Tropical Journal of Natural Product Research

Original Research Article

Antihypertensive Activity Prediction of *Physalis angulata* L. Through Computational Analysis

Marisca E. Gondokesumo[1]*, Muhammad R. Rasyak[2], Mansur Ibrahim[3]

[1]Faculty of Pharmacy, University of Surabaya, Universitas Surabaya, Surabaya, Indonesia
[2]Eijkman Research Centre for Molecular Biology, National Research and Innovation Agency, Indonesia
[3]University Mega Rezky, Makassar, Indonesia

ARTICLE INFO	ABSTRACT
  	Traditional medicine is widely used in Indonesia to improve health, but research into the safety and efficacy of medicinal plants has not been widely conducted. This study focuses on *Physalis angulata* L., commonly used in Southeast Asia, to explore its potential as an Angiotensin-Converting Enzyme (ACE) and Beta-1 Adrenergic Receptor (ADRB1) protein inhibitor. Utilizing a machine learning approach, compounds from *Physalis angulata* L. were identified and analyzed through *in silico* methods, with the AdaBoost Classifier model proving most effective. The identified active compounds, particularly Physanolide A, demonstrated superior binding affinity compared to reference ligands. Further analysis using Lipinski's rule of five confirmed the compliance and potential efficacy of Physanolide A and other compounds. While some compounds exhibited one violation of Lipinski's rule, they still presented promising binding interactions. This study emphasizes the complexity of protein-ligand interactions and the alternative binding strategies employed by natural compounds. Despite the limitations of *in silico* methods, the findings provide valuable insights into the potential of natural products as drug candidates. The study highlights the need for further *in vitro* and *in vivo* validation to ensure the safety and efficacy of these compounds. By addressing these limitations, this research contributes to the development of evidence-based natural therapies and a deeper understanding of protein-ligand interactions, paving the way for future investigations into the medicinal properties of *Physalis angulata* L. *Keywords*: *Physalis angulata,* Angiotensin-Converting Enzyme inhibitor, *in silico*, Machine Learning, Molecular docking.

Introduction

Hypertension is recognized as a significant public health concern globally, with approximately 1.13 billion people affected worldwide.[1] It is considered a major risk factor for various cardiovascular diseases, including heart attack, stroke, kidney failure, eye damage, and is attributed to a substantial portion of mortality in many countries.[2] Hypertension has damaged the blood vessels, leading to conditions such as atherosclerosis, which increases the risk of heart attacks and strokes.[3] Additionally, kidney damage, known as nephropathy, can be caused by hypertension, which can progress to kidney failure if not treated. Furthermore, eye damage, known as hypertensive retinopathy, can also be caused by hypertension, potentially leading to vision loss and even blindness. The severe consequences of hypertension can significantly impact an individual's quality of life and overall health.[2] Plants have long been utilized in traditional medicine for treating various ailments, including hypertension. *Physalis angulata* L., commonly known as ciplukan, has been used in Indonesian traditional medicine for centuries for a variety of ailments, including body aches, asthma, diabetes, chickenpox, cough, fever, diarrhea, back pain, and hypertension.[4]

*Corresponding author. E mail: marisca@staff.ubaya.ac.id
Tel: +62 878-5136-7988

This plant's extracts have been found to exhibit significant antioxidant and anti-inflammatory activities, which can help mitigate the oxidative stress and inflammation associated with various diseases.[5,6] *P. angulata* L contains active ingredients with medicinal properties. One that has been researched is Physalis D in the seeds, identified for antidiabetic and antihypertensive properties.[7] The efficacy of *Physalis angulata* L. in reducing blood pressure in animal models of hypertension has been demonstrated by several studies. For instance, in a study using L-NAME-induced hypertensive rats, it was found that the plant's ethanol leaf extract significantly reduced systolic and diastolic blood pressures, as well as the percentage of endothelial progenitor cells in the blood.[5] The angiotensin-converting enzyme (ACE) and beta-adrenergic receptor 1 (ADRB1) are two key proteins involved in the regulation of blood pressure. ACE is responsible for the conversion of angiotensin I to angiotensin II, a potent vasoconstrictor that increases blood pressure.[8-9] ADRB1, on the other hand, is a receptor that binds to beta-adrenergic agonists, such as epinephrine and norepinephrine, which stimulate the heart to beat faster and stronger, leading to increased blood pressure.[10-11] Both ACE and ADRB1 play crucial roles in the development and maintenance of hypertension, making them important targets for the treatment of hypertension.

In this study, a machine learning approach was chosen to predict the activity of compounds from *P. angulata* L. based on pharmacological data. Molecular docking was also used to evaluate compound interactions and binding affinities to ACE and ADRB1 receptors. These two methods can be the right initial choice for selecting drug candidates *in silico*.[12]

This study aims to identify potential inhibitors from *P. angulata* L. with potential for ACE and ADRB1 inhibitors using machine-learning approaches and molecular docking. This study makes a novel contribution by analyzing the dual inhibitory potential of ACE and

ADRB1 from *P. angulata* L. using the integration of machine learning and molecular docking, which has not been reported previously. The goal of this research is to obtain candidate compounds with high binding affinity for both protein targets, which provides a basis for the development of natural-based antihypertensive agents.

Materials and Methods

Data mining and fingerprint extraction
Machine learning was employed in this study to predict inhibitors of the proteins ACE and ADRB1 within *P. angulata* L. Keyword searches using *Physalis angulata* L. were conducted within the Knapsack database core system to retrieve compound data for this plant.[13] SMILES structures were extracted from all identified compounds using RDKit v2023.09.1, an open-source Python software package. These structures were used to generate Klekota-Roth fingerprints.[14] Each fingerprint, comprising 4860 substructures, was assigned a binary score of either one or zero. Subsequently, Dudedocking (https://dude.docking.org/) was leveraged to generate a decoy dataset, including known antihypertensive small molecule protein inhibitors commercially available as drugs for antihypertension.[15] A decoy compound was additionally provided by the platform to serve as a non-active control. Before constructing the model, each molecule was converted into a chemical fingerprint representation using RDKit. The datasets containing active molecules (ACE and ADRB1 inhibitors) and inactive molecules were combined and used as the basis for building the machine learning model.

Machine learning model development
The framework for constructing various machine learning models was provided by the Python library Scikit-learn v1.3.0.[16] The development environment for writing and implementing the model code was facilitated by Jupyter Notebooks v7.0.[17] The evaluation metric used to identify the most effective model for this specific dataset was the Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) curve. A convenient tool for automating the evaluation process was provided by Lazypredict v0.2.12 (https://lazypredict.readthedocs.io), which allowed the comparison of twenty-seven different classification models on the dataset. This chosen model was then used to predict potentially active compounds to inhibit ACE and ADRB1 within *P. angulata* L.

Molecular docking of the predicted active compound and interaction analysis
Following the machine learning prediction of potentially active ACE and ADRB1 inhibitor substances within *P. angulata* L., a molecular docking was conducted to assess binding affinity using PLANTS 1.1 software[18,] which employs an innovative "artificial ant colony". Ultimately, the conformation with the lowest energy state is identified.[19] Two scoring functions (PLANTSCHEMPLP and PLANTSPLP) were utilized by PLANTS to evaluate the docked poses, balancing accuracy with computational efficiency.[18] Subsequently, the binding affinity (ΔG) of the docked poses was calculated using PRODIGY software.[20] This program analyzes protein-ligand interactions at the atomic level (https://wenmr.science.uu.nl/prodigy/lig). PRODIGY-LIG offers several key advantages, including user-friendliness, broad applicability to various protein-ligand complexes, and demonstrated effectiveness.[21]

Druglikeness and ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) analysis
Lipinski's rule of five is a commonly used metric for assessing drug-likeness. This rule evaluates four key properties of potential drug candidates: a maximum of five hydrogen bond donors, ten or fewer hydrogen bond acceptors, a molecular weight less than 500 Daltons, and a LogP value no greater than 4.15.[22] Through analysis of these properties, insights into a drug candidate's potential for absorption, distribution, metabolism, and excretion within the body are gained. To evaluate the toxicity of each compound, Toxtree software v3.1 is

employed.[23] The Threshold of Toxicological Concern (TTC) defines a safe exposure level for all substances that are often utilized during toxicity assessments.[23] Human intestinal absorption (HIA) analyses are frequently incorporated alongside toxicity assessments. This analysis is used to predict the number of compounds that can be absorbed by the gastrointestinal tract (GI) and estimate how many compounds can be absorbed from the. This approach assists in the identification of potential drug candidates.[24]

Tanimoto similarity
Following successful molecular docking simulations, the potential ACE and ADRB1 inhibitor compounds identified from *P. angulata* L. underwent further analysis. The fingerprint of each selected compound was then used to calculate similarity structures using Klekota-Roth fingerprinting in RDKit.[14]
PyPLIF v1.4, a Python library, was employed to convert the interaction information obtained from the docking simulations into a format suitable for similarity analysis.[25] This process involved generating a "fingerprint" for each compound using interaction fingerprinting Python (IFP).
Subsequently, another Python library, Pyplif-HIPPOS v1.0, was used to compare the fingerprints of the potential ACE and ADRB1 inhibitor compounds against known reference ligands that are already bound to the protein.[26] A similarity score was calculated for each comparison by Pyplif-HIPPOS. This method enabled the identification of potential ACE and ADRB1 inhibitor compounds that exhibited interaction patterns similar to those of established ACE and ADRB1 inhibitor agents.

Results and Discussion

Data mining and fingerprint extraction
To predict potential ACE and ADRB1 protein inhibitors from natural sources, a machine learning approach was employed. *P. angulata L* was investigated using *in silico* methods to identify promising candidates. 49 compounds were retrieved using the keyword "*Physalis angulata* L". For training and testing purposes, a dataset of active and decoy small molecules was compiled from Dude docking. The entire dataset comprised 898 compounds (298 active compounds and 600 decoy compounds).

Machine learning model development
Evaluation by Lazypredict identified the AdaBoost Classifier model as the top performer for both of ACE and ADRB1 proteins compared to other models considered.[27] This model achieved impressive metrics, including accuracy (0.99), specificity (1.00), AUC/ROC (0.99), and F1 score (0.99) for the ACE protein (Table 1). While accuracy (0.99), specificity (0.99), AUC/ROC (0.99), and F1 score (0.99) for the ADRB1 protein (Table 2). To ensure model reliability, five-fold cross-validation was performed. The data was divided into five parts, and the model was trained and tested alternately on each combination of folds. The validation results showed stable performance, with AdaBoost producing training scores of 1.00 and testing scores of 0.997 for ACE, and training scores of 1.00 and testing scores of 0.981 for ADRB1. The machine learning model identified 21 compounds from *P. angulata L* as potential inhibitors for the ACE protein and 15 compounds for the ADRB1 inhibitor Protein.

Molecular docking of the predicted active compound
Molecular docking was used to assess the binding potential of compounds predicted to inhibit ACE and ADRB1 proteins. The ACE structure (PDB: 3BKL) was used with a protein-bound reference ligand (PubChem 24808493). Meanwhile, the ADRB1 structure (PDB: 7BVQ) was used with a protein-bound reference ligand (PubChem 13023332). Re-docking 1000 times was used first to validate the PLANTS 1.1 software with a known reference ligand (PubChem code: 24808493) that had been bound to the ACE protein (PDB: 3BKL) and the ADRB1 protein (PDB: 7BVQ) with 13023332 bound ligands. The Root Mean Square Deviation (RMSD) of the ligand's position in each result was calculated. A low RMSD (ideally below 2 Å) indicates that the software can accurately reproduce known binding affinity, thus providing confidence in its ability to assess predicted active compounds.[28]

Table 1: The Score of each of the models in Lazypredict for ACE Protein

Model	Accuracy	Balanced Accuracy	ROC AUC	F1 Score	Specificity
AdaBoost Classifier	0.995475	0.993333	0.993333	0.995468	1
Bagging Classifier	0.995475	0.993333	0.993333	0.995468	1
XGB Classifier	0.995475	0.993333	0.993333	0.995468	0.965753
LGBM Classifier	0.995475	0.993333	0.993333	0.995468	0.993151
Extra Trees Classifier	0.99095	0.989909	0.989909	0.99095	1
Ridge Classifier CV	0.99095	0.989909	0.989909	0.99095	1
Random Forest Classifier	0.99095	0.989909	0.989909	0.99095	0.979452
Decision Tree Classifier	0.99095	0.986667	0.986667	0.99092	0.993151
Ridge Classifier	0.986425	0.986484	0.986484	0.986447	0.965753
Calibrated Classifier CV	0.986425	0.983242	0.983242	0.986403	0.890411
Logistic Regression	0.986425	0.983242	0.983242	0.986403	1
SGD Classifier	0.9819	0.979817	0.979817	0.9819	1
Passive Aggressive Classifier	0.977376	0.979635	0.979635	0.97748	0.863014
Linear SVC	0.977376	0.979635	0.979635	0.97748	0.972603
Extra Tree Classifier	0.972851	0.969726	0.969726	0.972851	0.993151
Perceptron	0.968326	0.969543	0.969543	0.968472	0.965753
Bernoulli NB	0.968326	0.969543	0.969543	0.968472	0.986301
Nearest Centroid	0.968326	0.969543	0.969543	0.968472	0.972603
SVC	0.968326	0.956575	0.956575	0.968046	0.965753
KNeighbors Classifier	0.918552	0.931872	0.931872	0.919978	1
NuSVC	0.923077	0.893151	0.893151	0.921067	0.993151
GaussianNB	0.904977	0.87621	0.87621	0.902942	0.986301
Linear Discriminant Analysis	0.828054	0.811507	0.811507	0.828593	0.993151
Quadratic Discriminant Analysis	0.687783	0.54	0.54	0.584639	0.986301
Dummy Classifier	0.660633	0.5	0.5	0.525627	0.993151
Label Spreading	0.660633	0.5	0.5	0.525627	1
Label Propagation	0.660633	0.5	0.5	0.525627	1

Table 2: The Score of each of the models in Lazypredict for ADRB1 Protein

Model	Accuracy	Balanced Accuracy	ROC AUC	F1 Score	Specificity
AdaBoost Classifier	0.994652	0.995902	0.995902	0.994662	0.991803
LGBM Classifier	0.994652	0.992308	0.992308	0.994643	1
Logistic Regression	0.989305	0.984615	0.984615	0.989265	0.983607
Bagging Classifier	0.989305	0.984615	0.984615	0.989265	1
Decision Tree Classifier	0.989305	0.984615	0.984615	0.989265	1
XGB Classifier	0.989305	0.984615	0.984615	0.989265	1
Ridge Classifier CV	0.983957	0.980517	0.980517	0.983928	0.934426
Ridge Classifier	0.983957	0.980517	0.980517	0.983928	1
Passive Aggressive Classifier	0.983957	0.980517	0.980517	0.983928	0.959016
Linear SVC	0.983957	0.980517	0.980517	0.983928	0.786885
Random Forest Classifier	0.983957	0.976923	0.976923	0.983865	1
Extra Trees Classifier	0.983957	0.976923	0.976923	0.983865	1
Calibrated Classifier CV	0.983957	0.976923	0.976923	0.983865	0.827869
Nearest Centroid	0.97861	0.976419	0.976419	0.97861	0.991803

SGD Classifier	0.973262	0.97232	0.97232	0.973309	1
BernoulliNB	0.973262	0.968726	0.968726	0.973213	0.983607
Perceptron	0.957219	0.960025	0.960025	0.957501	1
SVC	0.957219	0.938462	0.938462	0.956496	0.991803
NuSVC	0.930481	0.9	0.9	0.928377	0.95082
Extra Tree Classifier	0.903743	0.89029	0.89029	0.903383	0.942623
KNeighbors Classifier	0.855615	0.88575	0.88575	0.859018	1
Linear Discriminant Analysis	0.855615	0.867781	0.867781	0.858354	0.991803
GaussianNB	0.850267	0.802585	0.802585	0.843379	0.991803
Quadratic Discriminant Analysis	0.770053	0.694388	0.694388	0.749255	0.97541
Label Spreading	0.657754	0.507692	0.507692	0.527375	1
Label Propagation	0.657754	0.507692	0.507692	0.527375	1
Dummy Classifier	0.652406	0.5	0.5	0.515169	1

The successful re-docking of the reference ligand for the ACE protein (performed 1000 times) validated the ability of PLANTS 1.1 software to accurately reproduce known binding affinities. In this process, 91% of the re-docked ligand poses showed RMSD values consistently below or equal to 2.0 Å, compared to the reference pose (Supplementary Figure 1-2), while the ADRB1 protein, 92% of the re-docked ligand poses showed RMSD values consistently below or equal to 2.0 Å. After successful validation, the active compounds were simulated by molecular docking. The docking scores and binding affinity scores of selected compounds from *P. angulata* L. are show in Tables 3 and 4. Several compounds selected from the machine learning process for both plants exhibited better binding interactions compared to the reference ligand. This result further supports the effectiveness of the model developed to predict active compounds from *P. angulata* L.

Table 3: Docking and binding affinity score (ΔG) for the predicted compounds from *Physalis angulata* L. as an inhibitor of the ACE Protein

| Compound	PubChem_ID	Binding affinity (ΔG) kcal/mol
Physagulin H	131752708	-14.01
Physalin K	102254401	-12.87
Physalin E	10009993	-12.85
Physalin H	155551	-12.61
Physagulin G	73816389	-12.39
Physagulin E	10100412	-12.36
Physagulin F	16757390	-12.26
Physagulin A	42620981	-12.24
Physagulin B	431071	-12.24
Withangulatin C	16082069	-12.19
Physanolide A	13962952	-12.17
Physalin V	11613161	-12.11
Physalin G	85110396	-12.06
Physalin D	74820162	-12.03
Physalin B	101273381	-11.85
Dihydrowithanolide E	131751517	-11.83
Withangulatin I	147647	-11.75
Withangulatin A	24814495	-11.74
Physagulin N	102254402	-11.69
Physagulin M	102257049	-11.69
Physagulin D	85083174	-11.59
Reference ligand	24808493	-11.83

Table 4: Docking and binding affinity score (ΔG) for the predicted compounds from *Physalis angulata* L. as an inhibitor of ADRB1 protein

Compound	PubChem_ID	Binding affinity (ΔG) Kcal/mol
Physagulin K	131752708	-12.57
Physagulin I	10009993	-12.34
Physagulin G	10100412	-12.1
Physagulin E	44426450	-11.69
Physagulin F	147647	-11.57
Physagulin B	102254401	-11.51
Physagulin C	85083174	-11.47
Physanolide A	16082069	-11.44
Physalin V	44426449	-11.42
Physalin U	85155112	-11.39
Physalin J	74820162	-11.37
Withangulatin A	102254402	-11.33
Physagulin N	85470089	-11.32
Physagulin L	16082070	-11.3
Physagulin D	101273380	-11.25
Reference ligand	13023332	-9.26

Thirteen compounds from *P. angulata* L. displayed more favorable binding affinity scores: Physalin H, Physagulin G, Physagulin E, Physagulin F, Physagulin A, Physagulin B, Withangulatin C, Physanolide A, Physalin V, Physalin G, Physalin D, Physalin B, and Dihydrowithanolide E compared to the reference ligand of ACE protein inhibitor (PubChem: 24808493) (Table 3). Similarly, fifteen compounds selected from *P. angulata* L. using machine learning showed lower binding affinity compared to the reference ligand: Physagulin K, Physagulin I, Physagulin G, Physagulin E, Physagulin F, Physagulin B, Physagulin C, Physanolide A, Physalin V, Physalin U, Physalin J, Withangulatin A, Physagulin N, Physagulin L, and Physagulin D compared to reference ligand of ADRB1 protein inhibitor (PubChem: 13023332) (Table 4).

The symbol ΔG represents the change in free energy during a molecular interaction. In docking, it specifically refers to the difference in free energy between the unbound state, where the ligand and protein are separate, and the bound state, where they form a complex. A negative ΔG value indicates that the bound state is more energetically favorable compared to the unbound state, signifying the release of energy upon ligand binding to the protein, which leads to a more stable interaction.[28] Binding affinity is defined as the strength of the interaction between a ligand and a protein, reflecting the tightness of their binding. It is measured in energy units, typically kilocalories per mole (kcal/mol). A lower binding affinity value indicates a stronger attraction between the ligand and the protein.[29] A lower binding affinity score was exhibited by these compounds compared to the reference ligand for *P. angulata* L. This finding highlights the effectiveness of the model in selecting promising candidates.
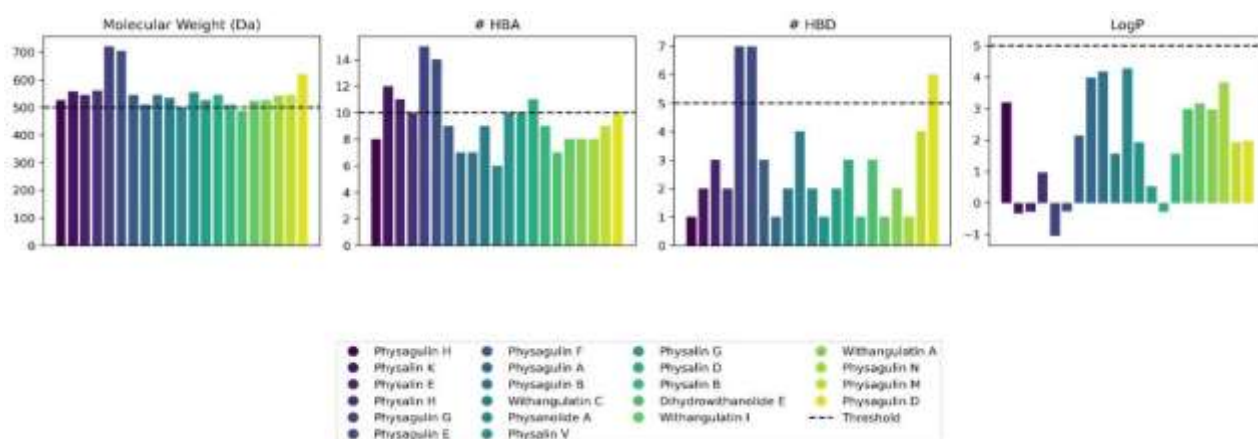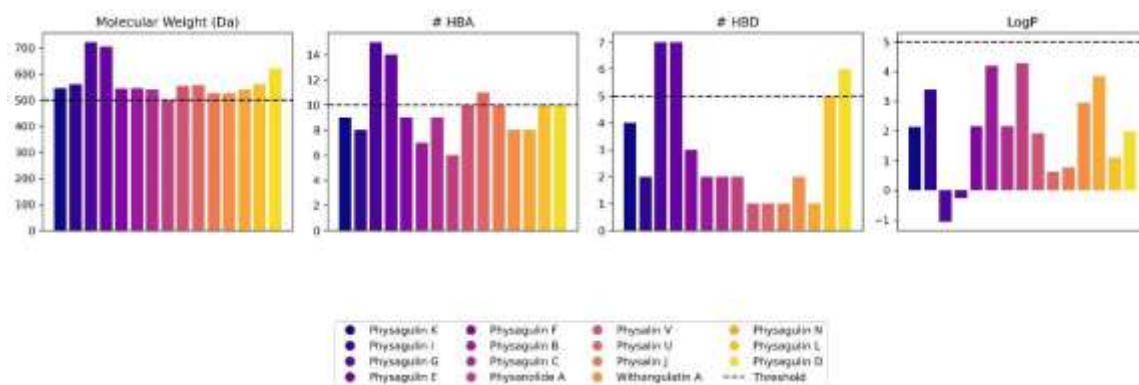

Figure 1: Lipinski rule of 5 for compounds from *Physalis angulata* L. as ACE protein inhibitor

Druglikeness and Absorption, Distribution, Metabolism, Excretion, and Toxicity Analysis
Analysis using Lipinski's rule of five identified the following compounds from *P. angulata* L. as compliant: Dihydrowithanolide E, and Physanolide A had passed Lipinski's rule (Figure 1) for ACE protein inhibitor and have better binding affinity compared to the reference ligand. Conversely, Physanolide A is the only selected compound for ADRB1 that has passed the Lipinski rule of five and has better binding affinity compared to its reference ligand (Figure 2).[22]

Figure 2: Lipinski rule of 5 for compounds from *Physalis angulata* L. as ADRB1 protein inhibitor

However, several drugs that are widely used might have one violation of the Lipinski rule. Interestingly, there are thirteen compounds from selected compound for ACE inhibitor that have one violation of lipinski rule including: Physalin B, Physagulin A, Withangulatin I, Physalin G, Physagulin H, Withangulatin A, Withangulatin C, Physagulin N, Physagulin F, Physagulin M, Physagulin B, Physalin V, and Physalin H. Over thirteen compounds, nine compounds including: Physalin B, Physagulin A, Physalin G, Physagulin H, Withangulatin C, Physagulin F, Physagulin B, Physalin V, and Physalin H have better lower binding affinity means better potential binding compared to reference ligand. While for ADRB1 inhibitor, ten compounds were identified that have one violation, including: Physalin J, Withangulatin A, Physagulin C, Physagulin N, Physagulin F, Physagulin B, Physagulin K, Physalin V, Physagulin I, Physagulin L., and all compounds have better binding affinity compared to the reference ligand.

This criterion was found important because compounds that adhere to the rule generally have better chances of being absorbed by the body when taken orally. This is because they possess characteristics that allow them to pass through the digestive system and reach the bloodstream effectively. By identifying compounds that might have issues with oral bioavailability early on, researchers can focus their resources on those more likely to succeed. This helps streamline the drug development process and avoids wasting time and money on candidates with potential absorption problems.[22,30]

In *P. angulata* L., Physagulin B, Physanolide A, Physalin V, and Physalin B were assessed for their potential for gastrointestinal (GI) absorption for an ACE inhibitor (Table 5). While in ADRB1 inhibitors, including Physagulin B, Physalin V, and Physanolide A (Table 6). These compounds were predicted to exhibit high GI absorption according to Daina et al. (2017). High GI absorption indicates good oral bioavailability, meaning the drugs can be efficiently absorbed from the gut and enter the bloodstream.[31] Following structural analysis using protox rules, Physagulin B, Physanolide A, Physalin V, and Physalin B were classified as class 4 and 5 toxicity, where class IV: harmful if swallowed ($300 < LD50 \leq 2000$), class V: may be harmful if swallowed ($2000 < LD50 \leq 5000$) for ACE protein inhibitor from *P. angulata* L.[32] While for ADRB1 inhibitor Physagulin B, Physalin V, and Physanolide A were classified in class IV in protox.

Furthermore, analysis with Toxtree's carcinogenicity and mutagenicity modules, which utilize discriminant analysis and structural rules, yielded negative results for all compounds selected mentioned above. Physagulin B, Physanolide A, Physalin V, and Physalin B for ACE inhibitors. Physagulin B, Physalin V, and Physanolide A as ADRB1 inhibitors.

Table 5: ADME and Toxicity prediction of predicted ligand inhibitor for ACE protein from *Physalis angulata* L. that maximum have one lipinski violations

Pubchem_ID	Compound Name	GI Absorption	Ld50	Toxicity Class	Carcinogenicity
101273381	Physagulin H	High	32	2	Alert for genotoxic carcinogenicity
155551	Physalin H	Low	2050	5	Negative
74820162	Physagulin F	High	7	2	Negative
85110396	Physagulin A	High	400	4	Alert for genotoxic carcinogenicity
85083174	Physagulin B	High	400	4	Negative
16757390	Withangulatin C	Low	7	2	Alert for genotoxic carcinogenicity
102254401	Physanolide A	High	1330	4	Negative
102254402	Physalin V	High	590	4	Negative
42620981	Physalin G	Low	2050	5	Negative
11613161	Physalin B	High	2050	5	Negative
131751517	Dihydrowithanolide E	High	7	2	Alert for genotoxic carcinogenicity
24814495	Withangulatin I	High	34	2	Alert for genotoxic carcinogenicity

147647	Withangulatin A	High	400	4	Alert for genotoxic carcinogenicity
16082069	Physagulin F	High	7	2	Alert for genotoxic carcinogenicity
102257049	Physagulin M	Low	400	4	Negative

Table 6: ADME and Toxicity prediction of predicted ligand inhibitor for ADRB1 protein from *Physalis angulata* L. that maximum have one Lipinski violations

Pubchem_ID	Compound Name	GI Absorption	Ld50	Toxicity Class	Carcinogenicity
85083174	Physagulin B	High	400	4	Negative
85155112	Physagulin C	High	34	2	Alert for genotoxic carcinogenicity
74820162	Physagulin F	Low	7	2	Alert for genotoxic carcinogenicity
101273380	Physagulin I	High	7	2	Alert for genotoxic carcinogenicity
85470089	Physagulin K	Low	400	4	Negative
16082070	Physagulin L	Low	400	4	Negative
16082069	Physagulin N	High	7	2	Alert for genotoxic carcinogenicity
44426449	Physalin J	High	2050	5	Alert for genotoxic
102254402	Physalin V	High	590	4	Negative
102254401	Physanolide A	High	1330	4	Negative
147647	Withangulatin A	High	400	4	Alert for genotoxic

Tanimoto similarity

Some ligands exhibited seemingly favorable profiles with a higher number of the strongest interaction types (hydrogen bonds and cationic bonds), while natural compounds exhibited interesting alternative binding strategies. These findings emphasize the complexity of protein-ligand interactions. Natural ligands exhibit potential trade-offs, with fewer hydrogen bonds and cationic interactions. Furthermore, several compounds from *P. angulata* L. exhibited aromatic interactions. Tanimoto similarity calculations using the Klekota-Roth fingerprinting tool in the RDkit package identified seven compounds in *P. angulata* L. with structural fingerprint similarities higher than 20% to the reference ligand (Pubchem ID: 24808493) (Table 7-8). The percentage structural fingerprint similarities for Physagulin B, Physanolide A, Physalin V, and Physalin B were 33%, 24%, 36%, and 36%, respectively. The highest similarity for compounds that passed Lipinski, ADME, and toxicity predictions was Physalin V and Physysalin B (36% structural similarity). Meanwhile, for the ADRB1 protein that passed Lipinski, ADME, and toxicity predictions, Physagulin B, Physalin V, and Physanolide A had structural similarities of 20%, 18%, and 22% to the reference ligand (pubchem_id:13023332).

Table 7: Tanimoto similarity of the structure of selected compounds from *Physalis angulata* L. with reference ligand (PubChem ID: 24808493) as the ACE protein inhibitor

Compound	PubChem_ID	Tanimoto similarity
Physagulin H	101273381	30%
Physalin H	155551	36%
Physagulin F	74820162	30%
Physagulin A	85110396	30%
Physagulin B	85083174	33%
Withangulatin C	16757390	29%
Physanolide A	102254401	24%
Physalin V	102254402	36%
Physalin G	42620981	36%
Physalin B	11613161	36%
Dihydrowithanolide E	131751517	28%
Withangulatin I	24814495	32%
Withangulatin A	147647	31%
Physagulin F	16082069	30%
Physagulin M	102257049	30%

Table 8: Tanimoto similarity of the structure of the selected compound from *Physalis angulata* L. with reference ligand (PubChem ID: 13023332) as an ADRB1 protein inhibitor

Compound	PubChem_ID	Tanimoto similarity
Physagulin K	85470089	20%
Physagulin I	101273380	21%
Physagulin F	74820162	20%
Physagulin B	85083174	20%
Physagulin C	85155112	21%
Physalin V	102254401	18%
Physanolide A	102254402	22%
Physalin J	44426449	22%
Withangulatin A	147647	20%
Physagulin N	16082069	20%
Physagulin L	16082070	20%

Although the structural similarity of natural compounds compared to the reference ligand showed lower similarity, analysis of amino acid residues involved in compound binding (IFP similarity) showed that the highest IFP similarity of the four compounds that passed Lipinski, ADME, and toxicity tests was Physanolide A. A total of 28% of the compound's amino acids showed similarity to the amino acid binding site of the reference ligand for ACE inhibitors (Table 9). Meanwhile, for the ADRB1 protein, the closest similarity is to Physanolide A, with a similarity of 36% (Table 10). This also demonstrates the potential of Physanolide A from *P. angulata* L. to inhibit ACE and ADRB1 proteins. Other compounds that passed the Lipinski test, ADME, and toxicity tests for ACE inhibitors from *P. angulata* L. were Physagulin B, Physalin V, and Physalin B, which showed interaction similarities of 20%, 12%, and 23%, respectively. Meanwhile, for ADRB1 inhibitors, Physagulin B and Physalin V showed interaction similarities of 20% and 18%, respectively. The potential of Physanolide A, along with other compounds, to inhibit ACE and ADRB1 proteins is emphasized by these findings.

Table 9: IFP similarity of the structure of the selected compound from *Physalis angulata* L. with reference ligand (PubChem ID: 24808493) as an ACE protein inhibitor

Pubchem_ID	Compound	IFP similarity (%)
101273381	Physagulin H	25%
102254401	Physanolide A	28%
102254402	Physalin V	12%
102257049	Physagulin M	13%
11613161	Physalin B	23%
131751517	Dihydrowithanolide E	33%
147647	Withangulatin A	9%
155551	Physalin H	24%
16082069	Physagulin F	24%
16757390	Withangulatin C	17%
24814495	Withangulatin I	10%
42620981	Physalin G	19%
74820162	Physagulin F	31%
85083174	Physagulin B	20%
85110396	Physagulin A	21%

Table 10: IFP similarity of the structure of the selected compound from *Physalis angulata* L. with reference ligand (PubChem ID: 24808493) as a ADRB1 protein inhibitor

Pubchem_ID	Compound	IFP similarity (%)
85470089	Physagulin K	25%
101273380	Physagulin I	29%
74820162	Physagulin F	39%
85083174	Physagulin B	16%
85155112	Physagulin C	39%
102254401	Physalin V	23%
102254402	Physanolide A	36%
44426449	Physalin J	29%
147647	Withangulatin A	23%
16082069	Physagulin N	31%
16082070	Physagulin L	25%

The importance of considering a broader spectrum of intermolecular interactions when evaluating protein-ligand binding is highlighted by this study. A deeper understanding of the interplay between different types of interactions and their contribution to binding affinity is needed. However, the reliance on *in silico* methods (machine learning and molecular docking) in this study is acknowledged as a limitation. While these methods provide valuable insights, they cannot perfectly replicate real-world biological systems. Further research through *in vitro and in vivo* studies is needed to validate the predicted activity and safety profiles of these compounds. Despite these limitations, valuable insights have been offered. Natural products may possess alternative binding strategies compared to traditional drug candidates, making a broader understanding of the various intermolecular interactions crucial for evaluating protein-ligand binding. This study paves the way for further research on natural products from *P. angulata* L. By addressing these limitations and conducting further research, this research may contribute to the development of evidence-based natural therapies and a more comprehensive understanding of protein-ligand interactions.

Conclusion

This study underscores the potential of *P. angulata* L. in providing antihypertensive treatments. The use of machine learning and *in silico* methods facilitated the identification of promising ACE and ADRB1 protein inhibitors, with Physanolide A emerging as a particularly promising candidate. This work not only contributes to the understanding of protein-ligand interactions but also paves the way for the development of evidence-based natural therapies, emphasizing the need for a more nuanced approach in evaluating the medicinal potential of natural products. Despite the limitations of these methods in fully replicating biological systems, the findings reveal alternative binding strategies of natural products and highlight the importance of considering a wider range of intermolecular interactions. Further validation through *in vitro* and *in vivo* studies is essential to confirm the safety and efficacy of these compounds.

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. Mills KT, Stefanescu A, He J. The global epidemiology of hypertension. Nat Rev Nephrol. 2020; 16:223–237.
2. Kario K, Okura A, Hoshide S, Mogi M. The WHO Global report 2023 on hypertension warning the emerging hypertension burden in globe and its treatment strategy. Hypertens Res. 2024; 47:1099–1102.
3. Chen Y, ZF Zhou, JM Han, X Jin, ZF Dong, L Liu, D Wang, TB Ye, BS Yang, YP Zhang, CX Shen. Patients with comorbid coronary artery disease and hypertension. Ann Transl Med. 2022; 10(7):97700.
4. Ambarwati A, Pujiati E. Antihypertensive Effect of Ciplukan (Physalis angulata L): A Literature Review. Menara J Health Sci. 2024; 3:1–9.
5. Nugrahenny D, Permatasari N, Soeharto S, Rahayu ID, Widodo E, Mintaroem K, Sujuti H, Ratnawati R, Mayangsari E, Irnandi DF, Ardani LS, Wardoyo TA, Claresta NB, Yoga NFC, Anindita AC, Siarnu NUS, Alia FP. Physalis angulata Leaf Ethanol Extract Reduces Oxidative Stress and Improves Endothelial Progenitor Cells in L-NAME-Induced Hypertensive Rats. Hayati J Biosci. 2023; 30:81–87.
6. Wiraswati HL, Ekawardhani S, Rohmawaty E, Laelalugina A, Zuhrotun A, Hendriani R, Wardhana YW, Bestari MB, Sahirdjan EH, Dewi S. Antioxidant, antiinflammation, and antifibrotic activity of ciplukan (Physalis angulata L). extract. J Inflam Res. 2024;6297-6306.
7. Novitasari A, Rohmawaty E, Rosdianto AM. Physalis angulata Linn. as a medicinal plant (Review). Biomed Rep. 2024; 20:47.
8. Herman LL, Padala SA, Ahmed I, Bashir K. Angiotensin-Converting Enzyme Inhibitors (ACEI). In: StatPearls. Treasure Island (FL): StatPearls Publishing; 2024.
9. Ahmad H, Khan H, Haque S, Ahmad S, Srivastava N, Khan A. Angiotensin-Converting Enzyme and Hypertension: A Systemic Analysis of Various ACE Inhibitors, Their Side Effects, and Bioactive Peptides as a Putative Therapy for Hypertension. J Renin-Angiotensin-Aldosterone Syst JRAAS. 2023;7890188.
10. Kelley EF, Snyder EM, Johnson BD. Influence of Beta-1 adrenergic receptor genotype on cardiovascular response to exercise in healthy subjects. Cardiol Res. 2018; 9(6):343.
11. Botzer A, Finkelstein Y, Unger R. Blood Pressure Regulation Evolved from Basic Homeostatic Components. Biomed. 2021; 9:469.
12. Dwira S, Tedjo A, Dharmawan MA, Erlina L, Fadilah F. Differentially Expressed Genes, and Molecular Docking and Dynamic Analysis Revealing the Potential of Compounds in *Zingiber officinale* Roscoe as Inhibitors of TP53-regulating Kinase (TP53RK) that Influence the p53 Signaling Pathway Related to Apoptosis and Cell Cycle. Trop J Nat Prod Res. 2024;8(8).
13. Mishra N, Tripathi S, Mishra N. Global insights into *Physalis* research: A bibliometric journey from origin to recent advances. Food Human. 2024; 3:100395.
14. Klekota J, Roth FP. Chemical substructures that enrich for biological activity. Bioinforma Oxf Engl. 2008; 24:2518–2525.
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; 55:6582–6594.
16. Raschka S, Patterson J, Nolet C. Machine learning in Python: main developments and technology trends in data science, machine learning, and artificial intelligence. Information. 2020; 11(4):193.
17. Kluyver T, Ragan-Kelley B, Pérez F, Granger B, Bussonnier M, Frederic J, Kelley K, Hamrick J, Grout J, Corlay S, Ivanov P, Avila D, Abdalla S, Willing C. Jupyter Notebooks – a publishing format for reproducible computational workflows. In: Positioning and Power in Academic Publishing: Players, Agents and Agendas. Berlin: IOS Press; 2016. p. 87–90.
18. Korb O, Stützle T, Exner TE. Empirical Scoring Functions for Advanced Protein−Ligand Docking with PLANTS. J Chem Inf Model. 2009; 49:84–96.
19. Korb O, Stützle T, Exner TE. PLANTS: Application of Ant Colony Optimization to Structure-Based Drug Design. In: Dorigo M, Gambardella LM, Birattari M, Martinoli A, Poli R, Stützle T, editors. Ant Colony Optimization and Swarm Intelligence, Lecture Notes in Computer Science. Berlin: Springer; 2006. p. 247–257.
20. Xue LC, Rodrigues JP, Kastritis PL, Bonvin AM, Vangone A. PRODIGY: a web server for predicting the binding affinity of protein–protein complexes. Bioinformatics. 2016; 32:3676–3678.
21. 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. Bioinforma Oxf Engl. 2019; 35:1585–1587.

22. Lipinski CA. Drug-like properties and the causes of poor solubility and poor permeability. J Pharmacol Toxicol Methods. 2000; 44:235–249.

23. Patlewicz G, Jeliazkova N, Safford RJ, Worth AP, Aleksiev B. An evaluation of the implementation of the Cramer classification scheme in the Toxtree software. SAR QSAR Environ Res. 2008; 19:495–524.

24. Damião MCFCB, Pasqualoto KFM, Polli MC, Parise Filho R. To be drug or prodrug: structure-property exploratory approach regarding oral bioavailability. J Pharm Pharm Sci Publ Can Soc Pharm Sci Soc Can Sci Pharm. 2014; 17:532–540.

25. Radifar M, Yuniarti N, Istyastono EP. PyPLIF: Python-based Protein-Ligand Interaction Fingerprinting. Bioinformation. 2013; 9:325–328.

26. Istyastono EP, Radifar M, Yuniarti N, Prasasty VD, Mungkasi S. PyPLIF HIPPOS: A Molecular Interaction Fingerprinting Tool for Docking Results of AutoDock Vina and PLANTS. J Chem Inf Model. 2020; 60:3697–3702.

27. Dara S, Dhamercherla S, Jadav SS, Babu CM, Ahsan MJ. Machine Learning in Drug Discovery: A Review. Artif Intell Rev. 2022; 55:1947–1999.

28. Liu K, Kokubo H. Exploring the Stability of Ligand Binding Modes to Proteins by Molecular Dynamics Simulations: A Cross-docking Study. J Chem Inf Model. 2017; 57:2514–2522.

29. Kontoyianni M. Docking and Virtual Screening in Drug Discovery. Methods Mol Biol Clifton NJ. 2017; 1647:255–266.

30. Kar S, Leszczynski J. Open access *in silico* tools to predict the ADMET profiling of drug candidates. Expert opin drug discov. 2020; 15(12):1473-1487.

31. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci. Rep. 2017; 7:42717.

32. Banerjee P, Kemmler E, Dunkel M, Preissner R. ProTox 3.0: a webserver for the prediction of toxicity of chemicals. Nucleic Acids Res. 2024; 52(W1):W513-20.