13888134	1,5-Dihydroxy-1,7-bis(4-hydroxy-3-methoxyphenyl) -4,6-heptadien-3-one	High	No	High (Class III)	Negative
9848818	Tauroursodeoxycholic acid	Low	No	High (Class III)	Negative
131752799	7-Hydroxy-1,7-bis(4-hydroxy-3-methoxyphenyl) -1-heptene-3,5-dione	High	No	High (Class III)	Negative
5312487	8,11-Octadecadienoic acid	High	Yes	Low (Class I)	Negative
348482	Gitoxigenin	High	No	High (Class III)	Negative
81573	Nerolidyl propionate	High	Yes	Low (Class I)	Negative

Table 5. ADME and Toxicity prediction of predicted ligand inhibitor for Cytochrome P450 EryK protein from *Piper betle*

Pubchem_id	compound name	GI Absorption	BBB	creamer_	Carcinogenicity
5281	octadecanoic acid	High	No	Low (Class I)	Negative
12535	Triacontane	Low	No	Low (Class I)	Negative
638024	Piperine	High	Yes	High (Class III)	Structural Alert for nongenotoxic carcinogenicity and Structural Alert for genotoxic carcinogenicity
222284	(-)-beta-sitosterol	Low	No	High (Class III)	Negative
94577	cepharadione a	High	Yes	High (Class III)	Structural Alert for nongenotoxic carcinogenicity and Structural Alert for genotoxic carcinogenicity
13747834	beta-sitosterol palmitate	Low	No	High (Class III)	Negative
19255	dotriacontanoic acid	Low	No	Low (Class I)	Negative

5320621	piperlonguminine	High	Yes	High (Class III)	Structural Alert for nongenotoxic carcinogenicity and Structural Alert for genotoxic carcinogenicity
5280489	beta-carotene	Low	No	Intermediate (Class II)	Negative
3032853	cadinene	Low	No	Low (Class I)	Negative
5281515	caryophyllene	Low	No	Low (Class I)	Negative
2758	cineole	High	Yes	High (Class III)	Negative
7127	eugenol-methyl-ether	High	Yes	Low (Class I)	Structural Alert for nongenotoxic carcinogenicity and Structural Alert for genotoxic carcinogenicity
493570	riboflavin	Low	No	High (Class III)	Structural Alert for genotoxic carcinogenicity
1130	thiamin	High	No	Low (Class I)	Structural Alert for genotoxic carcinogenicity

Tanimoto Similarity for Chemical Structure and Binding Site Similarity Calculation

Tanimoto similarity calculations using Python interaction fingerprinting (IFP) did not identify any compounds in turmeric or betel leaf with over 10% structural similarity to the reference ligand (PubChem ID: 2812) (Table 6-7). However, at least 50% of the amino acids involved in compound binding exhibited similarity to the binding site of the reference ligand (Table 8). Glycocholic acid and gitoxigenin showed the highest binding site similarity (50%) in turmeric, while cadinene and caryophyllene exhibited the same in betel leaf (Table 9). Linoleic acid, palmitic acid, alpha-linolenic acid, 8,11-octadecadienoic acid, and nerolidyl propionate displayed IFP similarities ranging from 18-20% compared to the reference ligand, while cadinene and caryophyllene, classified as low toxicity, exhibited 50% IFP similarity.

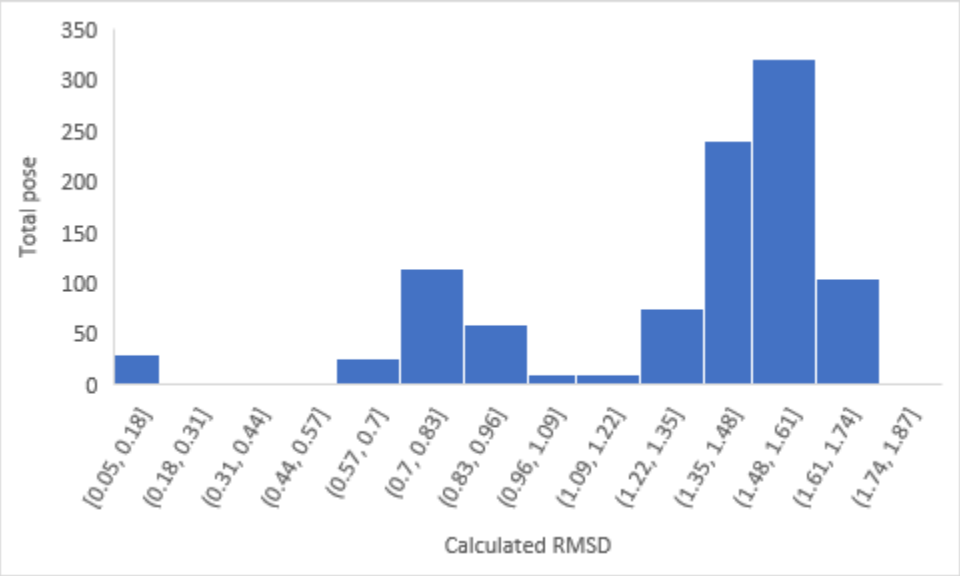
Table 6: Tanimoto similarity of structure selected compound from *Curcuma longa* with reference ligand (pubchem id: 2812)

Pubchem_id	Compound	Tanimoto similarity (%)
10467	Arachidic acid	2.1
5280450	Linoleic acid	3.6
445639	Oleic acid	1.9
985	Palmitic acid	2.1
5281	Octadecanoic acid	2.1
969516	Curcumin	6.5
222284	(-)-beta-Sitosterol	3.7
5280794	beta-Stigmasterol	2.4
5280934	alpha-Linolenic acid	3.6
91486	Dihydrosphingosine	0.0
984	Hexadecanal	0.0
10140	Glycocholic acid	1.1
31292	Octadecanamide	0.0
10429233	Dihydrocurcumin	5.7
124072	1,7-Bis(4-hydroxy-3-methoxyphenyl)-3,5-heptanedione	3.8
13888134	1,5-Dihydroxy-1,7-bis(4-hydroxy-3-methoxyphenyl)-4,6-heptadien-3-one	4.1
9848818	Tauroursodeoxycholic acid	0.0
131752799	7-Hydroxy-1,7-bis(4-hydroxy-3-methoxyphenyl)-1-heptene-3,5-dione	4.1
5312487	8,11-Octadecadienoic acid	3.6
348482	Gitoxigenin	5.1
81573	Nerolidyl propionate	1.5

The amino acids used in docking poses for cadinene and caryophyllene were found to be about 50% similar to the reference ligand (Table 9), indicating similar interactions with the protein. In turmeric, gitoxigenin and glycocholic acid exhibited more interactions, including additional

ISSN 2616-0684 (Print)
ISSN 2616-0692 (Electronic)

hydrogen bonds. Compounds showing better GI absorption, BBB permeability, and low toxicity—linoleic acid, palmitic acid, alpha-linolenic acid, 8,11-octadecadienoic acid, and nerolidyl propionate—demonstrated few differences in interactions. This study highlights the widespread use of traditional medicine within Indonesian civilization for managing health and treating illnesses. However, a critical gap in knowledge persists regarding the safety and efficacy of these medicinal plants. This lack of scientific understanding presents a significant challenge, as it hinders the development of evidence-based practices and raises concerns about potential adverse effects. The example of turmeric and meniran exemplifies this point. Though commonly used in Southeast Asia for its medicinal properties, the precise chemical compounds responsible for its potential health benefits remain unidentified.



Supplementary Figure 1: Calculated RMSD of 1000 times re-docking of the reference ligand

Further research is necessary to elucidate these mechanisms of action and conduct rigorous clinical trials to confirm the safety and efficacy of turmeric and betel leaf, particularly regarding its purported anticancer properties.

Table 7: Tanimoto similarity of structure selected compound from *Piper betle* with reference ligand (pubchem id: 2812)

Pubchem_id	Compound	Tanimoto similarity (%)
5281	octadecanoic acid	2.1
12535	triacontane	0.0
638024	piperine	6.0
222284	(-)-beta-sitosterol	3.7
94577	cepharadione a	8.8
13747834	beta-sitosterol palmitate	3.2
19255	dotriacontanoic acid	2.1
5320621	piperlonguminine	4.2
5280489	beta-carotene	1.6
3032853	cadinene	1.7
5281515	caryophyllene	3.3
2758	cineole	0.0
7127	eugenol-methyl-ether	5.6
493570	riboflavin	5.4
1130	thiamin	4.3

Employed a machine learning approach to predict potential Cytochrome P450 EryK protein inhibitors from natural sources. *t* and betel leaf were investigated using in silico methods to identify promising candidates. Following compound selection from public databases (Knapsack) and fingerprint generation using RDKit, machine learning models were developed and compared. The SVC model emerged as the most effective in predicting active compounds[36].

The molecular docking analysis conducted on compounds derived from turmeric and betel leaf has provided valuable insights into their binding affinities and potential therapeutic applications. The findings aim to

compare these results with existing literature to validate observations and highlight significant differences or similarities.

Binding Affinity Comparisons

It was observed that curcumin and caryophyllene exhibited higher binding affinity values compared to reference ligands, suggesting strong interactions with target proteins. These findings align with previous studies demonstrating the potent anti-inflammatory properties of both compounds[37]. Conversely, certain compounds like arachidic acid from turmeric displayed lower binding affinities but still showed therapeutic benefits through alternative pathways that have not yet been fully explored[39]. Further investigation into conformational dynamics could provide deeper insights into these complex interactions.

promise for therapeutic use due to strong interactions with other targets. Similarly, dotriacontanoic acid demonstrated potent inhibitory effects on inflammatory mediators, underscoring its potential therapeutic applications similar to those observed by Dar et al. [38]

Structural Variations Impact

The varied binding affinities could be explained by structural variations among compound molecules. For instance, beta-stigmasterol was found to display weaker binding compared to curcumin but may offer unique

Table 8: IFP similarity of structure selected compound from *Curcuma longa* with reference ligand (pubchem id: 2812)

Pubchem_id	Compound	IFP similarity (%)
984	Hexadecanal	27
985	Palmitic acid	18
5281	Octadecanoic acid	15
10140	Glycocholic acid	50
10467	Arachidic acid	20
31292	Octadecanamide	15
81573	Nerolidyl propionate	17
91486	Dihydrosphingosine	8
96516	Curcumin	27
124072	1,7-Bis(4-hydroxy-3-methoxyphenyl)-3,5-heptanedione	30
222284	(-)-beta-Sitosterol	36
348482	Gitoxigenin	50
445639	Oleic acid	20
5280450	Linoleic acid	18
5280794	beta-Stigmasterol	23
5280934	alpha-Linolenic acid	18
5312487	8,11-Octadecadienoic acid	20
9848818	Tauroursodeoxycholic acid	44
10429233	Dihydrocurcumin	33
13888134	1,5-Dihydroxy-1,7-bis(4-hydroxy-3-methoxyphenyl)-4,6-heptadien-3-one	33
131752799	7-Hydroxy-1,7-bis(4-hydroxy-3-methoxyphenyl)-1-heptene-3,5-dione	13

Table 9: IFP similarity of structure selected compound from *Piper betle* with reference ligand (pubchem id: 2812)

Pubchem_id	Compound	IFP similarity (%)
1130	thiamin	44
2758	cineole	25
5281	octadecanoic acid	18
7127	eugenol-methyl-ether	38
12535	triacontane	21
19255	dotriacontanoic acid	21
94577	cepharadione a	29
222284	(-)-beta-sitosterol	36
493570	riboflavin	40
638024	piperine	44

3032853	cadinene	50
5280489	beta-carotene	33
5281515	caryophyllene	50
5320621	piperlonguminine	33
13747834	beta-sitosterol palmitate	30

Conclusions

This study highlights the potential of natural products from turmeric and betel leaf as inhibitors of Cytochrome P450 EryK protein, utilizing a combined in silico approach for compound identification through molecular docking and machine learning. Although curcumin and caryophyllene showed strong binding interactions, variations in binding affinities among other compounds underscore the complexity of their mechanisms. Limitations regarding oral bioavailability were noted, along with the need for further investigations in in vitro and in vivo settings to validate efficacy and safety. Overall, this work contributes to the development of evidence-based natural therapies while enhancing our understanding of protein-ligand interactions. The discovery of novel protein targets by curcumin warrants further biochemical characterization. Long-term toxicity assays are recommended to enhance the understanding of safe usage in medical treatments, while also exploring potential synergistic effects when combined with other therapeutic agents like beta-caryophyllene, which showed promising anti-inflammatory properties despite relatively small differences in binding affinity scores (−8.92 kcal/mol for curcumin and −8.17 kcal/mol for caryophyllene).

Conflict of Interest

The authors declare no conflict of interest.

Authors' Declaration

The authors hereby declare that the work presented in this article is original and that any liability for claims relating to the content of this article will be borne by them.

References

1. Liu CX. Overview on development of ASEAN traditional and herbal medicines. Chin Herb Med 2021;13(4):441-450.
2. Park HL, Lee HS, Shin BC, Liu JP, Shang Q, Yamashita H, Lim B. Traditional medicine in China, Korea, and Japan: a brief introduction and comparison. Evidence-Based Complementary and Alternative Medicine. 2012;2012(1):429103.
3. Sumarni W, Sudarmin S, Sumarti SS. The scientification of jamu: A study of Indonesian's traditional medicine. J Phys Conf Ser 2019:1321(3):032057.
4. Ekor M. The growing use of herbal medicines: issues relating to adverse reactions and challenges in monitoring safety. Front Pharm 2014;4:66193.
5. Wahyuningsih R, Adawiyah R, Sjam R, Prihartono J, Ayu Tri Wulandari E, Rozaliyani A, Ronny R, Imran D, Tugiran M, Siagian FE, Denning DW. Serious fungal disease incidence and prevalence in Indonesia. Mycoses 2021;64(10):1203-1212.
6. Ramadan G, Al-Kahtani MA, El-Sayed WM. Anti-inflammatory and anti-oxidant properties of *Curcuma longa* (turmeric) versus Zingiber officinale (ginger) rhizomes in rat adjuvant-induced arthritis. Inflammation 2011;34:291-301.
7. Tomeh MA, Hadianamrei R, Zhao X. A review of curcumin and its derivatives as anticancer agents. Int J Mol Sci 2019;20(5):1033.
8. Abouali N, Moghimipour E, Mahmoudabadi AZ, Namjouyan F, Abbaspoor Z. The effect of curcumin-based and clotrimazole vaginal cream in the treatment of vulvovaginal candidiasis. J Family Med Prim Care 2019;8(12):3920-3924.
9. Varthya SB, Thangaraju P, Venkatesan S. Curcumin and fungal infection–commonly available herbs for common female infection. J Family Med Prim Care 2020;9(2):1272.
10. Sivareddy B, Reginald BA, Sireesha D, Samatha M, Reddy KH, Subrahamanyam G. Antifungal activity of solvent extracts of *Piper belte* and Ocimum sanctum Linn on Candida albicans: An: in vitro: comparative study. J oral Maxillofac Pathol 2019;23(3):333-337.
11. Selvaraj GK, Wilson JJ, Kanagaraj N, Subashini E, Thangavel S. Enhanced antifungal activity of *Piper belte* against candidiasis infection causing Candida albicans and in silico analysis with its virulent protein. Biomed Biotechnol Res J 2022;6(1):73-80.
12. Karo MB, Tambaip T, Hatta M, Simanjuntak TP, Irmawaty L, Rina T, Kamelia E, Rahmawati F, Bintang M. A mini review of Indonesian medicinal plants for Vulvovaginal candidiasis. Rasayan J Chem 2017;10(4):1280-1288.
13. Kumari M, Tiwari N, Chandra S, Subbarao N. Comparative analysis of machine learning based QSAR models and molecular docking studies to screen potential anti-tubercular inhibitors against InhA of mycobacterium tuberculosis. Int J Comput Biol Drug Des 2018;11(3):209-235.
14. Genc B, Tunc HÜ. Optimal training and test sets design for machine learning. Turk J Elec Eng Comp Sci 2019;27(2):1534-1545.
15. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK. Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. J Med Chem 2012;55(14):6582-6594.
16. Dupont S, Lemetais G, Ferreira T, Cayot P, Gervais P, Beney L. Ergosterol biosynthesis: a fungal pathway for life on land?. Evolution 2012;66(9):2961-2968.
17. Jordá T, Puig S. Regulation of ergosterol biosynthesis in Saccharomyces cerevisiae. Genes 2020;11(7):795.
18. Akins RA. An update on antifungal targets and mechanisms of resistance in Candida albicans. Medical mycology. 2005 Jun 1;43(4):285-318.
19. Onyewu C, Blankenship JR, Del Poeta M, Heitman J. Ergosterol biosynthesis inhibitors become fungicidal when combined with calcineurin inhibitors against Candida albicans, Candida glabrata, and Candida krusei. Antimicrob Agents Chemother 2003;47(3):956-964.
20. Montemiglio LC, Gianni S, Vallone B, Savino C. Azole drugs trap cytochrome P450 EryK in alternative conformational states. Biochemistry 2010;49(43):9199-9206.
21. Shinbo Y, Nakamura Y, Altaf-Ul-Amin M, Asahi H, Kurokawa K, Arita M, Saito K, Ohta D, Shibata D, Kanaya S. KNApSAcK: a comprehensive species-metabolite relationship database. Plant metabolomics 2006:165-181.
22. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA, Yu B, Zaslavsky L. PubChem 2023 update. Nucleic Acids Res 2023;51(D1):D1373-D1380.
23. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J. Scikit-learn: Machine learning in Python. J Mach Learn Res. 2011;12:2825-2830.
24. Kluyver T, Ragan-Kelley B, Pérez F, Granger BE, Bussonnier M, Frederic J, Kelley K, Hamrick JB, Grout J, Corlay S, Ivanov

P. Jupyter Notebooks-a publishing format for reproducible computational workflows. Elpub. 2016;2016:87-90.
25. Korb O, Stutzle T, Exner TE. Empirical scoring functions for advanced protein−ligand docking with PLANTS. J Chem Inform Model 2009;49(1):84-96.
26. Korb O, Stützle T, Exner TE. PLANTS: Application of ant colony optimization to structure-based drug design. InInternational workshop on ant colony optimization and swarm intelligence 2006: 247-258. Berlin, Heidelberg: Springer Berlin Heidelberg.
27. Xue LC, Rodrigues JP, Kastritis PL, Bonvin AM, Vangone A. PRODIGY: a web server for predicting the binding affinity of classification scheme in the Toxtree software. SAR QSAR Environ Res 2008;19(5-6):495-524.
31. Damião MC, Pasqualoto KF, Polli MC, Parise Filho R. To be drug or prodrug: structure-property exploratory approach regarding oral bioavailability. J Pharm Pharm Sci 2014;17(4):532-540.
32. Radifar M, Yuniarti N, Istyastono EP. PyPLIF: Python-based protein-ligand interaction fingerprinting. Bioinformation 2013;9(6):325.
33. Istyastono EP, Radifar M, Yuniarti N, Prasasty VD, Mungkasi S. PyPLIF HIPPOS: A molecular interaction fingerprinting tool protein–protein complexes. Bioinformatics 2016;32(23):3676-3678.
28. Vangone A, Schaarschmidt J, Koukos P, Geng C, Citro N, Trellet ME, Xue LC, Bonvin AM. Large-scale prediction of binding affinity in protein–small ligand complexes: The PRODIGY-LIG web server. Bioinformatics. 2019;35(9):1585-1587.
29. Lipinski CA. Drug-like properties and the causes of poor solubility and poor permeability. J Pharmacol Toxicol Methods 2000;44(1):235-249.
30. Patlewicz G, Jeliazkova N, Safford RJ, Worth AP, Aleksiev B. An evaluation of the implementation of the Cramer
for docking results of AutoDock Vina and PLANTS. J Chem Inform Model 2020;60(8):3697-3702.
34. Liu K, Kokubo H. Exploring the stability of ligand binding modes to proteins by molecular dynamics simulations: a cross-docking study. J Chem Inform Model 2017;57(10):2514-22.
35. 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(1):42717.
36. Dara S, Dhamercherla S, Jadav SS, Babu CM, Ahsan MJ. Machine learning in drug discovery: a review. Artif Intell Rev 2022;55(3):1947-1999.