International Journal of Applied Pharmaceutics

ISSN - 0975 - 7058 Vol 12, Special Issue 1, 2020

Research Article

IN SILICO APPROACH FOR SCREENING OF THE INDONESIAN MEDICINAL PLANTS DATABASE TO DISCOVER POTENTIAL DIPEPTIDYL PEPTIDASE-4 INHIBITORS

AULIA FARKHANI[1], RANI SAURIASARI[2], ARRY YANUAR[1]*

[1]Laboratory of Biomedical Computation and Drug Design, Faculty of Pharmacy, Universitas Indonesia, Depok, West Java, Indonesia.
[2]Laboratory of Clinical and Community Pharmacy, Faculty of Pharmacy, Universitas Indonesia, Depok, West Java, Indonesia.
Email: arry.yanuar@ui.ac.id

Received: 26 September 2019, Revised and Accepted: 17 December 2019

ABSTRACT

Background: Dipeptidyl peptidase-4 (DPP4) is an enzyme responsible for inactivating the hormone incretin, which potentiates insulin secretion and glucagon inhibition; inhibitors of DPP4 are used as therapeutic drugs for type-2 diabetes.

Objective: In this study, we evaluated potential DPP4 inhibitors from the Indonesian Medicinal Plants Database using an ***in silico*** approach.

Methods: A ligand-based pharmacophore model was used for screening the database using LigandScout 4.2. This model was validated using several parameters of enrichment metrics, including receiver operating characteristics, area under curve (AUC), and enrichment factor (EF). Hit compounds were also docked with DPP4 to calculate the free binding energy and analyze the interaction between the ligand and DPP4. In addition, bioavailability and medicinal chemistry predictions were performed for the hit compounds.

Results: The best pharmacophore model demonstrated $AUC_{100\%}$ and $EF_{1\%}$ values of 0.82 and 33.8, respectively. The pharmacophore features of the model included hydrogen bond donors, hydrogen bonds, hydrophobic interactions, and positive ionization areas. Based on our results of virtual screening and molecular docking, six hit compounds were ultimately identified, namely, L-noradrenaline, octopamine, Nb-demethylechitamine, alliin, isoalliin, and subaphylline.

Conclusion: Collectively, our findings indicate that subaphylline is the most promising compound for further studies, including *in vitro* and *in vivo* experiments and those focused on molecular dynamics and structural modification.

Keywords: Dipeptidyl peptidase-4, Virtual screening, Pharmacophore-based, Molecular docking, *In silico*, Diabetes.

 DOI: http://dx.doi.org/10.22159/ijap.2020.v12s1.FF008

INTRODUCTION

Dipeptidyl peptidase-4 (DPP4) has been identified as a potential new therapeutic target for reducing the rate of diabetes and its associated early mortality. DPP4 plays a role in glucose homeostasis by deactivating the incretin hormones glucose-dependent insulinotropic peptide and glucagon-like peptide-1 that potentiate insulin secretion from pancreatic β cells and inhibits glucagon secretion from pancreatic alpha cells [1,2]. However, these two hormones are short-lived because they are rapidly degraded by the DPP4 enzyme shortly after they are secreted. DPP4 inhibitors prevent the action of the enzyme, resulting in increased levels of active incretin hormones in the body, which serves to increase the body's ability to control blood glucose levels. However, DPP4 has an amino acid residue composition and active site pocket that are similar to other DPP isozymes [3]. Thus, there is the possibility of inhibiting other DPP4 isozymes with the use of non-selective DPP4 inhibitors. Based on a review by Drucker, non-selective DPP4 inhibitors can impact immune regulation, biological transplantation, cancer cell growth, and metastasis [2]. In addition, toxicity and tolerability studies with selective DPP4, DPP8, DPP9, and QPP inhibitors have revealed that DPP8/9 and QPP inhibitors produce toxicity in test animals, whereas selective DPP4 inhibitors do not, suggesting that selectivity assessments of potential clinical candidates are essential for an optimal safety profile [4].

"Gliptin" is a DPP4 inhibitor that was chemically synthesized and can be categorized based on the similarity of the chemical scaffold and its binding mode. Various heterocyclic frameworks have been reported to have inhibitory activity against DPP4 [5,6]. Virtual screening with a pharmacophore approach is one method that can be used in search of potential new compounds as DPP4 inhibitors with different structural frameworks have been reported. Pharmacophores are the steric and electronic feature ensembles needed to ensure supramolecular interactions with optimal specific biological targets and trigger or inhibit their biological responses [7]. Three-dimensional (3D) pharmacophore modeling is a technique that describes the interaction of small molecular ligands with macromolecular targets. This approach is considered intuitive and has been increasingly successful in the discovery of computational medicine in recent years [8-10]. In this study, virtual screening was performed on the database of Indonesian herbal plants using a 3D pharmacophore approach to obtain hit compound candidates that can selectively inhibit DPP4 activity.

METHODS

Virtual screening

Two-dimensional structures (.mol) of the ligand candidates were obtained from the Indonesian Herbal Database (www.herbaldb.farmasi.ui.ac.id) containing 1377 compounds. Structures were converted to one-dimensional (1D) structures (.smi) for screening purposes. The 1D structures (.smi) of the active compound and "gliptin" with 547 compounds were obtained from A Directory of Useful Decoys (DUD, www.dude.docking.org) and PubChem (www.ncbi.nlm.nih.gov/pccompound) databases. The 1D structures (.smi) of the decoys were obtained from A DUD (www.dude.docking.org) that contains 40,944 compounds.

A ligand-based pharmacophore model was created using LigandScout 4.2. The active compounds were divided into two datasets, namely, 17 "gliptin" compounds used as the training set and 530 compounds used as the test

set. The training set data were used to build the pharmacophore model. The test set data for both the active and decoy compounds were converted into the.ldb format for screening processes to validate the pharmacophore models that were developed. The best pharmacophore model was selected by calculating several validation parameters, including receiver operating characteristics (ROC), area under curve (AUC), and enrichment factor (EF). Virtual screening was performed on the Indonesian Herbal Database with the best pharmacophores model using LigandScout 4.2.

Molecular docking

The 3D structure of the *homo sapiens* DPP4 protein macromolecule was obtained from the Protein Data Bank (http://www.rcsb.org/pdb) as 5T4B, which is the DPP4 homodimer (chain A and B) with 728 amino acids, bound with the (**34a**) ligand (2-[(3R)-3-aminopiperidin-1-yl]-3-(but-2-yn-1-yl)-5-[(4-methylquinazolin-2-yl)methyl]-3H-imidazo[2,1-b]purin-4(5H)-one; 75N), N-acetyl-D-glucosamine and sodium ion, with a resolution value of 1.76 Å [11].

Hit compounds resulting from the *in silico* screening were docked into the crystal structure of DPP4 using Autodock4.2. The position of the (**34a**) ligand in the X-ray crystal structure was defined as the DPP4 active site and determined using LigandScout. The center of the ligand coordinates X, Y, and Z were 37.6567, 50.0321, and 40.1088 angstroms, respectively. The original ligand (**34a**) was also redocked with the DPP4 structure to verify the molecular docking procedure by measuring the root-mean-square deviation (RMSD) between the best pose obtained by docking and the X-ray crystal structure at <2.0 Å.

Bioavailability and medicinal chemistry predictions

We performed calculations related to various bioavailability and medicinal chemistry parameter predictions using SwissADME free web tools [12] (http://www.swissadme.ch). Various parameters were predicted, including physicochemical properties such as molecular weight (MW), count of specific atom types, molar refractivity (MR), polar surface area, lipophilicity (Log P), and solubility (Log S). These parameters were then interpreted in assessing the parameters of drug-likeness, lead-likeness, and pharmacokinetics. Other predictions, such as the tendency of a compound to become a substrate/non-substrate for glycoprotein permeability and its interaction with cytochrome P450, were also carried out to enrich the pharmacokinetic assessment of the compounds. Furthermore, potentially problematic fragments related to molecular accessibility, medicinal chemistry, and compound synthesis were also identified.

RESULTS

Virtual screening and validation

Pharmacophore model 9 was considered the best model and was chosen based on the ROC graph (Fig. 1) of the ten 3D pharmacophore models that were developed with $AUC_{100\%}$ and $EF_{1\%}$ values of 0.82 and 33.8, respectively.

The best pharmacophore model for the entire gliptin contains a consensus of four basic pharmacophore features, including hydrogen bond donors (HBD; marked in red), hydrogen bond acceptors (HBA; marked in green), hydrophobic interactions (marked in yellow), and positive ionized (PI) areas (marked in blue).

Virtual screening was performed on the Indonesian Herbal Database with the best pharmacophores model using LigandScout 4.2. Based on the results of the *in silico* screening, 12 hit compounds were obtained, namely, l-noradrenaline, which had the highest pharmacophore-fit score of 48.37, followed by octopamine, miraxanthin-ii, miraxanthin-i, miraxanthin-v, mimosine, Nb-demethylechitamine, l-histidine, alliin, isoalliin, subaphylline, and l-theanine (Table 1). Overall, the hit compounds matched with three types of pharmacophore features: Two HBA (red), one HBD (green), and one PI area (blue).

Molecular docking

Differing from the results of *in silico* screening where l-noradrenaline had the highest feature pharmacophore-fit score, the compound with the best free binding energy (ΔG value, in kcal/mol) in terms of molecular docking was subaphylline (−8.2), followed by Nb-demethylechitamine (−7.39), l-noradrenaline (−7.12), octopamine (−6.93), isoalliin (−6.81), alliin (−6.54), mimosine (−6.47), l-theanine (−6.03), miraxanthin-ii (−4.93), miraxanthin-v (−5.70), miraxanthin-I (−5.57), l-histidine (−5.46), and miraxanthin-ii (−4.93) (Table 1).

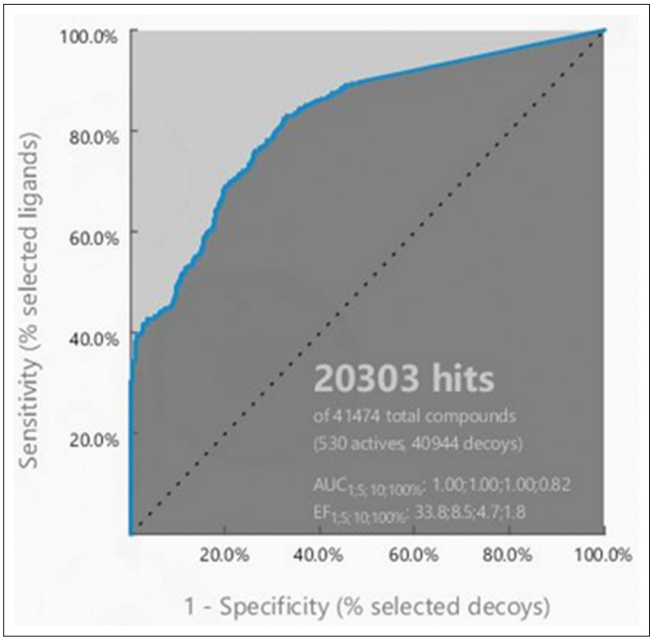


Fig. 1: Receiver operating characteristics graph of the best pharmacophore model

Overall, at least nine compounds interacted with the residues in the DPP4 active sites; these have been reported in previous studies [1,10-14], and eight of them have been predicted to be more selective toward DPP4 because they have interactions with the S1' subsite. The remaining three compounds with the lowest binding affinities (miraxanthin-i, ii, and v) do not appear to interact with the residues in the active sites of DPP4 (Table 2 and Fig. 2).

Bioavailability and medicinal chemistry predictions

Predictions were made for six hit compounds with interaction energy values <−6.50 kcal/mol and were suitable for interaction with the active site of DPP4. These compounds were subaphylline, Nb-demethylechitamine, l-noradrenaline, octopamine, isoalliin, and alliin. Overall, based on our prediction results, the six hit compounds had good oral bioavailability, were suitable for synthesis, and met the criteria of drug-likeness and lead-likeness (Table 3 and Fig. 4).

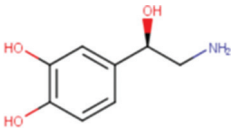
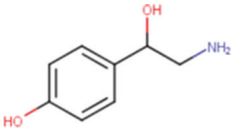
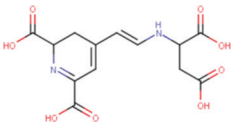
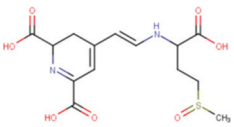
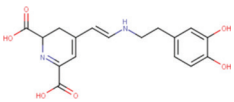
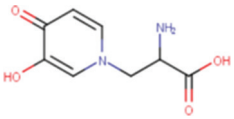
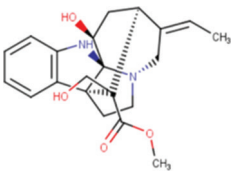
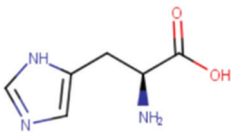
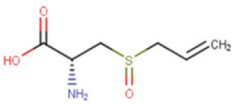
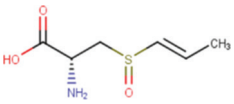
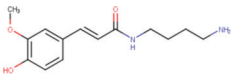
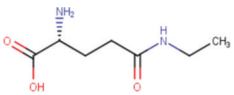
DISCUSSION

Virtual screening and molecular docking

The pharmacophore feature is a type of ligand-receptor interaction and includes HBD, HBA, positively and negatively charged groups, and hydrophobic regions [7]. The best pharmacophore model was built based on 17 gliptin compounds that have been previously reported. A pharmacophore model can have an AUC value between 0 (all inactive molecules first) and 1 (all active molecule first) an AUC value of 0.5 means that the method performed like a random selection [15]. The best pharmacophore model had an $AUC_{100\%}$ value of 0.82, which means the model is useful for virtual screening.

Virtual screening was performed on the Indonesian Herbal Database with the best pharmacophores model using LigandScout 4.2. As shown in Fig. 5, the best pharmacophore model for the entire gliptin contains a consensus of four basic pharmacophore features, including two HBD, one HBA, one hydrophobic interactions, and one

Table 1: Hit compounds from *in silico* screening using Ligandscout 4.2

Compound	Two-dimensional structure	Matching feature	Pharmacophore-fit score	ΔG* (kcal/mol)	Ki* (μM)
L-noradrenaline			48.37	−7.12	6.00
Octopamine			48.07	−6.93	8.39
Miraxanthin-II			47.75	−4.93	244.82
Miraxanthin-I			46.90	−6.03	82.41
Miraxanthin-V			46.36	−5.70	66.80
Mimosine			46.02	−6.47	18.04
Nb-demethylechitamine			45.95	−7.39	3.83
L-histidine			45.71	−5.46	99.92
Alliin			45.61	−6.54	16.02
Isoalliin			45.39	−6.81	10.15
Subaphylline			45.29	−8.20	0.98
L-theanine			45.11	−6.03	38.30

*Calculated from the molecular docking state by Autodock4

PI areas. 12 hit compounds were obtained, namely, l-noradrenaline, octopamine, miraxanthin-ii, miraxanthin-i, miraxanthin-v, mimosine, Nb-demethylechitamine, l-histidine, alliin, isoalliin, subaphylline, and l-theanine (Table 1). Overall, the hit compounds matched with three types of pharmacophore features: Two HBA (red), one HBD (green), and one PI area (blue).

We also analyzed the interactions of the entire list of hit compounds with DPP4 as the target and calculated the values of predicted ΔG and Ki using Autodock4. The parameters of the grid size box and the genetic algorithm runs used for molecular docking were set with an 80×80×80 grid point with an interval of 0.375 Å, and 100, which is determined after verification that the RMSD value obtained was 0.17 Å.

Overall, at least nine compounds interacted with the residues in the DPP4 active sites; and eight of them have been predicted to be more selective toward DPP4 because they have interactions with the S1' subsite (Table 2 and Fig. 2). Furthermore, hit compounds that had binding free energy <−6.5 kcal/mol may be more effective toward DPP4. These included alliin (−6.54), isoalliin (−6.81), l-noradrenaline (−7.12), Nb-demethylechitamine (−7.39), octopamine (−6.93), and subaphylline (−8.2).

Table 2: Interactions between the hit compounds and the active sites of dipeptidyl peptidase-4

Residue/compounds	Ali	Iso	His	Nor	The	Mim	MX1	MX2	MX5	Nde	Oct	Sub.
Tyr 48							v					
Arg 125	v	v	v	v	v	v				v	v	v
Glu 205	v	v	v	v	v	v				v	v	v
Glu 206	v	v	v	v	v	v				v	v	v
Val 207					v	v		v	v			
Ser 209												
Phe 357										v		
Arg 358												
Arg 429								v	v			
Tyr 456									v			
Asp 545							v					v
Val 546							v					v
Tyr 547	v	v		v			v	v	v	v	v	v
Cys 551								v	v			
Ser 552								v	v			
Gln 553								v	v			
Lys 554							v	v	v			v
Asp 556									v			
Arg 560							v	v	v			
Asn 562							v					
Trp 563							v					
Tyr 585								v	v			
Trp 627							v					
Gly 628												
Trp 629							v			v	v	v
Ser 630	v	v	v	v	v	v	v			v	v	v
Tyr 631	v	v	v	v		v			v	v	v	v
Val 656	v	v										
Trp 659	v	v										
Tyr 662	v	v	v	v	v	v				v	v	v
Tyr 666	v	v	v	v	v	v				v		
Arg 669					v	v						
Asp 708		v										
Asn 710	v	v		v	v	v				v	v	v
His 740		v	v							v		v
Tyr 752							v					
Subsites (in color):			S1	S2	S1’	S2’	S2 extensive					

Ali: Alliin, Iso: Isoalliin, His: L-histidine, Nor: L-noradrenaline, The: L-theanine, Mim: Mimosine, MX1: Miraxanthin-I, MX2: Miraxanthin-II, MX5: Miraxanthin-V, Nde: Nb-demethylechitamine, Ocp: Octopamine, Sub: Subaphylline. V: Hydrogen bond of hit compounds with interaction energy values <−6.50 kcal/mol (Fig. 3).

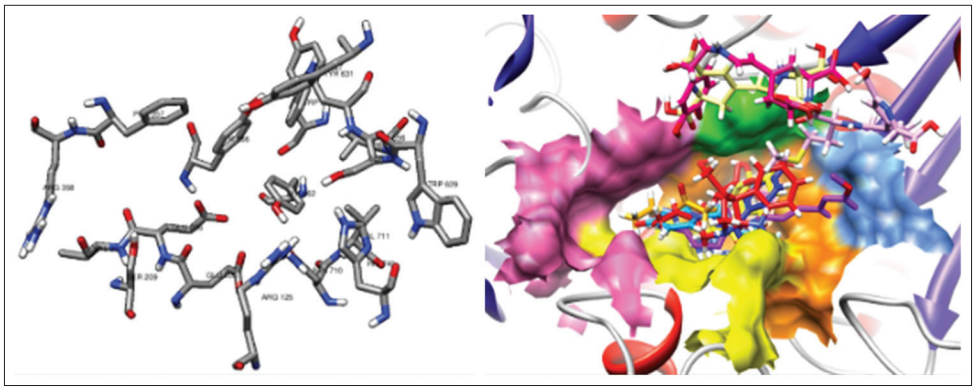


Fig. 2: (Left) Amino acid residues in the active site of dipeptidyl peptidase-4 (DPP4): (Right) Visualization of the interactions between the entire list of hit compounds with the S1 (orange), S2 (yellow), S1’ (green), S2’ (cyan), and S2 extensive (magenta) DPP4 subsites

Nabeno *et al*. classified the DPP4 enzyme active sites into five binding sites (subsites), namely, S1, S2, S1’, S2’, and the site beyond S2 as S2 extensive [10]. The S1 and S2 pocket sites are important interactions and are considered the basic binding mode in DPP4 activity. The S1 sites are the hydrophobic pocket consisting of the catalytic triad (Ser630, Asn710, and His 740) and the S2 sites are the ionic interaction sites with Glu205, Glu 206 [1], and Arg125 residues [13,14]. In addition, the S1 pocket is also formed by highly hydrophobic residues such as Tyr631, Val656, Trp659, Tyr662, Tyr666, and Val711 [13,14]. Arulmozhiraja *et al*. revealed that the hydrophobicity is also related to enzyme activity [13]. Interactions through the S1 ‘(Tyr547) [2,13,14], S2’ (Trp629) [13], and S2 (Val207, Ser209, Arg358, Phe357) [2,13,16] pockets are also important and related to increased activity and selectivity.

As shown in Fig. 3 and Table 2, subaphylline interacted with residues on the active site of DPP4 with the highest binding free energy. Four strong hydrogen bonds were revealed, with a distance of 2.5 - 3.2 Å [17] on subaphylline, and the Glu205 (2.99), Glu206 (2.79) – S2 pockets,

Table 3: Comparison between drug-likeness and pharmacokinetic predictions of hit compounds using SwissADME

Parameter	Alliin	Isoalliin	L-noradrenaline	Nb-demethylechitamine	Octopamine	Subaphylline
Physicochemical properties						
Formula	$C_6H_{11}NO_3S$	$C_6H_{11}NO_3S$	$C_8H_{11}NO_3$	$C_{22}H_{28}N_2O_4$	$C_8H_{11}NO_2$	$C_{14}H_{20}N_2O_3$
Molecular weight (g/mol)	177.22	177.22	169.18	384.47	153.18	264.32
Num. heavy atoms	11	11	12	28	11	19
Num. arom. heavy atoms	0	0	6	6	6	6
Fraction *Csp3*	0.5	0.5	0.25	0.59	0.25	0.36
Num. rotatable bonds (NRotBs)	5	4	2	3	2	8
NHBA	4	4	4	5	3	4
NHBD	2	2	4	3	3	3
MR	43.24	43.24	44.13	111.84	42.11	74.79
TPSA [22] ($Å^2$)	99.6	99.6	86.71	82.03	66.48	84.58
Lipophilicity (Log Po/w)						
iLOGP [23]	0.55	0.88	0.87	3.02	1.01	2.37
XLOGP3 [12,24]	−3.53	−3.45	−1.24	1.7	−0.9	1.03
WLOGP	0.2	0.55	−0.23	0.68	0.06	1.16
MLOGP [12,25,26]	−2.88	−2.88	−0.25	2.1	0.33	0.88
SILICOS-IT [12]	−1	−1.17	0.02	2.3	0.49	1.82
Consensus log Po/w [12]	−1.33	−1.21	−0.17	1.87	0.2	1.45
Water solubility (Log S), Log S Scale: Insoluble <−10<poorly<−6<moderatly<−4<soluble<−2<very<0						
ESOL [12,27]	1.62	1.5	−0.35	−2.98	−0.49	−1.83
Solubility (mg/ml; mol/l)	7.31e+03; 4.12e+01	5.59e+03; 3.15e+01	7.63e+01; 4.51e−01	3.99e−01; 1.04e−03	4.91e+01; 3.20e−01	3.88e+00; 1.47e−02
Class	Highly soluble	Highly soluble	Very soluble	Soluble	Very soluble	Soluble
Ali [12,28]	2.02	1.94	−0.09	−2.59	−0.01	−2.4
Solubility (mg/ml; mol/l)	1.86e+04; 1.05e+02	1.53e+04; 8.65e+01	1.39e+02; 4.51e−01	9.84e−01; 2.56e−03	1.49e+02; 9.70e−01	1.06e+0; 4.02e−03
Class	Highly soluble	Highly soluble	Very soluble	Soluble	Very soluble	Soluble
SILICOS-IT [12]	−0.21	0.16	−0.76	−4.34	−1.32	−3.3
Solubility (mg/ml; mol/l)	1.08e+02; 6.10e−01	2.56e+02; 1.44e+00	2.94e+01; 1.74e−01	1.75e−02; 4.55e−05	7.25e+00; 4.73e−02	1.32e−01; 4.98e−04
Class	Soluble	Soluble	Soluble	Moderately soluble	Soluble	Soluble
Pharmacokinetics						
GI absorption [21]	High	High	High	High	High	High
BBB permeant [21]	No	No	No	No	No	No
P-gp substrate [12]	No	No	No	Yes	No	No
CYP1A2 inhibitor [12]	No	No	No	No	No	No
CYP2C19 inhibitor [12]	No	No	No	No	No	No
CYP2C9 inhibitor [12]	No	No	No	No	No	No
CYP2D6 inhibitor [12]	No	No	No	Yes	No	No
CYP3A4 inhibitor [12]	No	No	No	No	No	No
Drug-likeness						
Lipinski (violation), MW ≤500, MLOGP ≤4.15, NHBA ≤10, NHBD ≤5	Yes	Yes	Yes	Yes	Yes	Yes
Ghose [12,29] (violation), 160 ≤MW ≤480,−0.4 ≤MLOGP ≤5.6, 40 ≤MR ≤130, 20 ≤num. atoms ≤70	Yes	Yes	Yes	Yes	No (1), MW<160	Yes
Veber [12,30] (violation), TPSA≤131.6, NRotBs≤10	Yes	Yes	Yes	Yes	Yes	Yes
Egan [12,31] (violation), WLOGP ≤5.88, TPSA ≤131.6	Yes	Yes	Yes	Yes	Yes	Yes
Muegge [12,32] (violation), 200 ≤MW ≤600, −2 ≤XLOGP ≤5, TPSA ≤150, NrotBs ≤15, num. ring≤7, num. carbons<4, num. heteroatoms >1	No (2), MW<200, XLOGP3 <−2	No (2), MW <200, XLOGP3 <−2	No (1), MW <200	Yes	No (1), MW <200	Yes
Bioavailability score [33]	0.55	0.55	0.55	0.55	0.55	0.55
Medicinal chemistry						
PAINS [18] (alert)	0	0	1 (Catechol A)	0	0	0
Brenk [19] (alert)	1 (Isolated alkenes)	0	2 (Catechol)	1 (Isolated alkenes)	0	1 (Michael acceptor 1)
Lead-likeness (violation), 250≤MW ≤350, XLOGP ≤3.5, NrotBs ≤7	No (1), MW<200	No (1), MW <200	No (1), MW <200	No (1) MW >350	No (1), MW <200	No (1), rotors>7
Synthetic accessibility [34], from 1 (very easy) to 10 (very difficult)	3.21	3.58	1.69	5.9	1.35	2.4

PAINS: Pan-assay interference compounds, MW: Molecular weight, TPSA: Topological polar surface area, NHBA: Num. H-bond acc., NHBD: Num. H-bond don., MR: Molar refractivity

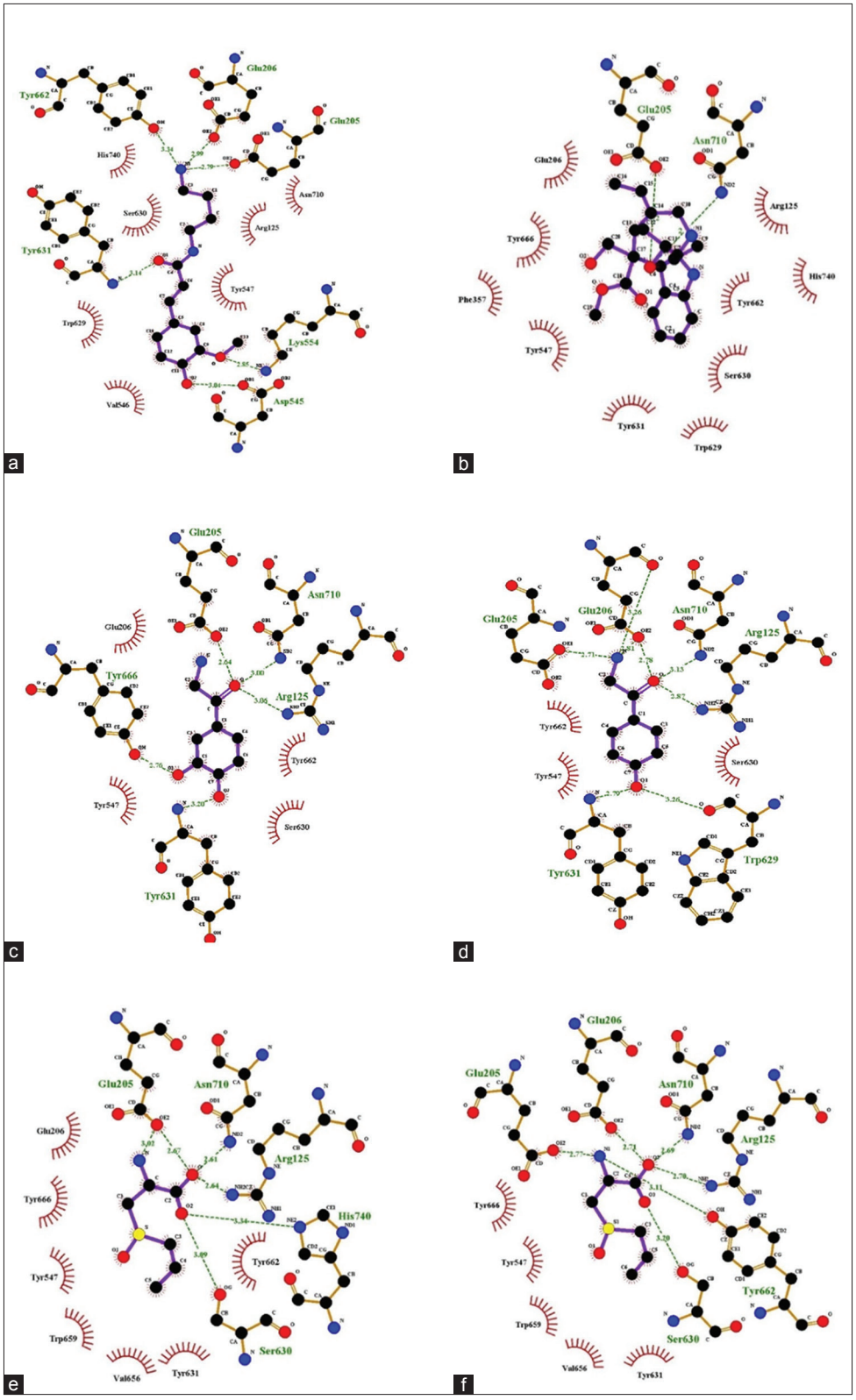


Fig. 3: Interaction of hit compounds with interaction energy values <−6.50 kcal/mol and the active site of dipeptidyl peptidase-4 (DPP4): Two dimensional visualization by Ligplot, the description referred to the web version; ligand bond, non-ligand bonds, hydrogen bond, and its length, non-ligand residues involved in hydrophobic contact(s), corresponding atoms involved in hydrophobic contact(s) (a) subaphylline (b) Nb-demethylechitamine (c) L-Noradrenaline (d) octopamine (e) isoalliin (f) alliin

Tyr631 (3.14) – hydrophobic cavity of S1 pocket, Asp545 (3.0) and Lys554 (2.85) residues, and one weak hydrogen bond with a distance of 3.0–4.0 Å [15] at the Trp662 (3.34) residue. Hydrophobic interactions also occurred between subaphylline with the S1 catalytic site (Ser630, Asn710, and His 740), S2 (Arg 125), Val546, and also at least at two hydrophobic pocket constituents (Tyr 547 and Trp629). Based on the results of molecular docking, we conclude that the compound has the potential to inhibit act as a DPP4 inhibitor.

Nb-demethylechitamine has two strong hydrogen bonds, as shown in Fig. 3, at Glu205 (2.72) and Asn710 (2.79). In addition, Nb-demethylechitamine has ten hydrophobic interactions in the active site of DPP4, namely, Ser630, His 740, Tyr631, Tyr662, and Tyr666

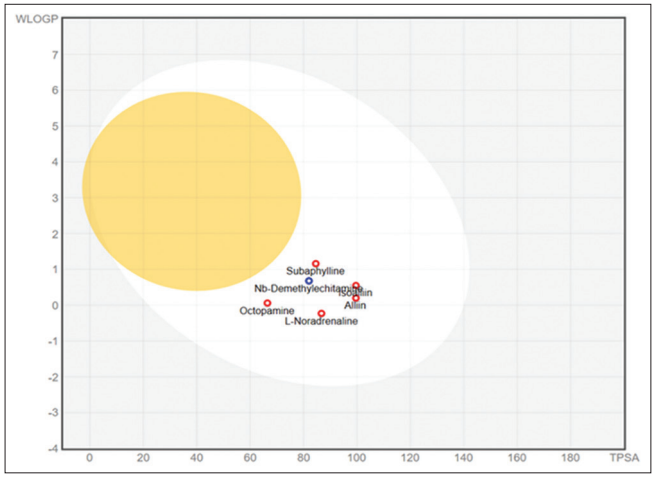


Fig. 4: Illustration of BOILED-Egg and P-gp substrates/non-subst rates for passive HIA and blood–brain barrier penetration A of hit compounds by SwissADME: Legend referred to the references [12,21]: BBB (yellow or yolk regions), HIA (elliptical region, Egan Egg), P-gp substrates, P-gp non-substrates

residues in the S1 pocket, Arg125 and Glu206 in the S2 subsites, Tyr 547 in the S1' subsites, and Trp629 in the S2' subsites. In accordance with the previous studies where Arulmozhiraja *et al*. revealed that the large-sized compounds that bind to hydrophobic pockets through hydrophobic interactions with lower interaction energy tend to have greater inhibitory activity than smaller compounds with higher binding energies [15]. Thus, we propose that Nb-demethylechitamine may also have potential DPP4 inhibitory activity.

In this study, l-noradrenaline emerged as a compound that could potentially inhibit DPP4 activity. L-noradrenaline has five strong hydrogen bonds, namely, at Arg125 (3.06) and Glu205 (2.64) in the S2 subsites, along with Ser630 (3.09), Asn710 (3.00), Tyr631 (3.20), and Tyr666 (2.76) in the S1 subsites. In addition, hydrophobic interaction was observed at Glu 206 in the S2 subsites, Tyr 547 in the S1' subsites, and Ser630 and Tyr662 in the S1 subsites.

As shown in Fig. 3 and 6, octopamine has a smaller size but has quite a lot of interaction with the active site of DPP4. It has six strong hydrogen bonds and two weak hydrogen bonds based on their distances, namely, Arg125 (2.87), Glu205 (2.71), and Glu206 (2.81 and 3.26) in the S2 subsites, Asn710 (3.13) and Tyr631 (2.79) in the S2 subsites, and Trp629 (3.26) in the S2' subsites. Hydrophobic interactions only occur at three residues in the S1 subsites (Ser630 and Tyr662) and S1' subsites (Tyr547); this is possible considering the small size of the compound.

For isoalliin, this compound has a small size but has many interactions with the active site of DPP4 (Table 2 and Figs. 3 and 6). Based on these

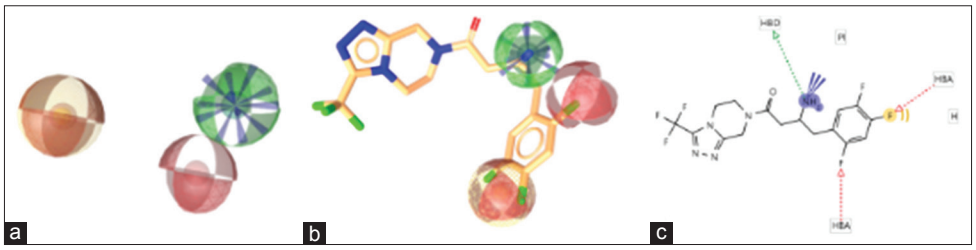


Fig. 5: (a-c) Pharmacophore features of the best model

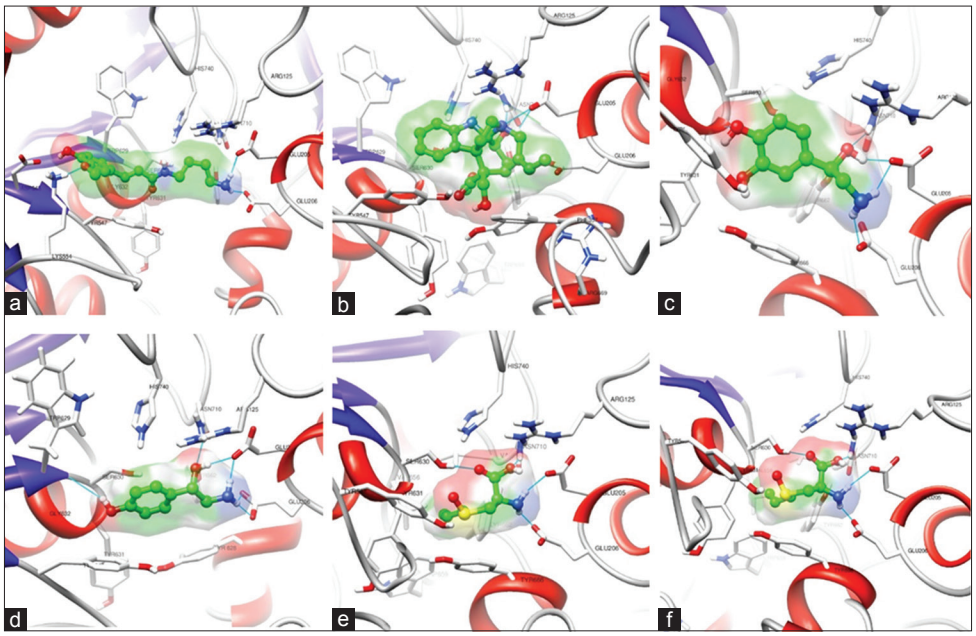


Fig. 6: Three-dimensional visualization of interactions between the hit compounds with energy values <−6.50 kcal/mol and dipeptidyl peptidase-4 (a) subaphylline (b) Nb-demethylechitamine (c) L-Noradrenaline (d) octopamine (e) isoalliin (f) alliin

findings, we believe this compound may be interesting to develop. As shown in Figs. 3 and 6, isoalliin has five strong hydrogen bonds and one weak hydrogen bond, namely, Arg125 (2.64) and Glu205 (2.67 and 3.02) in the S2 subsites, as well as Ser630 (3.09), Asn710 (2.61), and His740 (3.34), which are the triad catalytic residues in the S1 subsite. In addition, there were at least seven hydrophobic interactions identified at Glu 206 in the S2 subsites, Val656 and Tyr 547 in the S1' subsite, and the hydrophobic pocket constituent residues Tyr631, Val656, Trp659, Tyr662, and Tyr666.

Alllin has a molecular structure, type, and number of interactions almost the same as isoalliin. Alliin has six strong hydrogen bonds, namely, at Arg125 (2.70), Glu205 (2.77), and Glu 206 (2.71) in the S2 subsite, along with the triad catalytic residues Ser630 (3.20), Asn710 (2.69), and Tyr662 (3.11) in the S1 subsite. Hydrophobic interactions appear to occur at the S1' subsite with Tyr 547 and the S1 hydrophobic pocket constituent residues Tyr631, Val656, Trp659, and Tyr666.

Bioavailability and medicinal chemistry predictions

Drug activity is strongly influenced by bioavailability, which is one factor that must be considered in the search and development of a new drug compound. There are several concepts related to the assessment of bioavailability, including drug-likeness and lead-likeness, in terms of *in silico* studies and pharmacokinetic parameters. In addition, an equally important assessment is the possibility that a compound can be synthesized. Therefore, we performed various calculations related to bioavailability and medical chemistry parameter predictions using SwissADME free web tools [9] (http://www.swissadme.ch).

We further determined that subaphylline was the most potent compound; this result is well correlated with our previous molecular docking assessments (ΔG value and interaction with DPP4). This compound is in accordance with all the rules of the drug-likeness concept, which is a method developed to evaluate the relationship of the physicochemical properties of a compound using oral absorption and ADME parameters (Table 3), namely, MW (264.32 g/mol), consensus log Po/w (1.45), Number of H-bond acceptors (4), Number of H-bond donors (3), MR (74.79), num. atoms (39), topological polar surface area (TPSA) (84.58 $Å^2$), NRotBs (8), num. ring (1), and num. heteroatoms (5). This compound is also predicted to have good solubility in water (Table 3), thus it is preferred to be developed. For pharmacokinetic parameters, the BOILED-Egg prediction model (Fig. 4) shown that this compound has good absorption in the gastrointestinal (GI) tract and does not penetrate the blood-brain barrier (elliptical region, Egan egg). Further predictions demonstrated that subaphylline is not likely a P-gp substrate, but rather inhibits CYP2D6. For the identification of problematic fragments, pan-assay interference compounds (PAINS) [18] predictions shown that the compound does not contain fragments that have the potential to give false-positive results; however, Brenk [19] predictions showed that the compound contains a Michael acceptor (C=C C=O) that is reactive and potentially toxic [19,20]; thus it is essential to evaluate the compound with further testing. Finally, predictions regarding the accessibility of these compounds to be synthesized indicated that the compounds would likely be easy to synthesize (synthetic accessibility scale, 2.4). Overall, subaphylline has good oral bioavailability, is suitable for synthesis, and meets the criteria as a lead compound (lead-likeness; only one violation, NRotBs >7).

As shown in Table 3 and Fig. 4, similar to the predictions for subaphylline, Nb-demethylechitamine was also considered in accordance with all the rules of the drug-likeness concept, namely, MW (384.47 g/mol), consensus log Po/w (1.87), NHBA (5), NHBD (3), MR (111.84), num. atoms (54), TPSA (84.58 $Å^2$), NRotBs (3), num. ring (3), and num. heteroatoms (6). Nb-demethylechitamine has good solubility in water, good absorption in the GI tract and does not penetrate the blood-brain barrier. In addition, it is a P-gp substrate and does not inhibit cytochrome P450. For the prediction of fragments, PAINS predictions shown that the compound does not contain fragments that have the potential to give false positive results, but Brenk predictions demonstrated that the compound contains isolated alkenes (C=C C – Csp^3or H) that may be reactive and potentially toxic [19]. However, predictions regarding the accessibility of these compounds to be synthesized show that compounds are moderately synthesized (synthetic accessibility scale, 5.9). Overall, Nb-demethylechitamine has good oral bioavailability, suitable to be synthesized and meets the criteria as a lead compound (lead-likeness; only one violation, MW >350).

For isoalliin and alliin, overall, they met the rules of the drug-likeness concept from Lipinski *et al*., Ghose *et al*., Veber *et al*., Egan *et al*., and Muegge *et al*. (with only two violations; MW <200, XLOGP3 <−2). They have similar physicochemical properties; MW (177.22 g/mol), consensus log Po/w (−1.21 and −1.33), NHBA (4), NHBD (2), MR (43.24), num. atoms (22), TPSA (99.6 $Å^2$), NRotBs (4 and 5), num. ring (0), and num. heteroatoms (5). Isoalliin and alliin are highly soluble in water, have good absorption in the GI tract, and do not penetrate the blood-brain barrier. They may be P-gp substrates but do not inhibit cytochrome P450. For the prediction of problematic fragments, isoalliin does not have structural problems (no alert in the PAINS and Brenk predictions), but alliin showed that the compound contains isolated alkenes (C=C C–Csp^3or H) that may be reactive and potentially toxic [19]. However, for the accessibility of these compounds to be synthesized, our results indicated that both compounds can be synthesized (synthetic accessibility scale, 3.58 and 3.21). Overall, isoalliin and alliin have good oral bioavailability, are suitable for synthesis, and meet the criteria as lead compounds (lead-likeness; only one violation, MW <200), Table 3 and Fig. 4.

Similar to the other compounds, as shown in Table 3 and Fig. 6, octopamine meets the rules of the drug-likeness concept from Lipinski *et al.*, Ghose *et al*. (with only one violation; MW <160), Veber *et al*., Egan *et al*., and Muegge *et al*. (with only one violation; MW <200). Octopamine has physicochemical properties of MW (153.18 g/mol), consensus log Po/w (0.2), NHBA (3), NHBD (3), MR (42.11), num. atoms (22), TPSA (66.48 $Å^2$), NRotBs (2), num. ring (1), and num. heteroatoms (3). Octopamine is very soluble in water, has good absorption in the GI tract, and does not penetrate the blood–brain barrier. It is not a likely a P-gp substrate and does not inhibit cytochrome P450. Overall, octopamine is considered a chemically good compound; it does not have problematic fragments (no alert in the PAINS and Brenk predictions) and has excellent accessibility to be synthesized (synthetic accessibility scale, 1.35). In addition, octopamine has good oral bioavailability and meets the criteria as a lead compound (lead-likeness; only one violation, MW <200).

For l-noradrenaline, we do not discuss the results of bioavailability and medicinal chemistry predictions because this compound is already available as a drug. Further development can be carried out with or without structural modification related to the possibility that l-noradrenaline can inhibit DPP4 based on the results of this study.

CONCLUSION

Based on the results of the *in silico* screening we performed, we conclude that the six hits obtained, namely, l-noradrenaline, octopamine, Nb-demethylechitamine, alliin, isoalliin, and subaphylline have the potential to be further developed or investigated related to their activities as inhibitors of DPP4. Moreover, combined with the other analyses that we have performed with regard to molecular docking, bioavailability, and medicinal chemistry predictions, we propose that subaphylline is the most recommended compound for further studies focused on molecular dynamics, *in vitro* and *in vivo* experiments, and structural modification.

ACKNOWLEDGMENTS

The authors would like to thank the Publikasi Internasional Terindeks untuk Tugas Akhir Mahasiswa (PITTA) 2019 grant from the Universitas Indonesia for providing the funds used for this study.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

1. Tolba MK, Khashab KA, Said AS. The effect of dipeptidyl peptidase-4 inhibitors on cardiovascular disease risk in type 2 diabetes mellitus. Int J Pharm Pharm Sci 2016;9:254-9.
2. Drucker DJ. Dipeptidyl peptidase-4 inhibition and the treatment of type 2 diabetes: Preclinical biology and mechanisms of action. Diabetes Care 2007;30:1335-43.
3. Kang NS, Ahn JH, Kim SS, Chae CH, Yoo SE. Docking-based 3D-QSAR study for selectivity of DPP4, DPP8, and DPP9 inhibitors. Bioorg Med Chem Lett 2007;17:3716-21.
4. Lankas GR, Leiting B, Roy RS, Eiermann GJ, Beconi MG, Biftu T, *et al*. Dipeptidyl peptidase IV inhibition for the treatment of type 2 diabetes: Potential importance of selectivity over dipeptidyl peptidases 8 and 9. Diabetes 2005;54:2988-94.
5. Patel BD, Ghate MD. Recent approaches to medicinal chemistry and therapeutic potential of dipeptidyl peptidase-4 (DPP-4) inhibitors. Eur J Med Chem 2014;74:574-605.
6. Sneha P, Doss CG. Gliptins in managing diabetes reviewing computational strategy. Life Sci 2016;166:108-20.
7. Wermuth CG. Pharmacophores: Historical perspective and viewpoint from a medicinal chemist. In: Langern T, Hoffmann RD, editors. Pharmacophores and Pharmacophore Searches. Weinheim, Germany: Wiley-VCH; 2006. p. 3-13.
8. Wolber G, Langer T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. J Chem Inf Model 2005;45:160-9.
9. Seidel T, Ibis G, Bendix F, Wolber G. Strategies for 3D pharmacophore-based virtual screening. Drug Discov Today Technol 2010;7:221-8.
10. Umashankar V, Gurunathan S. Drug discovery: An appraisal. Int J Pharm Pharm Sci 2015;7:59-66.
11. Pissarnitski DA, Zhao Z, Cole D, Wu WL, Domalski M, Clader JW, *et al*. Scaffold-hopping from xanthines to tricyclic guanines: A case study of dipeptidyl peptidase 4 (DPP4) inhibitors. Bioorg Med Chem 2016;24:5534-45.
12. Daina A, Michielin O, Zoete V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep 2017;7:e42717.
13. Arulmozhiraja S, Matsuo N, Ishitsubo E, Okazaki S, Shimano H, Tokiwa H. Comparative binding analysis of dipeptidyl peptidase IV (DPP-4) with antidiabetic drugs an ab initio fragment molecular orbital study. PLoS One 2016;11:e0166275.
14. Nojima H, Kanou K, Terashi G, Takeda-Shitaka M, Inoue G, Atsuda K, *et al*. Comprehensive analysis of the Co-structures of dipeptidyl peptidase IV and its inhibitor. BMC Struct Biol 2016;16:11.
15. Shamsara J. Correlation between virtual screening performance and binding site descriptors of protein targets. Int J Med Chem 2018;2018:3829307.
16. Nabeno M, Akahoshi F, Kishida H, Miyaguchi I, Tanaka Y, Ishii S, *et al*. A comparative study of the binding modes of recently launched dipeptidyl peptidase IV inhibitors in the active site. Biochem Biophys Res Commun 2013;434:191-6.
17. Desiraju GR, Steiner T. The Weak Hydrogen Bond. Oxford: Oxford University Press; 2001. p. 1-28.
18. Baell JB, Holloway GA. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. J Med Chem 2010;53:2719-40.
19. Brenk R, Schipani A, James D, Krasowski A, Gilbert IH, Frearson J, *et al*. Lessons learnt from assembling screening libraries for drug discovery for neglected diseases. ChemMedChem 2008;3:435-44.
20. Schultz TW, Yarbrough JW, Hunter RS, Aptula AO. Verification of the structural alerts for Michael acceptors. Chem Res Toxicol 2007;20:1359-63.
21. Daina A, Zoete V. A BOILED-egg to predict gastrointestinal absorption and brain penetration of small molecules. ChemMedChem 2016;11:1117-21.
22. Ertl P, Rohde B, Selzer P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. J Med Chem 2000;43:3714-7.
23. Daina A, Michielin O, Zoete V. iLOGP: A simple, robust, and efficient description of n-octanol/water partition coefficient for drug design using the GB/SA approach. J Chem Inf Model 2014;54:3284-301.
24. Cheng T, Zhao Y, Li X, Lin F, Xu Y, Zhang X, *et al*. Computation of octanol-water partition coefficients by guiding an additive model with knowledge. J Chem Inf Model 2007;47:2140-8.
25. Moriguchi I, Hirono S, Liu Q, Nakagome I, Matsushita Y. Simple method of calculating octanol/water partition coefficient. Chem Pharm Bull 1992;40:127-30.
26. Moriguchi I, Hirono S, Nakagome I, Hirano H. Comparison of reliability of log p values for drugs calculated by several methods. Chem Pharm Bull 1994;42:976-8.
27. Delaney JS. ESOL: Estimating aqueous solubility directly from molecular structure. J Chem Inf Comput Sci 2004;44:1000-5.
28. Ali J, Camilleri P, Brown MB, Hutt AJ, Kirton SB. Revisiting the general solubility equation: *In silico* prediction of aqueous solubility incorporating the effect of topographical polar surface area. J Chem Inf Model 2012;52:420-8.
29. Ghose AK, Viswanadhan VN, Wendoloski JJ. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. J Comb Chem 1999;1:55-68.
30. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem 2002;45:2615-23.
31. Egan WJ, Merz KM Jr, Baldwin JJ. Prediction of drug absorption using multivariate statistics. J Med Chem 2000;43:3867-77.
32. Muegge I, Heald SL, Brittelli D. Simple selection criteria for drug-like chemical matter. J Med Chem 2001;44:1841-6.
33. Martin YC. A bioavailability score. J Med Chem 2005;48:3164-70.
34. Ertl P, Schuffenhauer A. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. J Cheminform 2009;1:8.