Pharmaceutical Sciences Asia

Pharm Sci Asia 2025; 52(3), 396-416
DOI:10.29090/psa.2025.03.25.3874

Research Article

Exploration of potential natural dual inhibitors of SGLT1 and SGLT2 for the treatment of type 2 diabetes

Khoi Anh Nguyen, Phuong Thuy Viet Nguyen*, Kiet Hoang Anh Nguyen

Faculty of Pharmacy, University of Medicine and Pharmacy at Ho Chi Minh City, Ho Chi Minh City 700000, Viet Nam

ABSTRACT

Dual inhibitors of both sodium-glucose cotransporters (SGLT1 and 2) is of great interest due to combination of blocking renal SGLT2 as well as intestinal SGLT1 leading to decreases of glucose absorption and HbA1C and also increase of GLP-1 and beneficial effect on cardiovascular system[1]. Thus, this study aimed to explore potential natural drugs targeting both SGLT1 and SGLT2 proteins for the treatment of type 2 diabetes. The screening process through combining pharmacophore models drived from molecular dynamics simulations (MDs) and molecular docking for both structures of SGLT1 (PDB: 7WMV) and SGLT2 (PDB: 7VSI) were successfully applied. Seven potential natural dual SGLT1 and SGLT2 inhibitors were obtained including neodiosmin, apigenin-7-O-neohesperidosid, neobudoffieid, pinoresinol diflucosid, hispidulin-7-O-rutinosid, apigenin-7-O-rutinosid, syringaresinol-4-O-β-D-glucopyranosid). They satisfied the five features of two pharmacophore models for SGLT1 and SGLT2 proteins for their inhibitory bioactivities. Docking results showed that they fitted well in the binding sites of both proteins by forming key interactions similar to a reference dual inhibitor, sotagliflozin with their binding affinities ranging from -10.07 to -13.67 kcal.mol^{-1} for SGLT1 and from -10.56 to -13.90 kcal.mol^{-1} for SGLT2. Further experimental assays are required for testing bioactivities of these seven compounds as dual SGLT1 and 2 inhibitors.

Keywords:
sodium-glucose cotransporter, dual SGLT1 and SGLT2 inhibitors, pharmacophore models, molecular dynamics simulation, molecular docking, virtual screening

1. INTRODUCTION

Diabetes mellitus is a common chronic disease worldwide with the high levels of blood glucose in the body. So far, there have been multiple drug classes used for the treatment of diabetes, such as insulin or inhibitors of alpha-glucosidase (acarbose, miglitol), biguanides (metformin), dipeptidyl peptidase-4 (alogiptin, linagliptin), sulfonylureas (glimepiride, gliclazide, glipizide, glyburide, etc), thiazolidinediones (rosiglitazone, pioglitazone), glucagon-like peptide 1 (GLP-1) receptor agonists (dulaglutide, semaglutide), etc[1]. However, most antidiabetic drugs are associated with significant adverse effects: insulin and sulfonylureas can cause hypoglycemia and weight gain; metformin may lead to gastrointestinal disturbances and a risk of lactic acidosis; GLP-1 receptor agonists and alpha-glucosidase inhibitors often cause gastrointestinal side effects, etc[1]. These adverse effects may impact treatment outcomes and should be carefully considered when selecting appropriate therapeutic regimens.

SLGT inhibitors (Figure 1) were approved as a new class of glucose-lowering agents, including selective SGLT2 inhibitors (based on the C-glycosylated diarylmethane pharmacophore) like canagliflozin, empagliflozin, dapagliflozin, ertugliflozin, exagliflozin; and inhibitors of both SGLT1 and SGLT2(dual inhibitors) like sotagliflozin[1]. Among these, sotagliflozin stands out due to its dual mechanism of action, inhibiting both SGLT2 in the kidneys and SGLT1 in the intestines. This dual inhibition not only reduces renal glucose reabsorption, intestinal glucose absorption leading to

***Corresponding author:**
* Phuong Thuy Viet Nguyen Email: ntvphuong@ump.edu.vn

lower HbA1c and enhances endogenous GLP-1 secretion but also provides additional cardiovascular effects[2]. Moreover, compared to selective SGLT2 inhibitors such as canagliflozin, it has shown a lower risk of bone fractures, amputations, and urinary tract infections[2]. However, sotagliflozin presents several notable limitations, including an increased risk of diarrhea and hypotension[3]. Therefore, the search for novel compounds capable of dual SGLT1/SGLT2 inhibition with reduced side effects represents a promising research direction to enhance therapeutic efficacy and improve safety.

Currently, natural products are often used as lead compounds to design and synthesize novel drugs targeting both SGLT1 and SGLT2 proteins. These compounds are generally fewer gastrointestinal side effects (such as diarrhea, dehydration, and malabsorption) due to their typically moderate inhibitory activity[4]. Therefore, exploring natural products that act as dual inhibitors of SGLT1 and SGLT2 is considered a promising strategy to discover novel structural scaffolds for semisynthetic drug development and to reduce undesirable side effects in clinical.

Many natural compounds were isolated and tested for *in vitro* inhibitory bioactivity on SGLT1 and SGLT2[5] such as: (−)−kurarinon in *Sophora flavescens*, Fabaceae (IC_{50} SGLT1 = 10.4 µM, IC_{50} SGLT2 = 1.7 µM)[5] and sophoraflavanon G (IC_{50} SGLT1 = 18.7 µM, IC_{50} SGLT2 = 4.1 µM)[2], 10-Methoxy-N(1)-methylburnamin-17-O-veratrat (IC_{50} SGLT1 = 4.0 µM, IC_{50} SGLT2 = 0.5 µM and Alstiphyllanin D (IC_{50} SGLT1 = 5.0 µM, IC_{50} SGLT2 = 2.0 µM) in *Alstonia macrophylla*, Apocynaceae[5], etc. However, there are still not available natural drugs in the market as dual inhibitors for SGLT1 and SGLT2.

To accelerate the drug discovery process and reduce time and cost, *in silico* approaches have been increasingly applied in the research of antidiabetic agents targeting both SGLT1 and SGLT2. Some common techniques, molecular docking and molecular dynamics simulations and pharmacophore modelling are often used to achieve the success. Some potential natural compounds were identified such as bavachromanol in *Psoralea corylifolia,* Fabaceae fitting well in SGLT1 and SGLT2 with good binding affinities[6]; trilobatin derived from *Lithocarpus polystachyus,* Fagaceae showing its binding affinities similar to phlorizin towards both SGLT1 and SGLT2 proteins[7]; 14-Dexo-14-O-acetylorthosiphol Y from *Orthosiphon stamineus Benth* binding well with SGLT1 and SGLT2 protein considered as a potent anti-diabetic drug[8], etc. Although the available X-ray crystal structures of SGLT1 and SGLT2 provide valuable static snapshots of protein-ligand interactions, they represent only a single conformation and may not fully reflect the dynamic nature of the binding sites. Therefore, we employed a molecular dynamics (MD)-based pharmacophore modeling approach to better account for protein flexibility and binding pocket dynamics. This strategy offers a more realistic basis for pharmacophore generation and virtual screening[9]. While *in silico* methods such as molecular docking and MD simulations have been applied previously, the integration of MD-derived conformations into pharmacophore modeling

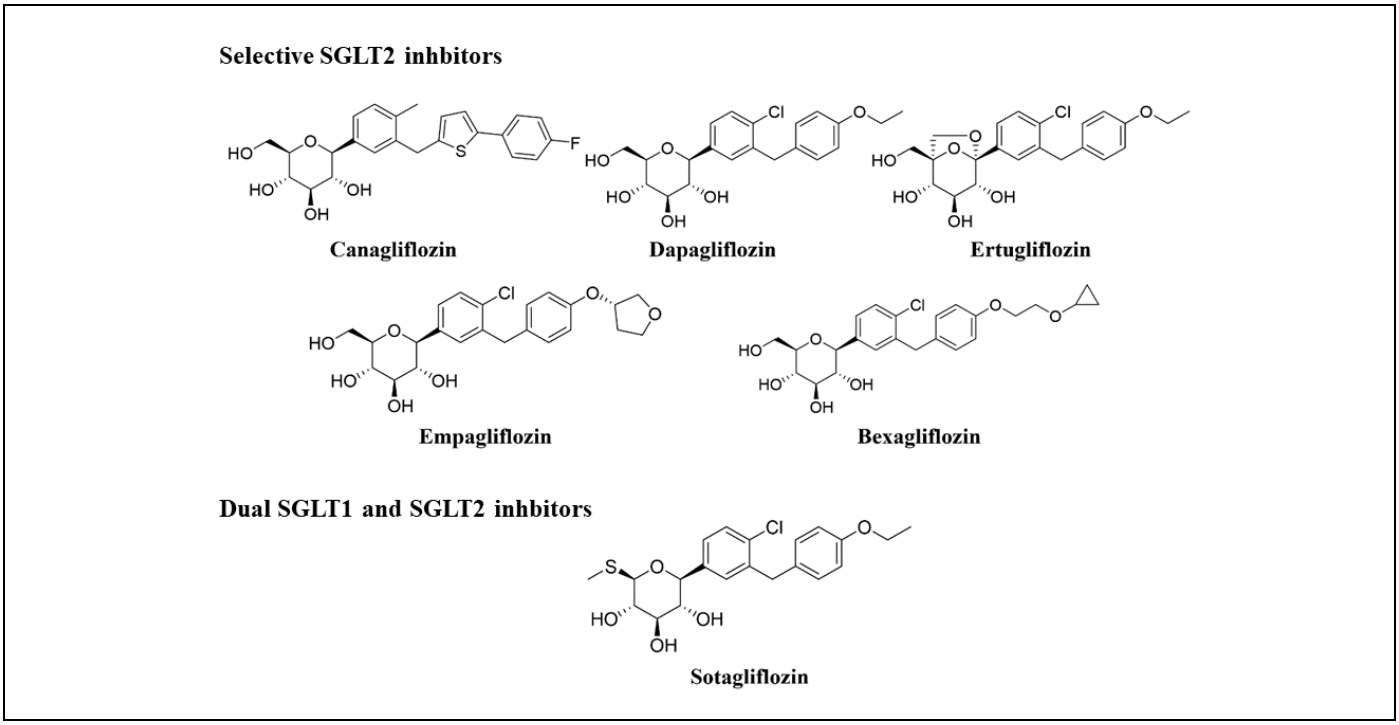

Figure 1. SGLT1 inhibitors and dual SGLT1 and SGLT2 inhibitors

remains limited, particularly in the context of dual SGLT1 and SGLT2 inhibitor discovery. Some recent studies have demonstrated that combining MD simulations with pharmacophore modeling can significantly improve the screening performance and enrich the identification of active compounds[10,11].

In this study, the objective was to explore potential natural dual inhibitors of SGLT1 and SGLT2. The novelty of this study lies in the hypothesis that representative conformations extracted from molecular dynamics (MD) simulations provide a more accurate basis for pharmacophore model generation compared to the conventional single-structure approach. Therefore, this study focused on application of molecular dynamics simulations (MDs) in generation of pharmacophore models more accurately compared to the common approach for inhibitors of both SGLT1 and SGLT2 using their PDB structures (SGLT1 with PDB ID: 7WMV and SGLT2 with PDB ID: 7VSI). Through cluster analysis, representative conformational structures obtained from MDs showed more advantages in generation of pharmacophore models. Subsequently, virtual screening using pharmacophore models and induced-fit molecular docking were carried out to identify potential natural compounds from our in-house database. Further biological testings are required to confirm bioactivities of these compounds as dual SGLT1 and SGLT2 inhibitors.

2. MATERIALS AND METHODS

2.1. Protein preparation

The structures of SGLT1 and SGLT2 were retrieved from Protein Data Bank (https://www.rcsb.org/) with the human SGLT1 (PDB code: 7WMV[12], its resolution of 3.20 Å) in complexed the co-crystallised ligand LX2761 as the selective inhibitor of SGLT1 and SGLT2 (PDB code: 7VSI[13], its resolution of 2.95 Å) complexed with empagliflozin as the selective inhibitor of SGLT2. 7WMV[12] is the tructure of human SGLT1-MAP17 complex bound with LX2761 and 7VSI[13] is Structure of human SGLT2-MAP17 complex bound with empagliflozin[7].

2.2. Database for screening

In-house phytochemical database includes 274 natural products which were collected from Faculty of Pharmacy at University of Medicine and Pharmacy at Ho Chi Minh city in Viet Nam (Supplementary material).

2.3. Molecular dynamic simulations

MDs aimed to investigate the fluctuation of protein structure during simulations as well as to select the representative equilibrium conformations of structures for pharmacophore model generations. MDs were conducted for both SGLT1 (PDB ID: 7WMV) and SGLT2 (PDB ID: 7VSI) proteins for 200 ns using GROMACS 2023.2 software[14] and employing the CHARMM27 force field[15]. Subsequently, cluster analysis was carried out for choosing the representative structural conformations of SGLT1 and SGLT2, pharmacophore models for SGLT1 and SGLT2 inhibitors, respectively were developed by MOE version 2022.02[16].

Solvent and co-crystallised ligands were removed from the structures of SGLT1 and SGLT2. Ligands were protonated and added hydrogen atoms using the UCSF Chimera 1.16[17] and then parameterized using the SwissPARAM[18] server (https://www.swissparam.ch/). MDs were run for both complexes of SGLT1 and SGLT2 with the co-crystallised ligands and the apo-proteins for 200 ns. Each ligand–protein complex, along with the corresponding apo form, was solvated in a cubic dodecahedron box using the TIP3P water model and neutralised by adding 4 Cl^- ions for SGLT1 and 7 Cl^- ions for SGLT2, respectively. Energy minimization ran using the steepest descent algorithm. System equilibration was conducted in two phases: a constant-volume (NVT) phase and a constant-pressure (NPT) phase, each lasting 100 ps. Subsequently, 200 ns production MD simulations were run for each system. The simulation setup and key parameters are summarised in **Table 1** below.

Production stage were run with a timestep of 2 fs after achieving the equilibrium. The MDs trajectories were recorded every 0.01 ns and visualized using the

Table 1. Summary of Molecular Dynamics Simulation Parameters

Stage	Parameter
Force field	CHARMM27
Simulation software	GROMACS 2023.2
Water model	TIP3P
Box type	Cubic dodecahedron
System neutralization	4 Cl^- (SGLT1), 7 Cl^- (SGLT2)
Energy minimization	Steepest descent algorithm
Equilibration	NVT: 100 ps at 300 K
	NPT: 100 ps at 1 atm
Production run	200 ns

VMD 1.9.4 program[19]. Criteria parameters for MDs included the values of RMSD (root-mean-square deviation), RMSF (root-mean-square fluctuation), Rg (radius of gyration), and SASA (solvent accessible surface area) of the protein backbone or the heavy atoms of the ligands during the MDs. Furthermore, the hydrogen bonds were determined to occur if the bonding angle between the hydrogen donor (D) and acceptor (A) larger than 120^o with the distance between D and A exceeding 3.5 Å.

Cluster analysis

Clustering methods have been widely used in MDs to group similar conformational structures from the trajectories of MDs.

In this study, the neighbor-based algorithm[20] was chosen as the clustering algorithm. This clustering process included many steps as followings[20]:

- Step 1: the values of RMSD of atoms positions between all pairs of conformational structures were calculated.
- Step 2: every conformation was counted for the number of neighbor conformations around this beginning conformation within the values of RMSD.
- Step 3: The conformations with the highest number of adjacent conformations was selected as the center of cluster and forming the cluster containing all the adjacent conformations. The conformation of this cluster was then removed out of the pool of structures of conformations.
- Step 4: The process was repeated until all of the conformations were arranged into the cluster and the pool of the conformations were empty.

Clustering analysis: Clustering analysis was performed based on the value of RMSD using the stable conformational structures. Subsequently, the module *gmx_cluster* was used for clustering every 0.01 ns and determining the representative conformations.

Criteria for cluster evaluations

The value of optimized cut-off was the important criteria to ensure the clustering analysis correctly and meaningfully. The value of the optimized one must have the number of the managed cluster (not devided the conformations into many clusters), must have the large number of similar conformations grouped into the smaller cluster and minimized the number of clusters possessing one conformation. Thus, the cut-off was chosen which satisfied three criterion as followings[21]: the first cluster has at least 70% conformations; the total of first ten clusters must have at least 80% of all formed conformations; and each cluster of the first ten clusters must have at least 20 conformations. With the criteria, RMSD cutoff was chosen of 1.15 Å for SGLT1 and of 1.25 Å for SGLT2. The representative conformations of more than 80% of all the formed conformations were exported and used for pharmacophore generation.

2.4. Structure-based pharmacophore models derived from MDs

Pharmacophore model generation

Pharmacophore models were generated from the results of MDs after selecting the representative conformations of structures of SGLT1 and SGLT2 by using MOE software. Furthermore, structure-based pharmacophore models using the experimental crystal structures of SGLT1 (PDB ID: 7WMV) and SGLT2 (PDB ID: 7VSI) which retrieved from Protein data bank were also constructed to compare each other.

Energy minimization for all conformations of the compounds was conducted by using the Conformation Import tool. Pharmacophore Elucidation tool was used for generation of a 3D-pharmacophore model, and the Pharmacophore Search for Criteria was for assessment process. Initially, the conformations representated from MDs were inputed in the MOE and then they were prepared by removing water, protonization and also modification of the structure and energy minimization. The conformations were then aligned and arranged by the tool *Align/Superpose* and *Pocket Residues* was used to priority arrangement of amino acids at the locations of binding sites.

Pharmacophore model was generated from the all of representative conformations using the tool *Pharmacophore Consensus* in the MOE program with the distance of tolerance of 1.2 Å and the threshold[22] of 50%. This tool automatedly suggested the common characteristics of all the conformations ensuring the distance from the common characteristic to the confor-

Table 2. Key features related to inhibitory bioactivities of SGLT1 and SGLT2

SGLT1	SGLT2
Hydrogen bonds[12] forming from C2-OH, C3-OH glycosid of ligand with Asn78, Phe101, Glu102, Trp291, Lys321	Hydrogen bonds[13] forming from C2-OH, C3-OH glycosid of ligand with Asn75, Phe98, Glu99, Trp291, Lys321
π-π (stacking)[12] with His83	π-π (stacking)[13] with His80
Hydrophobic interactions[12] with Ile98, Phe101	π-π (T-shape)[13] with Phe98 and hydrophobic interactions with Leu84, Val95, Phe453
Amid-π interactions[12] with Asp454; Leu274, Asp454 or hydrophobic interactions between methylsulfanyl group with Met283, Thr460 (regarding to biological acitivty with SGLT1)	Hydrophobic interactions[13] with Phe98, Leu274

Table 3. Testing database for SGLT1 and SGLT2.

Number of compounds in the testing database		SGLT1	SGLT2
Number of active compounds		60 ($IC_{50} \leq 1$ µM)	332 ($IC_{50} \leq 100$ nM)
Number of inactive compounds	Based on IC_{50}	107 ($IC_{50} > 1$ µM)	103 ($IC_{50} > 1$ µM)
	Using DUD-E	556	750
	Total of inactive compounds	663	853

mational charactetristic of not more than 1.2 Å and the common characteristic representing for more than 50% the characteristics of conformations. Among the forming common characteristics, some have inhibitory bioactivity effects on both SGLT1 and SGLT2 (Table 2). Moreover, hydrogen bonds had the high occupancy during MDs were selected for generation of pharmacophore models based on default settings.

Pharmacophore model validation

Two pharmacophore models were evaluated through using testing database by the criteria such as accuracy, overlap, sensitivity (Se), and specificity (Sp), precision (Acc), Güner-Henry score and enrichment factor. If a greater number of active substances satisfy the model, the model's reliability increases, and vice-versa. The test set was evaluated based on the two pharmacophore models, including the active and inactive test datasets[23-52] (Table 3).

The inactive dataset included the compounds posessing the IC_{50} values and decoy compounds belonging to DUD-E[53]. The values of IC_{50} of compounds targeting SGLT1 and SGLT2 were collected[23-52] and converted in terms of a reference compound, sotagliflozin (SGLT1: the IC_{50} = 34.6 nM and SGLT2: the IC_{50} = 1.5 nM)[40]. To perform the conversion, a threshold is established using the IC_{50} of the reference compound. Compounds with IC_{50} values equal to or lower than that of the reference are classified as active, while those with higher IC_{50} values are considered inactive. The IC_{50} values of compounds in the test set were determined through *in vitro* testing for inhibitory ability of absorption of [14*C*] methyl D-glucopyranoside (AMG) in Chinese hamster (CHO) or monkey kidney cells (COS) expressing on human SGLT1 and SGLT2.

Low conformations of the compounds in test set were constructed using *Conformation Import* in the MOE with below settings[54]: Conformation: 10000; Descriptor Filter: Clear; Stochastic Search Iteration Limit): 1000; Energy Minimization Iteration Limit: 1000; Energy Minimization Gradient Test: 0,0001.

2.5. Virtual screening

Pharmacophore Search in the MOE was used for screening process with the database of 274 compounds by the two generated pharmacophore models. All compounds were energy minimized and their conformations were generated prior to screening. The compounds satisfied both these pharmacophore models were chosen for further docking.

2.6. Molecular docking

MOE software was used for redocking and docking, with scores calculated using the London dG scoring function.

Redocking

Redocking was conducted for the complexes of SGLT1 and SGLT2 obtained from clustering analysis of MDs results to evaluate the reproducability of binding interactions as well as to compare to the redocking results of the initial structures from the PDB (PDB codes of SGLT1: 7WMV and SGLT2: 7VSI). Redocking criteria were docking scores ($kcal.mol^{-1}$), binding interactions with RMSD values < 2 Å ligand after docking and the initial ligand.

The representative conformations of the first cluster on the two targets were the input imported into the MOE. Using *Sequence Editor* to remove the ligand and solvents and using *Protonate 3D* in the MOE to protonate the structures of SGLT1 and SGLT2 in the condition of pH = 7.5 and T = 300 K were conducted. The Ligand Atom was used instead of Site Finder in the MOE to define the binding sites of SGLT1 and SGLT2 with the pocket sizes of 235 $Å^3$ and 115 $Å^3$, respectively.

Docking

The ligands for screening from the pharmacophore models were imported and energy minimized and also added hydrogens using *Energy Minimized* with the Gradient of 0.0001. Induced-Fit Docking was conducted through using developed pharmacophore models and the *Dock* MOE was used for docking with the default parameters, as followings: Receptor: MOE; Site: Dummy Atoms; Placement: Pharmacophore – Timeout (Second): 3600; Number of Return Poses: 50000; Refinement: Induced Fit – Side chains: Free. The docking was also conducted for the reference compound, sotagliflozin (a dual inhibitor of SGLT1 và SGLT2) with the same parameters. Docking criteria included docking score ($kcal.mol^{-1}$), binding interactions, in particular binding interactions with the

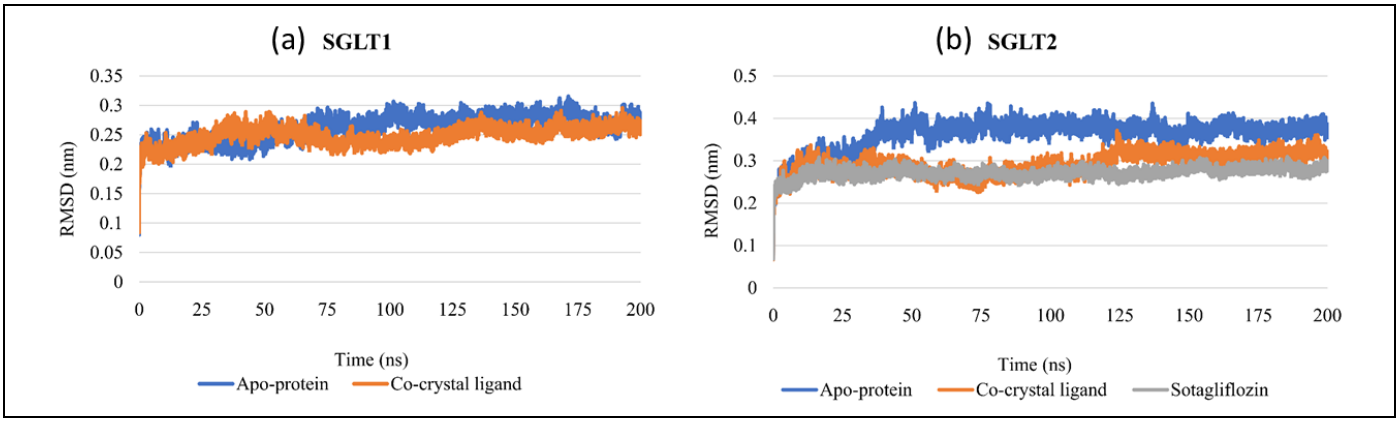

Figure 2. RMSD plots of backbone proteins in the apo structure and complex with the co-crystallised ligands: (a) SGLT1 (PDB code: 7WMV) and (b) SGLT2 (PDB code: 7VSI) during 200 ns of MDs.

key residues of the structures of SGLT1 and SGLT2. The compounds were evaluated the absorption-delivery-metabolism-excretion and toxicity (ADMET) using SwissADME (http://www.swissadme.ch/) with the parameters such as logP, logS, Lipinski's rule of five and toxicity.

3. RESULTS AND DISCUSSION

3.1 Pharmacophore model generations derived from MDs

Structure-based pharmacophore models, one of the common *in silico* techniques play important roles in drug discovery. Traditionally, structure-based pharmacophore models rely on the initial structure of protein target taken from Protein data bank. However, there are some concerns which affects the protein structure such as static structure, the co-crystallised ligand, non-physiological contacts between protein and ligand, solvent effects, etc. Thus, the correct structure of protein-ligand complex could be obtained by refinement process using molecular dynamics simulations. Based on the movements of protein or complex of protein-ligand during the MDs trajectory, the important protein-ligand interactions for dynamic structure were obtained for generating pharmacophore models. Pharmacophore derived from MDs structure of a protein-ligand complex ensures covering all dynamic of features patterns of structure.

In this study, pharmacophore models were developed from cluster analysis of structures taken from MDs and also compared with the corresponding pharmacophore models generated from the initial experimental protein–ligand structure retrieved from PDB.

MDs results: During 200 ns MDs, the co-crystallised complexes of SGLT1 and SGLT2 with their ligands shown to be stable by the evaluation of parameters including RMSD, RMSF, Rg and SASA. The RMSD plots in Figure 2 demonstrated the complexes of SGLT1 and SGLT2 with their co-crystallised ligands as well as the apo proteins (free ligands) reached a stable plateau quickly after 1 ns of MDs (RMSD < 2 Å). This rapid attainment of equilibrium indicated the structural stability of both the apo-protein and complex structures within the simulated environment. Furthermore, it suggested that the interactions between the protein and the co-crystallized ligand were robust and capable of maintaining the essential structural features of the complex, thereby enabling the selection of stable protein

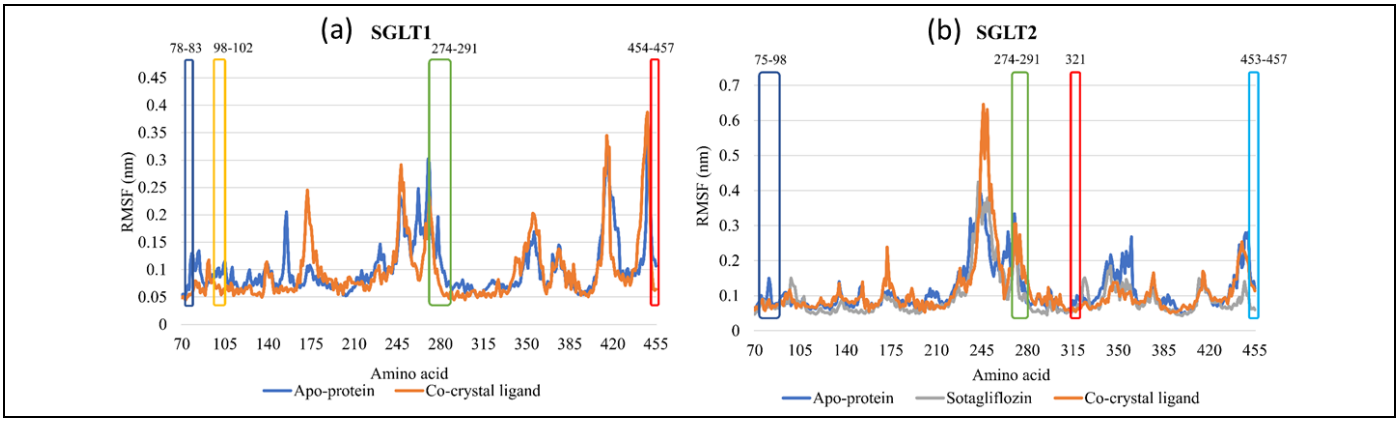

Figure 3. RMSF plots of the apo protein and the complexes with the co-crystallised ligands: (a) SGLT1 (PDB code: 7WMV) and (b) SGLT2 (PDB code: 7VSI).

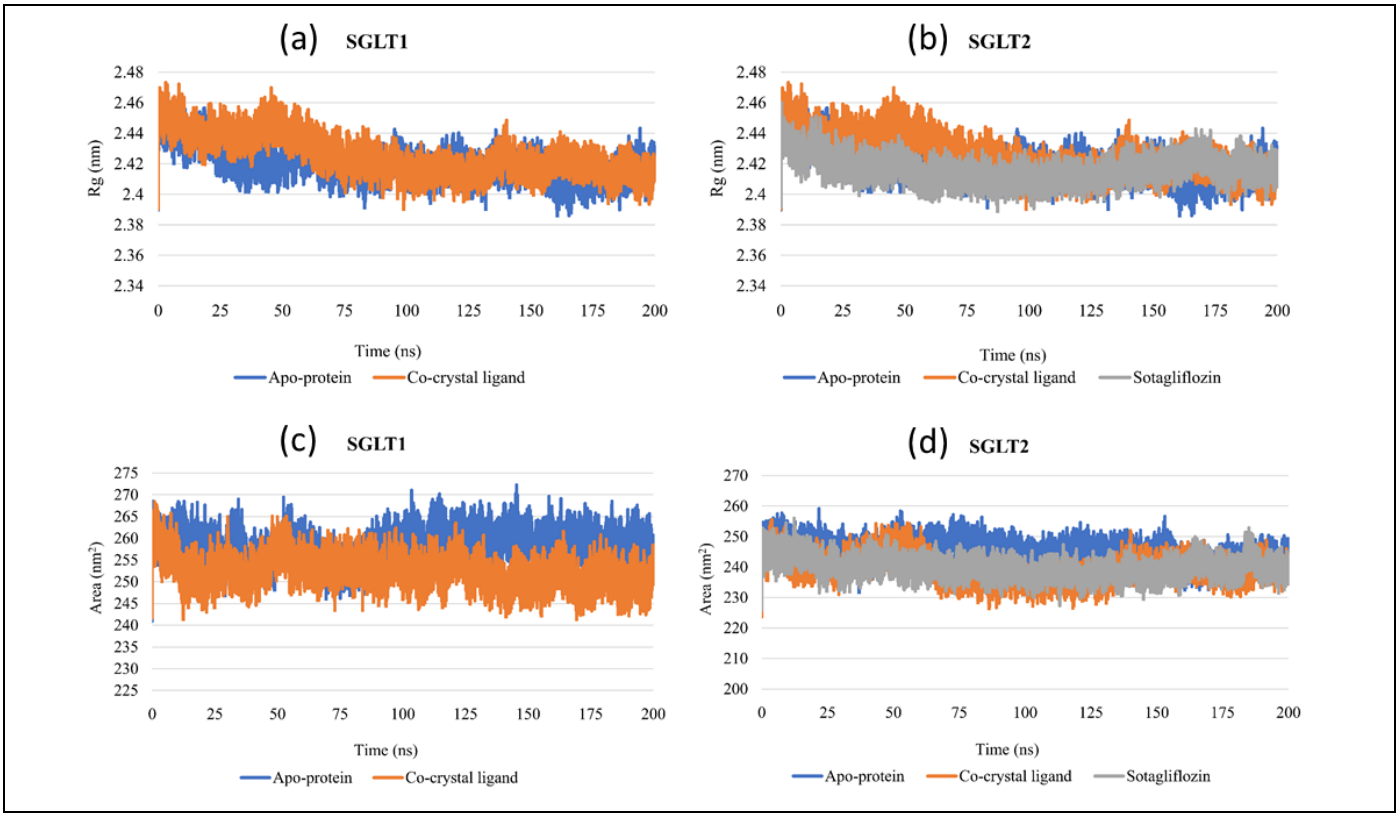

Figure 4. Rg and SASA plots of the apo protein and the complex with the co-crystallised ligand: (a) and (c): Rg and SASA values of SGLT1 (PDB code: 7WMV); and (b) and (d) Rg and SASA values of of SGLT2 (PDB code: 7VSI).

conformations for both therapeutic targets. This stability was further corroborated by the subsequent analyses of RMSF, Rg, SASA, and Percentages of hydrogen bond occupancy." Furthermore, it suggested that the interactions between the protein and the co-crystallized ligand were robust and capable of maintaining the essential structural features of the complex, thereby enabling the selection of stable protein conformations for both therapeutic targets. This stability was further corroborated by the subsequent analyses of RMSF, Rg, SASA, and percentages of hydrogen bond occupancy.

For flexibility of each residue, in both structures of SGLT1 and SGLT2, the values of RMSF of the Cα atoms of all of the acid amines at the binding pocket of these structures were lower than the values in the apo forms (Figure 3). That means the SGLT1 and SGLT2 were stable during 200 ns of MDs after binding ligands.

The values of Rg of the apo protein and the co-crystallised complex structures of SGLT1 và SGLT2 showed these structures were stable within 2.45 nm (Figure 4). In both structures of SGLT1 and SGLT2, the SASA values of the apo-protein and the complexes ranged from 258 nm² and 253.5 nm² for SGLT1 and 246.8 nm² và 242.8 nm² for SGLT2 (Figure 4).

Percentages of hydrogen bond occupancy of protein-ligand interactions during MDs were also calculated. For SGLT1, the co-crystallised ligand formed many hydrogen bonds with the amino acid residues of the binding site such as Glu102 at carbonyl group (with percentage of 90.64% at C3-OH and 41.48% at C2-OH) as well as hydroxyl group (with percentage of 34.91% at C2-OH và 23.09% at C3-OH).

These interactions were compatible with the key binding interactions regarding the inhibitory activity of SGLT1. However, the native ligand in the crystal complex structure, LX2761 also interacted with SGLT1 by hydrogen bonds with the high occupancy at Ser94 (81.23%) and Thr287 (69.89%) (Table 4).

For SGLT2, similar to LX2761, the co-crystallised ligand, empagliflozin also bonded many hydrogen bonds with high occupancy at C2-OH and C3-OH glycosid of empagliflozin with Asn75 (71.33%),

Table 4. Percentages of hydrogen bond occupancy of the co-crystallised ligand with the residues of the structure of SGLT1 (PDB code: 7WMV).

No.	Hydrogen bond		Percentage of hydrogen bond occupancy (%)
	Donor	Acceptor	
1	Ligand-Side	Glu102-Side	378.52
2	Lys321-Side	Ligand-Side	182,33
3	Ser94-Side	Ligand-Side	144.49
4	Ligand-C4 (-OH)	Thr287 (-OH)	95.21
5	Ligand-Side	Lys321-Side	75.95

Table 5. Percentages of hydrogen bond occupancy of the co-crystallised ligand with the acid amines of the structure of SGLT2 (PDB code: 7VSI).

No.	Hydrogen bond		Percentage (%)
	Donor	Acceptor	
1	Lys321-Side	Ligand-Side	146.61
2	Ligand-Side	Asn75-Side	71.81
3	Ligand-Side	Glu99-Side	65.77
4	Ligand-Side	Phe98-Main	53.67
5	Ligand-Side	Lys321-Side	53.49
6	Asn75-Side	Ligand-Side	50.08

Glu99 (56.64%) and Phe98 (51.08%) of SGLT2 (Table 5. Percentages of hydrogen bond occupancy of the co-crystallised ligand with the acid amines of the structure of SGLT2 (PDB code: 7VSI).

Cluster analysis

Clustering analysis was conducted to identify the representative conformations for SGLT1 và SGLT2 using the equilibrium structure with the values of RMSD of backbone from 1 ns to 200 ns and the conformations were saved every 0.01 ns. The conformations of SGLT1, SGLT2 were devided into 17 clusters and 20 clusters, respectively. Of which, 4 first clusters of SGLT1 and SGLT2 contains 86.51% and 81.33% (> 80%) over all the conformations, respectively (Figure 5). As the results, 4 conformations representating for each clusters of SGLT1 and SGLT2 were extracted and compared with the original profile presented in Figure 6 and Figure 7.

For SGLT1, in conjunction with the hydrogen bond occupancy results (Table 4), the initial hydrogen bonds particularly those with Glu102 and Lys321 were consistently maintained across most of the representative conformations from the four MD clusters. These interactions have previously been associated with enhanced binding affinity toward SGLT1. Notably, additional hydrogen bonds emerged and were stably maintained during simulations, such as the interaction with Ser94, which appeared in all four clusters, and with Thr287, observed in three out of four clusters. These newly identified interactions, revealed through molecular dynamics simulations, suggested their potential role in stabilizing the ligand–protein complex and contributing to the inhibitory activity.

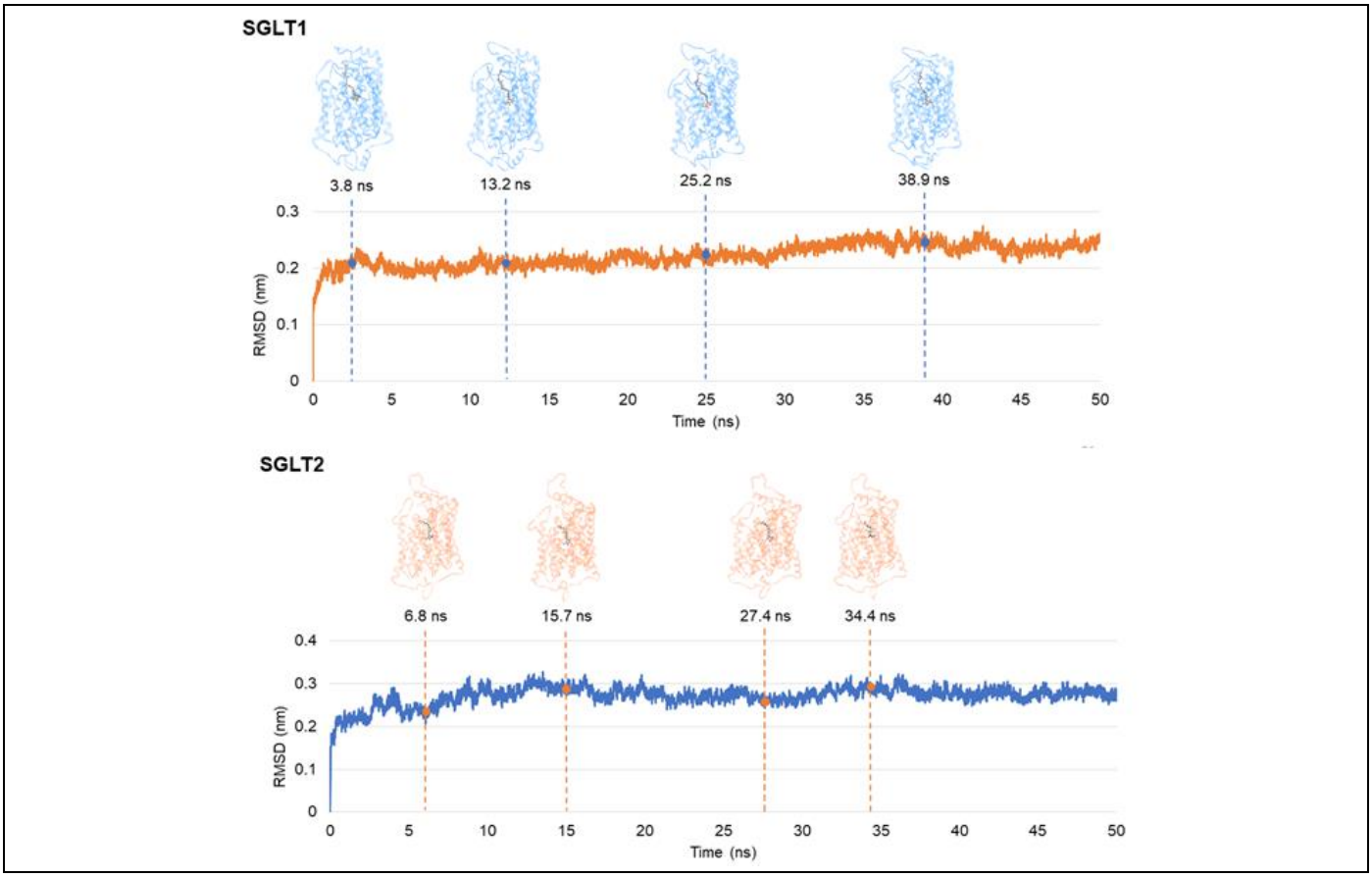

Figure 5. Representative conformations for 4 clusters with cluster 1 (A), cluster 2 (B), cluster 3 (C) and cluster 4 (D) of the SGLT1 (blue color) and SGLT2 (orange color).

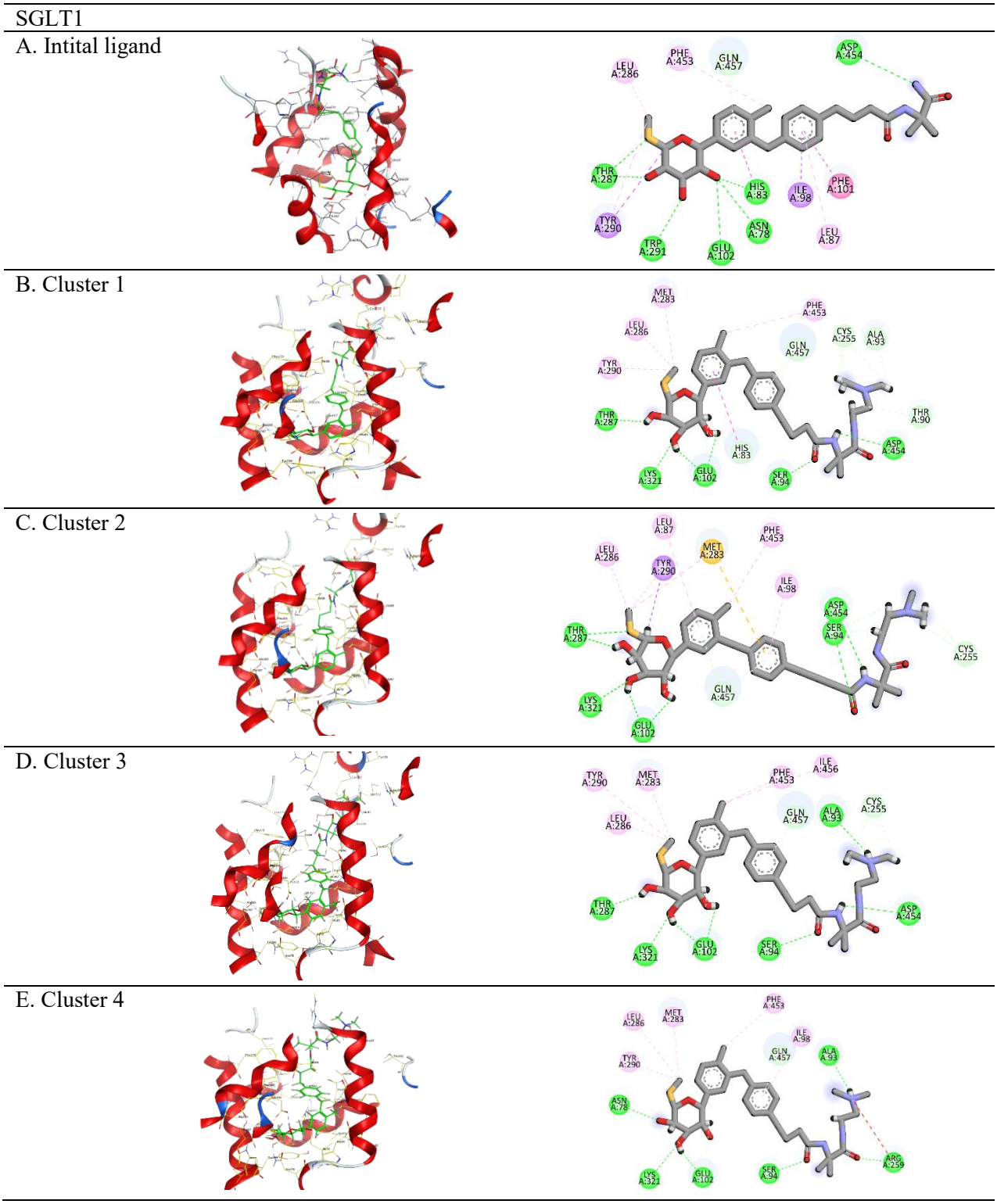

Figure 6. Comparison of interactions between the initial conformation SGLT1 (A) and the representative conformations SGLT1 for Cluster 1 (B), Cluster 2 (C), Cluster 3 (D), Cluster 4 (E). In which, hydrogen bonds were marked green, π-π interactions were marked purple/pink, π - sulfur interactions were marked yellow

For SGLT2, based on the hydrogen bond occupancy analysis (Table 5), the typical hydrogen bonds involving key residues such as Asn75, Phe98, Glu99, and Lys321 were consistently maintained with high occupancy and observed across all four representative conformations from the MD-derived clusters. These interactions had been reported to play crucial roles in stabilizing the ligand within the binding pocket and locking the active conformation of SGLT2, thereby preventing glucose reabsorption.

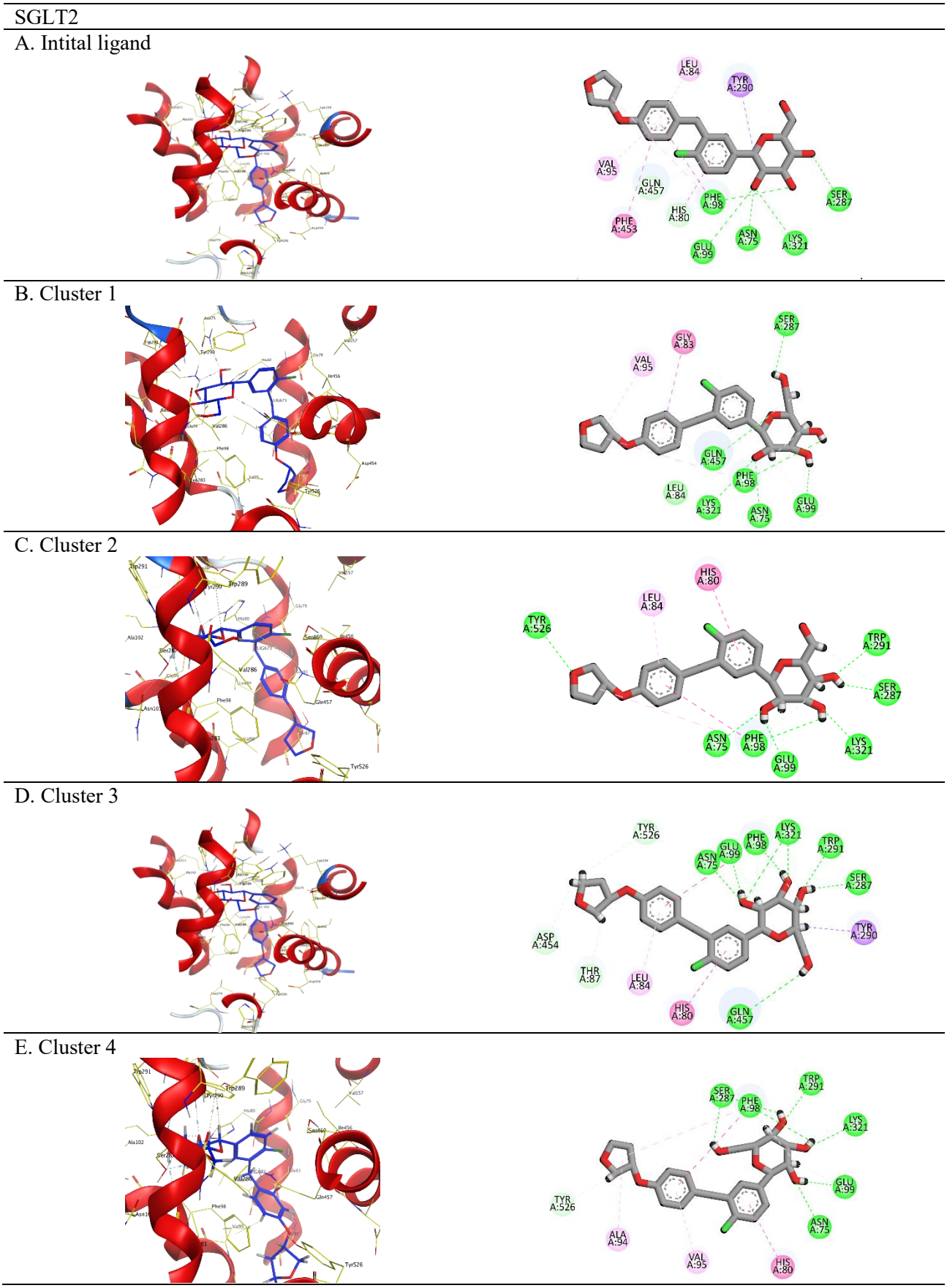

Figure 7. Comparison of interactions between the initial conformation SGLT2 (A) and the representative conformations SGLT2 for Cluster 1 (B), Cluster 2 (C), Cluster 3 (D), Cluster 4 (E). In which, hydrogen bonds were marked green, π-π interactions were marked purple/pink, π - sulfur interactions were marked yellow

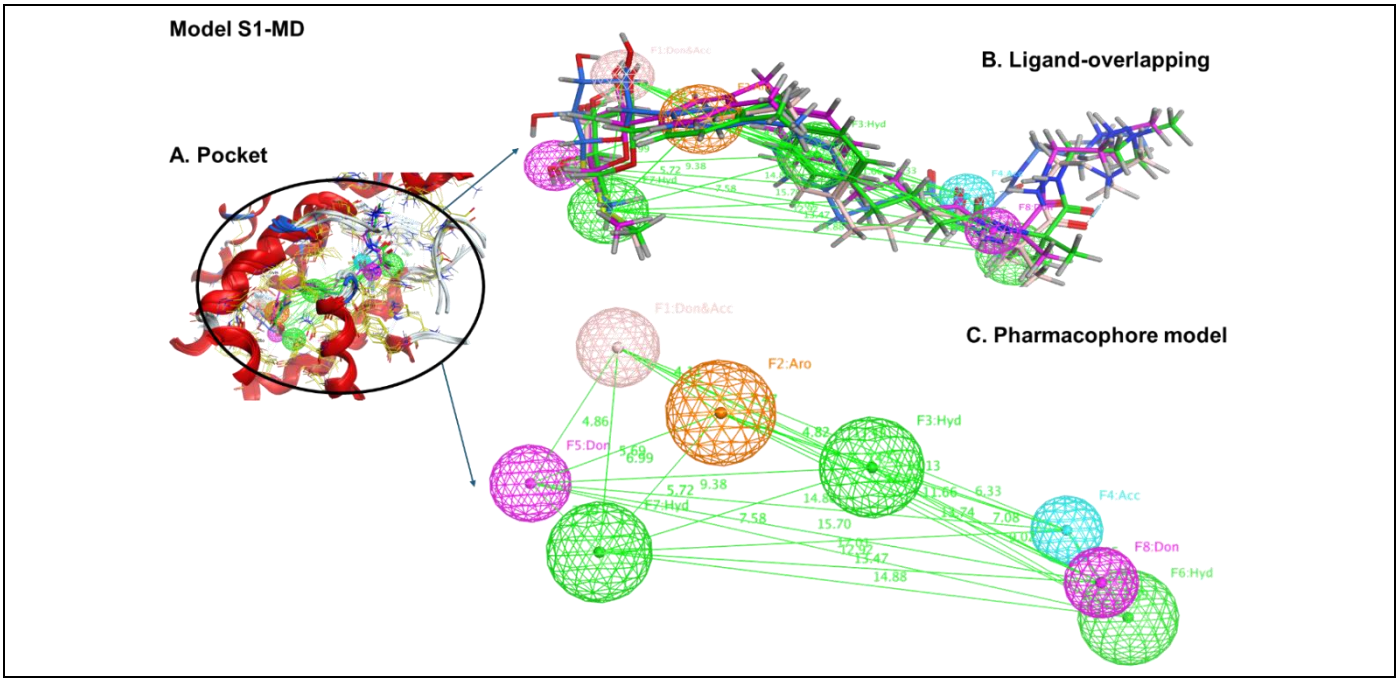

Figure 8. MD-derived pharmacophore model for SGLT1 (S1-MD), including: (A) pharmacophore model within the binding pocket; (B) pharmacophore model overlaid with ligand conformations from the four representative MD clusters; (C) refined pharmacophore model consisting of three essential features and two constraint features selected from five identified pharmacophoric features. The light pink sphere represented hydrogen bond donors/acceptors, purple spheres indicate hydrogen bond donors, light blue spheres was hydrogen bond acceptors, green spheres represented hydrophobic features, and orange sphere corresponded to aromatic ring features.

Pharmacophore generation derived MDs

With four representative conformations for each structure of SGLT1 and SGLT2 obtained from MD simulations, pharmacophore models were generated for both targets. The pharmacophoric features derived from MD-based models were also utilized to construct pharmacophore models based on the original PDB structures

For SGLT1: Pharmacophore model for SGLT1 was generated in two versions: one derived from MD simulations (S1-MD, Figure 8) and one from the original

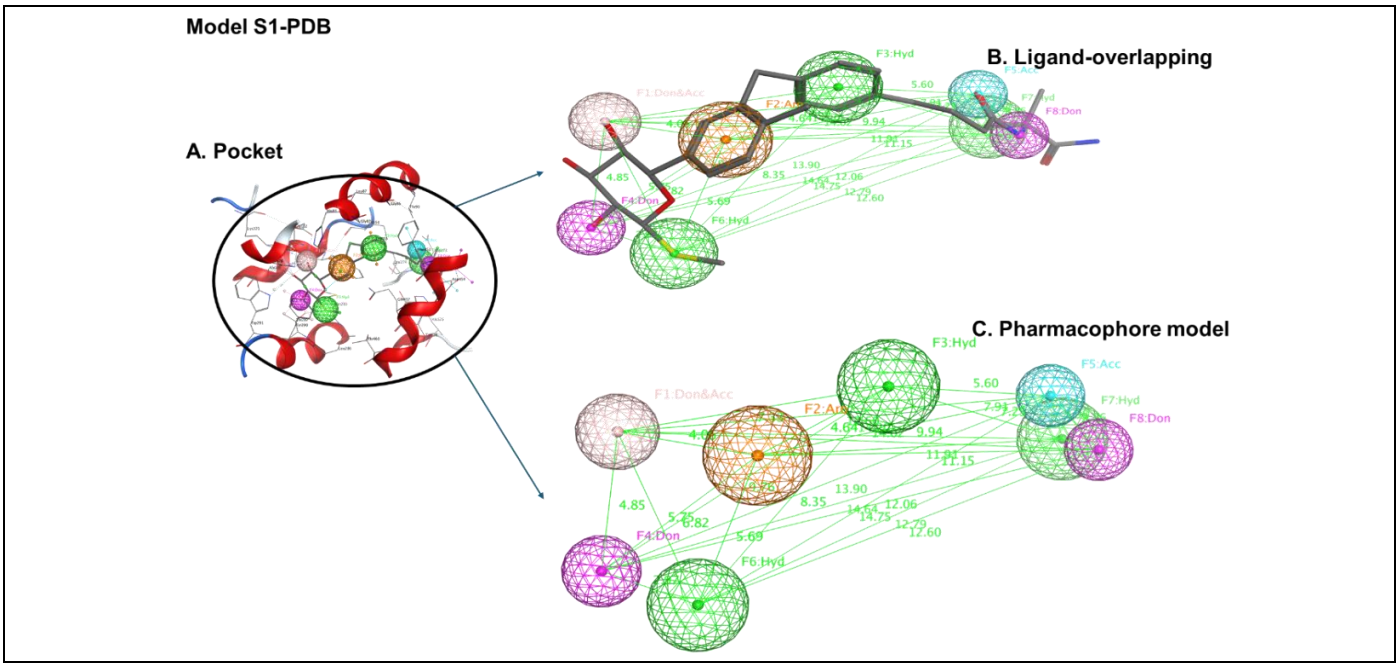

Figure 9. Pharmacophore model from the original crystal structure (S1-PDB) including: (A) pharmacophore model within the binding pocket; (B) pharmacophore model overlaid with ligand conformations from the four representative MD clusters; (C) refined pharmacophore model consisting of three essential features and two constraint features selected from five identified pharmacophoric features. The light pink sphere represented hydrogen bond donors/acceptors, purple spheres indicate hydrogen bond donors, light blue spheres was hydrogen bond acceptors, green spheres represented hydrophobic features, and orange sphere corresponded to aromatic ring features.

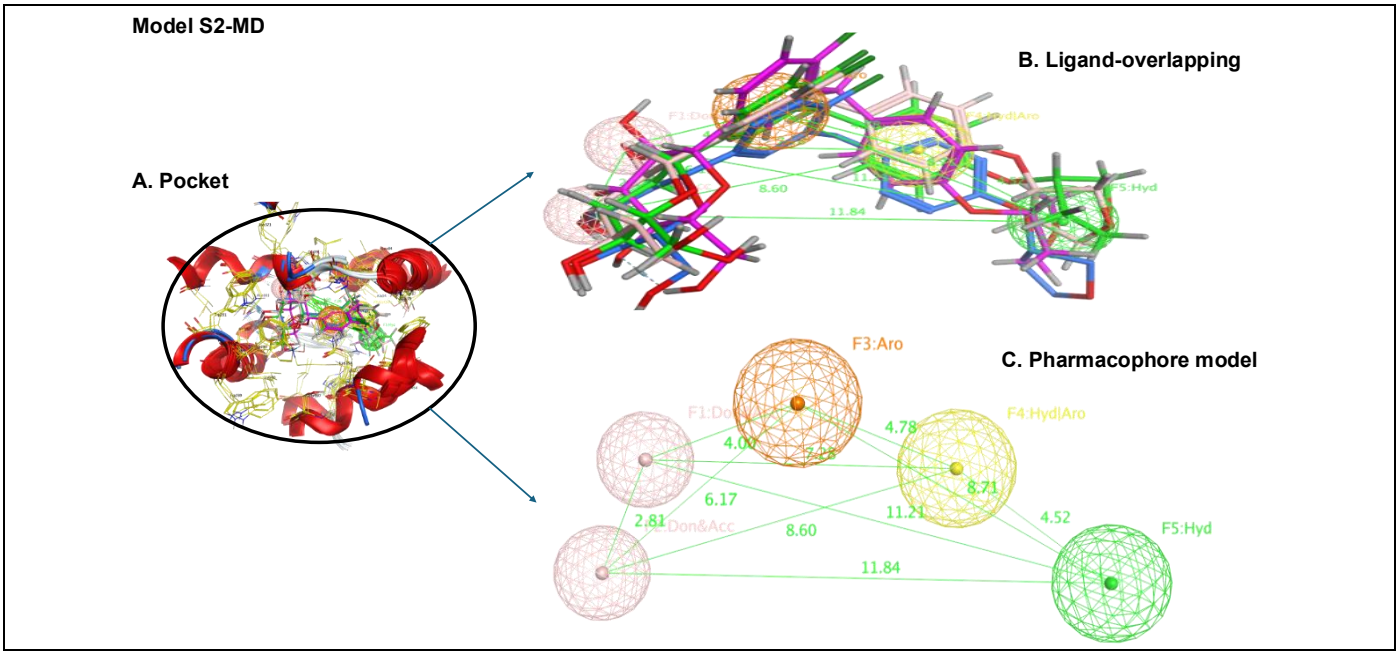

Figure 10. MD-derived pharmacophore model for SGLT2 (S2-MD), including: (A) pharmacophore model within the binding pocket; (B) pharmacophore model overlaid with ligand conformations from the four representative MD clusters; (C) refined pharmacophore model consisting of five features. The light pink spheres represented hydrogen bond donors/acceptors, yellow sphere was aromatic ring features or hydrophobic features, green sphere represented hydrophobic features, and orange sphere corresponded to aromatic ring features.

crystal structure (S1-PDB, Figure 9). Both models shared the same five pharmacophoric features, consisting of three essential points and five constraint points. Of which, 3 essential points included F1 (Don and Acc) for hydrogen bonding groups such as C2-OH, C3-OH glycosid of ligand to acid amines Glu102 and Lys321; F2 (Aro) for aromatic interactions like π-π interactions of benzene group of ligand to His83; and F3 (Hyd) hydrophobic interactions between benzene with Ile98 and Phe101. Two constraint points were as follows: one of two points including F4 (Acc) for hydrogen acceptor like -OH group of Ser94 and F5 (Don) for hydrogen donor with Thr287; and one of three points including F6 (Hyd) for hydrophobic interactions with Leu274; F7 (Hyd) for hydrophobic interactions with methylsulfanyl group of Leu268, Met283, Thr460

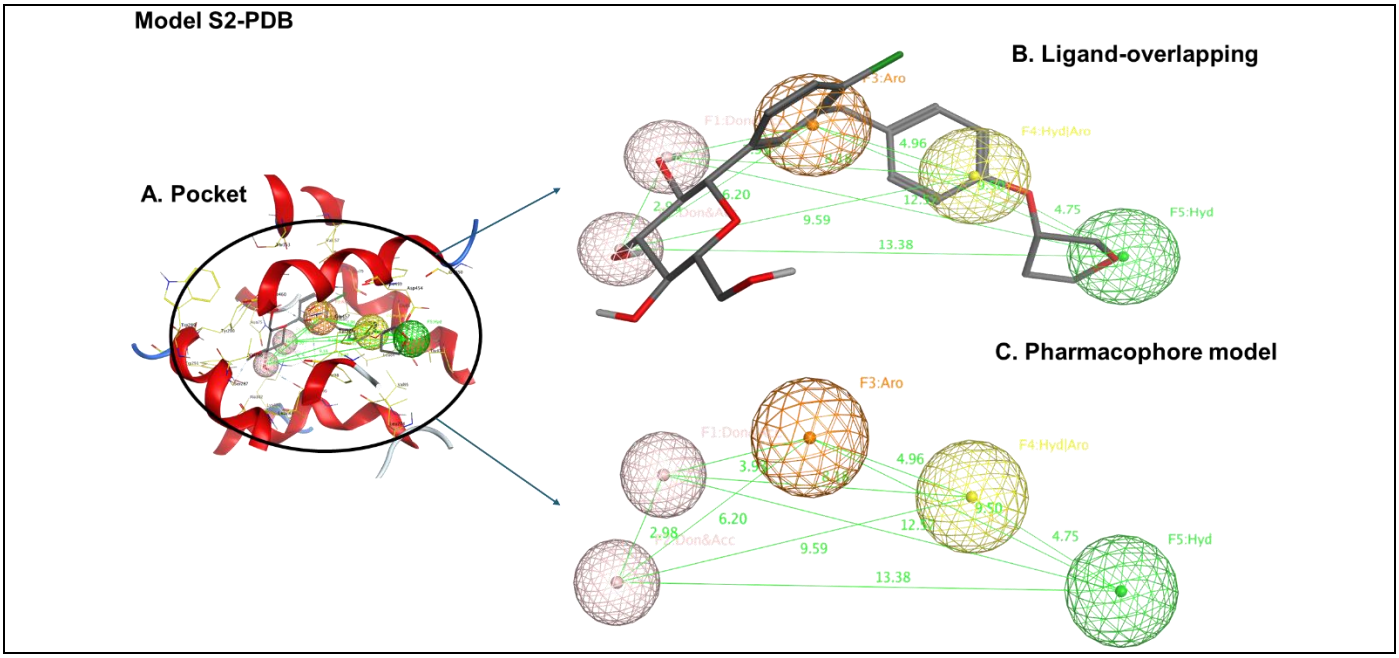

Figure 11. Pharmacophore model from the original crystal structure (S2-PDB) including: (A) pharmacophore model within the binding pocket; (B) pharmacophore model overlaid with ligand conformations from the four representative MD clusters; (C) refined pharmacophore model consisting of three essential features and two constraint features selected from five identified pharmacophoric features. The light pink spheres represented hydrogen bond donors/acceptors, yellow sphere was aromatic ring features or hydrophobic features, green sphere represented hydrophobic features, and orange sphere corresponded to aromatic ring features.

Table 6. Pharmacophore evaluation for SLGT1 and SGLT2 models (S1-MD, S2-MD and S1-PDB, S2-PDB) with test dataset compounds for SGLT1 (active compounds: 60 and inactive compounds: 663) and for SGLT2 (active compounds: 332 and inactive compounds: 853).

Pharmacophore model	SGLT1		SGLT2	
	S1-MD	S1-PDB	S2-MD	S2-PDB
Compounds safisfying the model	142	150	349	343
Active compounds safisfying the model	56	55	297	289
Inactive compounds safisfying the model	86	95	52	54
Inactive compounds not satisfying the model	577	568	801	799
Active compounds not safisfying the model	4	5	35	54
Sensitivity	93.33%	91.67%	89.46%	87.05%
Specificity	87.03%	86.43%	93.90%	93.79%
Accuracy	87.55%	86.86%	92.66%	91.90%
Enrichment factor	4.75	4.42	3.04	3.01
GH score	0.46	0.43	0.81	0.80

and F8 (Don) for hydrogen donor with amide group of Asp454. Each model was designed to select compounds that match all three essential features and at least two of the constraint features (Partial Match: At least 5).

For SGLT2: Pharmacophore model for SGLT1 was generated in two versions: one derived from MD simulations (S2-MD, Figure 10) and one from the original crystal structure (S2-PDB, Figure 11). Both models shared the same five pharmacophoric features, including F1 (Don and Acc: hydrogen bonding forming from C2-OH of ligand with acid amines Phe98 and Lys321); F2 (Don and Acc: hydrogen bonding forming at the site of C3-OH of ligand with acid amines Asn75 and Glu99); F3 (Aro: hydrophobic interactions (π-π) between chlorophenyl group of ligand with acid amines His80); F4 (Aro|Hyd: hydrophobic interactions π-π with Phe98 and hydrophobic interactions with Leu84, Val95 and Phe453); F5 (Hyd: hydrophobic interactions between tetrahydrofuran group of ligand with Phe98 and Leu274).

Pharmacophore evaluations: For S1-MD and S1-PDB models, testing group included 60 active compounds and 663 inactive compounds while for S2-MD and S2-PDB, testing group included 332 active compounds and 853 inactive compounds (Table 6). Although both MD-derived and PDB-based pharmacophore models share identical feature types, the spatial arrangement and inter-feature distances differ due to the conformational flexibility captured through MD simulations (Figure 8, 9, 10, 11). This conformational diversity allowed MD-based models (S1-MD and S2-MD) to more accurately represent the dynamic nature of the binding pocket, which might explain their improved performance in pharmacophore-based virtual screening compared to static PDB-derived models[10,11].

Table 6 demonstrated that the pharmacophore models built from the clusters of a MDs simulation (S1-MD và S2-MD) were chosen as they showed better ability to distinguish between active and decoy ligand structures, better than the pharmacophore models obtained from the PDB structures.

For SGLT1: the model S1-MD better than the S1-PDB model in terms of sensitivity (93.33% > 91.67%), specificity (87.03% > 86.43%), accuracy (87.55% > 86.86%), Enrichment factor (4.75 > 4.42) and GH score (0.46 > 0.43).

For SGLT2: the model S2-MD better than the S2-PDB model in terms of sensitivity (89.46% > 87.05%), specificity (93.90% > 93.79%), accuracy (92.66% > 91.90%), Enrichment factor (3.04 > 3.01) and GH score (0.81 > 0.80).

Although the differences between MD- and PDB-based models are modest, the consistent improvement across all key performance metrics suggests that incorporating dynamic conformational information contributes meaningfully to virtual screening quality and hit identification.

Table 7. Ten compounds satisfied the pharmacophore features of both SGLT1 and SGLT2

Number	Name	Classify
C155	Neodiosmin	Flavonoid glycosid
C156	Apigenin-7-O-neohesperidosid	Flavonoid glycosid
C157	Neobudofficid	Flavonoid glycosid
C165	Pinoresinol diglucosid	Lignan glycosid
C186	Hispidulin-7-O-rutinosid	Flavonoid glycosid
C203	Daidzin	Flavonoid glycosid
C241	Acid 1,5-dicaffeoylquinic	Organic acid
C258	Apigenin-7-O-rutinosid	Flavonoid glycosid
C261	3',4',5,7-tetrahydroxyflavanon 7-O-(α-L-rhamnopyranosyl-(1->3)-β-D-glucopyranosid	Flavonoid glycosid
C270	Syringaresinol-4-O-β-D-glucopyranosid	Lignan glycosid

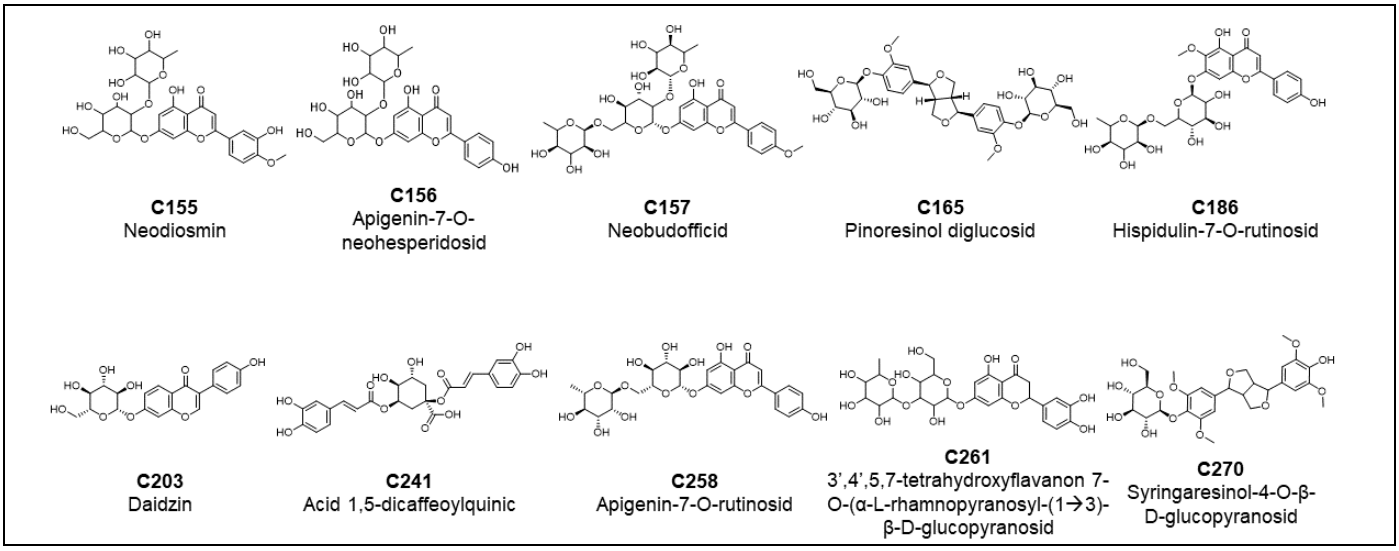

Figure 12. Ten compounds satisfied the pharmacophore features of both SGLT1 and SGL

Virtual screening

Two selected pharmacophore models, S1-MD và S2-MD were used for screening Vietnamese phytochemical compounds database. There were 39 compounds met the criteria of the S1-MD and 12 compounds satisfying the S2-MD models. Of which, 10 compounds (Figure 12 and Table 7) including 7 flavonoid glycosid, 2 lignan glycosid and 1 organic acid) satisfied all of pharmacophore features of both SGLT1 and SGLT2. The flavonoid were also proven in different studies to possess inhibitory acitivity on SGLT1 and SGLT2 *in vitro* as follows: (−)−kurarinon (IC_{50} = 10.4 µM for SGLT1 and IC_{50} = 1.7 µM for SGLT2)[5] and sophoraflavanon G (IC_{50} = 18.7 µM for SGLT1 và IC_{50} = 4.1 µM for SGLT2)[5] while lignan glycosid and organic acids have not been any bioassays for the inhibitory activity on SGLT1 and SGLT2.

Induced-fit molecular docking

Before docking, the process of redocking were conducted using MOE program in order to evaluate the compatability of structures and the programs. The results showed that the MOE program was suitable for molecular docking for SGLT1 and SGLT2 structures with the good compatability and high reproduction of the initial structure.

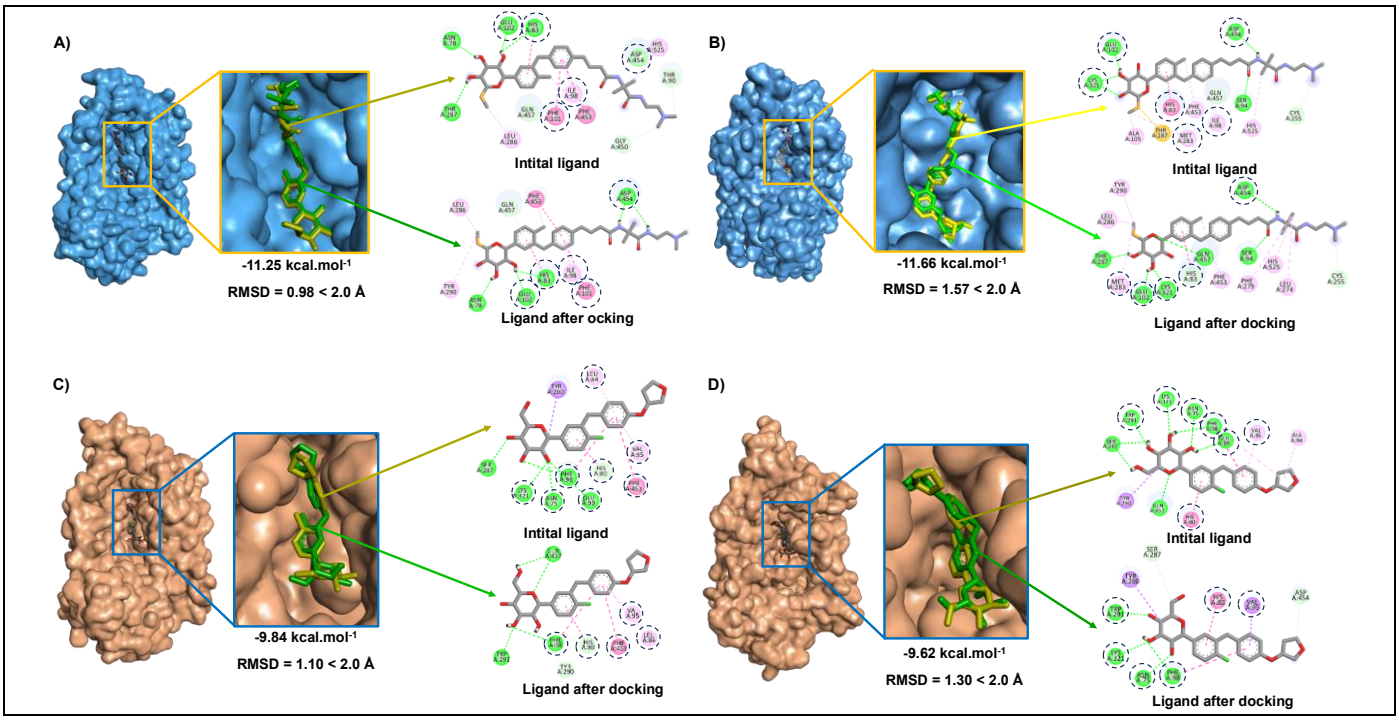

Figure 13. Redocking for the structure obtained from the PDB structures: A) SGLT1 with PDB code: 7WMV and B) the molecular dynamics simulation of SGLT1; and C) SGLT2 with PDB code: 7VSI and D) the molecular dynamics simulation of SGLT2.

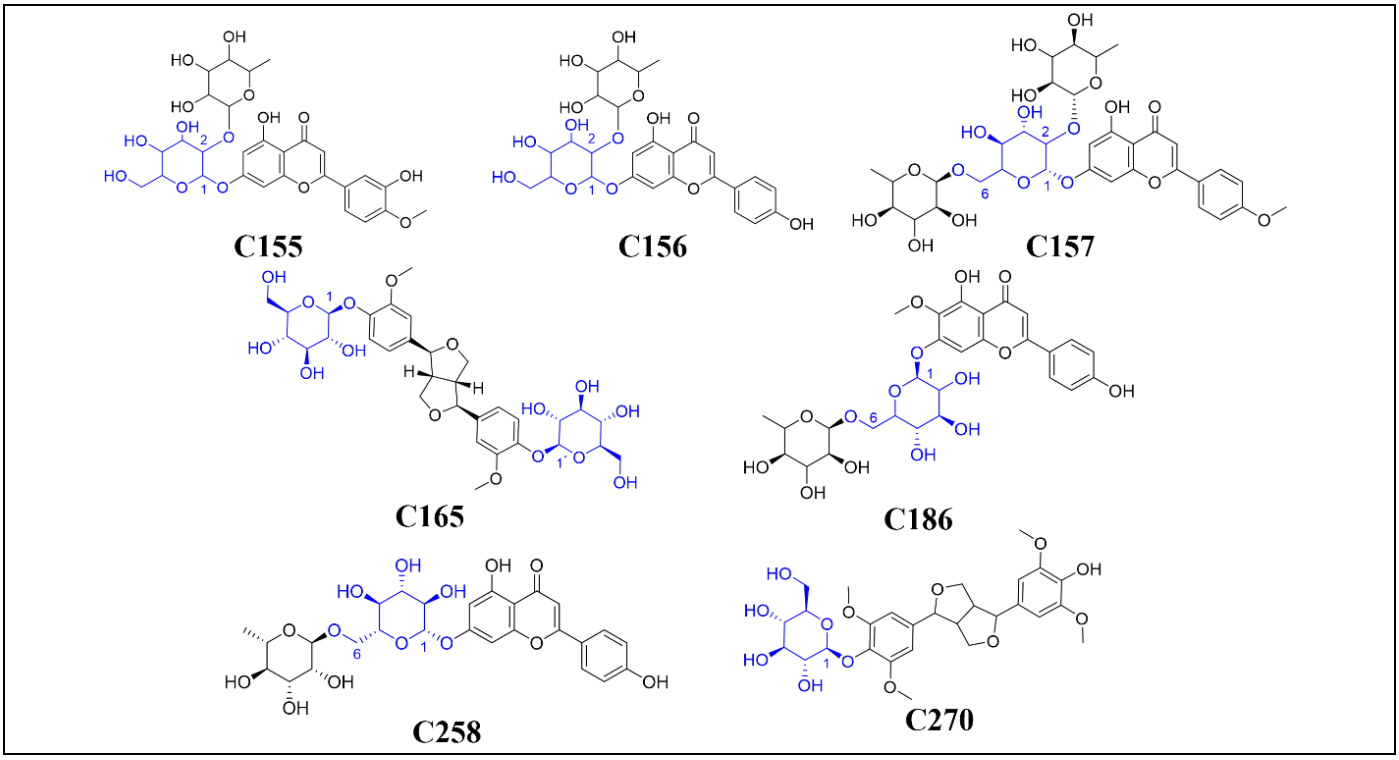

Figure 14. Molecular structure of 7 potential compounds as dual-inhibitors of SGLT1, SGLT2

For SGLT1, in the both structures of SGLT1 obtained from MDs and the PDB structure (Figure 13), the co-crystallised ligand (LX2761) fitted well in the binding pocket of the protein with good binding affinity and binding interactions, particularly the value of RMSD between the ligand after docking and the co-crystallised ligand of less than 2,0 Å (RMSD = 1.57 Å, docking score = -11.66 kcal.mol^{-1}) and for PDB

Table 8. Docking scores and binding interactions of 7 potential compounds and the reference compound into the structures of SGLT1 and SGLT2, with HB: hydrogen bond.

Compound	Binding affinity (kcal.mol^{-1})		Binding interactions	
	SGLT1	SGLT2	SGLT1	SGLT2
C155	-11.75	-11.94	HB: Glu102, Thr287 π-π: His83	HB: Asn75, Glu99 π-π: His80, Phe98 π-alkyl: Leu84, Val95
C156	-10.07	-11.25	HB: Glu102, Lys321, Thr287 Alkyl interaction: Met283 π-alkyl: Ile98	HB: Phe98, Glu99, Lys321 π-π: Phe98 π-alkyl: Leu84, Val95
C157	-13.67	-13.90	HB: Glu102, Lys321 π-π: His83	HB: Phe98, Glu99, Lys321 π-π: His80, Phe98 π-σ: Leu84, Val95
C165	-11.64	-11.23	HB: Glu102 π-π: His83 alkyl interactions: Ile98, Phe101, Met283	HB: Asn75, Glu99 π-π: Phe98 alkyl interactions: Val95
C186	-11.34	-10.56	HB: Glu102, Lys321, Thr287 π-π: His83 π-σ: Ile98	HB: Asn75, Glu99 π-π: His80, Phe98 π-alkyl: Val95 π-σ: Leu84
C258	-10.43	-12.20	HB: Glu102, Lys321 π-π: Phe101 π-alkyl: Ile98, Met283	HB: Asn75, Glu99 π-π: His80 π-alkyl: Phe98
C270	-11.47	-11.61	HB: Ser94, Glu102, Lys321, Thr287 π-σ: Ile98	HB: Asn75, Phe98, Glu99, Lys321 π-π: Phe98, Phe453 alkyl interactions: Val95, Leu274 π-alkyl: Phe98
Sotagliflozin	-9.50	-10.26	HB: Glu102, Lys321, Thr287 π-π: His83 π-alkyl: Ile98	HB: Asn75, Phe98, Glu99, Lys321 π-π: His80, Phe98 π-σ: Leu84 alkyl interactions: Val95

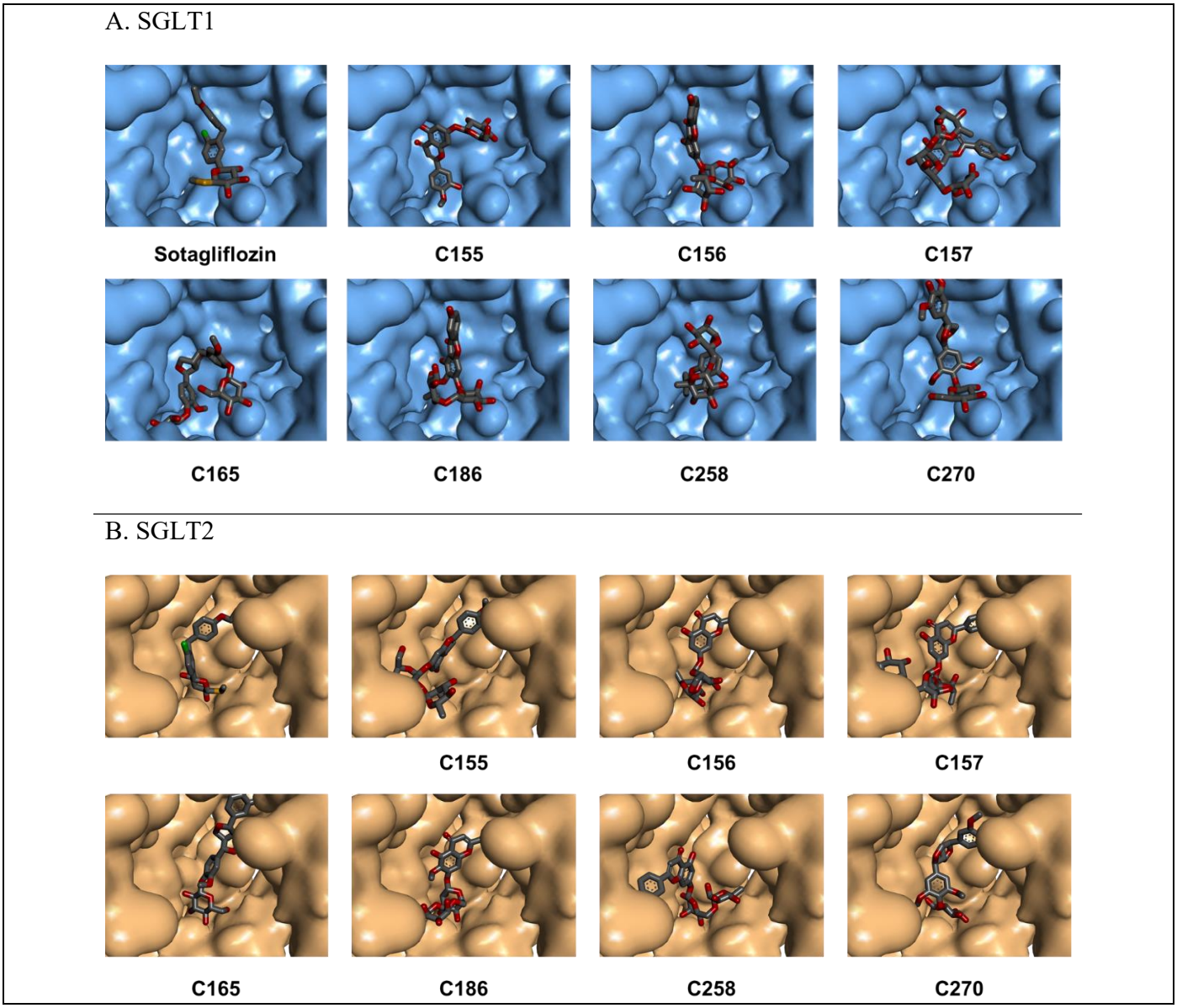

Figure 15. Reference compound and seven potential compounds in the binding site of SGLT1 (A) and SGLT2 (B)

structure (RMSD = 0.98 Å, docking score = -11.25 kcal.mol^{-1}). Similarly, in the Figure 13, the co-crystallised ligand (empagliflozin) bound well in the binding pocket of the protein SGLT2 with good docking scores for both the structures and the value of RMSD between the ligand after docking and the co-crystallised ligand of less than 2.0 Å: with structure obtained from MDs (RMSD = 1.30 Å, docking score = -9.62 kcal.mol^{-1}) and the structure retrieved from PDB (RMSD = 1.10 Å, docking score = -9.84 kcal.mol^{-1}). Therefore, the structures obtained from MDs for both SGLT1 and SGLT2 proteins were chosen for further screening process.

Among 10 compounds docked into the binding sites of SGLT1 and SGLT2, respectively, there were 7 compounds (C155, C156, C157, C165, C186, C258 và C270-Figure 14 and Table 8) fitted well in the structure of SGLT1 and SGLT2 with their good binding affinities

Table 9. The values of logP, logS and the evaluation results of these seven compounds following ADMET and Lipinski's rule.

Compound	Parameters			
	LogP	LogS	Lipinski's rule of five	Toxicity
C155	-0.41 (hydrophilic)	-5.27 (poorly soluble)	Violation of 3	No
C156	-0.66 (hydrophilic)	-4.19 (poorly soluble)	Violation of 3	No
C157	-1.04 (hydrophilic)	-4.74 (poorly soluble)	Violation of 3	No
C165	-1.35 (hydrophilic)	-3.12 (poorly soluble)	Violation of 3	No
C186	-0.45 (hydrophilic)	-4.70 (poorly soluble)	Violation of 3	No
C258	-0.41 (hydrophilic)	-4.54 (poorly soluble)	Violation of 3	No
C270	0.65 (hydrophobic)	-3,66 (highly soluble)	Violation of 2	No

ranged from -10.07 to -13.67 kcal.mol^{-1} (for SGLT1) and -10.56 đến -13.90 kcal.mol^{-1} (for SGLT2). These compounds were selected as potential dual inhibitors for SGLT1 and SGLT2. Three remaining compounds C203, C241 và C261 could not get into the two structures. This could be explained that due to their nature structure, such as C203 (flavonoid daidzein) possessing isoflavone instead of flavanone like the seven compounds; C241 having acid quinic in the center but the caffeoyl group owning the long chain carbon, so the benzene in this group not being in the convenient area for forming π-π interactions with the acid amins residues of the binding pocket; C261 with glucopyranose haiving both tetrahydroxyflavanon and rhamnopyranose groups at the positions of C1 and C3, hindrance in the adjacent position leading not to form the important interactions with the binding pocket of structures.

Docking result for the reference compound (sotagliflozin as dual SGLT1 and SGLt2 inhibitor) into the SGLT1 and SGLT2 was carried out with the same docking parameters of redocking. The results showed that sotagliflozin could attach the binding pocket of SGLT1 with good docking score (-9.50 kcal.mol^{-1}) and of SGLT2 (docking score = -10.26 kcal.mol^{-1}). Sotagliflozin formed key interactions with both SGLT1 and SGLT2, including hydrogen bonds with Glu102, Lys321, and Thr287, and hydrophobic interactions with His83 and Ile98 in SGLT1, which helped anchor the glucose moiety of the inhibitor within the active site[13]. These interactions ensured tight and specific binding, particularly the hydrophobic contacted that stabilize the inactive conformation of SGLT1, thereby preventing glucose reabsorption. Similarly, for SGLT2, sotagliflozin established hydrogen bonds with Asn75, Phe98, Glu99, and Lys321, as well as hydrophobic interactions with His80, Leu84, Phe98, and Val95. These interactions were critical for maintaining strong binding affinity and effectively locking SGLT2 in its inactive state, contributing to the overall inhibitory effect on glucose[12].

After docking, seven compounds also demonstrated these compounds bound in similar pose with the reference compound (Figure 14 and 15). Seven compounds (Figure 14) have similar characteristics including glycoside possessing many hydroxyl groups (-OH), in particular having at least two adjacent hydroxyl groups helping the structures attached inside the binding pockets leading to form hydrogen bonds and hydrophobic interactions with the key residues of SGLT1 and SGLT2.

Seven compounds are required to be modified for synthesis to be orally administration as they did not comply with the Lipinski's rule of five due to possessing glycoside group. Table 9 performed the values of logP, logS and the evaluation results of these seven compounds following ADMET and Lipinski's rule.

The identification of seven natural compounds as potential dual inhibitors of SGLT1 and SGLT2 aligns with previous research on the therapeutic benefits of simultaneously targeting these both transporters. They exhibited binding affinities and key interactions within the active sites of SGLT1 and SGLT2 that are comparable to sotagliflozin, a known dual SGLT1/2 inhibitor, demonstrated superior glycemic control[1-2], thus providing mechanistic support for their potential efficacy. Notably, the incorporation of pharmacophore models derived from MD simulations ensures that the compounds were screened against dynamic, biologically relevant protein conformations, increasing the likelihood of identifying true binders.

Currently, there have not been any research about bioactivity as well as evaluation of inhibitory bioactivity of these seven compounds (Figure 10) regarding their inhibitory activity on SGLT1 and SGLT2. However, a preliminary biological relevance was explored based on traditional uses and existing pharmacological studies of the medicinal plants from which these compounds were isolated. These include reported hypoglycemic, cardioprotective, and vasodilatory effects, which suggest potential relevance to SGLT modulation. These compounds were isolated from the medicinal plants with some publications related to these compounds. For example, ethanol extract of *Valeriana officinalis* - a medicinal herb with the same properties as *Valeriana hardwickii* had the effect of increasing coronary circulation, reducing heart rate and lowering blood pressure in cats[55]. Aqueous extract of *Eucommia ulmoides* leaves reduced significantly blood sugar levels and increased insulin and peptide-C concentrations in plasma[56]. Besides, compound C165 (pinoresinol diglucoside) had the effect of inhibiting cyclic AMP, dilating blood vessels and increasing coronary blood flow, helping to lower blood pressure *in vitro*[57]. Extract of leaves (*Scoparia dulcis*) had the effect of lowering blood sugar and HbA1c in diabetic mice[48]. Extract of artichoke leaves (*Cynara scolymus*) helped reduce blood glucose levels in mice[54] and type 2 diabetic patients2[55].

Although no experimental data have yet confirmed their direct inhibitory activity against SGLT1 and SGLT2, several of these compounds such as apigenin derivatives, pinoresinol diflucoside, and syringaresinol glucoside have been previously reported to exhibit pharmacological effects such as hypoglycemic, vasodilatory, or cardioprotective properties, which may support their therapeutic potential. The multitarget activity of these compounds may enhance their therapeutic value by addressing not only hyperglycemia but also the oxidative stress and low-grade inflammation commonly associated with type 2 diabetes. Thus, integrating this study's findings

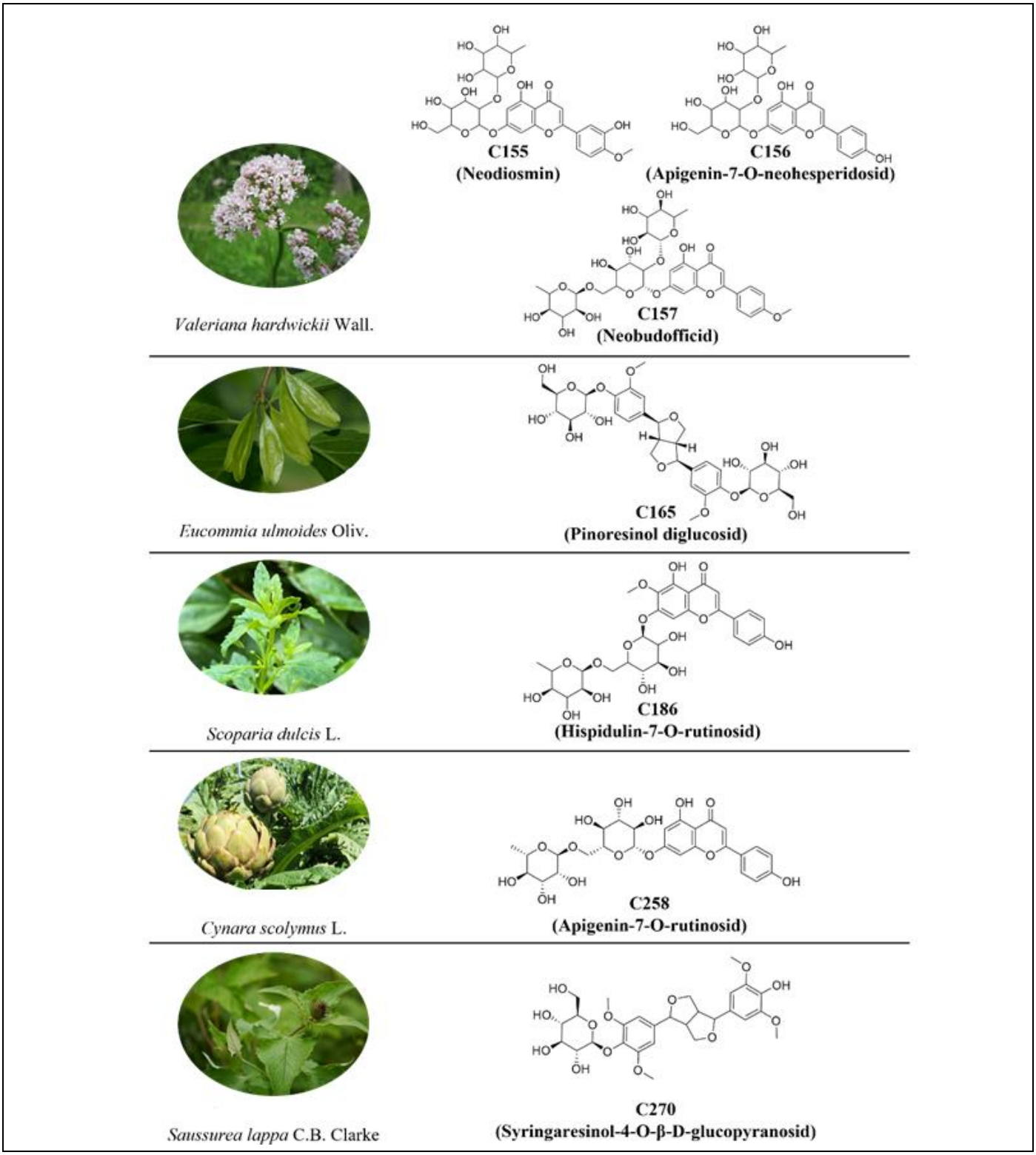

Figure 16. Seven compounds isolated from the medicinal plants as potential inhibitors for SGLT1 và SGLT2

with the existing literature supports the scientific rationale for further investigating these compounds as promising candidates for dual SGLT inhibition, and highlights the value of combining in silico pharmacophore modeling with molecular docking for natural product-based drug discovery. In future work, predictive tools such as QSAR models, cell-based assays, and off-target interaction analysis (e.g., using SwissTargetPrediction) will be incorporated to strengthen the biological validation before conducting wet-lab experiments. This stepwise integration can provide more confidence in candidate prioritization and reduce experimental costs. Further wet-lab experiments are required to evaluate the bioactivity and also inhibitory activities of SGLT1 and SGLT2 of the seven compounds.

4. CONCLUSION

In this study, exploration of potential natural drugs targeting both SGLT1 and SGLT2 proteins for the treatment of type 2 diabetes was of interest. Pharmacophore models derived from molecular dynamics simulations (MDs) for both structures of SGLT1 were successfully constructed to overcome the issues of generating traditional structure-based pharmacophore from PDB structures. Through a virtual screening workflow that included pharmacophore matching and induced-fit docking, seven potential dual SGLT1 and SGLT2 inhibitors were identified. These compounds showed good inhibitories and satisfied the five features of two pharmacophore models for SGLT1 and SGLT2. They also bound well in the binding sites of both proteins by forming key interactions similar to a reference compound, sotagliflozin with their binding affinities varied from -10.07 to -13.67 kcal.mol^{-1} for SGLT1 and from -10.56 to -13.90 kcal.mol^{-1} for SGLT2. Further *in vitro* assays are needed to confirm their bioactivities as dual SGLT1 and SGLT2 inhibitors.

5. ACKNOWLEDGEMENTS

We would like to thank the Department of Pharmaceutical Information Technology at the School of Pharmacy, University of Medicine and Pharmacy at Ho Chi Minh City, Viet Nam for their support.

Author contribution

Khoi Anh Nguyen: Molecular dynamics simulation, pharmacophore model

Phuong Thuy Viet Nguyen: Study design, manuscript editing

Kiet Hoang Anh Nguyen: Molecular docking, virtual screening

Conflict of interest

None to declare.

Funding

None to declare.

Ethics approval

None to declare.

Article info:

Received February 01, 2025

Received in revised form July 05, 2025

Accepted July 11, 2025

REFERENCE

1. Shi Q, Nong K, Vandvik PO, Guyatt GH, Schnell O, Rydén L, et al. Benefits and harms of drug treatment for type 2 diabetes: systematic review and network meta-analysis of randomised controlled trials. BMJ. 2023;381:e074068.
2. Rieg T, Vallon V. Development of SGLT1 and SGLT2 inhibitors. Diabetologia. 2018;61(10):2079-86.
3. Azizogli AR, Vitti MR, Mishra R, Osorno L, Heffernan C, Kumar VA. Comparison of SGLT1, SGLT2, and Dual Inhibitor biological activity in treating Type 2 Diabetes Mellitus. Adv Ther. 2023;6(12).
4. Oranje P, Gouka R, Burggraaff L, Vermeer M, Chalet C, Duchateau G, et al. Novel natural and synthetic inhibitors of solute carriers SGLT1 and SGLT2. Pharmacology research & perspectives. 2019;7(4):e00504.
5. Choi CI. Sodium-Glucose Cotransporter 2 (SGLT2) Inhibitors from Natural Products: Discovery of Next-Generation Antihyperglycemic Agents. Molecules. 2016;21(9).
6. Ramu P, Usharani s, Priyadarshini JF. Isolation, Identification and molecular docking of phytochemical ingredients of Psorlea corylifolia as anti-diabetic agent. JOURNAL OF CRITICAL REVIEWS. 2023;10:159-73.
7. Wang L, Liu M, Yin F, Wang Y, Li X, Wu Y, et al. Trilobatin, a Novel SGLT1/2 Inhibitor, Selectively Induces the Proliferation of Human Hepatoblastoma Cells. Molecules. 2019;24(18).
8. Subbiah U, Athira A, and Venkata Subbiah H. Molecular docking and dynamics simulation of Orthosiphon stamineus against SGLT1 and SGLT2. Journal of Biomolecular Structure and Dynamics. 2023;41(23):13663-78.
9. Schaller D, Šribar D, Noonan T, Deng L, Nguyen TN, Pach S, et al. Next generation 3D pharmacophore modeling. WIREs Comput Mol Sci. 2020;10(4):e1468.
10. Sohn YS, Park C, Lee Y, Kim S, Thangapandian S, Kim Y, et al. Multi-conformation dynamic pharmacophore modeling of the peroxisome proliferator-activated receptor γ for the discovery of novel agonists. Journal of molecular graphics & modelling. 2013;46:1-9.
11. Paul D, Basu D, Ghosh Dastidar S. Multi-conformation representation of Mpro identifies promising candidates for drug repurposing against COVID-19. Journal of Molecular Modeling. 2021;27(5):128.
12. Niu Y, Cui W, Liu R, Wang S, Ke H, Lei X, et al. Structural mechanism of SGLT1 inhibitors. Nature Communications. 2022;13(1):6440.
13. Niu Y, Liu R, Guan C, Zhang Y, Chen Z, Hoerer S, et al. Structural basis of inhibition of the human SGLT2–MAP17 glucose transporter. Nature. 2022;601(7892):280-4.
14. Team GD. GROMACS 2023.2 Source Code. Zenodo. 2023.
15. Bjelkmar P, Larsson P, Cuendet MA, Hess B, Lindahl E. Implementation of the CHARMM Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models. Journal of chemical theory and computation. 2010;6(2):459-66.
16. Environment MO. 2022.02 Chemical Computing Group ULC, 910-1010 Sherbrooke St. W.,. Montreal, QC H3A 2R7, Canada. 2023.
17. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. Journal of computational chemistry. 2004;25(13):1605-12.
18. Zoete V, Cuendet MA, Grosdidier A, Michielin O. SwissParam: A fast force field generation tool for small organic molecules. Journal of computational chemistry. 2011;32(11):2359-68.
19. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. Journal of molecular graphics. 1996;14(1):33-8, 27-8.
20. Daura X, Gademann K, Jaun B, Seebach D, van Gunsteren WF, Mark AE. Peptide Folding: When Simulation Meets Experiment. Angewandte Chemie International Edition. 1999;38(1-2):236-40.
21. Gan JL, Kumar D, Chen C, Taylor BC, Jagger BR, Amaro RE, et al. Benchmarking ensemble docking methods in D3R Grand Challenge 4. J Comput Aided Mol Des. 2022;36(2):87-99.
22. Gentile F, Fernandez M, Ban F, Ton AT, Mslati H, Perez CF, et al. Automated discovery of noncovalent inhibitors of SARS-CoV-2

main protease by consensus Deep Docking of 40 billion small molecules. Chemical science. 2021;12(48):15960-74.
23. Lee SH, Kim MJ, Lee SH, Kim J, Park HJ, Lee J. Thiazolylmethyl ortho-substituted phenyl glucoside library as novel C-aryl glucoside SGLT2 inhibitors. Eur J Med Chem. 2011;46(7):2662-75.
24. Jesus AR, Vila-Viçosa D, Machuqueiro M, Marques AP, Dore TM, Rauter AP. Targeting Type 2 Diabetes with C-Glucosyl Dihydrochalcones as Selective Sodium Glucose Co-Transporter 2 (SGLT2) Inhibitors: Synthesis and Biological Evaluation. J Med Chem. 2017;60(2):568-79.
25. Kang SY, Song KS, Lee J, Lee SH, Lee J. Synthesis of pyridazine and thiazole analogs as SGLT2 inhibitors. Bioorganic & medicinal chemistry. 2010;18(16):6069-79.
26. Putapatri SR, Kanwal A, Banerjee SK, Kantevari S. Synthesis of novel l-rhamnose derived acyclic C-nucleosides with substituted 1,2,3-triazole core as potent sodium-glucose co-transporter (SGLT) inhibitors. Bioorganic & medicinal chemistry letters. 2014;24(6):1528-31.
27. Song KS, Lee SH, Kim MJ, Seo HJ, Lee J, Lee SH, et al. Synthesis and SAR of Thiazolylmethylphenyl Glucoside as Novel C-Aryl Glucoside SGLT2 Inhibitors. ACS Med Chem Lett. 2011;2(2):182-7.
28. Xiaoyu Z, Bin S, Hongbo Z, Jun L, Lilin Q, Xiaoning W, et al. Synthesis and biological evaluation of 6-hydroxyl C-aryl glucoside derivatives as novel sodium glucose co-transporter 2 (SGLT2) inhibitors. Bioorganic & medicinal chemistry letters. 2018;28(12):2201-5.
29. Fushimi N, Fujikura H, Shiohara H, Teranishi H, Shimizu K, Yonekubo S, et al. Structure-activity relationship studies of 4-benzyl-1H-pyrazol-3-yl β-d-glucopyranoside derivatives as potent and selective sodium glucose co-transporter 1 (SGLT1) inhibitors with therapeutic activity on postprandial hyperglycemia. Bioorganic & medicinal chemistry. 2012;20(22):6598-612.
30. Lee J, Kim JY, Choi J, Lee SH, Kim J, Lee J. Pyrimidinylmethylphenyl glucoside as novel C-aryl glucoside SGLT2 inhibitors. Bioorganic & medicinal chemistry letters. 2010;20(23):7046-9.
31. Xu B, Feng Y, Lv B, Xu G, Zhang L, Du J, et al. ortho-Substituted C-aryl glucosides as highly potent and selective renal sodium-dependent glucose co-transporter 2 (SGLT2) inhibitors. Bioorganic & medicinal chemistry. 2010;18(12):4422-32.
32. Du X, Lizarzaburu M, Turcotte S, Lee T, Greenberg J, Shan B, et al. Optimization of triazoles as novel and potent nonphlorizin SGLT2 inhibitors. Bioorganic & medicinal chemistry letters. 2011;21(12):3774-9.
33. Xu B, Lv B, Feng Y, Xu G, Du J, Welihinda A, et al. O-Spiro C-aryl glucosides as novel sodium-dependent glucose co-transporter 2 (SGLT2) inhibitors. Bioorganic & medicinal chemistry letters. 2009;19(19):5632-5.
34. Lee SH, Song K-S, Kim JY, Kang M, Lee JS, Cho S-H, et al. Novel thiophenyl C-aryl glucoside SGLT2 inhibitors as potential antidiabetic agents. Bioorganic & medicinal chemistry. 2011;19(19):5813-32.
35. Kim MJ, Lee SH, Park SO, Kang H, Lee JS, Lee KN, et al. Novel macrocyclic C-aryl glucoside SGLT2 inhibitors as potential antidiabetic agents. Bioorganic & medicinal chemistry. 2011;19(18):5468-79.
36. Lee J, Lee SH, Seo HJ, Son EJ, Lee SH, Jung ME, et al. Novel C-aryl glucoside SGLT2 inhibitors as potential antidiabetic agents: 1,3,4-Thiadiazolylmethylphenyl glucoside congeners. Bioorganic & medicinal chemistry. 2010;18(6):2178-94.
37. Park EJ, Kong Y, Lee JS, Lee SH, Lee J. Exploration of SAR regarding glucose moiety in novel C-aryl glucoside inhibitors of SGLT2. Bioorganic & medicinal chemistry letters. 2011;21(2):742-6.
38. Lv B, Xu B, Feng Y, Peng K, Xu G, Du J, et al. Exploration of O-spiroketal C-arylglucosides as novel and selective renal sodium-dependent glucose co-transporter 2 (SGLT2) inhibitors. Bioorganic & medicinal chemistry letters. 2009;19(24):6877-81.
39. Ohtake Y, Sato T, Kobayashi T, Nishimoto M, Taka N, Takano K, et al. Discovery of tofogliflozin, a novel C-arylglucoside with an O-spiroketal ring system, as a highly selective sodium glucose cotransporter 2 (SGLT2) inhibitor for the treatment of type 2 diabetes. J Med Chem. 2012;55(17):7828-40.
40. Li Y, Shi Z, Chen L, Zheng S, Li S, Xu B, et al. Discovery of a Potent, Selective Renal Sodium-Dependent Glucose Cotransporter 2 (SGLT2) Inhibitor (HSK0935) for the Treatment of Type 2 Diabetes. J Med Chem. 2017;60(10):4173-84.
41. Mascitti V, Maurer TS, Robinson RP, Bian J, Boustany-Kari CM, Brandt T, et al. Discovery of a clinical candidate from the structurally unique dioxa-bicyclo[3.2.1]octane class of sodium-dependent glucose cotransporter 2 inhibitors. J Med Chem. 2011;54(8):2952-60.
42. Xu G, Lv B, Roberge JY, Xu B, Du J, Dong J, et al. Design, synthesis, and biological evaluation of deuterated C-aryl glycoside as a potent and long-acting renal sodium-dependent glucose cotransporter 2 inhibitor for the treatment of type 2 diabetes. J Med Chem. 2014;57(4):1236-51.
43. Wang Y, Lou Y, Wang J, Li D, Chen H, Zheng T, et al. Design, synthesis and biological evaluation of 6-deoxy O-spiroketal C-arylglucosides as novel renal sodium-dependent glucose cotransporter 2 (SGLT2) inhibitors for the treatment of type 2 diabetes. Eur J Med Chem. 2019;180:398-416.
44. Guo C, Hu M, DeOrazio RJ, Usyatinsky A, Fitzpatrick K, Zhang Z, et al. The design and synthesis of novel SGLT2 inhibitors: C-glycosides with benzyltriazolopyridinone and phenylhydantoin as the aglycone moieties. Bioorganic & medicinal chemistry. 2014;22(13):3414-22.
45. Robinson RP, Mascitti V, Boustany-Kari CM, Carr CL, Foley PM, Kimoto E, et al. C-Aryl glycoside inhibitors of SGLT2: Exploration of sugar modifications including C-5 spirocyclization. Bioorganic & medicinal chemistry letters. 2010;20(5):1569-72.
46. Xu B, Feng Y, Cheng H, Song Y, Lv B, Wu Y, et al. C-Aryl glucosides substituted at the 4′-position as potent and selective renal sodium-dependent glucose co-transporter 2 (SGLT2) inhibitors for the treatment of type 2 diabetes. Bioorganic & medicinal chemistry letters. 2011;21(15):4465-70.
47. Ohtake Y, Sato T, Matsuoka H, Kobayashi T, Nishimoto M, Taka N, et al. C-Aryl 5a-carba-β-d-glucopyranosides as novel sodium glucose cotransporter 2 (SGLT2) inhibitors for the treatment of type 2 diabetes. Bioorganic & medicinal chemistry. 2012;20(13):4117-27.
48. Kakinuma H, Oi T, Hashimoto-Tsuchiya Y, Arai M, Kawakita Y, Fukasawa Y, et al. (1S)-1,5-anhydro-1-[5-(4-ethoxybenzyl)-2-methoxy-4-methylphenyl]-1-thio-D-glucitol (TS-071) is a potent, selective sodium-dependent glucose cotransporter 2 (SGLT2) inhibitor for type 2 diabetes treatment. J Med Chem. 2010;53(8):3247-61.
49. Li AR, Zhang J, Greenberg J, Lee T, Liu J. Discovery of non-glucoside SGLT2 inhibitors. Bioorganic & medicinal chemistry letters. 2011;21(8):2472-5.
50. Goodwin NC, Ding Z-M, Harrison BA, Strobel ED, Harris AL, Smith M, et al. Discovery of LX2761, a Sodium-Dependent Glucose Cotransporter 1 (SGLT1) Inhibitor Restricted to the Intestinal Lumen, for the Treatment of Diabetes. Journal of Medicinal Chemistry. 2017;60(2):710-21.
51. Ohtake Y, Sato T, Matsuoka H, Nishimoto M, Taka N, Takano K, et al. 5a-Carba-β-d-glucopyranose derivatives as novel sodium-dependent glucose cotransporter 2 (SGLT2) inhibitors for the treatment of type 2 diabetes. Bioorganic & medicinal chemistry. 2011;19(18):5334-41.
52. Xu G, Du F, Kuo GH, Xu JZ, Liang Y, Demarest K, et al. 5,5-Difluoro- and 5-Fluoro-5-methyl-hexose-based C-Glucosides as

potent and orally bioavailable SGLT1 and SGLT2 dual inhibitors. Bioorganic & medicinal chemistry letters. 2020;30(17):127387.
53. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. Journal of Medicinal Chemistry. 2012;55(14):6582-94.
54. Chen IJ, Foloppe N. Conformational sampling of druglike molecules with MOE and catalyst: implications for pharmacophore modeling and virtual screening. Journal of chemical information and modeling. 2008;48(9):1773-91.
55. Li J, Li X, Wang C, Zhang M, Ye M, Wang Q. The potential of Valeriana as a traditional Chinese medicine: traditional clinical applications, bioactivities, and phytochemistry. Front Pharmacol. 2022;Volume 13 - 2022.
56. He X, Wang J, Li M, Hao D, Yang Y, Zhang C, et al. Eucommia ulmoides Oliv.: ethnopharmacology, phytochemistry and pharmacology of an important traditional Chinese medicine. Journal of ethnopharmacology. 2014;151(1):78-92.
57. Pari L, Venkateswaran S. Hypoglycaemic activity of Scopariadulcis L. extract in alloxan induced hyperglycaemic rats. Phytotherapy research : PTR. 2002;16(7):662-4.