Computational Analysis of *Gynura bicolor* Bioactive Compounds as Dipeptidyl Peptidase-IV Inhibitor

**Lina Rozano,1 Muhammad Redha Abdullah Zawawi,2 Muhamad Aizuddin Ahmad,1 and Indu Bala Jaganath3**

1Agri-Omics and Bioinformatics Programme, Biotechnology and Nanotechnology Research Centre, Malaysian Agricultural Research and Development Institute, 43400 Selangor, Malaysia
2Institute of Systems Biology (INBIOSIS), Universiti Kebangsaan Malaysia, 43600 Bangi, Malaysia
3Director-General’s Office, Malaysian Agricultural Research and Development Institute, 43400 Selangor, Malaysia

Correspondence should be addressed to Lina Rozano; rozalina@mardi.gov.my

Received 14 March 2017; Revised 11 June 2017; Accepted 22 June 2017; Published 8 August 2017

**1. Introduction**

Type-2 diabetes is a chronic metabolic impairment that affects the quality of life. Currently, diabetes is ranked as the eighth leading cause of death with 1.5 million deaths, and 90% of these are from type-2 diabetes [1]. The main cause of type-2 diabetes is excessive blood glucose and the inability of the body to produce enough insulin, also known as insulin resistance in insulin-targeting tissues such as liver, skeletal muscle, and adipocytes. The body’s resistance to insulin causes glucose to remain in the blood, further damaging other organs owing to the high level of sugar, which leads to loss of vision, kidney failure, and cardiovascular diseases.

One way of controlling blood glucose levels is through the inhibition of dipeptidyl peptidase-IV (DPPIV), a serine peptidase responsible for transforming incretins into their inactive metabolites. Incretins or glucagon-like peptide-1 (GLP1) have a role in stimulating glucose-dependent insulin secretion and regulate glycaemia but are short-lived because of DPPIV catalytic activity. Because of this, inhibition of dipeptidyl peptidase-IV increases the level of circulating GLP-1, which then stimulates insulin biosynthesis and secretion, which can reverse the hyperglycemic condition in type-2 diabetes.

The introduction of gliptin-based drugs in 2006 for the treatment of type-2 diabetes has changed the pattern of diabetes medication usage among type-2 diabetes patients [2, 3]. Gliptin drugs increase the concentration of incretin hormones, increasing insulin level in a glucose-dependent manner and decreasing glucagon levels in the circulation. Most diabetic patients opt for gliptin-based pills because they have similar efficacy as sulfonylurea drugs such as...
metformin. Up until now, eight synthetically developed compounds in the gliptin class have been approved for the treatment of diabetes: sitagliptin, anagliptin, linagliptin, saxagliptin, alogliptin, vildagliptin, teneligliptin, gemigliptin, and dutagliptin [4]. However, wide application among type-2 diabetes patients has led to fatal side effects that relate to high risk of cardiovascular diseases, inflammation of pancreas, allergic reactions, and rheumatoid arthritis [5–9].

In parallel with the discovery and development of chemically synthesized DPPIV inhibitors such as tricyclic heterocycles and fungal synthetic (+)-antroquinonol, the exploitation of plant bioactive compounds for DPPIV inhibitory properties is also underway [10–13]. Novel synthetic compounds have been derived from plant backbone structures, such as compound 55P0110 from quinozolidine alkaloids. Novel synthetic compounds are also investigated for their antidiabetic properties, specifically, compounds from the plant species that have undergone in vitro validations. This includes compounds such as resveratrol, luteolin, apigenin, flavone, and cyanidin 3,5-diglucoside, which can be found in citrus, grapes, soybeans, and aronia berries [15–17]. Other plant species that have DPPIV inhibitory properties have been demonstrated through in vitro studies are Urena lobata, Fagonia cretica L., Hedera nepalensis K. Koch, Senna nigricans, Commiphora mukul, Emblica officinalis, Terminalia arjuna, and Smilax china [18–22].

Traditionally, Gynura species have been widely studied for their antidiabetic properties, specifically, Gynura procumbens [23–25]. Besides lowering blood glucose levels, it does possess other beneficial physiochemical properties such as anti-inflammatory, antihypertensive, antiulcerogenic, and chemopreventative actions [26–32]. However, studies on G. bicolor are not as extensive as G. procumbens, but it is reported to have high antihyperglycemic properties because of the presence of flavonoid compounds such as dichaeoylquinic acid and caffeic acid groups [33, 34]. G. bicolor also has anti-inflammatory protection, and chemopreventive properties [35–37]. Because of the mass availability of DPPIV inhibitory compounds in plants, dependency on in silico screening for DPPIV inhibitor becomes a crucial part of the discovery of potential DPPIV inhibitors before proceeding to the next stage in the development of drug lead compounds [38, 39].

The aim of this study was to evaluate bioactive compounds in G. bicolor as potentially potent inhibitors of DPPIV through molecular docking analysis. The candidate agents discovered can then be further developed as robust DPPIV inhibitors.

2. Materials and Method

2.1. Plant Extracts and Identification of Bioactive Compounds. G. bicolor leaves were collected from the Biotechnology and Nanotechnology Research Centre, Malaysian Agricultural Research and Development Institute (MARDI), Selangor, Malaysia. Plant identification was conducted by Mohd Norfaizal Ghazalli (MARDI) and a voucher specimen of G. bicolor (MDI 12809) was deposited in MDI Herbarium, MyGenebank™ Complex, Malaysian Agricultural Research and Development Institute, Selangor, Malaysia. The extraction was performed on ground and freeze-dried samples using methanol extraction. In the methanol extraction, 20 mL of methanol was added to the freeze-dried sample (0.5 g) and the mixture was homogenized for 1 minute followed by vortexing for 30 minutes. The mixture was then centrifuged at 8,900 rpm for 5 minutes at 4°C. The supernatant was filtered with Whatman, number 40 filter paper to remove solid particles from the sample, and 10 mL methanol was added without homogenization. The extracted sample (2.0 mL) was transferred into microcentrifuge tube and dried by vacuum concentration for LCMS-MS/HPLC analysis. Samples were introduced to HPLC for chemical profiling at wavelength 280 nm and 360 nm. This was followed by quantitative identification of compounds in G. bicolor using LCMS-MS.

2.2. Ligand Preparation. 2D and 3D structure of G. bicolor bioactive compounds: 5-O-cafeoylquinic acid, trans-5-p-coumarylquinic acid, cis-5-p-coumarylquinic acid, 3,4-dicaffeoylquinic acid, and 3-coumarylquinic acid were generated using ChemDraw (PerkinElmer Inc, Massachusetts, USA) based on LCMS-MS data (Figure 1). All ligand structures file conversions were performed using BIOVIA Discovery Studio Visualizer (Accelrys Software Inc. San Diego, CA) followed by geometrical cleansing. Comparative analysis of the ligands was conducted against five selected gliptin drugs obtained from PubChem database (https://pubchem.ncbi.nlm.nih.gov) [40]: Sitagliptin (PubChem CID: 4369359), Linagliptin (PubChem CID: 10096344), Anagliptin (PubChem CID: 44515473), Saxagliptin (PubChem CID: 11243969), and alogliptin (PubChem CID: 11450633) based on structures obtained from PubChem. Diprotin-A (PubChem CID: 94701), inhibitor of DPPIV receptor, was included in the analysis.

2.3. Receptor Preparation. Six 3D protein structure files of dipeptidyl peptidase-IV (DPPIV) with PDB ID: 3WQH, 3WT2, 4A5S, 4FPW, 4PNZ, and 4PV7 were obtained from RCSB Protein Data Bank (http://www.rcsb.org) [41]. The overall stereochemical properties of each DPPIV were assessed based on the information obtained from PDB X-ray Structure Validation Report for each PDB structure, which includes the crystal structure resolution, Wilson B-factor, R-value, stereochemical parameters, overall percentiles scores, and the MolProbity Ramachandran analysis (http://molprobity.biochem.duke.edu/) [42]. The resolution and R-values showed the goodness of the protein model being used. The X-ray crystal structure with resolution values of 2.0 Å or less and R-values of 0.2 or less are considered acceptable. Structural similarity measurement of the six DPPIV protein structures were conducted using mpLBPA web server (http://www.dsimb.inserm.fr/dsimb_tools/mlpbpa/index.php) based on similarity involving the local backbone [43] and Partial Order Structure Alignment (POSA) web server (http://posa.sanfordburnham.org/) to study structural divergence of the protein structures [44]. The crystal structure of human DPP IV in complex with a novel heterocyclic DPPIV inhibitor with PDB ID: 4A5S was selected for molecular docking analysis [45]. The active site of DPPIV (PDB ID: 4A5S) was
predicted using CASTp Server (http://sts.bioe.uic.edu/castp/) [46], where it scans the protein surfaces for pockets and also interior of proteins for voids followed by further protein functional surfaces identification and spatial pattern characterization using SplitPocket web server (http://pocket.med.wayne.edu/patch/) [47]. Prediction of glycosylation groups for DPPIV (PDB ID: 4A5S) was performed using NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/) and YinOYang 1.2 Server (http://www.cbs.dtu.dk/services/YinOYang/) [48].

2.4. Molecular Docking Simulation. Docking studies were performed using Lead IT software of BioSolve Gmbh FlexX package (http://www.biosolveit.de/FlexX/) [49] and iGEMDOCK software (version 2.1) (http://gemdock.life.nctu.edu.tw/dock/igemdock.php) [50]. Under the FlexX package, the energy minimized DPPIV (PDB ID: 4A5S) receptor and ligands underwent flexible molecular docking analysis. FlexX predicts the geometry of the complex as well as an estimate for the strength of binding by fragmenting the ligand at rotatable bonds and reassembling it within a binding pocket. It contains an optimizer that allows off-grid torsional positions of ligand placement during docking process. Within the torsional degrees of freedom, the procedure moves the atom away from the MIMUMBA grid; the cost function to be minimized is the currently valid scoring function. The protein was prepared in Lead IT using default settings. FlexX binding site analysis included all complete residues with at least one atom within a distance of up to 6.5 Å with respect to the reference ligand. Prior to the docking process, both ligands and DPPIV were prepared and were assigned bonds, bond orders, explicit hydrogens, charges, and flexible torsions. Using a buildup algorithm, the ligands were flexibly located into the protein active site. This is being done through the superposing interaction points of the selected base fragment and the protein active site. The clash factor was set to 0.6. Other parameters were kept as default. The base fragment was then incrementally built up to the complete compound by modeling the ligand flexibility with a torsion library for the added components. The correctness of the protein preparation step was checked by a self-docking process in which the cocrystallized ligand was redocked in the receptor. The root mean square deviation (RMSD) was less than 2.0 Å compared to the reference structure. Up to 200 poses were generated for each compound using FlexX package. The next round of docking studies involves the use of iGEMDOCK software, an integrated virtual screening environment which utilizes postscreening analysis with pharmacological interactions through Generic evolutionary algorithm (GA) and an empirical scoring function. In iGEMDOCK, standard flexible docking (normal) option was selected to perform molecular docking analysis with population size of 200, 70 generations, and 2 solutions. A comparative analysis of the plant metabolites was conducted with the molecular docking scores of DPPIV with commercial drugs. The potent inhibitors for DPPIV were selected based on having the least free-binding energy values.

2.5. Drug-Likeness and ADME. The metabolites were analyzed for drug-relevant properties based on “Lipinski’s rule of five” and bioactivity prediction using the Molinspiration web server [51]. Further ADME prediction was also conducted using the PreADMET web server (http://preadmet.bmdrc.kr) [52], where the risk of toxicity upon consumption of compounds can be predicted. Four ADME properties of G. bicolor bioactive compounds were tested: blood-brain barrier (BBB), human intestinal absorption, Caco-2 cell model, and
plasma protein binding ability. BBB was represented as $BB = \frac{\text{Brain}}{\text{Blood}}$ or log BB in predicting whether compounds pass across the blood-brain barrier [53]. Prediction of HIA used chemical structures at pH 7.4 and shows the sum of bioavailability and absorption evaluated from the ratio of excretion (or cumulative excretion) in urine, bile, and feces based on percentage values (% HIA) [54, 55]. The Caco-2 cell model is derived from human colon adenocarcinoma and possesses multiple drug transport pathways through the intestinal epithelium. We applied it as a reliable in vitro model for the prediction of oral drug absorption ($P_{\text{Caco-2}}$ (nm/sec)) [56]. PPB predicts the percentage of compounds bound in plasma protein as in vitro data in humans, which influences the action, disposition, and efficacy of compounds (% PPB).

3. Results

3.1. Profiling and Identification of G. bicolor Bioactive Compounds. To profile and identify bioactive compounds in G. bicolor, plant extracts were subjected to HPLC and LCMS-MS analysis. HPLC analysis resulted in the observation of five major peaks: A, B, C, D, and E, with peak A being the most abundance followed by peaks E, D, B, and C (Figure 2). Five major peaks were also observed in LCMS-MS chromatogram with detection at 280 nm (Figure 3). Elution time for peaks A, B, C, D, and E were recorded at $t_R = 14.30, 18.46, 20.47, 25.51,$ and 25.80 minutes, respectively, as described in Table 1. G. bicolor compounds were identified based on $m/z$ values of each peak. Peaks A, B, C, D, and E were identified as 5-O-caffeoylquinic acid, trans-5-p-coumaroylquinic acid, cis-5-p-coumaroylquinic acid, 3,4-dicaffeoylquinic acid, and 3-caffeoylquinic acid, respectively (Table 1).

3.2. DPPIV Receptor Selection for Docking. Determinations of DPPIV receptor for molecular docking process were based on the evaluation of six protein crystal structures obtained from the PDB with PDB ID: 3WQH, 3W2T, 4A5S, 4FFW, 4PNZ, and 4PV7. Table 2 displays comparison of the six DPPIV receptors according to their crystal structure resolution, Wilson $B$-factor, $R$-value, and Ramachandran plot values. This also includes information on Ramachandran plots for each structure (Figure 4). Analysis results for the six DPPIV receptors crystal structure similarity, which measures the alignment score, $N_{\text{rms}}, N_{\text{gdt}},$ RMSD of core, and $N_{\text{r,5}}$, were displayed in Supplementary Figures 1 and 2, in Supplementary Material, available online at https://doi.org/10.1155/2017/5124165. The DPPIV structure with PDB ID: 4A5S was selected for further molecular docking with G. bicolor compounds. Identification of 4A5S ligand binding site of 4A5S was made according to the measurement of the largest identified pocket (pocket 184) with volume of 19238 $\text{Å}^3$ and area 6863.7 $\text{Å}^2$ obtained from CASTp analysis. This was further confirmed with SplitPocket analysis (Supplementary Table 1).

3.3. Free-Binding Energy of G. bicolor Compounds. To obtain free-binding energy between G. bicolor compounds and DPPIV receptor, molecular docking analysis was conducted using Lead IT and iGEMDOCK. Glititin drugs and diprotin-A, known inhibitors of DPPIV receptor, were also included in the docking analysis for comparison. Molecular docking analysis for all compounds resulted in free-binding energy ranging from as low as $-31.8807 \text{KJ/mol}$ to $-22.2267 \text{KJ/mol}$ using Lead IT, as presented in Table 3.
Table 2: Criteria for selection of DPPIV receptor for molecular docking.

<table>
<thead>
<tr>
<th>PDB ID</th>
<th>Resolution (Å)</th>
<th>Wilson plot B factor (Å²)</th>
<th>R-value</th>
<th>Phi psi in most favored regions (%)</th>
<th>Ramachandran outliers (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3WQH</td>
<td>2.85</td>
<td>57.7</td>
<td>0.227</td>
<td>92.1</td>
<td>0.5</td>
</tr>
<tr>
<td>3W2T</td>
<td>2.36</td>
<td>41.9</td>
<td>0.180</td>
<td>95.9</td>
<td>0.1</td>
</tr>
<tr>
<td>4A5S</td>
<td>1.62</td>
<td>20.8</td>
<td>0.164</td>
<td>97.0</td>
<td>0.0</td>
</tr>
<tr>
<td>4FFW</td>
<td>2.90</td>
<td>74.4</td>
<td>0.287</td>
<td>91.9</td>
<td>1.0</td>
</tr>
<tr>
<td>4PNZ</td>
<td>1.90</td>
<td>24.2</td>
<td>0.168</td>
<td>96.8</td>
<td>0.0</td>
</tr>
<tr>
<td>4PV7</td>
<td>3.24</td>
<td>114.9</td>
<td>0.223</td>
<td>92.7</td>
<td>0.2</td>
</tr>
</tbody>
</table>

As for iGEMDOCK analysis, 3,4-dicaffeoylquinic acid is the top most in the rank with $-149.9$ kcal/mole followed by 5-O-cafeoylquinic acid, trans-5-p-coumaroylquinic acid, and 3-cafeoylquinic acid (Table 4). For both docking analysis, highest negative scores indicate a better active compound. All G. bicolor compounds and the reference inhibitor, diprotin-A, show electrostatic interaction ranging from $-2.81$ Kcal/mol to $2.17$ Kcal/mol with no electrostatic energy seen in any of the gliptin drugs.

To understand the mode of action on DPPIV inhibition through the molecular docking process, binding poses for each structure were studied. Binding pose of G. bicolor compounds, together with gliptin drugs and diprotin-A, is indicated in Figures 5 and 6, showing that all compounds interact closely with key residues of sites S1, S2, and S3 within DPPIV receptor pockets.

Results showed that the binding energies of G. bicolor compounds with the DPPIV receptor ranged from as low as $-29.0750$ KJ/mol up to $-22.2267$ KJ/mol (Table 3). The data were compared with binding energies of gliptin drugs and diprotin-A with the DPPIV receptor. Anagliptin had the lowest binding energy of $-31.8807$ KJ/mol, while cis-5-p-coumaroylquinic acid has the highest binding energy of $-22.2267$ KJ/mol. In our study, molecular docking of G. bicolor compounds with 4A5S showed that binding interaction formed with quinic acid of the plant compounds through establishment of hydrogen bond with residues in any of the three DPPIV binding pocket regions, S1, S2, and S3. Plant compound structures were also observed to be located in the hydrophobic regions of DPPIV pockets as observed by the green line (Figure 5).

As observed in Figure 5(A2), 5-O-cafeoylquinic acid of G. bicolor forms 8 hydrogen bonds with DPPIV residues. The quinic acid structure resides in S1 pocket of Ser630, while the caffeic acid structure in S3 region interacts with residue Phe357. 3,4-Dicaffeoylquinic acid docking involves
Figure 4: Ramachandran plot analysis for dipeptidyl peptidase-IV receptors obtained from Protein Data Bank based on PDB ID. (a) 3WQH, (b) 3W2T, (c) 4A5S, (d) 4FFW, (e) 4PNZ, and (f) 4PV7.
DPPIV receptor unlike 3,4-dicaffeoylquinic acid and 3-caffeoylquinic acid docked well in the S1 and S3 region of Ser630 and His740 residues. Of Phe357, and quinic acid overlaps in the S1 pockets of both coumaric acid structures resides in the S3 hydrophobic region formed between the compounds of DPPIV residues. Its cis-p-coumaroylquinic acid. As for the least ranked compound, average conpair, and rank (based on total energy). Total energy = VDW + H-Bond + Elec.

<table>
<thead>
<tr>
<th>Ligands</th>
<th>Score</th>
<th>Match</th>
<th>Lipo</th>
<th>Ambig</th>
<th>Clash</th>
<th>Rot</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-O-Caffeoylquinic acid</td>
<td>−27.3288</td>
<td>−33.6824</td>
<td>−7.5070</td>
<td>−7.2838</td>
<td>5.9444</td>
<td>9.800</td>
</tr>
<tr>
<td>trans-5-p-Coumaroylquinic acid</td>
<td>−27.1177</td>
<td>−31.2300</td>
<td>−7.3174</td>
<td>−7.9231</td>
<td>2.7528</td>
<td>11.200</td>
</tr>
<tr>
<td>cis-5-p-Coumaroylquinic acid</td>
<td>−22.2267</td>
<td>−32.6442</td>
<td>−3.4542</td>
<td>−4.6060</td>
<td>3.2777</td>
<td>9.800</td>
</tr>
<tr>
<td>3,4-Dicaffeoylquinic acid</td>
<td>−27.1703</td>
<td>−33.4586</td>
<td>−8.8033</td>
<td>−8.8360</td>
<td>3.1276</td>
<td>15.400</td>
</tr>
<tr>
<td>3-Caffeoylquinic acid</td>
<td>−29.0750</td>
<td>−33.9425</td>
<td>−6.8680</td>
<td>−7.5393</td>
<td>2.6748</td>
<td>11.200</td>
</tr>
<tr>
<td>Linagliptin</td>
<td>−27.5774</td>
<td>−26.3546</td>
<td>−10.3421</td>
<td>−9.2890</td>
<td>6.1882</td>
<td>7.000</td>
</tr>
<tr>
<td>Saxagliptin</td>
<td>−23.4595</td>
<td>−23.8361</td>
<td>−6.4052</td>
<td>−8.3097</td>
<td>4.0915</td>
<td>5.600</td>
</tr>
<tr>
<td>Diprotin-A</td>
<td>−30.6185</td>
<td>−36.7587</td>
<td>−7.6816</td>
<td>−8.5830</td>
<td>4.4048</td>
<td>12.600</td>
</tr>
</tbody>
</table>

Score: the total score of the docking solution. Match: contribution of the matched interacting groups. Lipo: contribution of the lipophilic contact area. Ambig: contribution of the lipophilic-hydrophilic or ambiguous contact area. Clash: contribution of the clash penalty. Rot: ligand conformational entropy score (Rot).

interaction of S1 and S3 pockets involving Ser630, His740, and Phe357 residues with its quinic acid structure, while two of its caffeic acid structures interact with the S1 hydrophobic region and Asn710 of S1 pocket. A total of 8 hydrogen bonds was formed. As for *trans*-5-*p*-coumaroylquinic acid, 8 hydrogen bonds were also formed. Its structure differs from the caffeoylquinic acid backbone, with caffeic acid being replaced by coumaric acid. In this case, the coumaric acid did not interact with any of its major pocket regions, but quinic acid resides in S1 hydrophobic region and interacts with Phe357 from S3 pocket region. It was observed that 5-O-cafeoylquinic acid, *trans*-5-*p*-coumaroylquinic acid, and 3-cafeoylquinic acid docked well in the S1 and S3 region of DPPIV receptor unlike 3,4-dicaffeoylquinic acid and *cis*-5-*p*-coumaroylquinic acid. As for the least ranked compound, *cis*-5-*p*-coumaroylquinic acid, only 4 hydrogen bonds were formed between the compounds of DPPIV residues. Its coumaric acid structure resides in the S3 hydrophobic region of Phe357, and quinic acid overlaps in the S1 pockets of both Ser630 and His740 residues.

In comparison with the amino acid interaction of gliptin drugs, anagliptin with the lowest free-binding energy of −31.8807 KJ/mol forms a total of 8 hydrogen bonds with DPPIV residues, and the structure forms interactions with all three active site pockets, S1, S2, and S3, through interaction with residues Phe357, Ser630, and Tyr662 (Figure 6).

### 3.4. Drug-Likeness, ADME, and Bioactivity Prediction

To assess drug-likeness properties of *G. bicolor* compounds, each compound was analyzed using the Molinspiration web server (http://www.molinspiration.com/). Two of the compounds, *trans*-5-*p*-coumaroylquinic acid and *cis*-5-*p*-coumaroylquinic acid, fulfill drug-relevant properties based on "Lipinski’s rule of five" as having molecular mass less than 500 Daltons, high lipophilicity (Log *P* less than 5), less than 5 hydrogen bond donors, less than 10 hydrogen bond acceptors, and molar refractivity between 40 and 130.

5-O-cafeoylquinic acid and 3-cafeoylquinic acid have one violation each while 3,4-dicaffeoylquinic acid is with three violations (Table 5).
Figure 5: Continued.
To predict ADME properties of each *G. bicolor* compound, the compounds were applied to preADMET web server (https://preadmet.bmdrc.kr/). The four ADME results are displayed in Table 6.

### 4. Discussion

This study reported three major findings: (1) the identification of three major compounds, 5-O-caffeoylquinic acid, 3,4-dicaffeoylquinic acid, and *cis*-5-*p*-coumaroylquinic acid, not yet reported before in *G. bicolor*; (2) the prediction of lower free-binding energy scores of four of *G. bicolor* compounds, 3-caffeoylquinic acid, 5-O-caffeoylquinic acid, 3,4-dicaffeoylquinic acid, and *trans*-5-*p*-coumaroylquinic acid when compared to the commercially available gliptin inhibitors; and (3) the computational investigation of DPPIV receptors in relation to protein-ligand binding.

*G. bicolor* is a local herb widely grown in tropical climate countries. Its major compound comprises caffeoylquinic
Figure 6: Continued.
acid backbone, extensively studied for its antidiabetic properties [57, 58]. The identification of 3-caffeoylquinic acid and trans-5-p-coumaroylquinic acid in G. bicolor had been previously reported, while the other three compounds, 5-O-caffeoylquinic acid, 3,4-dicaffeoylquinic acid, and cis-5-p-coumaroylquinic acid, have not been reported before in G. bicolor (Table 1) [33]. 5-O-Caffeoylquinic acid and 3caffeoylquinic acid, being two of the most abundant compounds discovered in G. bicolor, are a well-known chlorogenic acids that have also been observed in coffee and are well studied for their antidiabetic properties [59–61]. However, unlike in coffee, there is no caffeine in G. bicolor which makes it free from the side effects that comes with caffeine.

In this study, the molecular docking procedure was aimed at identifying individual poses and the free-binding energy of G. bicolor compounds that may bind to the DPPIV active site. Based on the free-binding energy generated from Lead IT, the gliptin drug anagliptin ranked as having the best inhibitory effects towards DPPIV, while cis-5-p-coumaroylquinic acid ranked as having the least inhibitory effects with free-binding energy of $-31.8807$ KJ/mol and $-22.2267$ KJ/mol, respectively. The second-best free-binding energy score was for diprotin-A, followed by alogliptin, 3-caffeoylquinic acid, linagliptin, 3,4-dicaffeoylquinic acid, 5-O-coumaroylquinic acid, trans-5-p-coumaroylquinic acid, sitagliptin, saxagliptin, and cis-5-p-coumaroylquinic acid. These data showed that G. bicolor bioactive compounds have comparable free-binding energy as observed for gliptin drugs. Gliptin drugs have been widely used as a positive control for DPPIV inhibitory experiments [10, 62]. Among the gliptin drugs, anagliptin is believed to have the best half-maximal inhibitory concentration (IC$_{50}$) values when
which are lower than relationship based on protein 3D structure alignments. DPPIV receptors shows that each one exhibits very close results. Measurement of structural similarity among the six properly glycosylated to obtain intrinsic dynamic property receptor selected for molecular docking should preferably be thus increasing protein stability [65]. Because of this, the protein structure but, rather, may decrease protein dynamics, however, the N-glycans did not induce significant changes in the molecular docking process (Supplementary Figure 3).

Gliptin compounds, such as 3-caffeoylquinic acid, and 3,4-dicaffeoylquinic acid, were included during the molecular docking process of DPPIV, and these were identified as active sites (Table 3). While diprotin-A is a potent DPPIV inhibitor with an Ile-Pro-Ile sequence commonly used as reference compound [17, 64].

The application of iGEMDOCK highlights the various binding energies, such as hydrogen bond (H-Bond), van der Walls (VDW) interaction, and electrostatic energy that occurs between ligand and receptor. The H-Bond interaction in ligand is related to the interaction of the hydrophilic group or the presence of atom with lone pair electron, while VDW interaction is related to lipophilic groups such as aromatic ring, or methyl group. Majority of the G. bicolor compounds and gliptin drugs possess hydrophilic group in the form of hydroxy indicated by lower H-Bond energy than VDW energy (Table 4). The comparison of binding scores from Lead IT and iGEMDOCK shows similarities in the pattern of the free-binding energy for the top four G. bicolor compounds, 3-caffeoylquinic acid, 5-O-caffeoylquinic acid, 3,4-dicaffeoylquinic acid, and trans-5-p-coumaroylquinic acid, which are lower than cis-5-p-coumaroylquinic acid (Table 4).

In general, the DPPIV receptor is a cell surface glycoprotein receptor. There are nine major N-glycans and eight O-glycan identified in DPPIV, and these were included during the molecular docking process (Supplementary Figure 3). However, the N-glycans did not induce significant changes in protein structure but, rather, may decrease protein dynamics, thus increasing protein stability [65]. Because of this, the receptor selected for molecular docking should preferably be properly glycosylated to obtain intrinsic dynamic property results. Measurement of structural similarity among the six DPPIV receptors shows that each one exhibits very close relationship based on protein 3D structure alignments. N

<table>
<thead>
<tr>
<th>Compounds</th>
<th>Molecular weight</th>
<th>Hydrogen bond donor</th>
<th>Hydrogen bond acceptor</th>
<th>Log P</th>
<th>Molecular polar surface area (PSA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-O-Caffeoylquinic acid</td>
<td>354.31</td>
<td>6</td>
<td>9</td>
<td>-0.45</td>
<td>164.74</td>
</tr>
<tr>
<td>trans-5-p-Coumaroylquinic acid</td>
<td>338.31</td>
<td>5</td>
<td>8</td>
<td>0.04</td>
<td>144.52</td>
</tr>
<tr>
<td>cis-5-p-Coumaroylquinic acid</td>
<td>338.31</td>
<td>5</td>
<td>8</td>
<td>0.04</td>
<td>144.52</td>
</tr>
<tr>
<td>3,4-Dicaffeoylquinic acid</td>
<td>516.46</td>
<td>7</td>
<td>12</td>
<td>1.21</td>
<td>211.28</td>
</tr>
<tr>
<td>3-Caffeoylquinic acid</td>
<td>354.31</td>
<td>6</td>
<td>9</td>
<td>-0.45</td>
<td>164.74</td>
</tr>
</tbody>
</table>

Table 5: Lipinski’s rule of five analysis for ligands used in the study.

<table>
<thead>
<tr>
<th>ADME</th>
<th>5-O-Caffeoylquinic acid</th>
<th>trans-5-p-Coumaroylquinic acid</th>
<th>cis-5-p-Coumaroylquinic acid</th>
<th>3,4-Dicaffeoylquinic acid</th>
<th>3-Caffeoylquinic acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBB</td>
<td>0.033661</td>
<td>0.0315336</td>
<td>0.0315336</td>
<td>0.0360627</td>
<td>0.033661</td>
</tr>
<tr>
<td>PPB</td>
<td>41.96179</td>
<td>46.246716</td>
<td>46.246716</td>
<td>87.772544</td>
<td>41.96179</td>
</tr>
</tbody>
</table>

Table 6: ADME properties of G. bicolor bioactive compounds.

compared with sitagliptin and alogliptin [63]. This information correlates with the obtained free-binding energy where anagliptin has the lowest binding score among other gliptin classes (Table 3). While diprotin-A is a potent DPPIV inhibitor with an Ile-Pro-Ile sequence commonly used as reference compound [17, 64].

The application of iGEMDOCK highlights the various binding energies, such as hydrogen bond (H-Bond), van der Walls (VDW) interaction, and electrostatic energy that occurs between ligand and receptor. The H-Bond interaction in ligand is related to the interaction of the hydrophilic group or the presence of atom with lone pair electron, while VDW interaction is related to lipophilic groups such as aromatic ring, or methyl group. Majority of the G. bicolor compounds and gliptin drugs possess hydrophilic group in the form of hydroxy indicated by lower H-Bond energy than VDW energy (Table 4). The comparison of binding scores from Lead IT and iGEMDOCK shows similarities in the pattern of the free-binding energy for the top four G. bicolor compounds, 3-caffeoylquinic acid, 5-O-caffeoylquinic acid, 3,4-dicaffeoylquinic acid, and trans-5-p-coumaroylquinic acid, which are lower than cis-5-p-coumaroylquinic acid (Table 4).

In general, the DPPIV receptor is a cell surface glycoprotein receptor. There are nine major N-glycans and eight O-glycan identified in DPPIV, and these were included during the molecular docking process (Supplementary Figure 3). However, the N-glycans did not induce significant changes in protein structure but, rather, may decrease protein dynamics, thus increasing protein stability [65]. Because of this, the receptor selected for molecular docking should preferably be properly glycosylated to obtain intrinsic dynamic property results. Measurement of structural similarity among the six DPPIV receptors shows that each one exhibits very close relationship based on protein 3D structure alignments. N

<table>
<thead>
<tr>
<th>Compounds</th>
<th>Molecular weight</th>
<th>Hydrogen bond donor</th>
<th>Hydrogen bond acceptor</th>
<th>Log P</th>
<th>Molecular polar surface area (PSA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-O-Caffeoylquinic acid</td>
<td>354.31</td>
<td>6</td>
<td>9</td>
<td>-0.45</td>
<td>164.74</td>
</tr>
<tr>
<td>trans-5-p-Coumaroylquinic acid</td>
<td>338.31</td>
<td>5</td>
<td>8</td>
<td>0.04</td>
<td>144.52</td>
</tr>
<tr>
<td>cis-5-p-Coumaroylquinic acid</td>
<td>338.31</td>
<td>5</td>
<td>8</td>
<td>0.04</td>
<td>144.52</td>
</tr>
<tr>
<td>3,4-Dicaffeoylquinic acid</td>
<td>516.46</td>
<td>7</td>
<td>12</td>
<td>1.21</td>
<td>211.28</td>
</tr>
<tr>
<td>3-Caffeoylquinic acid</td>
<td>354.31</td>
<td>6</td>
<td>9</td>
<td>-0.45</td>
<td>164.74</td>
</tr>
</tbody>
</table>

Table 5: Lipinski’s rule of five analysis for ligands used in the study.

<table>
<thead>
<tr>
<th>ADME</th>
<th>5-O-Caffeoylquinic acid</th>
<th>trans-5-p-Coumaroylquinic acid</th>
<th>cis-5-p-Coumaroylquinic acid</th>
<th>3,4-Dicaffeoylquinic acid</th>
<th>3-Caffeoylquinic acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBB</td>
<td>0.033661</td>
<td>0.0315336</td>
<td>0.0315336</td>
<td>0.0360627</td>
<td>0.033661</td>
</tr>
<tr>
<td>PPB</td>
<td>41.96179</td>
<td>46.246716</td>
<td>46.246716</td>
<td>87.772544</td>
<td>41.96179</td>
</tr>
</tbody>
</table>

Table 6: ADME properties of G. bicolor bioactive compounds.

that shows that the structures are very similar (Supplementary Figure 1).

Comparison of the crystal structure of DPPIV does provide information on the flexibility of the side chains possibly involved in ligand stabilization. Nevertheless, it should be understood that the distribution of active site residues for the different DPPIVs differs in where the selected 4A5S has the highest distribution of active sites, as observed in Supplementary Table 2. This is due to structural resolution differences that affect the accuracy of docking prediction, since side chains placements depend on the protein structure resolution. Better resolution means more accurate side chain placement. Side chains play important roles in ligand binding because they can cause steric hindrance, and incorrect placement can form false cavities or pockets [66]. With high resolution structures, the atoms are highly ordered and easy to be observed from the electron density map, while at low resolution, 3 Å or higher, only basic contours of protein chains may be observed and atomic structure must be inferred. Flexibility of the side chains in each DPPIV differs since the distribution of active site during the docking process is affected by overall conformational of the protein and ligand to reach a state of stabilization.

There are three major ligand binding subdomains in the DPPIV structure, identified as S1, S2, and S3 [15]. Each site comprises different amino acid residue positions; subdomain S1 with residues Ser630, Asn710, and His740; subdomain S2 with Lys250, Gly260, and Tyr662; and subdomain S3 with residues Ser209, Phe357, and Arg358. The S1 hydrophilic pocket includes catalytic residues and is the primary determinant of substrate specificity. The selection of the largest binding pocket of DPPIV inhibitor PDB ID: 4A5S for ligand docking was supported by the physicochemical features of the pocket itself. SplitPocket analysis had identified the presence of a split pocket which is detectable when a functional pocket binds to a ligand, substrate, or other proteins or peptides
causing the interaction between heterogeneous atoms to reduce the empty space of the pocket and disrupt the integrity of its surface wall (Supplementary Table 1).

During molecular docking simulation, G. bicolor compounds and gliptin drugs occupied the same binding pocket of DPPIV PDB ID: 4A5S formed by residues Val656, Val711, Tyr662, Glu205, Tyr547, Trp629, Tyr631, Ser630, and His 740 (Figures 5 and 6). Ser630 and His 740 interact via hydrogen bonding with the compounds while Val656, Val711, Tyr662, Glu205, Tyr547, Trp629, and Tyr631 formed hydrophobic interactions with the compounds [67]. The main interactions in the active site of DPPIV, which is highly contributed by the hydroxyl coordination between Tyr547 and Ser630 by the water molecule, are highly important for the coordinated interactions in the active site [68]. The location of the binding site appears to be highly conserved across the six DPPIVs being investigated. However, disruption of the binding positions in DPPIV had been reported to impact DPPIV enzyme activity [69]. Reported experimental mutations on active site residues Ser630 resulted in negative activity of DPPIV, while mutating residue His740 greatly reduces inhibitor binding ability to the active sites [70].

The precise position of each G. bicolor bioactive compound in DPPIV receptors would reveal the points of interaction with DPPIV residues. The binding score seems to be affected by the number of hydrogen bonds formed between the plant compounds and DPPIV residues. Compared to other compounds found in G. bicolor, 3-coumaroylquinic acid has the lowest free-binding energy and is considered to have the most potent DPPIV inhibitory effect. This suggests that the position and the interactions of the hydroxyl groups in caffeic acid, quinic acid, and ester structures with DPPIV binding sites are crucial in determining its bioactivity, as revealed in Figure 5. The main backbone of the 3-coumaroylquinic acid structure resides but does not fully enclose itself in the hydrophobic region and forms 11 hydrogen bonds with DPPIV residues. The quinic acid structure of this molecule resides in the S1 pocket at residue Ser630.

In G. bicolor, the major bioactive compounds as indicated in this study have a quinic acid structure of either caffeic or coumaric acid structures attached to either a coumaryl or a caffeoyl group. This structure influences interactions with the DPPIV receptor. The inhibitory mechanisms of compounds from G. bicolor exerted on DPPIV receptors were proposed to be involved in competitive binding at the same active site engaged by gliptin drugs which are known to be highly selective and competitive DPPIV inhibitors [71]. If a larger portion of the quinic acid group resides in the hydrophobic region, it could stERICly hinder the binding formation involving the active sites within the DPPIV receptor, resulting in higher free-binding energy [15]. This can be observed in the molecular docking interaction between cis-5-p-coumaroylquinic acid and DPPIV as seen in Figure 5. Here, the whole structure that consists of one quinic acid group and the benzene ring of the coumaryl group is fully enclosed in the hydrophobic regions. However, the free-energy binding for trans isomer of 5-p-coumaroylquinic acid was reported to be lower compared with its cis form from both Lead IT and iGEMDOCK (Tables 3 and 4), as only a small section of the functional groups interacts with the hydrophobic region. This finding correlates with reported bioactive form of trans and cis chemical compounds where the latter are considered as bioinactive forms [72]. It is believed that cis-5-p-coumaroylquinic acid is less stable than trans-5-p-coumaroylquinic acid due to increased steric interaction of the substituents in the cis isomer.

In the drug-likeness analysis, all G. bicolor compounds passed the minimum cut-off value except for 3,4-dicaffeoylquinic acid, as shown in Table 3. All compounds molecular weight ranged from approximately 330 kDa except for 3,4-dicaffeoylquinic acid with 516 kDa owing to the presence of three phenyl rings. In the ADME analysis (Table 6), all five compounds were CNS-inactive and did not have the ability to cross the blood-brain barrier, which is essential, to avoid any CNS-related effects. HIA analysis revealed that all the studied compounds have the ability to be moderately absorbed from the intestine to the bloodstream and have moderate permeability. Most of the bioactive compounds are weakly bound to plasma protein except for 3,4-dicaffeoylquinic acid. 3,4-Dicaffeoylquinic acid possesses a stronger binding capacity to plasma proteins, making it less preferable as a drug candidate as this characteristic would affect diffusion or transport across cell membranes, limiting pharmacological actions.

5. Conclusion

The study demonstrated that 3-caffeoylquinic acid, 5-O-cafeoylquinic acid, 3,4-dicaffeoylquinic acid, and trans-5-p-coumaroylquinic acid compounds isolated from G. bicolor possess DPPIV inhibitory activity. These compounds are able to dock well to two of the DPPIV receptor active sites, S1 and S2. Molecular docking evaluation of G. bicolor compounds suggested 3-caffeoylquinic acid as a promising candidate with free-binding energy of −29.0750 KJ/mol, which is better than three of the commercially available gliptin drugs, sitagliptin, saxagliptin, and linagliptin. These data suggested the ability of 3-caffeoylquinic acid to dock well in inhibiting the action of the DPPIV receptor in the treatment of type-2 diabetes. Drug-likeness analysis supports the use of 3-caffeoylquinic as a drug lead compound. ADME properties can be taken as best-hit molecule and can be considered for further studies such as QSAR and molecular dynamics. Despite that, results obtained from this study require further validation of the inhibition action of G. bicolor compounds towards the DPPIV receptor using in vitro approach. The computational data supports the efficacy of G. bicolor compounds as naturally occurring DPPIV inhibitors and could be considered for development as a potent antidiabetic drug.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Authors’ Contributions

Author contributions to the study and manuscript preparation are as follows. Lina Rozano contributed to conception
and design. Muhamad Aizuddin Ahmad performed wet lab work. Lina Rozano, Muhammad Redha Abdullah Zawawi, and Muhamad Aizuddin Ahmad acquired data. Lina Rozano and Indu Bala Jaganath were involved in analysis and interpretation. Lina Rozano drafted the article. Indu Bala Jaganath supervised the study. All authors reviewed and approved the final manuscript.

Acknowledgments

The authors would like to acknowledge the Ministry of Agricultural (MOA) for the financial support and Biotechnology Research Centre, Malaysian Agricultural Research and Developmental Institute, for providing the research facilities. This work was supported by Malaysian Ministry of Agriculture under Tenth Malaysia Plan Grant no. 21003002500001.

References


[65] H. S. Lee, Y. Qi, and W. Im, "Effects of N-glycosylation on protein conformation and dynamics: Protein Data Bank analysis and molecular dynamics simulation study," *Scientific Reports*, vol. 5, article 8926, 2015.


