© 2025 Journal of Pharmacy & Pharmacognosy Research, 13 (3), 716-728, 2025
ISSN 0719-4250
https://jppres.com

DOI: https://doi.org/10.56499/jppres24.2104_13.3.716

Original Article

In silico discovery of potential sodium-glucose cotransporter-2 (SGLT-2) inhibitors from *Smallanthus sonchifolius* (Poepp.) H.Rob. via molecular docking and molecular dynamics simulation approach

[Descubrimiento *in silico* de potenciales inhibidores del cotransportador de sodio-glucosa-2 (SGLT-2) de *Smallanthus sonchifolius* Poepp.) H.Rob. mediante un enfoque de acoplamiento molecular y simulación de dinámica molecular]

Andri Prasetiyo[1*], Esti Mumpuni[1], Dewi Luthfiana[2,3], Rina Herowati[4], Galih Satrio Putra[5]

[1]Faculty of Pharmacy, Pancasila University, Jakarta, Indonesia.
[2]Master Student at Graduate School of Bioagricultural Sciences, Department of Applied Biosciences, Nagoya University, Japan.
[3]Bioinformatics Research Center, Indonesian Institute of Bioinformatics (INBIO), Malang, Indonesia.
[4]Faculty of Pharmacy, Setia Budi University, Surakarta, Indonesia.
[5]Department of Chemistry, Faculty of Mathematics and Natural Sciences, State University of Malang, Indonesia.
*E-mail: andriprasetiyo@univpancasila.ac.id

Abstract

Context: *Smallanthus sonchifolius* or yacon has been reported to have hypoglycemic effects. Inhibition of sodium-glucose cotransporter 2 (SGLT-2) is one of the mechanisms of antidiabetic drugs.

Aims: To evaluate potential compounds in *S. sonchifolius* as SGLT-2 inhibitors using molecular docking analysis and dynamic simulation, as well as predict the pharmacokinetic profiles of the potent compounds using ADMETlab 3.0.

Methods: Fourteen molecules from *S. sonchifolius* were docked with SGLT-2 (PDB Code 7VSI) simultaneously as empagliflozin and canagliflozin as standard ligands. Three compounds, they are 3,5-dicaffeoylquinic acid (3,5-DCQA); and 3,4-dicaffeoylquinic acid (3,4-DCQA), 1-kestose, showed lower rerank scores than empagliflozin and canagliflozin. The analysis proceeded to molecular dynamic simulation by investigating the RMSD and binding energy of SGLT-2 protein in complexes with the ligands to evaluate the dynamic behavior of protein-ligand complexes.

Results: All ligands illustrate similar patterns with two ligand controls, suggesting its acceptable movement from the binding site of SGLT-2. 1-Kestose exhibited stronger binding affinity with SGLT-2 during the period of simulation and continued to interact with the binding pocket of SGLT-2 until the end of the simulation, which involved some critical residues, including Phe98, Glu99, and Ser287, and mainly contributed to hydrogen bond interactions. However, 1-kestose was predicted to have low gastrointestinal absorption due to many polar substituents, while 3,4-DCQA and 3,5-DCQA have higher gastrointestinal absorption.

Conclusions: One phytochemical present in *S. sonchifolius* is 1-kestose, which is thought to have action as an SGLT-2 inhibitor for the treatment of diabetes based on the findings of molecular docking and molecular dynamics investigations. This compound falls into BCS Class III, which means it has good solubility but poor permeability.

Keywords: docking; molecular dynamics; SGLT-2; *Smallanthus sonchifolius*; yacon.

Resumen

Contexto: Se ha informado que *Smallanthus sonchifolius* o yacón tiene efectos hipoglucémicos. La inhibición del cotransportador de sodio-glucosa 2 (SGLT-2) es uno de los mecanismos de los fármacos antidiabéticos.

Objetivos: Evaluar compuestos potenciales en *S. sonchifolius* como inhibidores de SGLT-2 mediante análisis de acoplamiento molecular y simulación dinámica, así como predecir los perfiles farmacocinéticos de los compuestos potentes utilizando ADMETlab 3.0.

Métodos: Catorce moléculas de S. sonchifolius se acoplaron con SGLT-2 (código PDB 7VSI) simultáneamente como empagliflozina y canagliflozina como ligandos estándar. Tres compuestos, son ácido 3,5-dicafeoilquínico (3,5-DCQA); y ácido 3,4-dicafeoilquínico (3,4-DCQA), 1-kestosa, mostraron puntuaciones de rerank más bajas que empagliflozina y canagliflozina. El análisis procedió a la simulación dinámica molecular mediante la investigación de la RMSD y la energía de unión de la proteína SGLT-2 en complejos con los ligandos para evaluar el comportamiento dinámico de los complejos proteína-ligando.

Resultados: Todos los ligandos ilustran patrones similares con dos controles de ligando, lo que sugiere su movimiento aceptable desde el sitio de unión de SGLT-2. La 1-kestosa exhibió una afinidad de unión más fuerte con SGLT-2 durante el período de simulación y continuó interactuando con el bolsillo de unión de SGLT-2 hasta el final de la simulación, lo que involucró algunos residuos críticos, incluidos Phe98, Glu99 y Ser287, y contribuyó principalmente a las interacciones de enlaces de hidrógeno. Sin embargo, se predijo que la 1-kestosa tendría una baja absorción gastrointestinal debido a muchos sustituyentes polares, mientras que 3,4-DCQA y 3,5-DCQA tienen una mayor absorción gastrointestinal.

Conclusiones: Un fitoquímico presente en S. sonchifolius es la 1-kestosa, que se cree que tiene acción como inhibidor de SGLT-2 para el tratamiento de la diabetes según los hallazgos de las investigaciones de acoplamiento molecular y dinámica molecular. Este compuesto se encuentra en la clase III de BCS, lo que significa que tiene buena solubilidad pero poca permeabilidad.

Palabras Clave: acoplamiento; dinámica molecular; SGLT-2; *Smallanthus sonchifolius*; yacón.

ARTICLE INFO
Received: May 27, 2024.
Accepted: November 25, 2024.
Available Online: December 14, 2024.

AUTHOR INFO
ORCID: 0000-0002-9936-6058 (AP) 0009-0004-6423-7979 (RH)
0000-0002-9208-8040 (EM) 0000-0003-4192-3621 (GSP)
0000-0001-6109-5283 (DL)

INTRODUCTION

The prevalence of diabetes, a serious health problem, is frightening. Today, diabetes affects more than 500 million individuals around the globe (International Diabetes Federation, 2021). Diabetes mellitus is a long-term disease characterized by high blood sugar levels in the body, where the body is unable to produce enough insulin or the body is unable to use insulin effectively (Ramani et al., 2022). The most common oral antidiabetic drugs are sulfonylureas, meglitinides, biguanides, thiazolidinediones, alpha-glucosidase inhibitors, dipeptidyl peptidase-IV (DPP-IV) inhibitors, and sodium-glucose cotransporter 2 (SGLT-2) inhibitors. Currently, SGLT2 inhibitors are being widely developed to increase the level of patient compliance (Choi, 2016)

In 1988, SGLT-2 was identified using homology screening (Santer and Calado, 2010; Vallon and Thomson, 2017). According to reports, SGLT-2 mediates over 90% of renal glucose reabsorption (Hummel et al., 2011). SGLT-2 inhibitors work as antidiabetics by preventing glucose reabsorption in the proximal convoluted tubule so that blood sugar levels decrease, and by inhibiting the SGLT-2 protein, it also promotes renal glucose excretion (Abdul-Ghani et al., 2011). For the treatment and control of diabetes, there are numerous therapeutic and targeted techniques available (Nauck et al., 2021), one of which is the suppression of glucose reabsorption in the kidney via SGLT-2, a fresh approach to helping type 2 diabetics lower their blood sugar. When treating T2DM, SGLT-2 inhibitors are a great option because they reduce blood glucose levels without impairing insulin production (Miller and Shubrook, 2015)

The first antidiabetic drug with an SGLT2 inhibitor mechanism of action discovered and approved by the FDA was canagliflozin, followed by dapagliflozin, empagliflozin, and ertugliflozin (Dipiro et al., 2020). The initial concept for these drugs was to exploit how the kidneys metabolize glucose by preventing SGLT-2 from taking up glucose from the proximal renal tubule. This would promote glycosuria and regulate hyperglycemia. However, it is now recognized that these insulin-independent, antihyperglycemic drugs are pleiotropic substances with important advantages for the kidneys, heart, and metabolism (Toyama et al., 2019). SGLT-2 inhibitors reduce blood pressure, body weight, fasting and postprandial plasma glucose levels, HbA1c, and a number of cardiovascular and renal consequences without increasing the risk of hypoglycemia (DeFronzo et al., 2017).

Yacon is a common name of *Smallanthus sonchifolius* (Poepp.) H.Rob. (family *Asteraceae*) and has been reported to have antidiabetic activity or hypoglycemic effects (Contreras-Puentes and Alvíz-Amador, 2021). Numerous chemical constituents of *S. sonchifolius* have been found and classified according to different plant parts. 3,5-DCQA: *S. sonchifolius* leaves contain a number of compounds with anti-diabetic properties, including 3,4-DCQA, 1-kestose, rosmarinic acid, enhydrin, kaempferide, ferulic acid, nystose, quercetin, chlorogenic acid, caffeic acid, p-coumaric acid, protocatechuic acid, and gallic acid (Aziz et al., 2021; Contreras-Puentes and Alvíz-Amador, 2021; Pahlawan and Dwita, 2016).

Computational engineering-based drug design (CADD) is one approach to designing drugs using computer devices (Yu and Mackerel, 2017). Structure-based drug design is the most powerful technique for discovering guide compounds to become drug candidates (Batool et al., 2019). This technology's advantages are its affordability, speed, and effectiveness (Lavecchia and Giovanni, 2013). Computational techniques (molecular docking and molecular dynamic simulation) can help with drug design in several ways.

Phytochemical isolates from *S. sonchifolius* leaves have been found and reported to have antidiabetic activity, but the mechanism of action is still unclear. In this research, we will use a computational approach to search for the most potential compounds found in *S. sonchifolius* as SGLT2 inhibitors and explore their pharmacokinetic and toxicity parameters.

MATERIAL AND METHODS

Preparation of protein

SGLT-2 inhibitors' three-dimensional protein structures were discovered in the Protein Data Bank (7VSI), which may be accessed directly at the URL (http://www.rscb.org/pdb), as per the 2.95Å resolution after the PDBs were imported into the MVD (Molegro Virtual Docker version 6.0.0 software), water molecules were manually eliminated additionally. The missing charges, protonation states, and assignment of polar hydrogen were carried out using a particular Molegro algorithm.

Selection and detection of cavities

A built-in component of the MVD v.6.0.0 software is the cavity discovery approach, which automatically locates potential binding sites (cavities) inside the 15 × 15 × 15 $Å^3$ cube used for ligand binding**.** After the cavity identification algorithm has detected cavities, the algorithm in this approach concentrates on pursuing the specific area or volume during the simulation

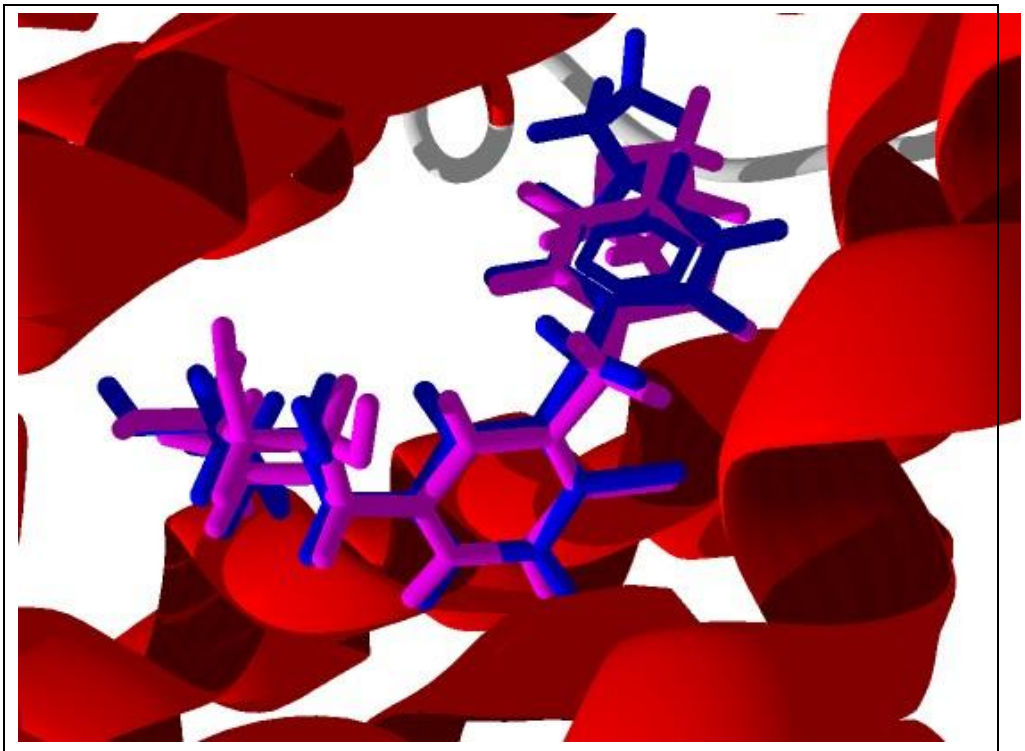


Figure 1. Validation docking protocol (redocking): native ligand before docking (blue) and ligand after docking (magenta) into active site SGLT-2 inhibitor receptor (protein code 7VSI).

phase. In this case, the algorithm found five unique binding sites for the SGLT-2 protein structure (Fig. 1). One of these five predicted cavities with the biggest volume (967.181Å^3), surface area (2149.12Å^2), and volume coordinates (X = 38.46Å; Y = 50.43Å, and Z = 46.16Å). Consequently, the ligand-binding process had been thought of for this specific cavity.

Preparation of ligand

To record the 3D structures of 3,5-caffeoylquinic (3,5-DCQA), the SDF format (*sdf) was selected: Phenhydrin, kaempferide, ferulic acid, nystose, quercetin, 1-kestose, rosmarinic acid, p-coumaric acid, caffeic acid, protocatechuic acid, gallic acid, and empagliflozin as reference standard obtained. Before ligands are imported for the docking process, it is necessary to convert 2D to 3D structures using the Chembiodraw v.16.0 program and minimize energy using the MMF94 force field approach. Table 1 illustrates the 2D and 3D structures of each ligand. To get the best results from the docking research, the imported structures must be appropriately set up, including having the right atom connections and bond order. The designed algorithm of the MVD is used for H-atom addition and the required valency tests.

Molecular docking study

Before ligand screening, the native ligand 7VSI was reinserted into the inner binding site of the SGLT-2 inhibitor crystals to confirm the docking process and produce docked poses and RMSD (Niu et al., 2022). The crystal structure of the SGLT-2 inhibitor (PDB Code 7VSI) was docked with 14 molecules from *Smallanthus sonchifolius* plants at the same time as Empagliflozin (standard ligands). During docking, the GRID resolution was set to 0.30Å, and the coordinates of the center of the active sites were X = 38.46 Å, Y = 50.43 Å, and Z = 46.16 Å. By default, 1.500 iterations and fifty population maximum were used as the maximum number of iterations. Next, ten separate docking runs of the differential evolution method were performed, each returning a single docking pose. Ten results were reranked using a more intricate scoring system after the ten independent docking runs to improve the docking accuracy. After the docking run, the resultant conformations were grouped, and only the negative lowest-energy representation from each cluster was returned. One pose with the lowest Re-rank value will be used for analysis. The score was determined to be each complex's ideal resolution. Molecular docking results were subjected to three experimental repetitions (n = 3). Accuracy calculations were derived from the averages, while precision was determined from the standard deviation (SD).

Table 1. Interpretation of pharmacokinetic parameters prediction results.

Pharmacokinetic phase	Pharmacokinetic parameter	Interpretation
Absorption	Caco2 permeability (log Papp in 10^{-6} cm/s)	A compound is considered to have high Caco-2 permeability if has Papp >8 × 10^{-6} cm/s
	Human intestinal absorption (%)	A molecule with HIA <30% is considered to be poorly absorbed
	P-glycoprotein substrate	The model predicts whether or not a compound is likely to be a substrate of Pgp.
Distribution	Human VDss (log L/kg)	Log VDss is considered low if <0.15 L/kg
	Fraction unbound (human)	The predicted fraction that would be unbound in plasma
	BBB permeability (log BB)	Compounds with log BB >0.3 are considered to readily cross the blood-brain barrier, while log BB <-1 is poorly distributed to the brain.
	CNS permeability (log PS)	A compound with log PS >-2 is considered to penetrate the CNS, while log PS <-3 is considered unable to penetrate the CNS.
Metabolism	CYP2D6 substrate	A compound is determined as likely to be metabolized by CYP2D6 or not
	CYP3A4 substrate	A compound is determined as likely to be metabolized by CYP3A4 or not.
Excretion	Total clearance (log mL/min/kg)	The prediction of total clearance in log (mL/min/kg)

Molecular docking study

Before ligand screening, the native ligand 7VSI was reinserted into the inner binding site of the SGLT-2 inhibitor crystals to confirm the docking process and produce docked poses and RMSD (Niu et al., 2022). The crystal structure of the SGLT-2 inhibitor (PDB Code 7VSI) was docked with 14 molecules from *Smallanthus sonchifolius* plants at the same time as Empagliflozin (standard ligands). During docking, the GRID resolution was set to 0.30Å, and the coordinates of the center of the active sites were X = 38.46 Å, Y = 50.43 Å, and Z = 46.16 Å. By default, 1.500 iterations and fifty population maximum were used as the maximum number of iterations. Next, ten separate docking runs of the differential evolution method were performed, each returning a single docking pose. Ten results were reranked using a more intricate scoring system after the ten independent docking runs to improve the docking accuracy. After the docking run, the resultant conformations were grouped, and only the negative lowest-energy representation from each cluster was returned. One pose with the lowest Re-rank value will be used for analysis. The score was determined to be each complex's ideal resolution. Molecular docking results were subjected to three experimental repetitions (n = 3). Accuracy calculations were derived from the averages, while precision was determined from the standard deviation (SD).

Simulation of molecular dynamics

For both the SGLT-2 protein system and the protein-ligand complexes that emerged from the docking simulations, we performed 100 ns molecular dynamics (MD) simulations. The YASARA structural program and the AMBER14 forcefield were utilized in the MD investigations. The simulations started with the md_run macro being executed. This macro also predicted the pKa of the protein residues to precisely control their protonation states at the desired pH of 7.4 and optimized the hydrogen bonding network to maximize the solute's stability. 0.9% physiological concentration of NaCl ions was supplied, with an excess of Na or Cl to neutralize the cell. The images were taken every ten ps during this investigation's 100 ns MD run. AMBER's default limit for Van der Waals forces was 8Å. The equations of motion were integrated with a multiple-time step of 1.25 fs for bound interactions and 2.5 fs for non-bonded interactions at a temperature of 310 K and a pressure of 1 atm (NPT ensemble). Running the previously modified md_analyze binding energy macro with the PBS method at 310K enabled the binding energy analysis utilizing the MM-PBSA method.

Prediction of drug likeness, pharmacokinetic, and toxicity properties

Screening of drug-likeness was conducted by inputting the structure of potential compounds in sdf format into the SCF Bio application (http://www.scfbio-iitd.res.in/software/drugdesign/lipinski.jsp) to obtain predicted Lipinski rule values. To predict pharmacokinetic and toxicity properties data, the SMILES code of the potential compounds was inputted into the online pkCSM (https://biosig.lab.uq.edu.au/pkcsm/prediction). The pharmacokinetic properties calculated were absorption (Caco2 permeability, human intestinal ab-

sorption, and P-glycoprotein substrate), distribution (human VDss, fraction unbound, BBB permeability, and CNS permeability), metabolism (CYP2D6 substrate and CYP3A4 substrate), as well as excretion (total clearance and renal OCT2 substrate). The toxicity properties obtained were including AMES toxicity, maximum tolerated dose, hERG I inhibitor, hERG II inhibitor, oral rat acute toxicity (LD_{50}), oral rat chronic toxicity (LOAEL), hepatotoxicity, skin sensitization, *T. pyriformis* toxicity, Minnow toxicity (Jayarayam et al., 2012; Pires et al., 2015; Sulistyowaty et al., 2021).

Data analysis

This study carried out molecular docking and molecular dynamics simulation in triplicate (n = 3). The standard deviation (SD) was used to evaluate precision, and the mean values were used to compute accuracy. To decide the most potential ligand candidate in molecular docking, we assessed the rerank score and hydrogen bond similarities between the compounds and the native ligands on the target protein. Ligand, with the lowest rerank score and high hydrogen bond similarity, was categorized as the best candidate. We analyzed the RMSD pattern and binding energy values in molecular dynamics simulation. The higher the RMSD Cα value, the higher the protein structural changes due to ligand binding. On the other hand, a high RMSD ligand movement suggested significant movement of the ligand from the protein binding site. The strong binding of ligand to the protein was assessed from the binding energy value. The more positive value indicated that the ligand has the strongest binding.

A molecule is considered drug-like if it meets at least two of the following criteria of the Lipinski rule: molecular mass less than 500 Dalton, high lipophilicity (expressed as LogP less than 5), less than five hydrogen bond donors, less than ten hydrogen bond acceptors, and molar refractivity should be between 40-130. Tables 1 and 2 present the interpretation of pharmacokinetics and toxicity.

RESULTS AND DISCUSSION

The result of the molecular docking study

The purpose of docking procedure validation is to evaluate how well the ligand is positioned at the binding site (Wang et al., 2017). By redocking, we validated the docking protocol using mvd software on the SGLT2 inhibitor receptor (protein code 7VSI). To assess the effectiveness of validation, we apply the RMSD value criterion. The most popular quantitative method for comparing two overlaid atomic coordinates' similarity is Root Mean Square Deviation, or RMSD (Castro-Alvarez et al., 2017; Prasetiyo et al., 2022). The docking protocol method is valid if the RMSD value is below 2Å (López-Camacho et al., 2016). In this investigation, the coordinate similarity between the empagliflozin ligand's native crystallographic position before docking and its position following docking by mvd. A similar structural orientation is seen between the docking posture and the crystallized ligand (RMSD 1.05 Å), according to the validation results of the docking process (Fig. 1). Emphagliflozin, the original ligand, was successfully repositioned at the original binding site using the docking protocol approach employing MVD. On this basis, the docking protocol can be used for the molecular docking of compounds from *S. sonchifolius* to the SGLT2 inhibitory receptor (protein code 7VSI).

Table 2. Interpretation of toxicity prediction results.

Toxicity parameter	Interpretation
AMES toxicity	Predict whether a compound is likely to be Ames positive and hence mutagenic.
Max. tolerated dose (human)	A MTRD of less than or equal to 0.477 log (mg/kg/day) is considered low, and high if greater than 0.477 log (mg/kg/day)
hERG I inhibitor	The model predicts whether a compound will likely be a HERG I inhibitor.
hERG II inhibitor	The model predicts whether a compound will likely be a HERG II inhibitor.
Oral rat acute toxicity (LD_{50})	The predicted LD50 in mol/kg
Oral rat chronic toxicity (LOAEL)	Lowest Observed Adverse Effect (LOAEL) in log (mg/kg bw/day)
Hepatotoxicity	It predicts whether a compound is likely to be associated with disrupted normal function of the liver.
T. pyriformis toxicity	pIG_{50} (negative logarithm of the concentration required to inhibit 50% of growth in log µg/L), with a value >-0.5 µg/L, is considered toxic
Minnow toxicity	LC_{50} values below 0.5 mM (log LC_{50} <-0.3) are regarded as high acute toxicity

Fourteen compounds isolated from *S. sonchifolius* and standard compounds (empagliflozin and canagliflozin) were docked together on the SGLT-2 receptor inhibitor (protein code 7VSI) using mvd software with a validated protocol. We use the rerank score value criterion to evaluate compound activity predictions (Bitencourt-Ferreira and de Azevedo, 2019). The lower value of the rerank indicated the strength of the ligand-receptor bond. The rerank score is derived from the linear combination of the E-intra of the ligands and the E-inter between the ligand and the protein (Prasetyo et al., 2022). The findings of the molecular docking investigation (Table 1) show that there are three compounds, namely 3,5-DCQA, 3,4-DCQA, and 1-kestose, which have a rerank score of -138.0 to 163.8 kcal/mol, where the rerank score is lower than with comparators (empagliflozin and canagliflozin) of -133.2 to 138.2 kcal/mol. The ligand-receptor bond will be more stable if the rerank score is more minor. Based on the rerank score of 3,5-DCQA, 3,4-DCQA; 1-ketosis has the potential to bind to the SGLT-2 receptor more strongly than the reference (empagliflozin and canagliflozin). In addition, Table 3 and Fig. 2 show the binding site of standard compounds (empagliflozin and canagliflozin) to SGLT-2 inhibitors involving key residues. The five hydrogen bonds observed in the SGLT-2 active site are Asn75, Phe98, Ser287, Trp291, and Gln457. These five bonds are essential in SGLT-2 inhibition (Ni et al., 2022). Table 2 and Fig. 2 show that 3,4-DCQA, 3,5-DCQA, and 1-kestose have hydrogen bond similarities compared to 40, 60, and 80% standards, respectively. Based on the similarity of hydrogen bonds of amino acid residues with empagliflozin, 1-kestose is the most potential ligand as an SGLT-2 inhibitor.

Results of molecular dynamics SGLT-2 inhibitors

Molecular dynamic simulation was used to assess the dynamic behavior of protein-ligand complexes. This study investigated the RMSD and binding energy of SGLT-2 protein in complexes with 1-kestose, 3,4-DCQA, and 3,5-DCQA, while Empagliflozin and Canagliflozin were used as standards RMSD and RMSD-Cα.

Fig. 3 displays the RMSD-Cα plots for SGLT-2, both as a native protein and in complexes with ligands, to help understand the conformational change of SGLT-2 complex structure. For about 30 ns, most complexes had patterns identical to those of the natural protein. However, in the subsequent simulation, the complexes showed different patterns with several fluctuations, indicating that ligand interaction affected protein intervention. This significant change in SGLT-2 structure affected by ligand binding may suggest its importance as a ligand inhibitor.

Results of molecular dynamics of RMSD-ligand movement

This study also looked at RMSD ligand mobility to learn more about the ligand movement from the binding site. A more significant RMSD value indicates a considerable ligand displacement from the binding pocket. Fig. 4 shows that, although 3,5-DCQA fluctuated more in the first 60 ns of the experiment, the plots for 3,4-DCQA, 1-kestose, and 3,5-DCQA stayed around roughly 3 Å throughout the simulation time. Ligands with less movement change are 3,4-DCQA, 1-kestose, and 3,5-DCQA, with average values of 2.224, 2.784, and 3.120 Å, respectively. However, all ligands illustrate similar patterns with two ligand controls, suggesting its acceptable movement from the binding site of SGLT-2.

Results of binding energy

This study also examined binding energy analysis to understand the ligands' strong affinity with SGLT-2. Fig. 5 shows the complexes' binding energy results over time. The complexes with the highest positive value indicate stronger binding with SGLT-2. The analysis demonstrated that the complexes with the highest average positive values are 1-kestose (-290.48 kJ/mol), 3,5-DCQA (-295.02 kJ/mol), and 3,4-DCQA (-407.47 kJ/mol), respectively.

This means 1-kestose exhibited stronger binding affinity with SGLT-2 during the simulation period. As illustrated in Fig. 6, 1-kestose persisted in interacting with the SGLT-2 binding pocket until the simulation's conclusion. This interaction involved a few crucial residues, such as Phe98, Glu99, and Ser287, and was primarily responsible for hydrogen bond interactions.

Result of **absorption** prediction

The absorption prediction results obtained through the open-source SCF-Bio application (http://www.scfbio-iitd.res.in/software/drugdesign/lipinski.jsp) indicate that the compounds 3,4-DCQA, 3,5-DCQA, and 1-kestose do not meet the criteria outlined by Lipinski's Rule of Five (Table 4). The compounds 3,4-DCQA, 3,5-DCQA, and 1-kestose must meet molar refractivity (MR), hydrogen bond donors (HBD), and hydrogen bond acceptors (HBA)requirements due to their complex structures with many polar atoms (O and N).

The pharmacokinetic parameter prediction results using the open-source pKCSM application (https://biosig.lab.uq.edu.au/pkcsm/prediction) yield consistent findings with SCF-Bio, with 3,4-DCQA, 3,5-DCQA, and 1-kestose demonstrating low permeability, specifically $<8 \times 10^{-6}$ cm/s (Table 5).

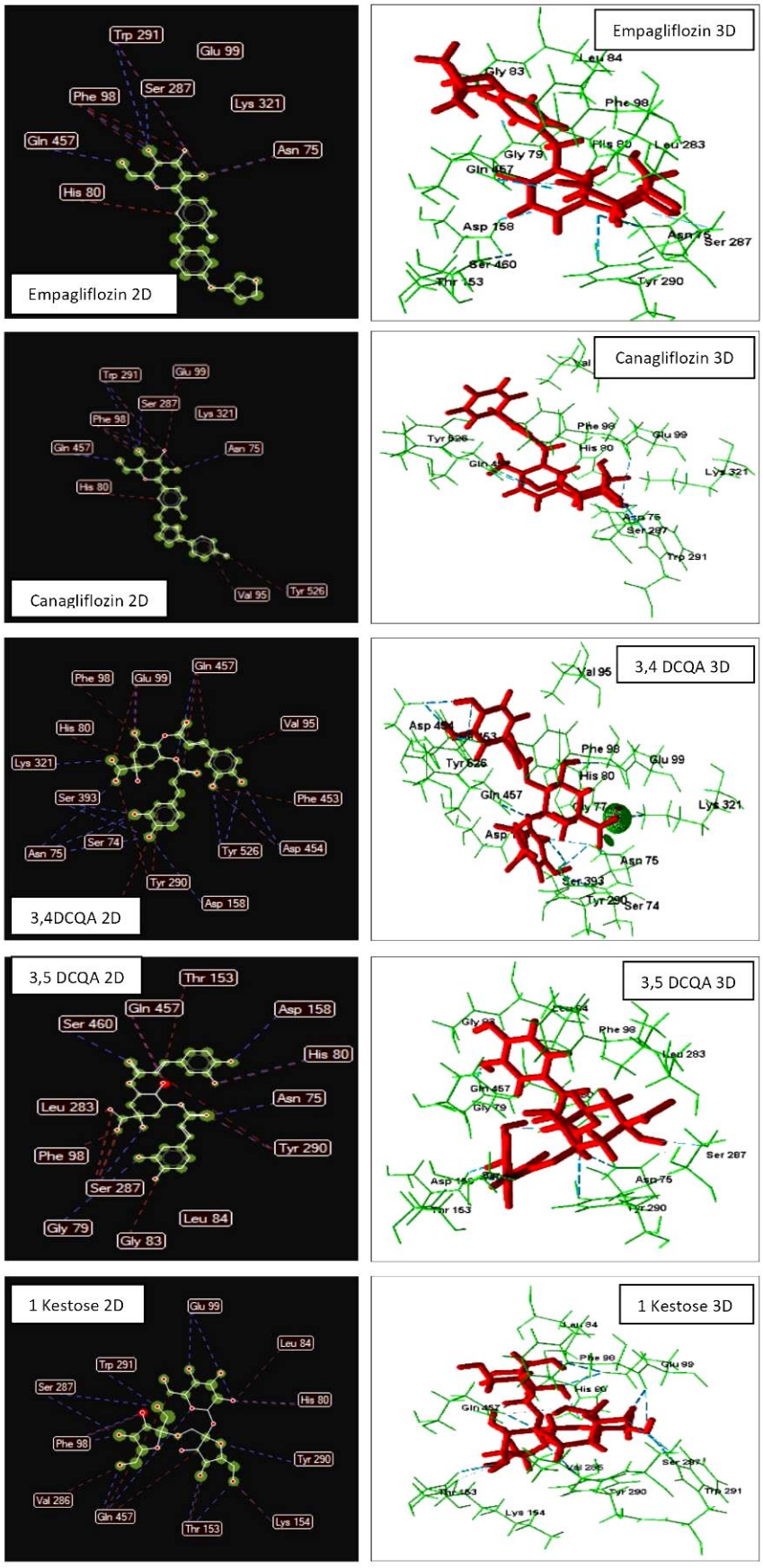


Figure 2. Molecular docking result of empagliflozin, canagliflozin, 3,4-DCQA, 3,5-DCQA, and 1-kestose on the active site of SGLT-2.

Table 3. Molecular docking results of standard ligands (empagliflozin and canagliflozin) and the best phytochemicals from *Smallanthus sonchifolius.*

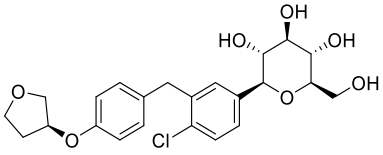
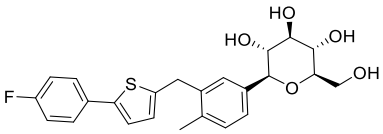
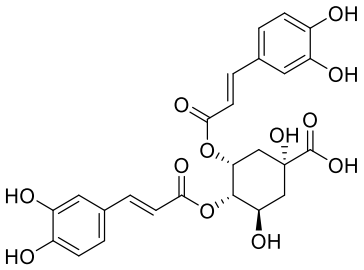
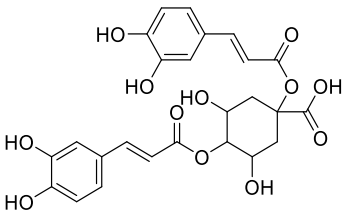
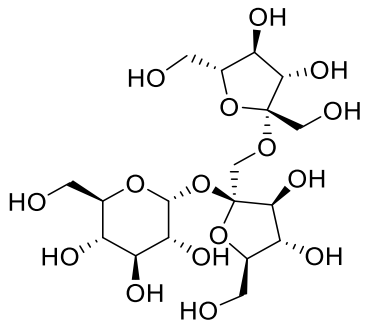
Compound	Rerank score (kcal/mol)	Hydrogen bond	Steric interaction	Similarity hydrogen bond
 **Empagliflozin (standard)**	**-137.82 ± 1.16**	**Asn 75** **Phe 98** **Ser 287** **Trp 291** **Gln 457**	Asn 75 His 80 Phe 98 Glu 99 Ser 287 Trp 291 **Lys 321** **Gln457**	**100%**
 **Canagliflozin (standard)**	-134.10 ± 1.54	**Asn 75** **Phe 98** **Ser 287** **Trp 291** **Gln 457**	Asn 75 His 80 Val 95 Phe 98 Glu 99 Ser 287 Trp 291 Lys 321 Gln 457 Trp 526	100%
 **3,4-DCQA**	-164.84 ± 1.44	Ser 74 **Asn 75** Glu 99 Asp 158 Lys 321 Ser 393 Asp 454 **Gln457**	Asn 75 Gly 77 Ser 74 His 80 Val 95 Phe 98 Glu 99 Asp 158 Tyr 290 Lys 321 Ser 393 Phe 453 Asp 454 Gln 457 Tyr 526	40%
 **3,5-DCQA**	-155.92 ± 1.34	**Asn 75** Gly 79 His 80 Asp 158 **Ser 287** Tyr 290 Ser 460 **Gln 457**	Asn 75 Gly 79 His 80 Gly 83 Leu 84 Phe 98 Thr 143 Asp 158 Leu 283 Ser 287 Tyr 290 Gln 457 Ser 460	60%

Table 3. Molecular docking results of standard ligands (empagliflozin and canagliflozin) and the best phytochemicals from *Smallanthus sonchifolius* (continued…)

Compound	Rerank score (kcal/mol)	Hydrogen bond	Steric interaction	Similarity hydrogen bond
 **1-Kestose**	-138.02 ± 2.35	His 80 **Phe 98** Glu 99 Thr 153 Lys 154 **Ser 287** Tyr 290 **Trp 291** **Gln 457**	His 80 Leu 84 Glu 99 Phe 98 Thr 153 Lys 154 Tyr 290 Trp 291 Val 286 Ser 287 Gln 457	80%

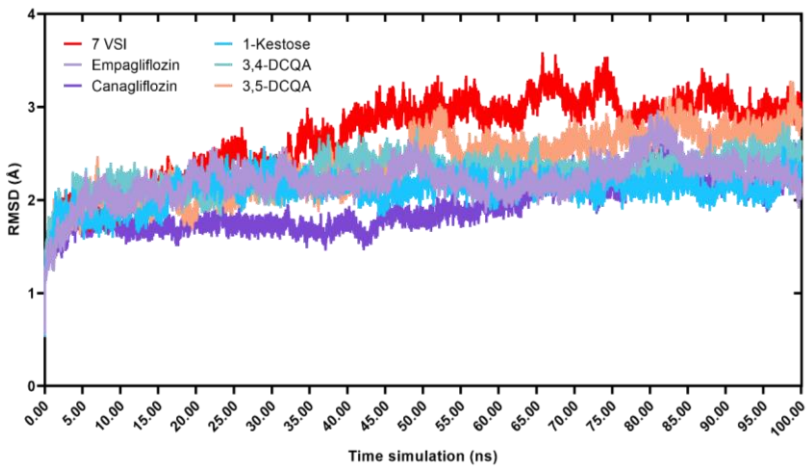


Figure 3. Result of molecular dynamics simulation showing standards RMSD and RMSD-Cα.

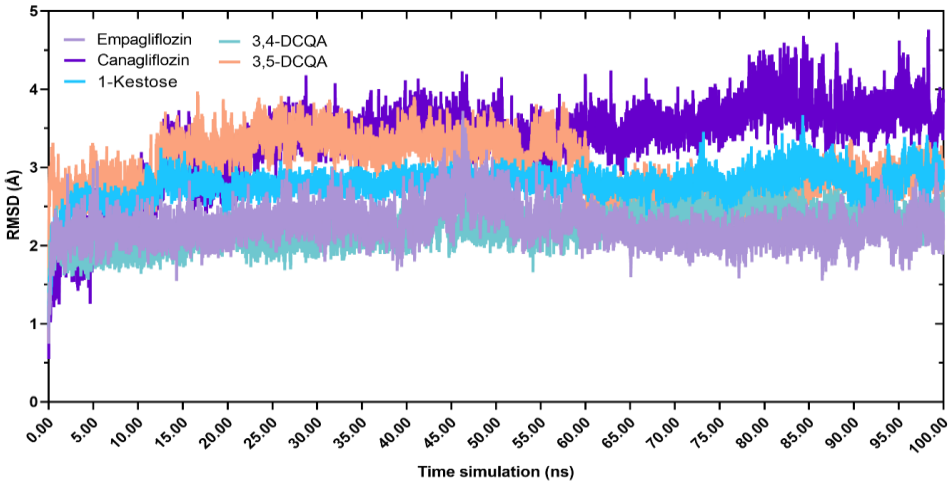


Figure 4. The result of the molecular dynamics simulation showed RMSD-ligand movement from 0 ns to 100 ns.

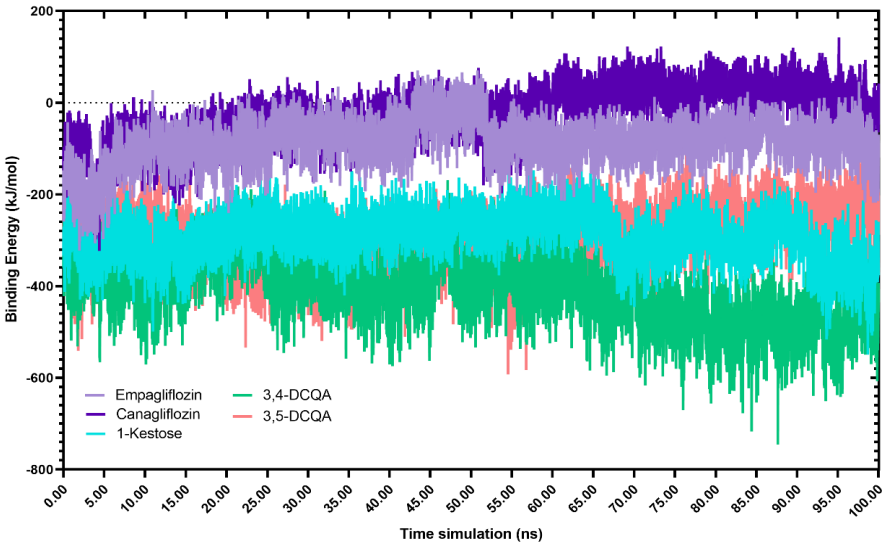


Figure 5. The result of molecular dynamics simulation shows energy binding from 0 ns until 100 ns.

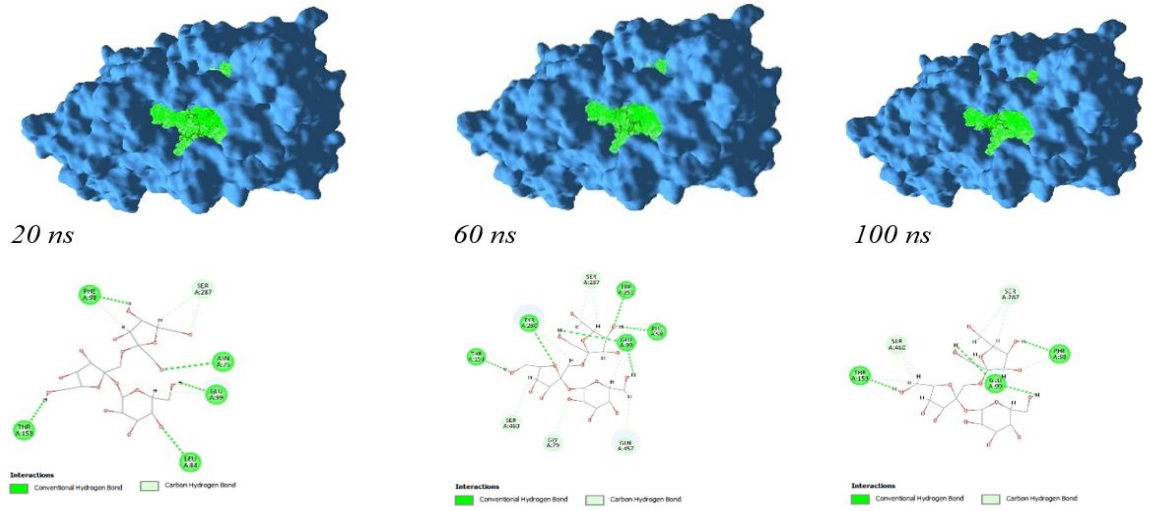


Figure 6. Result of molecular dynamics simulation showing 1-kestose interaction with some amino acids residual into at the active site of SGLT-2

Table 4. The result of drug-likeness screening.

Compound	Lipinski Rule of Five				
	MW	HBD	HBA	Log P	MR
3,4-DCQA	516	7	12	1.03	125.20
3,5-DCQA	516	7	12	1.03	125.20
1-Kestose	504	11	16	-7.57	101.30

MW: molecular weight; HBD: hydrogen bond donor; HBA: hydrogen bond acceptor; MR: molar refractivity.

Table 5. The result of pharmacokinetics profile predictions.

Pharmacokinetic parameter	3,5-Dicaffeoylquinic acid (3,5 DCQA)	3,4-Dicaffeoylquinic acid (3,4 DCQA)	1-Kestose
Caco2 permeability (log Papp in 10^{-6} cm/s)	-1.147	-1.203	-0.49
Human intestinal absorption (%)	44.225	29.037	0
P-glycoprotein substrate	Yes	Yes	Yes
Human VDss (log L/kg)	1.7	1.633	-0.248
Fraction unbound (human)	0.28	0.294	0.538
BBB permeability (log BB)	-2.069	-2.08	-1.469
CNS permeability (log PS)	-3.822	-3.804	-5.113
CYP2D6 substrate	No	No	No
CYP3A4 substrate	Yes	Yes	No
Total clearance (log mL/min/kg)	-0.044	-0.042	1.562

Table 6. The result of toxicity predictions.

Toxicity parameter	3,5-Dicaffeoylquinic acid (3,5 DCQA)	3,4-Dicaffeoylquinic acid (3,4 DCQA)	1-Kestose
AMES toxicity	No	No	No
Max. tolerated dose (human)	0.393	0.393	0.733
hERG I inhibitor	No	No	No
hERG II inhibitor	No	No	Yes
Oral rat acute toxicity (LD_{50})	2.643	2.626	2.800
Oral rat chronic toxicity (LOAEL)	4.374	4.131	4.162
Hepatotoxicity	No	No	No
T. pyriformis toxicity	0.285	0.285	0.285
Minnow toxicity	5.538	4.942	15.565

According to the Biopharmaceutical Classification System (BCS), these three phytochemicals from *S. sonchifolius* fall into BCS Class III, indicating excellent solubility but limited permeability (Ku, 2008; Papich and Martinez, 2015). From the human intestinal absorption parameter, 1-kestose is considered to be poorly absorbed. All three compounds are likely to be P-glycoprotein substrates, so they tend to be expelled from the cell, reducing absorption as they pass through the gastrointestinal tract lining (Lin and Yamazaki, 2003).

The volume of distribution is the theoretical volume that the total dose of a drug would need to be uniformly distributed with the same concentration as in blood plasma. VDss of 1-kestose is considered low, which may be caused by a large number of polar groups in the compound. All three compounds are poorly distributed to the brain and are considered unable to penetrate the CNS. Both 3,4-DCQA and 3,5-DCQA are likely to be metabolized by CYP3A4.

Table 6 shows the toxicity prediction results. Based on the results of all toxicity parameters, the three potential compounds from *S. sonchifolius* are generally considered to be safe.

CONCLUSION

One phytochemical present in *Smallanthus sonchifolius* is 1-kestose, which is thought to act as an SGLT-2 inhibitor for treating diabetes based on the findings of molecular docking and molecular dynamics investigations. This compound falls into BCS Class III, which means it has good solubility but poor permeability. Overall, this compound's distribution, metabolism, excretion, and safety profile are classified as good.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

ACKNOWLEDGMENTS

The authors express their gratitude to Prof. Dr. Siswandono of the Department of Pharmaceutical Chemistry at Airlangga University for providing permission and assistance with the software Virtual Docker Molegro. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

Abdul-Ghani MA, Norton L, DeFronzo RA (2011) Role of sodium-glucose cotransporter 2 (SGLT 2) inhibitors in the treatment of type 2 diabetes. Endocr Rev 32(4): 515–531. https://doi.org/10.1210/er.2010-0029

Aziz Z, Yuliana ND, Simanjuntak P, Rafi M, Mulatsari E, Abdillah S (2021) Investigation of yacon leaves (*Smallanthus sonchifolius*) for α-glucosidase inhibitors using metabolomics and *in silico* approach. Plant Foods Hum Nutr 76(4): 487–493. https://doi.org/10.1007/s11130-021-00926-3

Batool M, Ahmad B, Choi S (2019) A structure-based drug discovery paradigm. Int J Mol Sci 20(11): 2783 https://doi.org/10.3390/ijms20112783

Bitencourt-Ferreira G, de Azevedo WF (2019) Docking with SwissDock. In: de Azevedo Jr., W. (eds) Docking Screens for Drug Discovery. Methods in Molecular Biology, vol 2053. New York, NY: Humana. https://doi.org/10.1007/978-1-4939-9752-7_12

Castro-Alvarez A, Costa AM, Vilarrasa J (2017) The performance of several docking programs at reproducing protein-macrolide-like crystal structures. Molecules 22(1): 136. https://doi.org/10.3390/molecules22010136

Choi CI (2016) Sodium-glucose cotransporter 2 (SGLT2) inhibitors from natural products: Discovery of next-generation antihyperglycemic agents. Molecules 21(9): 1136. https://doi.org/10.3390/molecules21091136

Contreras-Puentes N, Alvíz-Amador A (2021) Hypoglycaemic property of yacon (*Smallanthus sonchifolius* (Poepp. and Hendl.) H. Robinson): A review. Pharmacogn Rev 14(27): 37–44. https://doi.org/10.5530/phrev.2020.14.7

DeFronzo RA, Norton L, Abdul-Ghani M (2017) Renal, metabolic and cardiovascular considerations of SGLT2 inhibition. Nat Rev Nephrol 13(1): 11–26. https://doi.org/10.1038/nrneph.2016.170

Dipiro TJ, Yee GC, Posey M, Haines ST, Nolin TD, Ellingrod V (2020) Pharmacotherapy: A Pathophysiologic Approach. 8th ed. New York: McGraw Hill.

Hummel CS, Lu C, Loo DDF, Hirayama BA, Voss AA, Wright EM (2011) Glucose transport by human renal Na+/D-glucose cotransporters SGLT1 and SGLT2. Am J Physiol Cell Physiol 300(1): C14–C21. https://doi.org/10.1152/ajpcell.00388.2010

International Diabetes Federation (2021) International diabetes federation. In: Boyko EJ, Magliano DJ, Karuranga S, Piemonte L, Riley P, Saeedi P, Hong S (eds.). IDF Diabetes Atlas, 10th edition. https://doi.org/10.1016/j.diabres.2013.10.013

Jayaram B, Singh T, Mukherjee G, Mathur A, Shekhar S, Shekhar V (2012) Sanjeevini: A freely accessible web-server for target directed lead molecule discovery. BMC Bioinformatics 13(Suppl 17): S7. https://doi.org/10.1186/1471-2105-13-S17-S7

Ku MS (2008) Use of the biopharmaceutical classification system in early drug development. AAPS J 10(1): 208–212. https://doi.org/10.1208/s12248-008-9020-0

Lavecchia A, Giovanni C (2013) Virtual screening strategies in drug discovery: A critical review. Curr Med Chem 20: 2839–2860. https://doi.org/10.2174/09298673113209990001

Lin JH, Yamazaki M (2003) Role of P-glycoprotein in pharmacokinetics: Clinical implications. Clin Pharmacokinet 42(1): 59–98. https://doi.org/10.2165/00003088-200342010-00003

López-Camacho E, García-Godoy MJ, García-Nieto J, Nebro AJ, Aldana-Montes JF (2016) A New Multi-objective Approach for Molecular Docking Based on RMSD and Binding Energy. In: Botón-Fernández M, Martín-Vide C, Santander-Jiménez S, Vega-Rodríguez MA (eds) Algorithms for Computational Biology. AlCoB 2016. Lecture Notes in Computer Science, vol 9702. Springer, Cham. https://doi.org/10.1007/978-3-319-38827-4_6

Miller E, Shubrook JH (2015) Sodium glucose co-transporter 2 inhibitors in the treatment of type 2 diabetes mellitus. Osteopath Fam Physician 7(5): 10–30.

Nauck MA, Wefers J, Meier JJ (2021) Treatment of type 2 diabetes: Challenges, hopes, and anticipated successes. Lancet Diabetes Endocrinol 9(8): 525–544. https://doi.org/10.1016/S2213-8587(21)00113-3

Niu Y, Liu R, Guan C, Zhang Y, Chen Z, Hoerer S, Nar H, Chen L (2022) Structural basis of inhibition of the human SGLT2–MAP17 glucose transporter. Nature 601: 280–284. https://doi.org/10.1038/s41586-021-04212-9

Pahlawan PP, Dwita O (2016) The effect of insulin leaves (*Smallanthus sonchifolius*) as antidiabetic. J Majority 5(4): 133–137.

Papich MG, Martinez MN (2015) Applying biopharmaceutical classification system (BCS) criteria to predict oral absorption of drugs in dogs: Challenges and pitfalls. AAPS J 17(4): 948–964. https://doi.org/10.1208/s12248-015-9743-7

Pires DEV, Blundell TL, Ascher DB (2015) pkCSM: Predicting small molecule pharmacokinetic and toxicity properties using graph based signatures. J Med Chem 58(9): 4066–4072. https://doi.org/10.1021/acs.jmedchem.5b00104

Prasetiyo A, Kumala S, Mumpuni E, Tjandrawinata RR (2022) Validation of structural-based virtual screening protocols with the PDB Code 3G0B and prediction of the activity of *Tinospora crispa* compounds as inhibitors of dipeptidyl-peptidase-IV. Res Results Pharmacol 8(1): 95–102. https://doi.org/10.3897/rrpharmacology.8.76237

Ramani J, Shah H, Vyas VK, Sharma M (2022) A review on the medicinal chemistry of sodium glucose co-transporter 2 inhibitors (SGLT2-I): Update from 2010 to present. Eur J Med Chem Rep 6: 100074. https://doi.org/10.1016/j.ejmcr.2022.100074

Santer R, Calado J (2010) Familial renal glucosuria and SGLT2: From a Mendelian trait to a therapeutic target. Clin J Am Soc Nephrol 5(1): 133–141. https://doi.org/10.2215/CJN.04010609

Sulistyowaty MI, Widyowati R, Putra GS, Budiati T, Matsunami K (2021) Synthesis, ADMET predictions, molecular docking studies, and in-vitro anticancer activity of some benzoxazines against A549 human lung cancer cells. J Basic Clin Physiol Pharmacol 32(4): 385–392. https://doi.org/10.1515/jbcpp-2020-0433

Toyama T, Neuen BL, Jun M, Ohkuma T, Neal B, Jardine MJ, Heerspink HL, Wong MG, Ninomiya T, Wada T, Perkovic V (2019) Effect of SGLT2 inhibitors on cardiovascular, renal and safety outcomes in patients with type 2 diabetes mellitus and chronic kidney disease: A systematic review and meta-analysis. Diabetes Obes Metab 21(5): 1237–1250. https://doi.org/10.1111/dom.13648

Vallon V, Thomson SC (2017) Targeting renal glucose reabsorption to treat hyperglycaemia: The pleiotropic effects of SGLT2 inhibition. Diabetologia 60(2): 215–225. https://doi.org/10.1007/s00125-016-4157-3

Wang XX, Levi J, Luo Y, Myakala K, Herman-Edelstein M, Qiu L, Wang D, Peng Y, Grenz A, Lucia S, Dobrinskikh E, D'Agati VD, Koepsell H, Kopp JB, Rosenberg AZ, Levi M (2017) SGLT2 protein expression is increased in human diabetic nephropathy: SGLT2 protein inhibition decreases renal lipid accumulation, inflammation, and the development of nephropathy in diabetic mice. J Biol Chem 292(13): 5335–5348. https://doi.org/10.1074/jbc.M117.779520

Yu W, Mackerell Jr AD (2017) Computer-aided drug design methods. Methods Mol Biol 1520: 85–106. https://doi.org/10.1007/978-1-4939-6634-9_5

AUTHOR CONTRIBUTION:

Contribution	Prasetyo A	Mumpuni E	Luthfiana D	Herowati R	Putra GS
Concepts or ideas	x				
Design		x			
Definition of intellectual content		x			
Literature search	x	x	x	x	x
Experimental studies	x	x		x	x
Data acquisition			x		
Data analysis	x		x	x	
Statistical analysis				x	x
Manuscript preparation					x
Manuscript editing	x				x
Manuscript review	x	x	x	x	x

Citation Format: Prasetyo A, Mumpuni E, Luthfiana D, Herowati R, Putra GSG (2025) *In silico* discovery of potential sodium-glucose cotransporter-2 (SGLT-2) inhibitors from *Smallanthus sonchifolius* (Poepp.) H.Rob. via molecular docking and molecular dynamics simulation approach. J Pharm Pharmacogn Res 13(3): 716–728. https://doi.org/10.56499/jppres24.2104_13.3.716

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Open Access: This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/ licenses/by/4.0/), which permits use, duplication, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.