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

Andri Prasetiyo¹, Shirly Kumala¹, Esti Mumpuni¹, Raymond R. Tjandrawinata²

¹ Doctoral Program of Pharmaceutical Sciences, Faculty of Pharmacy, Pancasila University, Jalan Raya Lenteng Agung No.56-80, Srengseng Sawah, Jagakarsa, Jakarta 12640, Indonesia

² Atma Jaya Catholic University, Jalan Raya Cisauk Lapan, Cisauk, Tangerang, Banten 15345, Indonesia

Corresponding author: Shirly Kumala (fskumala@univpancasila.ac.id)

**Abstract**

**Introduction:** Brotowali (*Tinospora crispa*) has been traditionally used as an antidiabetic drug. DPP-IV inhibitor as an antidiabetic will increase insulin secretion. It indirectly escalates incretin hormones, such as Glucagon-Like peptide-1 (GLP-1) which depends on glucose. This study predicts potential compounds from the Brotowali plants, such as DPP-IV inhibitors, using the Molegro Virtual Docker (MVD).

**Materials and methods:** Before the molecular docking simulation, internal validation and external validation are necessary. Internal validation was carried out by re-docking the native ligands in the DPP-IV enzyme crystal structure (PDB codes 3G0B, 3W2T, and 3BJM). The external validation was carried out by simultaneous docking of 59 active compounds and 1918 inactive compounds (decoys) from the A Directory of Useful Decoys (DUD) database with PDB code 3G0B on 16 combinations, four search algorithms, and four functions scoring.

**Results and discussion:** The molecular docking simulation was carried out on 50 compounds from the Brotowali plant and alogliptin as standard compounds with PDB code 3G0B. The best results of the docking method validation yielded the RMSD values of 0.43 and EF1% of 20.34 and EF20% of 3.1 (the combination of search algorithm Moldock optimizer and scoring function Moldock score). The re-rank score of 5 compounds from the Brotowali plant (Rumphioside C, Borapetoside E, Borapetoside F, Rumphioside I, and 6'-O-Lactoyl Borapetoside B) were -107.7 kcal/mol; -105.4 kcal/mol; -104.2 kcal/mol; -102.8 kcal/mol. Alogliptin (standard ligands) had a re-rank score of -101.6 kcal/mol. The combination of search algorithms MolDock optimizer and scoring function MolDock score is a valid protocol with a good result. The similarity of the binding sites of Borapetoside E and 6'-O-Lactoyl Borapetoside B is 75% when compared to the alogliptin binding sites (Glu 205, Glu 206, Tyr 547).

**Conclusion:** Based on the re-rank score and binding sites similarity, Borapetoside E and 6'-O-Lactoyl Borapetoside B have potential as an antidiabetic drug with a mechanism of action of DPP-IV inhibitors.

**Keywords**

antidiabetic, DPP-IV inhibitors, Tinospora crispa, validation, Molegro Virtual Docker.
Introduction

Diabetes is a chronic condition where the glucose level in the blood increases because it cannot produce the hormone insulin or this hormone does not work effectively (IDF 2017). The incretin hormone, glucagon-like peptide-1 (GLP-1), stimulates insulin secretion in response to glucose from food, inhibits glucagon secretion, and promotes proliferation of pancreatic beta cells. Active GLP-1 can be degraded rapidly by DPP-IV. DPP-IV inhibitors increase the plasma concentration of active GLP-1 and induce insulin secretion in response to increased blood glucose levels (Yoshida et al. 2012). DPP-IV inhibitors have a favorable weight-neutral profile with the minimal hypoglycemia risk (Taylor and Lam 2020).

Brotowali has been used in traditional medicine to treat diabetes (Klangjareonchai et al. 2015; Ahmad et al. 2016). The in vitro test results on the Brotowali extract activity in DPP-IV inhibitors indicate that the extract has activities as a DPP-IV inhibitor (Riyanti et al. 2016). Studies have been carried out on the constituents from Brotowali stems plant. More than 50 compounds have been isolated and identified as terpenoids (Cycloeucalenol, Cycloeucalenone, Tinocrispol A, Borapetol A, Borapetol B, 2-O-Lactoylborapetoside B, 6’-O-Lactoylborapetoside B, Borapetoside B, Borapetoside C, Borapetoside D, Borapetoside E, Borapetoside F, Borapetoside G, Borapetoside H, Rumphioside A, Rumphioside B, Syringin, Columbin), flavonoids (Apigenin, Diosmetin, Genkwanin), alkaloids (Magnoflorin, N-Formylanonaine, N-Acetylanonaine, N-Formylnorcicerine, N-Acetylnorcicerine, Lysicaine, Tyramine, Higenamine, N-cis-Feruloyltamrine, N-trans-Feruloyltamrine, Paprazine, Columbamine, Di-hydrodiscretamin, Palmatine, Jatrorrhizine, Berberine, Salsolinol, (−)-Luteolin), lignan (Secoisolariciresinol, Syringaresinol, Adenosine, Uridine, Adenine), and sterol (β-sitosterol, Stigmasterol, Makisterone C) (Praman et al. 2012; Ahmad et al. 2016).

The DPP-4 inhibitors that have been FDA approved are vildagliptin, saxagliptin, teneligliptin, sitagliptin, alogliptin, and linagliptin. The co-crystal structures of five inhibitors DPP-IV with native ligand have been reported PDB: 1X70, 3G0B, 3BJM, 3W2T, and 3VJK (Nabeno et al. 2013). Virtual Screening based on molecular docking has been used in drug discovery design to understand drug-receptor interaction. However, many issues often occur in these studies, including errors in selected location of the binding site and docking poses (Bielska et al. 2011; Chen 2015). In some cases, docking accuracy may change from 0% to 92.66% (Chen 2015). The virtual screening predictions should be evaluated to ensure reliability and determine the proper docking protocols to use (Lagarde et al. 2015). The retrospective virtual screening method assesses two main criteria: the accuracy of the binding site predictions and the enrichment of the benchmarking data in the active compound (Lagarde et al. 2015).

Molegro Virtual Docker (MVD) is a protein-ligand docking simulation program that allows us to carry out docking simulations in a fully integrated computational package. MVD is proven to apply to hundreds of protein docking performances similar to other docking programs, such as AutoDock4 and AutoDock Vina (Bitencourt-Ferreira and de Azevedo 2019). The MVD version 6 has four search algorithms: MolDock Optimizer (MDO) (based on differential evolution), MolDock Simplex Evolution (MDSE) (a modified algorithm based on Nelder-Mead local search algorithm), Iterated Simplex (IS) (based on Nelder-Mead algorithm), and iterated simplex with ant colony optimization (ISACO). It is also possible to choose four scoring functions in each search algorithm (Bitencourt-Ferreira and de Azevedo 2019). In summary, we have 16 combinations of four search algorithms and four scoring functions.

This research aimed to validate docking protocol using MVD (Sixteen docking protocol combinations of four search algorithms and four scoring functions), SBVS, and to find dpp4 inhibitor compounds, such as antidiabetic agents from the Brotowali plants.

Materials and methods

Hardware, software, and webserver


Selection of proteins

We use the X-ray crystal structure of the DPP-IV enzyme (PDB code 3G0B, 3BJM, and 3W2T) was retrieved from the protein data bank https://www.rcsb.org/. Proteins were selected for the docking study by the following criteria: the obtained structure with X-ray diffraction from a human cell has a resolution below 2.5Å (Chakraborti et al. 2021) and has a native ligand that has been approved by FDA.

Internal validation

Validation of the docking (Castro-Alvarez et al. 2017) method was carried out by re-docking the native ligands in the DPP-IV enzyme crystal structure (PDB Code 3G0B, 3BJM, and 3W2T). We used Molegro Virtual Docking software with a default setting, whereby appropriate missing hydrogen atoms were added, missing bonds were assigned, partial charges were added if necessary, and
flexible torsions in ligands were detected. We identified five cavities as potential binding sites. However, only one cavity was used for the ligand-docking study. The maximum number of poses in re-docking was five with ten replications and it was run by sixteen combinations of four search algorithms and four scoring functions. The validation method was successful when the Root Mean Square Deviation (RMSD) value was less than 1 Å (Dhananjayan 2015).

**External validation**

The validation of the docking method was carried out by concurrent docking of 59 active compounds and 1918 inactive compounds in the DPP-IV enzyme crystal structure (The best DPP-IV enzyme crystal structure was obtained from the internal validation). The active and inactive compounds (decoys) were downloaded in the SMILES format from the DUD-E website and then stored on a local server with actives_final.ism and decoys_final.ism. These compounds (active and decoys) are then converted into forms *.smi and added hydrogen. Then, the preparation was conducted by adjusting the pH to 7.4 and all *.smi files are converted into 3D (in the form *.mol2). These stages were carried out using the Open Babel 3.11 program. The active binding site region was defined as a spherical region encompassing within 10 Å of bound crystallographic ligand atom with a size of X: 42.21, Y: 34.47, and Z: 14.97 axes, respectively. Default settings were used for all the calculations. The maximum number of poses in redocking was five with ten replications and it was run by sixteen combinations of four search algorithms and four scoring functions.

For the enrichment results, we were using three enrichment indicators: EFmax (maximum enrichment factor), EF1 (enrichment factor at 1% of the ranked database), and EF20 (enrichment factor at 20% of the database). EF max and EF1 present the early enrichment, while EF20 presents the late-stage database screening (Granchi et al. 2015).

**Ligand preparation**

The structures of the compounds in the Brotowali plant and standard compounds were drawn using the Chem 2D Professional 16.0 program or downloaded on Pubchem, then continued in 3D with the ChemDraw 3D 16.0 program. The energy was optimized using MMFF94 tools on Chem3D 16.0, aiming to obtain minimum energy. The file was stored in MDL MolFile (*.mol) and used for the docking process (Bitencourt-Ferreira and de Azevedo 2019).

**Molecular docking**

A total of fifty compounds in the Brotowali plants and alogliptin (standard ligands) were docked simultaneously in the DPP-IV enzyme crystal structure (PDB Code 3G0B). The active binding site region was defined as a spherical region encompassing within 10 Å of bound crystallographic ligand atom with a size of X: 42.21, Y: 34.47, and Z: 14.97 axes, respectively. Default settings were used for all the calculations. Docking was performed using MolDock Optimizer algorithms, and for each of the 50 independent runs, a maximum number of 2000 iterations were executed on a single population of 50 individuals. Re-docking five poses with fifty replications run best combinations of search algorithms MolDock Optimizer (based on differential evolution) and scoring function MolDock score. The resulting conformations were clustered, and only the negative lowest-energy representation from each cluster was returned when the docking run was completed. For analysis, one pose with the lowest value of Re-rank Score was selected as the best solution for each complex.

**Results and discussion**

**Internal validation**

This study employs three DPP-IV inhibitor crystal structures bound to different native ligands; PDB code 3G0B (Zhang et al. 2011), PDB code 3BJM (Metzler et al. 2008), and PDB code 3W2T (Nabeno et al. 2013). The internal validation of the docking method was done by re-docking the native ligands with the DPP-IV inhibitors (3G0B, 3BJM, and 3W2T), using 16 combinations of the four search algorithms and four scoring functions. Root Mean Square Deviation (RMSD) is the most commonly used quantitative measure of the similarity between two superimposed atomic coordinates (Kufareva and Abagyan 2012). We used the following the criterion to evaluate re-docking success: the root-mean-square deviation (RMSD) between the crystallographic position of the ligand and the pose (generated by MVD). The best results could be obtained when RMSD values were less than 1.0 Å° compared to crystallographic structures in docking simulations.

The RMSD value of re-docking the native ligand with proteins PDB code 3G0B showed the RMSD values < 1 Å in all the docking protocol combinations (Table 1). It is shown that the sixteen docking protocols can accurately position a DPP-IV inhibitor on the DPP-IV binding site. The RMSD value of re-docking the native ligand with PDB Code 3BJM re-docking showed an RMSD value < 1 Å on several docking protocols (Table 2). It is shown that the six docking protocols can accurately position a DPP-IV inhibitor on the DPP-IV binding site. We chose PDB Code 3G0B for external validation using MVD, considering that it has an RMSD value < 1 Å in all the docking protocol combinations and has a higher resolution value than PDB code 3BJM. High-resolution structures are highly ordered, and it is easy to see the accuracy of every atom in the electron density map level (Castro-Alvarez et al. 2017). We did not choose PDB Code 3W2T for external validation using MVD, considering that it has an RMSD value > 1 Å in all the docking protocol combinations.
Table 1. Root-Mean-Square Deviation value for Protein Data Bank Code 3G0B (native ligands saxagliptin, resolution 2.35 Å)

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Score</th>
<th>RMSD (Angstrom)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MolDock</td>
<td>0.43</td>
<td>0.43</td>
</tr>
<tr>
<td>MolDock (GRID)</td>
<td>0.68</td>
<td>0.84</td>
</tr>
<tr>
<td>PLANTS</td>
<td>0.45</td>
<td>0.48</td>
</tr>
<tr>
<td>PLANTS (GRID)</td>
<td>0.39</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Note: RMSD – Root-Mean-Square Deviation; PDB – Protein Data Bank; GPU – Graphics Processor Unit; CUDA – Compute Unified Device Architecture; SE – Simplex Evolution.

Table 2. Root-Mean-Square Deviation value for Protein Data Bank Code 3BJM (Native ligands saxagliptin, resolution 2.35 Å)

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Score</th>
<th>RMSD (Angstrom)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MolDock</td>
<td>0.45</td>
<td>1.32</td>
</tr>
<tr>
<td>MolDock (GRID)</td>
<td>0.88</td>
<td>0.88</td>
</tr>
<tr>
<td>PLANTS</td>
<td>1.36</td>
<td>1.23</td>
</tr>
<tr>
<td>PLANTS (GRID)</td>
<td>0.60</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Note: RMSD – Root-Mean-Square Deviation; PDB – Protein Data Bank; GPU – Graphics Processor Unit; CUDA – Compute Unified Device Architecture; SE – Simplex Evolution.

External validation

The DPP-IV PDB code 30B enzyme crystal structure was chosen in the external validation. Docking external validation was also performed on docking between 59 active compounds and 1918 decoys on 16 combinations of four search algorithms and four scoring functions at the DPP-IV PDB code 30B enzyme crystal structure.

In this study, we calculated three enrichment indicators: EFmax (maximum enrichment factor), EF1% (enrichment factor at 1% of the ranked database), and EF20% (enrichment factor at 20% of the database). EFmax and EF20% were used for the early enrichment, while EF1% was used for the late-stage database screening.

Based on EFmax and EF20%, (Table 3 a good validation result was obtained with six combinations of docking protocols because the EF Max value > 30 and EF20% value > 3 (Table 4).

Table 3. Enrichment performance for each matrix unit

<table>
<thead>
<tr>
<th>Enrichment performance</th>
<th>EFmax</th>
<th>EF20%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very good</td>
<td>≥30</td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>20–30</td>
<td>2.5–3</td>
</tr>
<tr>
<td>Medium</td>
<td>10–20</td>
<td>2–2.5</td>
</tr>
<tr>
<td>Poor</td>
<td>&lt; 10</td>
<td>≤2</td>
</tr>
</tbody>
</table>

Note: EF – Enrichment Performance.

We used an EF value of 1% to find the best combination of docking. The combination of docking protocols with MolDock Optimizer algorithms and scoring Function MolDock has an EF max value > 30, an EF20% value > 3, and the highest value EF1% compared to other docking protocol combinations. So, we chose the combination of docking protocol with the MolDock optimizer algorithm and scoring function MolDock for molecular docking simulations using MVD with PDB code 3G0B. Other research experiments show that MolDock has powerful properties of high docking accuracy to the identification of ligand-binding mode. MolDock was able to identify the correct binding sites of 87% of the complexes. In comparison, the accuracy of Glide and Surflex is 82% and 75%, respectively (Kufareva and Abagyan 2012).

Molecular docking simulation

The molecular docking simulation of fifty compounds from the Brotowali plant and alogliptin (standard compounds) was carried out using algorithms the MolDock Optimizer and the score function MolDock (derived from the Piecewise Linear Potential). Alogliptin is an antidiabetic drug with a DPP-IV inhibitor mechanism already on the market and approved by the Food Drug Association (FDA). Re-rank scores are calculated in Molegro Virtual
Docker to estimate ligand binding, where lower values are associated with higher affinity.

The seven compounds of the Brotowali plant had a lower rank than alogliptin as standard compounds (Table 5). The lower the re-rank score, the more stable and active bond between ligands and receptors. Based on the re-rank score, we predicted that the seven compounds were more active than alogliptin as the standard ligand.

The protein-ligand complex hydrogen interactions of the seven compounds after docking molecular simulations are shown in Fig. 1. Alogliptin (standard compound) has the binding sites on four amino acid residues (Glu 205, Glu 206, Tyr 547, and Tyr 631). Alogliptin has bridges with Glu 205 and Glu 206 in the S2 subsite which have vital roles in the inhibitory activity of enzyme DPP-IV (Nabeno et al. 2013). The similarity of Rumphioside C and Borapetoside F binding sites is 50% with the alogliptin binding sites (Glu 205 and Glu 206). Whereas, the similarity of Borapetoside E and 6'-O-Lactoyl Borapetoside B
<table>
<thead>
<tr>
<th>No</th>
<th>Compounds</th>
<th>Rerank Score (Kcal/mol)</th>
<th>Binding sites</th>
<th>Binding sites similar to standard ligands</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rumphioside C</td>
<td>-107.7</td>
<td>Ser 630, His 740, Arg 125, Glu 205, Glu 206, Ser 209</td>
<td>50%</td>
</tr>
<tr>
<td>2</td>
<td>Borapetoside E</td>
<td>-105.4</td>
<td>Gly 632, Glu 205, Glu 206, Tyr 547, Tyr 662</td>
<td>75%</td>
</tr>
<tr>
<td>3</td>
<td>Borapetoside F</td>
<td>-104.2</td>
<td>Lys 554, Tnp 629, Val 546, Gly 632, Tyr 547, Ser 209, Arg 125, Ser 630</td>
<td>25%</td>
</tr>
<tr>
<td>4</td>
<td>Rumphioside I</td>
<td>-102.8</td>
<td>Tyr 547, Tyr 662</td>
<td>25%</td>
</tr>
<tr>
<td>5</td>
<td>6'-O-Lactoyl borapetoside B</td>
<td>-102.8</td>
<td>Tnp 629, Tyr 547, Ser 630, Tyr 666, Glu 206, Glu 205</td>
<td>75%</td>
</tr>
<tr>
<td>6</td>
<td>Rumphioside F</td>
<td>-102.8</td>
<td>Tyr 662, Glu 205, Arg 669, Tyr 547</td>
<td>50%</td>
</tr>
<tr>
<td>No</td>
<td>Compounds</td>
<td>Rerank Score (Kcal/mol)</td>
<td>Binding sites</td>
<td>Binding sites similar to standard ligands</td>
</tr>
<tr>
<td>----</td>
<td>------------------------</td>
<td>------------------------</td>
<td>------------------------</td>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>7</td>
<td>Borapetoside D</td>
<td>-101.8</td>
<td>Gly 632, Glu 206, Glu 205, Arg 125, Tyr 662</td>
<td>50%</td>
</tr>
<tr>
<td>8</td>
<td>Alogliptin</td>
<td>-101.6</td>
<td>Glu 205, Glu 206, Tyr 547, Tyr 631</td>
<td>100%</td>
</tr>
</tbody>
</table>

binding sites is 75% with the alogliptin binding sites (Glu 205, Glu 206, and Tyr 547). Based on the re-rank score and binding sites similarity, we predicted that Borapetoside E and 6'-O-Lactoyl Borapetoside B are the best compounds as DPP-IV inhibitors.

**Conclusions**

Structure-Based Virtual Screening with the crystal structure of DPP-IV (PDB 3G0B) using a combination of the MolDock Optimizer algorithm and the MolDock score function on the MVD proved valid with excellent criteria. Based on re-rank score and binding sites similarity, Borapetoside E and 6'-O-Lactoyl Borapetoside B have the potential as antidiabetic drugs with a mechanism of action DPP-IV inhibitors.

**Conflict of interest**

The authors have declared that no competing interests exist.

**Acknowledgement**

The authors would like to thank Prof. Dr. Siswandono, Apt, Department of Pharmaceuticals Chemistry, Airlangga University, for permission and support in software facilities Molegro Virtual Docker.

**References**

Author contributions

Andri Prasetiyo, Doctoral Program in Pharmaceutical Sciences, Assistant Professor, Faculty of Pharmacy, Pancasila University, Jakarta, Indonesia; andriprasetiyo@univpancasila.ac.id, ORCID ID https://orcid.org/0000-0002-9936-6058. The main author played a major role in conducting the experiments, processing the data, and writing the article.

Shirly Kumala, Professor, Faculty of Pharmacy, Pancasila University, Jakarta, Indonesia; fskumala@univpancasila.ac.id; ORCID ID https://orcid.org/0000-0002-5444-4438. The author played a major role in providing ideas for writing, compiling and designing the article.

Esti Mumpuni, Associate Professor, Faculty of Pharmacy, Pancasila University, Jakarta, Indonesia; esti.mumpuni@univpancasila.ac.id, ORCID ID https://orcid.org/0000-0002-9208-8040. The author played a leading role in the design of the experiment, analysis of the data, and participated in conducting the experiments.

Raymond R Tjandrawinata, PhD in Biochemical Sciences, Associate Professor, Faculty of Biotechnology, Atma Jaya Catholic University of Indonesia, Tangerang, Banten, Indonesia; raytjan@yahoo.com, ORCID ID https://orcid.org/0000-0003-0461-9874. The author played a role in interpreting and analyzing the data, writing and editing the article.