Research article

Identification of potential dipeptidyl peptidase IV inhibitors from the ConMedNP library by virtual screening, and molecular dynamics methods

Hans Merlin Tsahnang Fofack [a,b,1], Maraf Mbah Bake [c,d,1,*], Simon Petry [e], Baruch A. Ateba [a,b], Pascal Amoa Onguéné [f], Haydar Mohammad-Salim [g,h], Fidele Ntie-Kang [i,j,k], Luc Meva'a Mbaze [l,**], Serhii Vakal [m,***], Cyril A Kenfack [a,****]

[a] Laboratoire Optique et Applications, Centre de Physique Atomique Moleculaire et Optique Quantique, Faculte des Sciences, Université de Douala, B.P. 8580, Douala, Cameroon
[b] Analytical, Structural and Materials Chemistry Laboratory, Department of Chemistry, Faculty of Science, University of Douala, B.P. 24157, Douala, Cameroon
[c] Physical and Theoretical Chemistry Unit, Laboratory of applied Physical and Analytical Chemistry, Faculty of Science, University of Yaoundé I, P.O. BOX 812, Yaoundé, Cameroon
[d] Computational Chemistry Laboratory, Department of Chemistry, Higher Teacher Training College, University of Yaoundé I, P. O. Box 47, Yaoundé, Cameroon
[e] Institute of Chemistry and Biochemistry, Freie Universität Berlin, Takustraße 3, 14195, Berlin, Germany
[f] University Institute of Wood technology, Mbalmayo, Cameroon
[g] Department of Chemistry, Faculty of Science, University of Zakho, Zakho, Duhok, 42001, Kurdistan Region, Iraq
[h] Molecular Topology and Drug Design Research Unit, Department of Physical Chemistry, Pharmacy Faculty, University of Valencia, 46100, Valencia, Spain
[i] Center for Drug Discovery, Faculty of Science, University of Buea, P. O. Box 63, Buea, Cameroon
[j] Department of Chemistry, Faculty of Science, University of Buea, P. O. Box 63, Buea, Cameroon
[k] Institute of Pharmacy, Martin-Luther University Halle-Wittenberg, Kurt-Mothes Strasse 3, 06120, Halle (Saale), Germany
[l] Physical and Theoretical Chemistry Laboratory, Faculty of Science, University of Yaoundé I, P.O. Box 812, Yaoundé, Cameroon
[m] Structural Bioinformatics Lab, Faculty of Science and Engineering, Åbo Akademi University, Tuomiokirkontori 3, 20500, Turku, Finland

ARTICLE INFO

ABSTRACT

In this study, we screened novel dipeptidyl peptidase IV (DPP4) inhibitors from the ConMedNP library consisting of 3507 molecules. Interestingly, molecular docking, ADMET, and the anti-diabetic activity predictions suggest that three molecules, namely OTH_UD_XX06_1, GB19, and BMC_000104, have a high binding affinity toward DPP4. The molecular dynamics (MD) simulation results suggest that these hit molecules have a stable binding pose and occupy the binding pockets throughout the 200 ns simulation. The presence of intermolecular H-bonding between the ligands and DPP4 was observed throughout the simulation period. Thus, docking and MD results, predicted that the three compounds were the most potent DPP4 inhibitors that could putatively

* Corresponding authors.
** Corresponding author.
*** Corresponding authors.
**** Corresponding author.
E-mail addresses: mbahbakemaraf@yahoo.com (M. Mbah Bake), fidele.ntie-kang@ubuea.cm (F. Ntie-Kang), lmbazze@yahoo.fr (L.M. Mbaze), serhii.vakal@gmail.com (S. Vakal), ckenf@yahoo.com (C.A. Kenfack).
[1] These authors contributed equally to this work.

bind to the DPP4 active site *via* both conventional H-bonding and hydrophobic interactions. These results could aid the discovery of new drugs to treat type 2 diabetes.

1. Introduction

Diabetes refers to a group of metabolic disorders distinguished by hyperglycemia in the absence of treatment [1]. The prevalence of diabetes is on the rise across the globe. In 2019, 463 million people suffered from the disease and this figure is predicted to rise to 578 million by 2023 and 700 million by 2045 [2]. Furthermore, 10 % of death cases has been accounted to diabetes [3]. The most widespread form of the disease is type 2 diabetes mellitus (T2DM), and it represents 90 % of diagnosed cases. T2DM is defined by insufficient insulin production in the pancreas and insulin resistance in peripheral tissues [4]. The chronic hyperglycemia associated to diabetes can lead to long-term damage and dysfunction of vital organs such as kidneys, eyes, heart and nerves [5,6]. Diabetes management and its related challenges impacts not just individuals with the ailment, but also their family, as it requires a reduction in their salary and affects the health system in their community [7].

To effectively address type 2 diabetes, a highly recommended approach involves inhibiting the degradation of incretin hormones by the dipeptidyl peptidase IV (DPP4) enzyme [8]. Incretin hormones encompass a group of hormones such as, glucagon-like peptide-1 (GLP-1) and glucose-dependent insulinotropic polypeptide (GIP), which are secreted from the intestine after a meal and stimulate the production of insulin in the pancreas [9]. Currently in the United States of America and Europe, four DPP4 inhibitors has been approved namely sitagliptin, saxagliptin, alogliptin, and vildagliptin (the latter being approved in Europe) [7]. However, DPP4 inhibitors are associated with various drawbacks such as arthritis, pancreatitis, diarrhea, and congestive heart failure, which restrict their usage [10–12].

The current pharmacotherapy for diabetes includes insulin administration, secretagogues, euglycemic agents, biguanides, thiazolidi-nediones, and glycosity inhibitors. Additionally, combinations of oral hypoglycemic drugs are accessible such as metformin with glibenclamide; rosiglitazone with metformin; and vildagliptin with metformin and nateglinide with metformin. However, these drugs can eventually lead to severe side effects including heart complications, bone density loss, fluid retention, weight gain, and digestive and urinary tract problems [13,14]. In light of this situation, it is crucial to search for novel medications that offer effectiveness with low and tolerable side effects, ideally at lower cost to ensure broader access to treatment. *In silico* approaches have proven their effectiveness in identifying new medications and contributing to their therapeutic development process [15,16]. This research aims to identify potential DPP4 inhibitors using virtual screening, molecular dynamics, and MM-PBSA binding energy computation methods from the ConMedNP library, a natural product library from central African medicinal plants. To our knowledge, this library has never been used to investigate potential DPP4 inhibitors.

2. Methods and materials

2.1. Protein preparation and grid generation

The 3D crystal structure of the human DPP4 enzyme (PDB ID: 1X70) was retrieved from the Protein Data Bank website (https://www.rcsb.org/) for SBVS with a resolution of 2.10 Å. The active site of 1X70 contains the co-crystallized inhibitor sitagliptin [17]. The protein structure was prepared for docking using UCSF Chimera [18] and AutoDock Tools (ADT) 1.5.6 [19]. The PDB structure of DPP4 contains two chains, A and B, chain A was considered for docking analysis. The co-crystallized sitagliptin, 2-acetamido-2-deoxy-γ-β-ᴅ-glucopyranose, water molecules, and sodium ions were removed from the protein structure. The protein target was then minimized using the force field Amberff14SB implemented in UCSF Chimera. Polar hydrogen atoms, as well as Gasteiger charges, were added, while steepest descent and conjugate gradient parameters were kept default. The 3D grid parameters of the active site were generated using Chimera with the following parameters: center X: −35.84, Y: 50.25, Z: 35.31, and size X: 25, Y: 22, Z: −18 with 0.375 Å point spacing [20]. The grid box was made to cover key amino acids residues TYR631, VAL656, TRP659, TYR662, TYR666, and VAL711 as well as ARG125, GLU205, GLU206, SER209, PHE357, and ARG358 which constitute the S1 and S2 pockets of the DPP4 enzyme, respectively [21]. Finally, AutoDock tools were used to assign united atom Kollman charges, fragmental volumes, solvation parameters, and Autodock atom type to the prepared protein file. The generated file was then saved in the PDBQT format for further docking analysis [22].

2.2. Ligand preparation

The ConMedNP library [23] consisting of 3507 molecules in mol2 format was screened for potential DPP4 inhibitors. The library was chosen due to its diversity in secondary metabolites, since it is composed of 376 distinct medicinal plant species belonging to 79 plant families from the Central African flora. OpenBabel software was used to minimize the molecular library by applying the mmff94 force field and to convert the mol2 format of the molecules into the AutoDock-compatible PDBQT format [24].

2.3. Virtual screening of the ConMedNP

Structure-based virtual screening campaign was performed against DPP4 enzyme using Autodock vina with docking parameters on

defaults mode. The software employs knowledge-based scoring functions to predict the compound's binding affinity with the receptor [25]. Before the virtual screening process was executed, the native ligand was removed and redocked in the protein's active site for validation of the docking method. Compounds with binding affinities less than or equal to −7.96 kcal/mol were chosen for further investigation [26].

2.4. Lipinski's rule of five and ADMET filtering

Drug-likeness (Lipinski rule of five) as well as pharmacokinetic and toxicity profiles (ADMET) of selected molecules from the docking stage were evaluated using ADMETlab 2.0 server online [27]. The canonical smile format of molecules served as input. Molecules with excellent Lipinski and ADMET profiles were selected for anti-diabetic predictions.

2.5. Anti-diabetic activity prediction of the hit compounds

Anti-diabetic prediction of the 15 hits with the best drug-likeness and ADMET profiles was performed using the PASS (Prediction of Activity Spectra for Substances) online server [28]. PASS employs a database of approximately 26000 molecules with known biological activity to predict the activity of the compounds being studied. For each molecule PASS forecast the probability of activity (P_a) and inactivity (P_i) on a scale from 0.000 to 1.000. Molecules with P_a values above 0.3 with P_i value less than 1 are deemed to warrant further exploration of the pharmacological activity.

2.6. Analysis of protein-ligand interactions

After the docking-based virtual screening, Lipinski's rule and ADMET filtering, and the anti-diabetic activity predictions, Discovery Studio Visualizer was used to visualize the 3D interactions of the top 15 hits were visualized as well as to generate detailed 2D and 3D interaction diagram of the best three hits with the predicted anti-diabetic activities. The 2D and 3D diagrams show various interactions such as conventional hydrogen bonds and hydrophobic interactions [29]. Ligplot + v.2.1 was used to further analyze the 2D interactions of the selected hits [30].

2.7. Molecular dynamics simulation of protein-ligand complexes

Pymol was used to prepare the protein-ligand complex structure of DPP4 and candidate molecules for 200ns MD simulations (MDs). GROMACS v2019.4 software was utilized to carry out MDs of the protein–ligand complexes [31]. The AMBER force field AMBER99SB-ILDN protein, nucleic AMBER94 was used to generate their protein topologies. Topologies parameters of ligands were generated using the AnteChamber PYthon Parser interfacE (ACPYPE) server [32]. The complexes were placed in a cubic box of 1.0 nm size and solvated using the TIP3P water model. The systems were then neutralized by addition of chloride (Cl$^-$) and sodium (Na$^+$) ions. Minimization of each complex energy was performed in 50000 steps using steepest descent algorithm and then equilibrated for 200 ps using NVT and NPT ensemble. After the equilibration phase of each system, a 200 ns production MD was conducted. The trajectories were set to be generated every 10 fs and saved every 10 ps. Finally the trajectories of the unbound proteins and selected complexes were analyzed using parameters such as: root-mean-square-deviation (RMSD), root-mean-square-fluctuation (RMSF), and radius of gyration (R_g).

2.8. Molecular mechanics/generalized born surface area (MM-GBSA) free binding energy calculation

The MM-GBSA free energy of the hit molecules (OTH_UD_XX06_1, BMC_000104, and GB19) with DPP4 was computed using gmx_MMPBSA tool. The free energy of binding, i.e ΔG_{bind} of a ligand (L) binding to a protein (P) to form the protein-ligand complex (PL) can be calculated using the molecular mechanics Poisson−Boltzmann surface area (MM-PBSA) approach. The equations of this approach are:

$$\Delta G_{bind} = \Delta G_{(PL)} - (\Delta G_{(P)} + \Delta G_{(L)}) \tag{1}$$

each energy term to the right in the equation can be written as:

$$\Delta G_{bind} = \Delta H - T\Delta S = \Delta E_{MM} + \Delta G_{sol} - T\Delta S \tag{2}$$

where ΔE_{MM} corresponds to the molecular mechanical energy changes in the gas phase, ΔG_{sol} represents the solvation energy difference, T denotes absolute temperature, and ΔS is the entropy change. ΔE_{MM} can further be divided into three components: internal energies (ΔE_{int}), electrostatic energies (ΔE_{ele}), and van der Waals energies (ΔE_{vdW}):

$$\Delta E_{MM} = \Delta E_{int} + \Delta E_{ele} + \Delta E_{vdW} \tag{3}$$

For each complex, snapshots were extracted at 1000 frame intervals at each ns from 10 to 200 ns of the MD simulation. The dielectric model for PB was set to 2, the aqueous solvent's dielectric constant was set to 80, and the interior dielectric constant was set to 4 for the gmx MM-PBSA run [33].
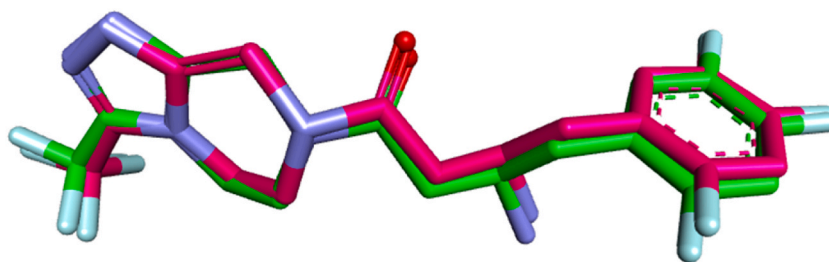
Fig. 1. Superimposition of Docking Validation protocol.

Table 1
ADMETlab2.0 physicochemical properties predictions of ligands.

Ligand	MW (g/mol)	LogP	LogS (logmol/L)	HBD	HBA	TPSA (Å^2)
OTH_AS_003_1	412.15	5.057	−2.837	4	7	120.36
ABD_UD_006	480.31	1.972	−3.29	6	7	138.45
JTM_UY_028	398.21	3.706	−3.744	2	5	91.67
OTH_UD_XX06_1	406.13	0.835	−3.259	4	9	138.82
UB_AYI_010	382.15	0.015	−1.934	4	7	110.1
17.3.1.7.9	472.32	4.484	−3.906	4	5	97.99
BMC_00020	404.15	0.853	−3.174	2	8	118.73
BMC_00021	416.18	1.99	−3.855	1	7	98.5
UB_AYI_011	366.16	0.421	−1.898	3	6	89.87
GB19	504.35	3.878	−3.561	5	6	118.22
BMC_000104	290.08	1.343	−2.739	5	6	110.38
PTA_UDS_126	350.21	2.364	−3.804	2	5	79.29
UB_AYI_022	386.23	0.632	−2.218	7	5	127.45
BMC_000116	364.22	2.402	−3.489	2	5	79.29
UB_AYI_020	350.21	2.364	−3.804	2	5	79.29

2.9. Estimation of scaffolds novelty of the selected hits

The estimation of the uniqueness of the scaffolds of the selected ligands was performed using the Schrodinger Canvas software [34]. The highest pairwise Tanimoto similarity of each ligand in relation to the 460 validated human DPP-IV inhibitors from the ChEMBL database was determined using the extended chemical fingerprints for four atoms (ECFP4). Compounds with the lowest Tanimoto coefficient (T_c) scores were identified as potential sources of new chemical scaffolds for the development of DPP4 inhibitors.

3. Results and discussion

3.1. Docking-based virtual screening

To validate the docking protocol, before performing the docking-based virtual screening, the co-crystallized sitagliptin was removed and re-docked into the active site of DPP4, and the results are depicted in Fig. 1. The redocked pose of the co-crystallized ligand, shown in magenta ball and stick representation, was superimposed with the original crystal ligand, depicted in a green ball and stick representation, to calculate the root mean square deviation (RMSD). The re-docked ligand reproduced the original pose with an RMSD value of 1.483 Å calculated using PyMol [35] and a binding affinity of −9.4 kcal/mol. In addition, the approved drugs sitagliptin and saxagliptin were docked in the protein target active site. The predicted binding affinity values were −9.4 and −6.4 kcal/mol, respectively. The average of their binding affinity values ($\Delta G \leq -7.9$ kcal/mol) was defined as a baseline energy value for the selection of molecules to be studied further. Out of the 3507 successfully screened phytochemicals, 1182 had a binding affinity less than −7.9 kcal/mol.

3.2. Post-processing, drug-likeness, and ADMET predictions

To ensure the safety and efficacy of the hit molecules, Lipinski's rule of five, pharmacokinetics, and toxicity profiles of the ligands were evaluated. Out of the 1182 molecules selected from docking-based virtual screening, 969 molecules obeyed the rule of five [i.e., molecular weight ≤500, number of hydrogen bond donors ≤5, number of hydrogen bond acceptors ≤10, and LogP ≤5] as can be seen in Table 1. Compounds with logS ranging from −4 to 0.5 log mol/L and TPSA values within the range from 0 to 140 were selected.

Furthermore, ADMET descriptors HIA, Caco-2 permeability, MDCK permeability, plasma protein binding (PPB), VD, BBB penetration, CYP450 inhibition, CL, $T_{1/2}$, AMES, and carcinogenicity modules were used to predict the pharmacokinetic and toxicity parameters of the 969 candidate molecules. We eliminated molecules with poor absorption and distribution parameters, easily crossing

Table 2
ADMETlab2.0 pharmacokinetic and toxicity predictions for the selected ligands.

Ligand	HIA	BBB	Caco-2	PPB (%)	hERG Blockers	H-HT	AMES Toxicity	Carcinogencity	CYP 1A2	CYP2C19	CYP2C9	CYP2D6	CYP3A4
OTH_AS_003_1	0.130	0.008	−4.865	87.661	0.010	0.439	0.466	0.209	0.379	0.141	0.764	0.584	0.12
ABD_UD_006	0.116	0.223	−5.115	47.727	0.086	0.277	0.091	0.834	0.345	0.005	0.009	0.000	0.154
JTM_UY_028	0.114	0.207	−4.802	85.356	0.025	0.546	0.065	0.340	0.147	0.076	0.028	0.015	0.165
OTH_UD_XX06_1	0.612	0.204	−5.291	79.093	0.460	0.173	0.152	0.504	0.254	0.038	0.009	0.032	0.038
UB_AYI_010	0.030	0.132	−5.249	30.452	0.028	0.135	0.066	0.527	0.008	0.017	0.013	0.006	0.099
17.3.1.7.9	0.025	0.614	−5.746	81.522	0.027	0.435	0.005	0.582	0.006	0.006	0.016	0.002	0.415
BMC_00020	0.029	0.188	−5.278	64.672	0.051	0.342	0.112	0.324	0.009	0.020	0.021	0.005	0.074
BMC_00021	0.006	0.188	−5.187	69.469	0.060	0.390	0.083	0.275	0.014	0.025	0.036	0.007	0.617
UB_AYI_011	0.026	0.265	−5.172	27.689	0.029	0.179	0.063	0.482	0.012	0.019	0.008	0.003	0.058
GB19	0.018	0.378	−5.646	88.056	0.005	0.220	0.007	0.012	0.009	0.004	0.017	0.003	0.119
BMC_000104	0.035	0.029	−6.052	89.230	0.033	0.103	0.604	0.185	0.360	0.038	0.290	0.184	0.432
PTA_UDS_126	0.003	0.392	−4.975	21.664	0.020	0.391	0.019	0.040	0.017	0.018	0.021	0.005	0.503
UB_AYI_022	0.010	0.150	−5.176	23.207	0.066	0.154	0.030	0.060	0.007	0.005	0.002	0.001	0.021
BMC_000116	0.003	0.409	−4.800	46.133	0.016	0.111	0.476	0.058	0.015	0.021	0.016	0.003	0.360
UB_AYI_020	0.003	0.392	−4.975	21.664	0.020	0.391	0.019	0.040	0.017	0.018	0.021	0.005	0.503

Table 3
Predicted anti-diabetic activity of ligands compared with known DPP4 inhibitors using way2drug server.

Ligand	P_a	P_i
OTH_AS_003_1	–	–
ABD_UD_006	–	–
JTM_UY_028	–	–
OTH_UD_XX06_1	0.661	0.008
UB_AYI_010	–	–
17.3.1.7.9	–	–
BMC_00020	–	–
BMC_00021	–	–
UB_AYI_011	–	–
GB19	0.462	0.029
BMC_000104	0.355	0.058
PTA_UDS_126	–	–
UB_AYI_022	–	–
BMC_000116	–	–
LBS_UY_155_1	–	–
Sitagliptin	0.595	0.013
Linagliptin	0.385	0.048
Saxagliptin	0.972	0.003
Alogliptin	0.322	0.071

P_a = Probability to be active, P_i = Probability to be inactive, – = no anti-diabetic activity found.

the blood-brain barrier (BBB), inhibiting cytochrome P450 isoenzymes (1A2, 3A4, 2C9, 2C19, and 2D6), high PPB, poor water solubility, and toxicity (high probability of carcinogenicity and mutagenesis). The remaining 14 candidates with good ADMET profiles are listed in Table 2. Human intestinal absorption (HIA) describes the absorption of orally administrated drugs from gut into the bloodstream. HIA is a necessary condition for the ligand's apparent efficacy and can be used as a substitute indicator for oral bioavailability to some extent. HIA scores greater than 30% are classified as HIA- (category 0), while HIA scores less than 30% are categorized as HIA+ (category 1), based on the HIA + scale [36]. Caco-2 cell permeability parameter plays an important role in predicting if an oral drug can enter systemic circulation via passive diffusion, carrier-mediated uptake, or active transport processes. Compounds with Caco-2 permeability parameter exceeding −5.15 log cm/s are deemed to be appropriate. PPB is a key process in drug absorption and distribution. PPB can greatly impact oral bioavailability as the drug's free concentration of the drug is affected when it binds to serum proteins. A compound is considered to have adequate PPB if it has a predicted score is ≤ 90%, and highly protein-bound drugs may have low therapeutic efficacy. Central nervous system (CNS) targeted drugs must cross the BBB to reach their intended molecular targets within the brain. On the other hand, peripheral targeted drugs may not require significant BBB penetration in order to avoid unwanted CNS-related side effects. Molecules with logBB greater than −1 are categorized as BBB+ (category 1), while those with logBB less than −1 are categorized as BBB- (category 0), according to the BBB + scale [36]. The metabolism reaction of drugs is divided into two categories: phase I (oxidative reactions) and phase II (conjugative reactions). This reaction is 80% attributed to human cytochrome P450 family isoenzymes, namely: isozymes 1A2, 3A4, 2C9, 2C19, and 2D6. These isozymes metabolize approximately two-thirds of known drugs in humans and are responsible for phase I reactions that take place in the liver. Inhibition of these isoenzymes by drugs will lead to the accumulation of these drugs and eventually cause harmful effects in the body.

We assessed the inhibition of these enzymes by molecules in our library, and those predicted to inhibit these enzymes were discarded. Non-substrates/non-inhibitors were put in category 0, while substrates/inhibitors – category 1. The output value is the probability of being substrate/inhibitor, within the range of 0–1. The human ether-a-go-related gene is important in the regulation of cardiac action potential and resting potential exchange. Long QT syndrome (LQTS), arrhythmia, and Torsade de Pointes (TdP) can result from hERG blockade, resulting in palpitations, fainting, or even sudden death. Molecules that possess IC_{50} values greater than Molecules that have an IC_{50} (half-maximal inhibitory concentration) greater than 10 μM or less than 50% inhibition at 10 μM are categorized as hERG-negative (category 0), indicating that they do not exert a substantial inhibitory effect on the hERG channel. In contrast, molecules displaying IC_{50} scores less than 10 μM or greater than 50% inhibition at 10 μM are categorized as hERG-positive (category 1), which translate they do block the hERG channel.

Human Hepatotoxicity (H-HT) implies liver damage caused by chemicals. Drug-induced liver injury is a significant factor leading to withdrawal of drugs from the market and a major cause of patient safety. Adverse hepatic effects in clinical assays frequently result in the costly and late termination of drug development programs. The result interpretation has two categories: category 0 - H-HT negative (−), -category 1 -H-HT positive (+). AMES prediction helps to estimate the mutagenicity of drugs. The mutagenic effect is closely related to carcinogenicity, and it is the most commonly used assay for evaluating compounds mutagenicity. AMES-negative (−) compounds are classified as category 0, while AMES-positive (+) ones are category 1.

Carcinogenicity is arguably the most concerning and serious toxicological effect, given its severe implications human health. Chemical carcinogenesis occurs when substances have the capacity to damage genetic material or disrupt cellular metabolic functions. A significant number of drugs that were previously approved have been withdrawn from the market after being found to be carcinogens in humans or animals. They are classified as category 1, while non-carcinogens are category 0. Chemicals are classified as either

Table 4
Docking analysis of the 3 hits.

Ligand	Interacting residues	Polar contact	Distance(Å)	Binding free energy(ΔG_{bind}) kcal/mol
OTH_UD_XX06_1	GLN553	Conventional hydrogen bond	2.03	−8.5
	HIS740	Conventional hydrogen bond	2.79	
	GLU205	Conventional hydrogen bond	2.25	
	PHE357	π-π stacked	4.27	
	TYR547	π-alkyl	4.49	
	TYR666	van der Waals		
	ARG125	van der Waals		
	ASN710	van der Waals		
	SER630	van der Waals		
	VAL656	van der Waals		
	VAL711	van der Waals		
	TYR662	van der Waals		
	TYR585	van der Waals		
	CYS551	van der Waals		
	TRP659	van der Waals		
GB19	GLN553	Conventional hydrogen bond	1.94	−8.1
	GLU205	Conventional hydrogen bond	2.16	
	PHE357	π-σ	3.74	
	PHE357	2 π-alkyl	5.15, 5.46	
	TYR666	π-alkyl	4.97	
	TYR585	van der Waals		
	TYR456	van der Waals		
	ARG429	van der Waals		
	SER552	van der Waals		
	CYS551	van der Waals		
	GLU206	van der Waals		
	SER209	van der Waals		
	ARG125	van der wails		
BMC_000104	ASN710	Conventional hydrogen bond	2.20	−8.0
	PHE357	π-π stacked	4.17	
	TYR662	π-π T-shaped	5.14	
	TYR666	π-π T-shaped	4.92	
	TYR666	π-alkyl	5.28	
	GLU206	van der Waals		
	HIS740	van der Waals		
	VAL711	van der Waals		
	VAL656	van der Waals		
	CYS551	van der Waals		
Sitaglipin	GLU205	Salt bridge	2.63, 2.72	−9.4
	GLU206	Salt bridge	2.64	
	ASP663	Attractive charge	4.88	
	GLU206	Attractive charge	2.64	
	TYR666	Pi-cation	4.61	
	ARG358	Conventional hydrogen bond	2.44	
	ARG125	Conventional hydrogen bond	2.68	
	TYR666	Conventional hydrogen bond	2.52, 2.54	
	VAL207	Halogen(Fluorine)	3.07, 3.50	
	GLU206	Halogen(Fluorine)	3.6	
	PHE357	π-π stacked	4.05	
	TYR662	π-π stacked	4.77	
	TYR666	π-π T-shaped	5.21	
	PHE357	π-alkyl	4.65	
Saxagliptin	GLU205	Attractive charge	3.10	−6.4
	GLU206	Salt bridge	4.70	
	ARG125	Conventional hydrogen bond	2.65	
	TYR547	π-alkyl	5.02	
	TYR666	π-alkyl	5.09, 5.50	
	HIS740	π-alkyl	5.49	
	TYR662	π-alkyl	5.05	
	PHE357	π-alkyl	4.45	

active (carcinogens) or inactive (non-carcinogens) based on their TD-50 values. The output value of ADMET parameters represents the probability of being HIA$^+$, BBB$^+$, T.$_{1/2}^+$, hERG$^+$, toxicity ranging from 0 to 1, with an empirical rule: 0–0.3 –excellent, 0.3–0.7 – medium, and 0.7–1.0 – poor.
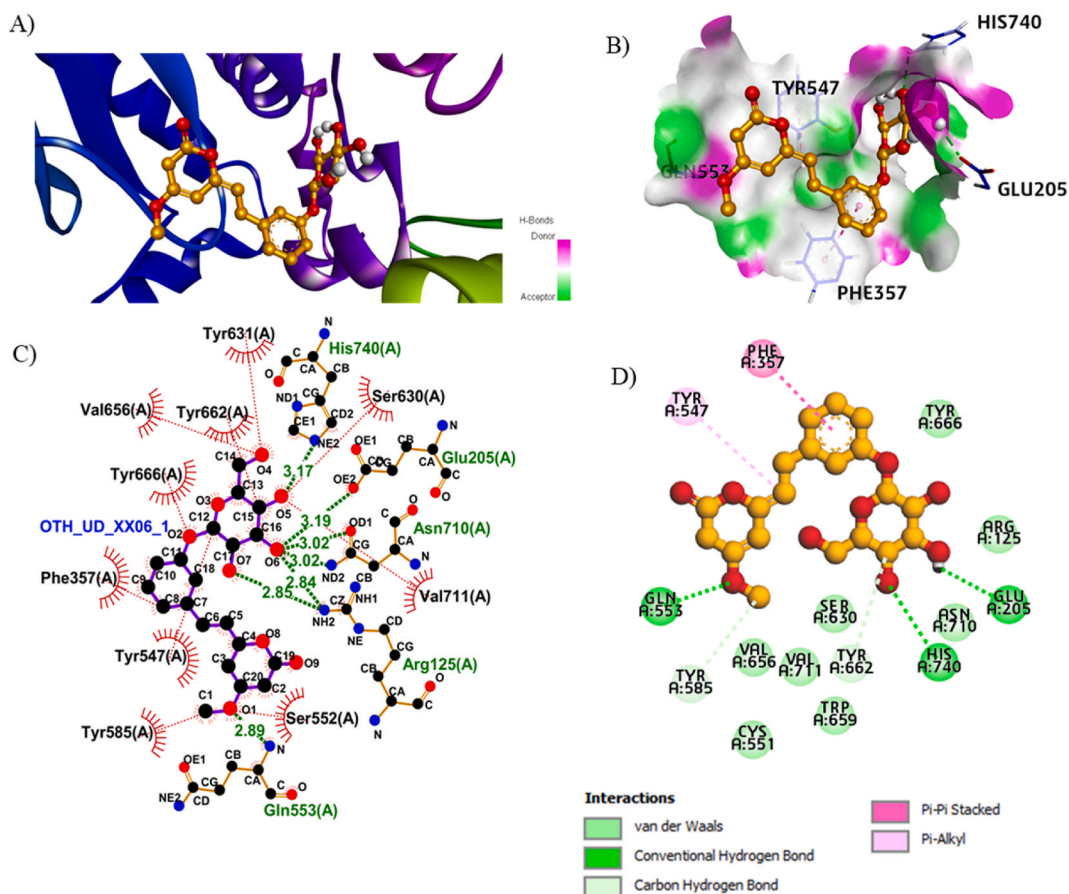
Fig. 2. Non-bonded interactions diagram of docked OTH UD_XX06_1. (A) Pose view representation, (B) Surface view representation, (C) Ligplot+ 2D interaction map, (D) Discovery studios 2D interaction map.

3.3. Anti-diabetic activity prediction

Anti diabetic activity of the 14 shortlisted compounds was invested, and the results were shown in Table 3. Compounds that exhibited anti-diabetic activity, with a $P_a > 0.3$ were selected for molecular dynamics studies. Among all these compounds, three, namely OTH_UD_XX06_1, GB19, and BMC_000104, demonstrated anti-diabetic activity with a $P_a > 0.3$ and $P_a > P_i$. OTH_UD_XX06_1 was predicted to have the highest anti-diabetic activity with a P_a of 0.661 and P_i value of 0.008, followed by GB19 with a P_a of 0.462 and P_i value of 0.029, while BMC_000104, possessed P_a and P_i values of 0.355 and 0.058. The predicted activity of the three compounds was compared with known DPP4 inhibitors, namely: Sitagliptin, Linagliptin, Saxagliptin, and Alogliptin. Saxagliptin has the highest predicted anti-diabetic activity with a P_a value of 0.972 and P_i value of 0.003, followed by Sitagliptin with a P_a value of 0.595 and P_i value of 0.013 and, lastly, Linagliptin and Alogliptin had the lowest predicted anti-diabetic activity with P_a values 0.385 & 0.322 and P_i values 0.048 & 0.071, respectively. OTH_UD_XX06_1 showed predicted activity that was greater than those of Sitagliptin, Linagliptin, and Alogliptin. GB19 also exhibited anti-diabetic activity higher than that of Linagliptin and Alogliptin, while BMC_000104 activity was greater than that of Alogliptin.

3.4. Docking studies analysis

The molecular docking results (see Table 4) of the three hit molecules showed their strong binding affinity with the DPP4 enzyme within the range from −8.5 kcal/mol to −7.9 kcal/mol, which was lower than that of Sitagliptin (−9.4 kcal/mol) but higher than that of Saxagliptin (−6.4 kcal/mol) used as a positive control. Among the three hits, OTH_UD_XX06_1 has the highest predicted affinity of −8.5 kcal/mol followed by GB19 with a binding affinity value of −8.1 kcal/mol and, lastly, BMC_000104 with an affinity value of −8.0 kcal/mol.

To further explore the inhibition mechanism of the three hit molecules, we predicted the 3D and 2D interactions of each, with the key amino acid residues playing a major role in DPP4 activity. DPP4 has various important subsites that contribute to the binding and effectiveness of inhibitors, including S1, S2, S1′, S2′, and the extensive S2 subsite. The S1 subsite is a hydrophobic pocket close to the active site and contains amino acids like SER630, VAL656, TRP659, TYR662, TYR666, ASN710, and VAL711. The S2 subsite is another
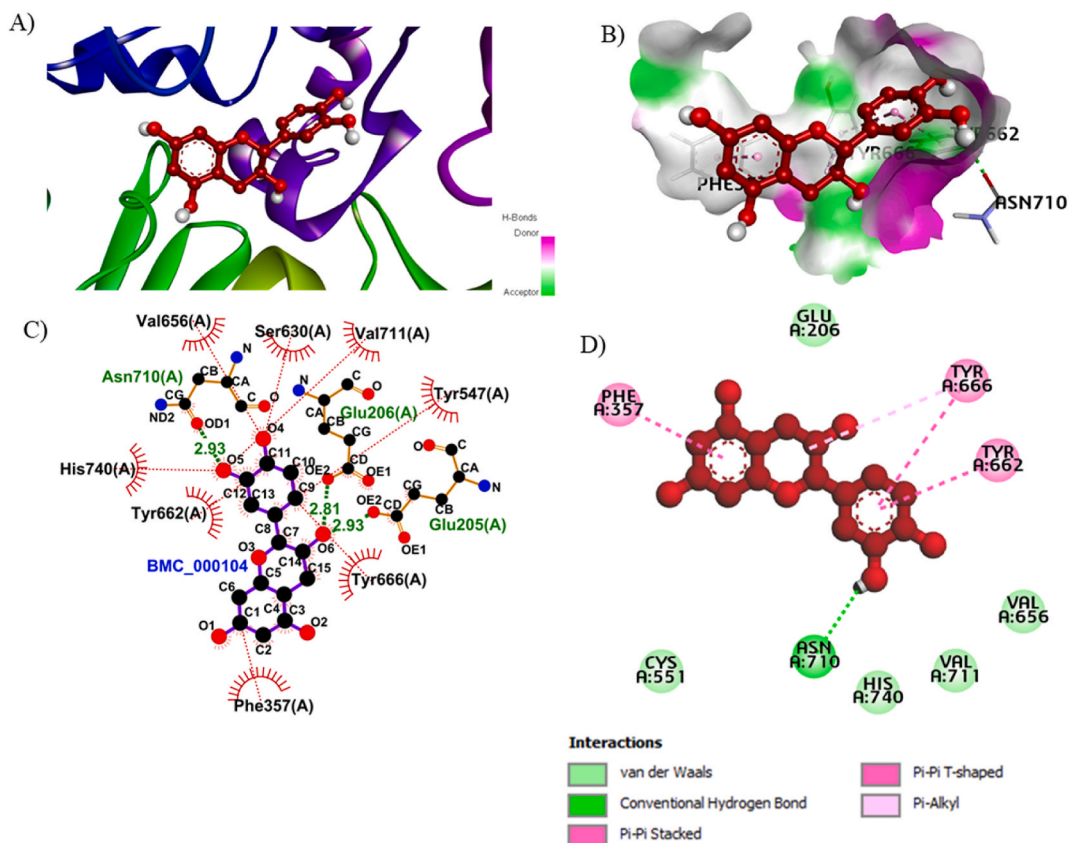
Fig. 3. Non-bonded interaction diagram of docked BMC_000104. (A) Pose view representation, (B) Surface view representation, (C) Ligplot+ 2D interaction map, (D) Discovery studios 2D interaction map.

nearby hydrophobic pocket made up of ARG125, GLU205, GLU206, PHE357, ARG358, and ARG669. The S1′ subsite includes PHE357, TYR547, PRO550, SER630, TYR631, and TYR666, while the S2′ subsite comprises TYR547, TRP629, SER630, and HIS740. The S2 extensive subsite involves VAL207, SER209, PHE357, and ARG358. Our docking analysis demonstrated that the three identified compounds, OTH_UD_XX06_1, BMC_000104, and GB19, exhibit essential interactions with the S1, S2, S1′, S2′, and S2 extensive subsites crucial for DPP4 inhibition at the binding site 1X70. Discovery Studio's 2D interaction diagram in Fig. 2 showed that OTH_UD_XX06_1 can form three conventional hydrogen bonds with GLN553, HIS740, and GLU205, and it can also form π-π stacked and π-alkyl interactions with PHE357 and TYR547, respectively. van der Waals interactions were also observed with residues TYR666, ARG125, ASN710, SER630, VAL656, VAL711, TYR662, TYR585, CYS551, and TRP659. On the other hand, the Ligplot+ 2D diagram (Fig. 2) revealed that OTH_UD_XX06_1 forms five hydrogen bonds with residues GLU205, HIS740, GLN553, ASN710, and ARG125, as well as hydrophobic interactions with PHE357, TYR631, TYR662, TYR666, TYR547, TYR585, VAL656, VAL711, SER630, and SER552. According to the Discovery Studios interaction map, BMC 000104 forms one conventional hydrogen bond with ASN710, one π-π stacking with PHE357, two π-π T-shaped interactions with TYR662 and TYR666, and one π-alkyl interaction with TYR666. It as also observed van der Waals interactions were formed with ARG125, GLU206, VAL711, CYS551, and SER209. Ligplot, on the other hand, revealed three conventional hydrogen bonds with GLU205, GLU206, and ASN710, as well as hydrophobic interactions with VAL656, SER630, VAL711, TYR547, HIS740, TYR662, PHE357, and TYR666 (Fig. 3). GB19 formed two conventional hydrogen bonds with GLN553 and GLU205, one π-sigma and two π-alkyl interactions with PHE357, and one π-alkyl interaction with TYR666, as shown on the discover studios 2D interaction map, and finally van der Waals interactions with residues ARG125, GLU206, ARG429, TYR585, TYR456, SER552, TYR666, CYS551, SER209 (Fig. 4). In contrast ligplot 2D interaction map (see Fig. 4) revealed three hydrogen bond interactions with GLU205, GLN555 and TYR585 as well as six hydrophobic interactions with ARG125, CYS551, GLU206, PHE357, TYR456, TYR547 and SER552. Supporting information of sitagliptin and saxagliptin interaction diagrams are provided in the supplementary materials figures S1 and S2.

3.5. Molecular dynamics analysis

Docking algorithms are effective at distinguishing between good and bad binders; however, they are not capable of estimating the dynamics and stability of binding. To address such challenges molecular dynamics simulation of the hit molecules complexed with
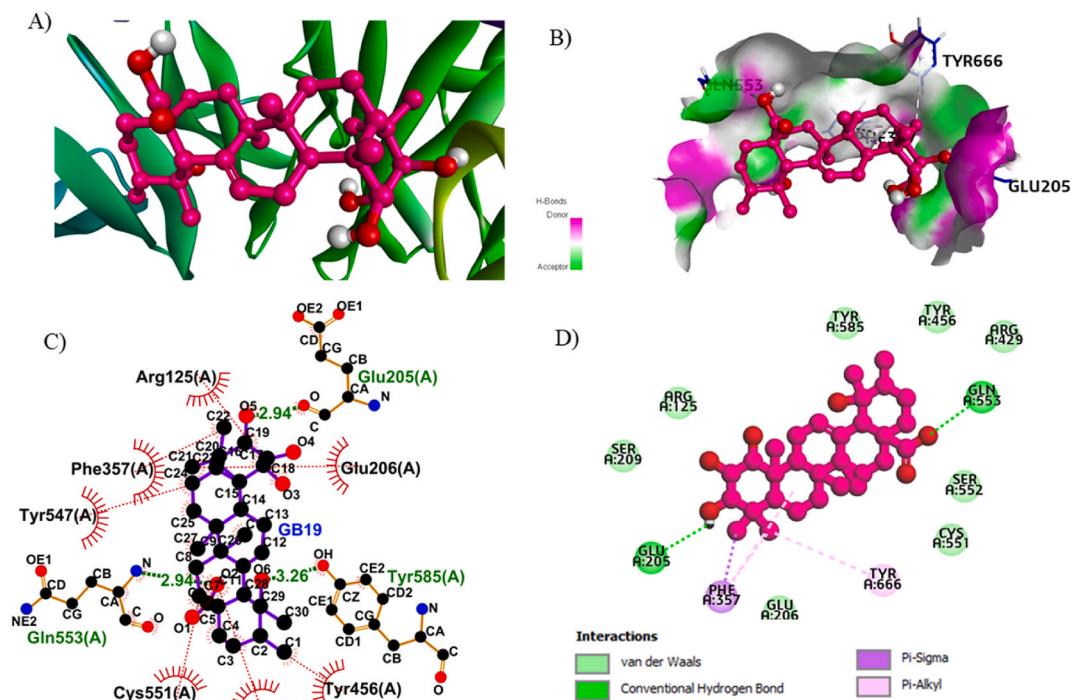
Fig. 4. Non-bonded interactions diagram of docked GB19. (A) Pose view representation, (B) Surface view representation, (C) Ligplot+ 2D interaction map, (D) Discovery studios 2D interaction map.
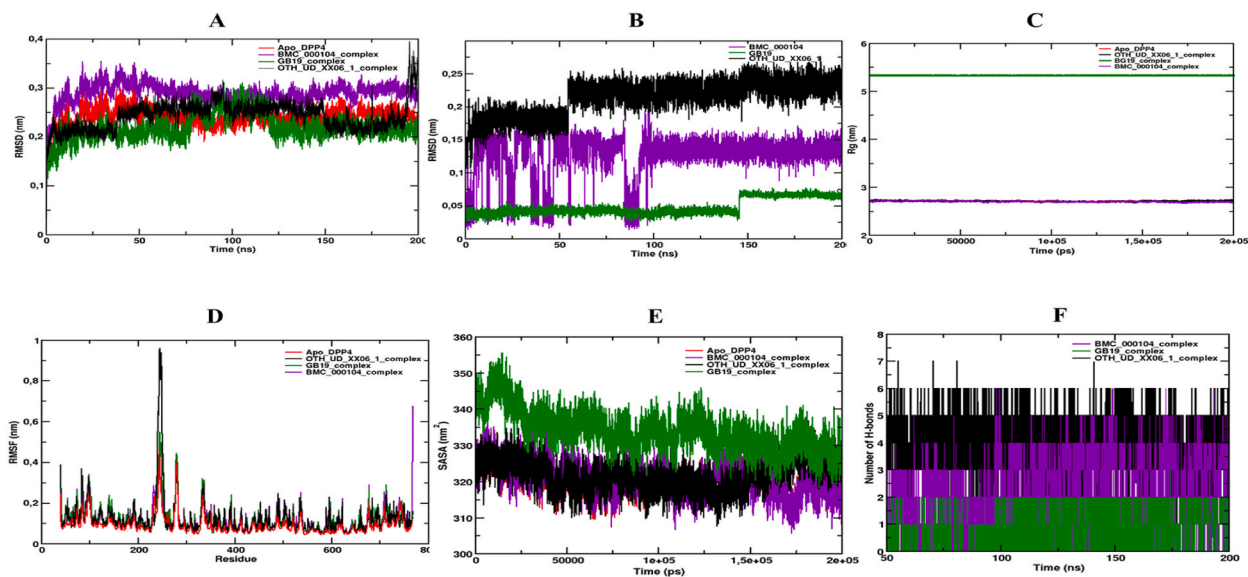

Fig. 5. Molecular dynamics simulation of apo-DPP4 (red) complex with OTH_UD_XX06_1 (black) and BMC 000104 (violet). (A) RMSD of OTH UD XX06_1, GB19 and BMC 000104; (B) Root mean square deviation (RMSD); (C) Radius of gyration (R_g); (D) Root mean square fluctuation (RMSF); (E) Solvent accessible surface area (SASA); (F) Number of hydrogen bond formed within simulation time. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

DPP4 was employed to validate the docking study, and the dynamic motion of the docked complexes was analyzed to understand their binding stability (refer to Fig. 5). The results depict the changes in RMSD values of the DPP4 backbone during the 200 ns MD simulation in the absence (red) and presence of OTH_UD_XX06_1 (black), BMC_000104 (violet) and finally GB19 (green) as shown in Fig. 5A. For the OTH_UD XX06_1 complex, the RMSD value was stable (0.21 nm) within the first 36 ns, and then exhibited

Table 5

Post molecular dynamics interaction analysis.

Ligand	Interacting residues	Polar contact	Distance (Å)
OTH_UD_XX06_1	GLU206	Conventional hydrogen bond	1.86, 1.87
	TYR547	Conventional hydrogen bond	1.99, 2.37
	TYR662	Conventional hydrogen bond	1.81
	TRP629	π-π stacked	4.72
	TYR547	π-alkyl	5.21
	GLY741	van der waals	
	HIS740	van der waals	
	SER630	van der waals	
	TRP659	van der waals	
	TYR666	van der waals	
	GLU206	van der Waals	
BMC_000104	GLU206	Conventional hydrogen bond	1.57, 1.69
	LYS554	Conventional hydrogen bond	3.08
	SER552	Conventional hydrogen bond	1.81
	TYR547	π-Alkyl	4.41
	TYR547	π-π stacked	4.33
	PHE357	π-π T-shaped	5.75
	TYP629	van der Waals	
	SER630	van der Waals	
	TYR666	van der Waals	
GB19	GLU205	Conventional hydrogen bond	1.94
	PHE357	van der Waals	
	ARG429	van der Waals	
	ASP556	van der Waals	
	VAL588	van der Waals	
	TYR456	van der Waals	
	LYS554	van der Waals	
	TYR585	π-Alkyl	5.14
	GLN553	van der Waals	
	SER209	van der Waals	
	TYR547	π-π stacked	3.96

conformational transition during 38–145 ns where the RMSD increased to a plateau value of 0.25 nm and then decreased back to 0.21 nm during 151–191 ns, and the final exhibited another transition state leading to a slight rise in RMSD value to 0.31 nm to the end of the simulation period. The BMC_000104 complex shows similar RMSD trend with the unbound DPP4, but the BMC_000104 complex fluctuated at slightly higher RMSD value of 0.27 nm compared to DPP4 RMSD value which oscillates around 0.25 nm. Furthermore, GB19_complex showed a stable RMSD value fluctuating around 0.20 nm during the first 78 ns, and exhibited a transition state within the time interval 78–120 ns where the RMSD value rose to 0.26 nm, and then dropped back to a plateau value of 0.22 ns till the end of the simulation time. This indicates that the backbone of the protein remained stable during the simulation process.

The RMSD values after the least squares fit were evaluated to better understand the binding stability of OTH_UD_XX06_1 (black) and BMC_000104 (violet) in the active site during the 200 ns simulation (Fig. 5B). The measured RMSD after the least squares fit revealed two conformation states for OTH_UD_XX06_1. First, during the initial 54 ns, the ligand remained stable and suddenly exhibited transition conformation states during 54 ns where the RMSD increased from 0.17 to 0.22 nm and remained stable throughout the simulation time. On the other hand, BMC_000104 fluctuated drastically during the first 96 ns suggesting a non-stable conformation during this period, while at 100–200 ns a stable conformation was reached with an RMSD value oscillating at 0.13 nm. Additionally, GB19 fluctuated slightly around the RMSD value 0.0418 nm during the first 145.5 ns and suddenly exhibited a transition conformation state at 145.5 ns period which led to an increased RMSD value from 0.0418 to 0.0673 ns and attained stability till the end of the simulation time. The RMSD value ≤ 0.3 nm of the conformations attained by OTH_UD_XX06_1, BMC 000104, and GB19 obtained after the least squares fit suggests stability of binding. From Fig. 5C, we can observe that the unbound DPP4 and OTH_UD_XX06_1 complex possessed nearly similar radius of gyration (R_g) profiles during the first 145 ns of simulation, and then the R_g value increased from 148 to 200 ns indicating a slight loss of system compactness. Moreover, BMC_000104 complex showed a high R_g profile during the first 61 ns and a decrease of R_g from 61 to 158 ns, a further decrease of R_g was observed from 158 to 200 ns indicating the tight packaging of the complex. On the other hand, the GB19 complex had the highest R_g value oscillating around 5.333 nm through the simulation period which is an indication of a great loss in the system compactness. The RMSF was used to analyze the flexibility of amino acid residues throughout the simulation time. Fig. 5D shows that the majority of the residues in the two complexes had low flexibility because they did not have high RMSF values except for regions with residues 39–40, 83–85, 95–100, 241-247, 279, and 767.

Furthermore, the solvent-accessible surface area (SASA) of the complexes was computed from the simulation trajectories (see Fig. 5E) to assess the complex volume change within the simulation time. Fig. 3E demonstrates that apo-DPP4 and BMC_000104 complex had the same SASA profile within the simulation time indicating no change in volume. The OTH_UD_XX06_1 complex had a similar SASA trend with the BMC_000104 complex during the first 129 ns, and thereafter the SASA increased during 160-200 ns indicating an expansion of the protein surface area. GB19 complex SASA profile differs significantly compared to that of the Apo DPP4. An increase in the SASA was observed during the first 26 ns, which then drops during the time interval 30–130 ns, and finally a slight
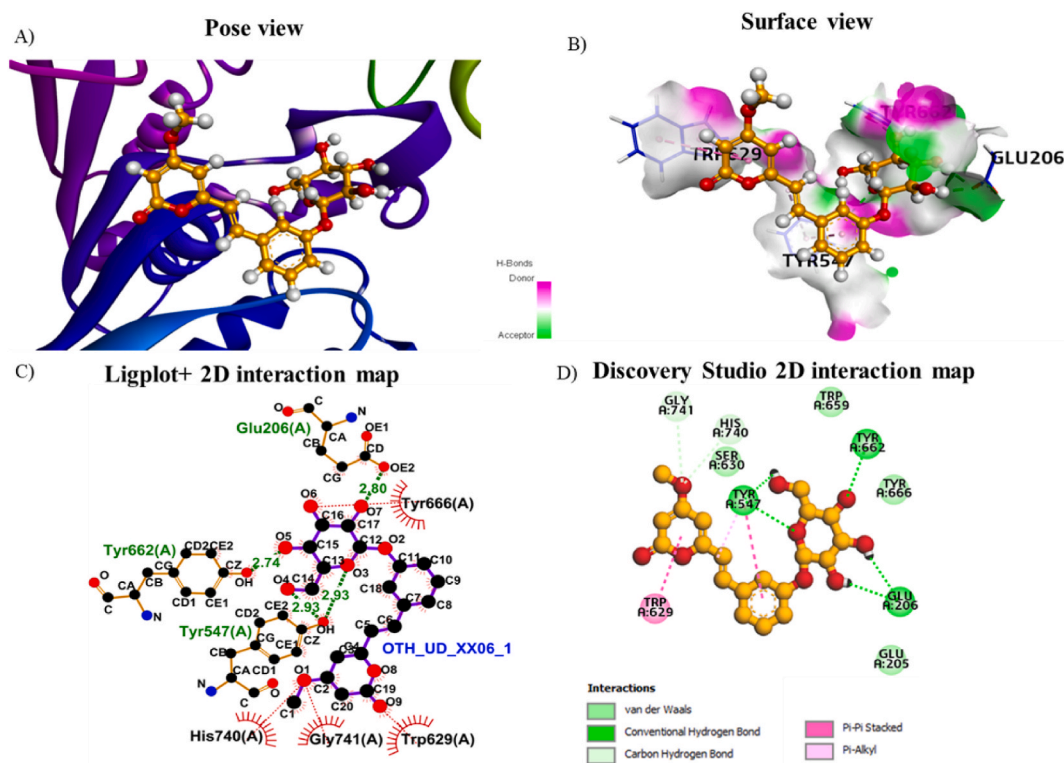
Fig. 6. Post-MD binding interactions of the DPP4 with OTH_UD_XX06_1 based on the last snapshot from the 200 ns MD simulation. (A) Pose view, (B) Surface view, (C) Ligplot+ 2D interaction diagrame and (D) Discovery studio 2D interaction diagram.

fall of the SASA value occurred during the period 132–200 ns. However, the overall SASA profile of the GB19 complex shows a considerable increase in the surface area of the system compared to the DPP4 apo protein.

To validate the stability of the docked complexes, hydrogen bonds paired within 0.35 nm between DPP4 and compounds OTH_UD_XX06_1, BMC_000104 and GB19 were assessed within the 200 ns MD simulation, as can be seen in Fig. 5E. OTH_UD_XX06_1 was strongly bound to the active site of DPP4 forming 2 to 6 hydrogen bonds, while BMC_000104 could from 2 to 5 hydrogen bonds with the active site and finally the number of hydrogen bonds formed between GB19 and DPP4 was observed to be within the range 1–2.

We examined the 200 ns frame from the MD simulation trajectory to determine if there was a change in the interaction dynamics of the docked complex. Ligplot+ and Discovery Studio tools were used to investigate the interaction changes in the active site of DPP4. The results are summarized in Table 5 showing that several interactions (Figs. 6–8) differed between the docked complexes. Discovery Studio 2D interaction diagram demonstrates that the stability of the OTH_UD_XX06_1-DPP4 complex is based on five conventional hydrogen bonds with GLU206 (two hydrogen bonds), TYR547 (two hydrogen bonds), and TYR662 (one hydrogen bond). Hydrophobic interactions were also observed with TRP629 (π-π stacking) and TYR547 (π-π stacked and π-alkyl interactions). Amino acids GLY741, HIS740, SER630, TRP659, TYR666, and GLU206 interacted with OTH_UD_XX06_1 through van der Waals contact. Also, the Ligplot+ 2D diagram showed that the OTH_UD_XX06_1-DPP4 complex was stabilized by four hydrogen bonds with the same residues that were outlined by Discovery Studio, but accounted for more hydrophobic interactions with residues HIS740, GLY741, TRP629, and TYR666 (as can be seen in Fig. 6).

Discovery Studio 2D analysis revealed that the BMC_000104-DPP4 complex was stabilized also by 4 conventional hydrogen bonds with residues SER552 (one hydrogen bond), LYS554 (one hydrogen bond), GLU206 (two hydrogen bonds), and TYR547 (see Fig. 7). The BMC_000104 complex also had hydrophobic interactions with residues TYR547 (π-π stacked and π-alkyl interactions) and PHE357 (π-π T-shaped interactions). It was observed that van der Waals interactions were also formed with the residues TYP629, SER630, and TYR666. Ligplot+ 2D analysis of this complex also revealed 4 hydrogen bonds with residues SER630 (one hydrogen bond), SER552 (one hydrogen bond), and GLU206 (two hydrogen bonds), while hydrophobic contacts were observed within residues LYS554, GLU205, TRP629, TYR547, and TYR666. On the other hand, GB19 was stabilized by forming one conventional hydrogen bond with GLU205, π-Alkyl and π-π stacked interactions were observed with TYR585 and π-π stacked respectively. Gb19 also formed van der Waals contacts with amino acids PHE357, ARG429, ASP556, VAL588, TYR456, and LYS554. The Ligplot+ 2D interaction map showed a conventional hydrogen bond with GLU205, as well as hydrophobic interactions with GLU206, PHE357, TYR585, GLN553, and ARG560 (see Fig. 8).

Furthermore, analyses of snapshots for each of the complexes (Figs. 9–11) were carried out at different time intervals (0 ns, 50 ns, 100 ns, 150 ns, and 200 ns) from the MD trajectories. The binding pose of OTH_UD_XX06_1 (see Fig. 9) showed some conformational
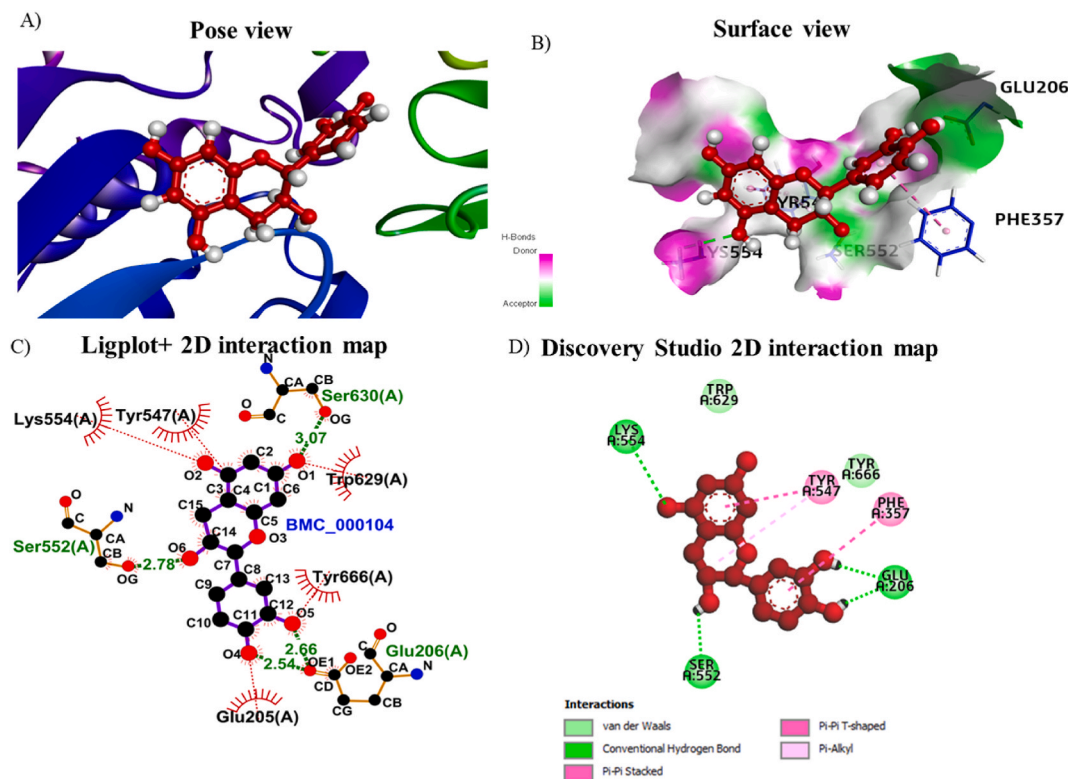
Fig. 7. Post-MD binding interactions of the DPP-4 with BMC_000104 based on the last snapshot from the 200 ns MD simulation. (A) Pose view, (B) Surface view, (C) Ligplot+ 2D interaction diagrame and (D) Discovery studio 2D interaction diagram.

changes of the methoxy group of the pyrone moiety during the time interval from 0 to 50 ns, and another conformational change of the methoxy pyrone group was observed at 100 ns time, and finally, the methoxy pyrone group maintained this conformation within the time interval from 100 ns to 200 ns. On the other hand, BMC_000104 exhibited different conformations in the active site of DPP4 during the first 50 ns. The 100 ns and 150 ns frames showed that BMC_000104 maintained a stable conformation in the active site, and lastly, no significant conformational changes were observed in the last frame (see Fig. 10). Lastly, snapshot of GB19 indicates slight conformational changes implicating OH and methyl groups (see Fig. 11).

Moreover, superimposition RMSD values between the pre- and post-MD structures (Fig. 12) calculated with Pymol after the 200 ns simulation were 1.562 Å (A), 1.855 Å (B), and 1.719 Å (C) respectively, indicating few deviations in the docked structures.

3.6. Binding energy computation

The MM-PBSA approach was used to calculate the binding free energy of OTH_UD_XX06_1, BMC_000104, and GB19 with the DPP4 enzyme using the gmx_MMPBSA tool. Snapshots for the MM-PBSA assessments for each complex were taken at 10 ns intervals during the 200 ns long MD simulations. The values of ΔG_{bind} for each complex are summarized in Tables 6–8 respectively. OTH_UD_XX06_1 had the lowest calculated free binding energy (−26.01 kcal/mol) compared to BMC_000104 (−16.92 kcal/mol) and GB19 (−14.81 kcal/mol). These results shows that OTH_UD_XX06_1 forms a more stable complex when bound to DPP4 compared to BMC_000104 and GB19.

3.7. Estimation of scaffold novelty of the selected hits

Based on our *in silico* analyses, the compounds named OTH_UD_XX06_1 ((3*R*, 4*S*, 5*S*, 6*R*)-2-{3-[(1*E*)-2-(4-methoxy-2-oxido-2*H*-pyran-2-ylium-6-yl)ethenyl]phenoxy}-6-(oxidomethyl)-1-lambda-3-oxan-2-ylium-1-ide-3,4,5-tris(olate)), GB19 ((1*R*, 2*S*, 3*R*, 4a*R*, 6a*R*, 6b*S*, 8a*S*, 11*R*, 12*R*, 12a*S*, 14a*S*, 14b*R*)-8a-[(hydrogeniohydroxy)(hydroxy)methyl]-4, 4, 6a, 6b, 11, 12, 14b-heptamethyl-2, 3, 4, 4a, 5, 6, 6a, 6b, 7, 8, 8a, 9, 10, 11, 12, 12a, 14, 14a, 14b-icosahydropicene-1, 2, 3, 12-tetrakis(olate)), and BMC_000104 ((2*R*, 3*S*)-2-(3, 4-dioxidophenyl)-3, 4-dihydro-2*H*-1-benzopyran-3, 5, 7-tris(olate)) were predicted to be the most promising hits against human DPP4 enzyme, so their structure novelty was estimated based on the 2D similarity to the known compounds deposited in the ChEMBL database with demonstrated *in vitro* K_i values ≤ 10 μM. The highest observed T_c value between OTH_UD_XX06_1, GB19, and BMC_000104 and the list of known inhibitors of DPP4 was 0.27, 0.28, and 0.33, respectively, suggesting a lack of structural similarity. Thus, we claim that OTH_UD_XX06_1 (phenyl-ethenylpyrane family), GB19 (icosahydropicene family), and BMC_000104 (phenyl-
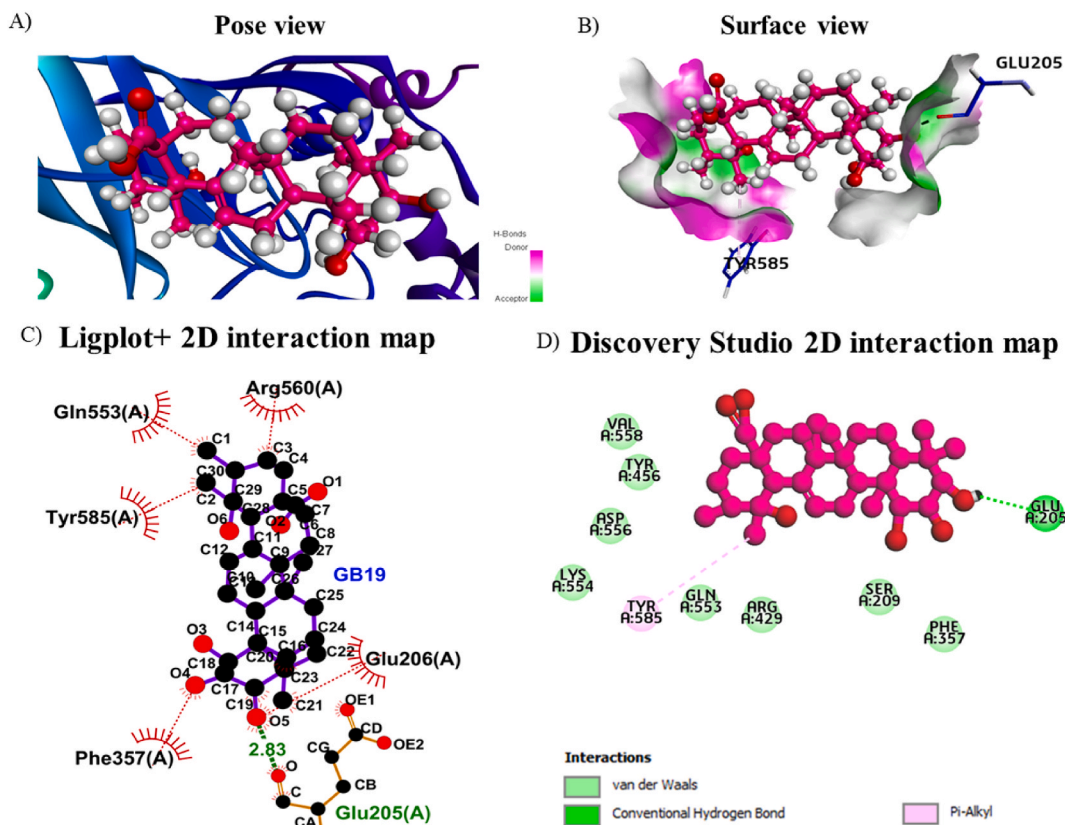
Fig. 8. Post-MD binding interactions of the DPP4 with GB19 based on the last snapshot from the 200 ns MD simulation. (A) Pose view, (B) Surface view, (C) Ligplot+ 2D interaction diagram and (D) Discovery studio 2D interaction diagram.

benzopyran family) have new chemical scaffolds not seen earlier for other potential inhibitors of human DPP4 enzyme.

4. Discussion

DPP4 inhibition improves insulin secretion by suppressing glucagon release, resulting in a decrease in blood glucose levels [37]. DPP4 inhibition is a novel strategy for T2DM treatment, and a multitude of phytochemicals have been identified as potential DPP4 inhibitors. In this work we have employed molecular docking, MD simulation, pharmacokinetic prediction, activity prediction and scaffold novelty to search for potential DPP4 inhibitors. The results obtained from the SBVS revealed that OTH_UD_XX06_1, BMC_000104, and GB19 as potent inhibitors of DPP4 from the ConMedNP library. To the best of our knowledge this database has not yet been the subject of a study concerning the search of DPP4 inhibitors. The selected hits were able to form hydrogen bond interactions, hydrophobic interactions, and van der Waals interactions with the amino acids that were reported to play an important role in the catalytic activity of the DPP4 enzyme [38], namely the amino acids that constitute the S1 pocket (TYR631, VAL656, TRP659, TYR662, and VAL711), the S2 pocket (ARG125, GLU205, GLU206, SER209, PHE357, and ARG358), the catalytic triad of DPP4 (Ser630, ASP708, ASN710, and HIS740), and finally residue TYR547 [70]. Amongst the hits, BMC_000104 was identified as catechin, catechine has been linked to antioxidant activity, antimicrobial activity, anti-aging, and anti-diabetic properties in numerous studies [39–41]. Molecular docking binding affinity results (−8.0 kcal/mol) of BMC_000104 were consistent with the study of Haron et al. Furthermore, similar interaction patterns of BMC_000104 with amino acids residues GLU205, SER209, PHE357, TYR547, SER630, TYR662, and TYR666 and BMC_000104 were also reported by Haron et al. in their study [42]. To the best of our knowledge, no report has been published on the anti-diabetic activity of OTH_UD_XX06_1 and GB19. Our study is the first to predict that these molecules could inhibit the activity of DPP4. OTH_UD_XX06_1 is a glycosylated desmethylyangonine from the pyrone family, and a study has shown that pyrones possess hypoglycemic activity [43]. On the other hand GB19 a pentacyclic triterpenoid (ursane-type) was identified to be 1α,2α,3β,19α-tetrahydroxyurs-12-en-28-oic acid, on like pyrones triterpenes have also been reported to be hypoglycemic agent [44]. Moreover, molecular dynamic analysis and MMPBSA analysis showed that among the hits, OTH_UD_XX06_1 forms the most stable complex with DPP4, followed by BMC_000104 and GB19. The anti-diabetic activity prediction also revealed OTH_UD_XX06_1 to be the most active molecule; it was predicted to be more active than reference DPP4 inhibitors Sitagliptin, Alogliptin, and Linagliptin. GB19 was predicted to have an activity greater than Linagliptin and Alogliptin, and finally, BMC_000104 was predicted to be the least active among the hits, but its activity was predicted to be greater than that of Alogliptin.
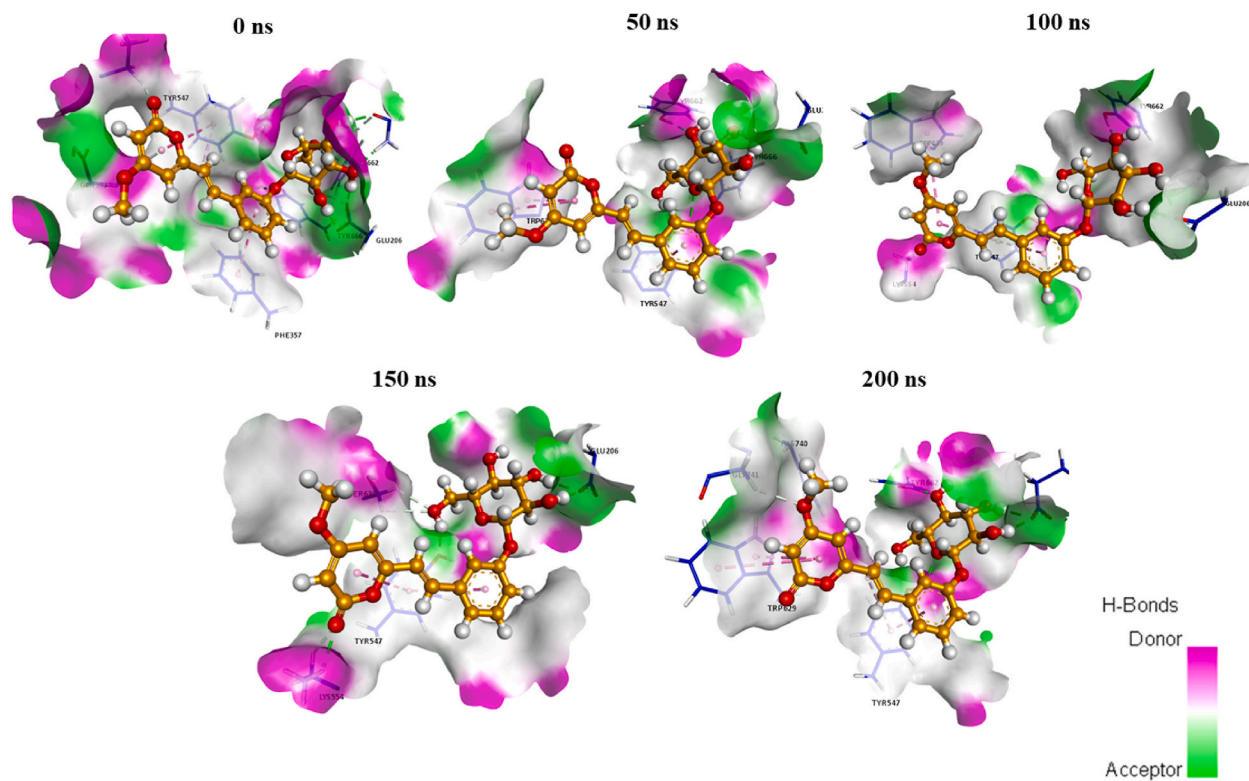
Fig. 9. Binding orientations of OTH_UD_XX06_1 in the active site of DPP4 at the different time frames: 0 ns, 50, 100, 150, and 200 ns.
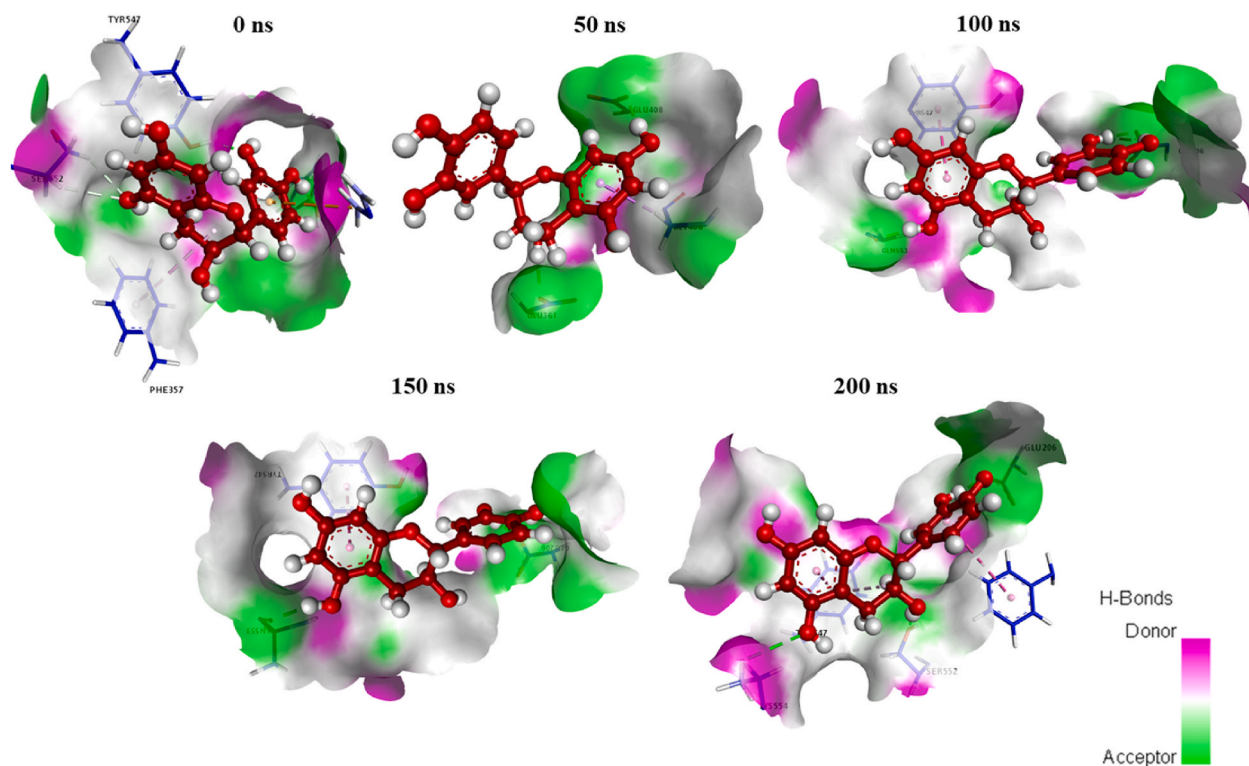

Fig. 10. Binding orientations of BMC_000104 in the active site of DPP4 at the different time frames: 0, 50, 100, 150, and 200 ns.
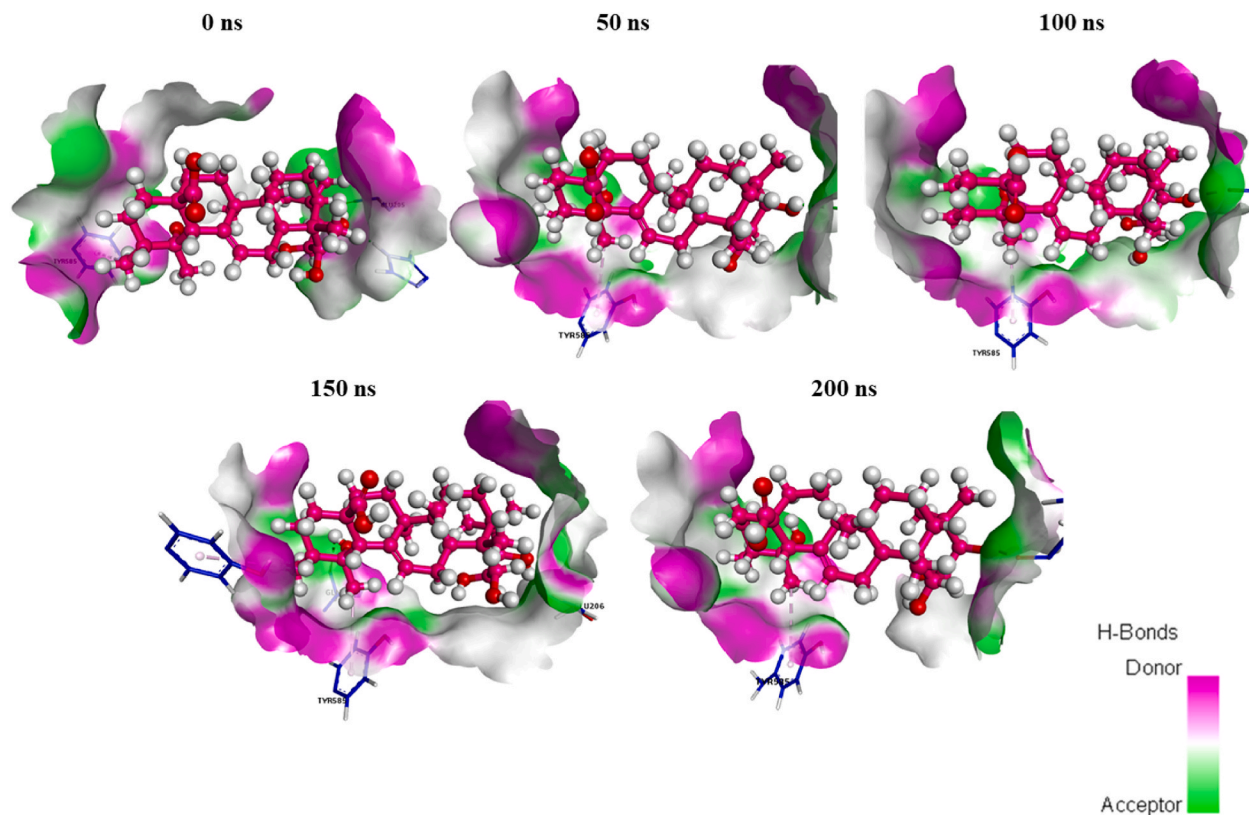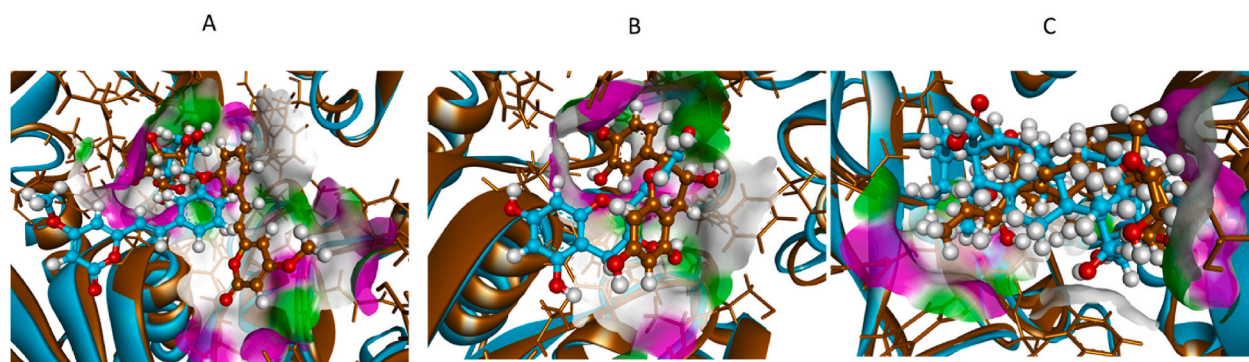
0 ns   **50 ns**   **100 ns**

150 ns   **200 ns**

Fig. 11. Binding orientations of GB19 in the active site of DPP4 at the different time frames: 0 ns, 50, 100, 150, and 200 ns.

Fig. 12. The superimposed pre-MD and post-MD structures of the OTH_UD_XX06_1-DPP4 (A), BMC_000104-DPP4 (B) and GB19-DPP4 (C) complexes. Brown color represents the pre-MD structure, while cyan color depicts the post-MD structure. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Furthermore, OTH_UD_XX06 1, BMC_000104, and GB19 have no structural similarity to previously identified DPP4 inhibitors, suggesting that these compounds could be investigated further as potential inhibitors against DPP4.

5. Conclusions

In this study, we have employed a computational drug design workflow to identify potential DPP4 inhibitors from the ConMedNP library consisting of 3507 molecules. The virtual screening survey helped in reducing the candidates from 3507 to 3 potential hits molecules namely OTH_UD_XX06_1, BMC_000104, and GB19. The ADMET and anti-diabetes activity predictions of the drug molecules confirm good anti-diabetes activity, a better absorption and metabolism profile with no toxicity risks. Furthermore, the docking study's binding pose and interactions were further evaluated using a molecular dynamics simulation study, where multiple descriptors,

Table 6
MM-PBSA binding free energy of the OTH_UD_XX06_1-DPP4 complex.

Energy Component (kcal/mol)	Average	SD (Prop.)	SD	SEM (Prop.)	SEM
ΔE_{bond}	−0.00	2.41	0	0.54	0
ΔE_{angle}	0.00	3.30	0	0.74	0
$\Delta E_{dihedral}$	−0.00	1.45	0	0.33	0
ΔE_{vdW}	−30.22	0.46	4.33	0.10	0.97
ΔE_{ele}	−60.04	0.07	17.35	0.01	3.88
$\Delta 1\text{-}4\ E_{vdW}$	0.00	0.70	0	0.16	0
$\Delta 1\text{-}4\ E_{ele}$	−0.00	1.49	0	0.33	0
ΔE_{PB}	68.12	4.03	14.88	0.90	3.33
$\Delta E_{non\text{-}polar}$	−3.88	0.02	0.14	0.00	0.03
ΔE_{disper}	0.00	0	0	0	0
ΔG_{gas}	−90.26	0.47	17.17	0.10	3.84
ΔG_{sol}	64.24	4.03	14.83	0.90	3.32
ΔE_{total}	−26.01	4.06	4.80	0.91	1.07

Table 7
MM-PBSA binding free energy of the BMC_000104-DPP4 complex.

Energy Component (kcal/mol)	Average	SD (Prop.)	SD	SEM (Prop.)	SEM
ΔE_{bond}	0.00	1.76	0.00	0.39	0.00
ΔE_{angle}	−0.00	1.66	0.00	0.37	0.00
$\Delta E_{dihedral}$	0.00	1.71	0.00	0.38	0.00
ΔE_{vdW}	−15.67	0.88	4.08	0.20	0.91
ΔE_{ele}	−46.49	4.00	8.90	0.89	1.99
$\Delta 1\text{-}4\ E_{vdW}$	−0.00	0.53	0.00	0.12	0.00
$\Delta 1\text{-}4\ E_{ele}$	0.00	2.56	0.00	0.57	0.00
ΔE_{PB}	47.66	0.04	10.71	0.01	2.40
$\Delta E_{non\text{-}polar}$	−2.42	0.13	0.37	0.03	0.08
ΔE_{disper}	0.00	4.09	0.00	0.00	0.00
ΔG_{gas}	−62.16	4.09	11.24	0.92	2.51
ΔG_{sol}	45.24	0.13	10.53	0.03	2.35
ΔE_{total}	−16.92	4.09	6.49	0.92	1.45

Table 8
MM-PBSA binding free energy of the GB19-DPP4 complex.

Energy Component (kcal/mol)	Average	SD (Prop.)	SD	SEM (Prop.)	SEM
ΔE_{bond}	0.00	2.03	0	0.45	0
ΔE_{angle}	0.00	5.20	0	1.16	0
$\Delta E_{dihedral}$	0.00	2.36	0	0.53	0
ΔE_{vdW}	−32.16	3.10	4.15	0.69	0.93
ΔE_{ele}	9.09	2.31	9.53	0.52	2.13
$\Delta 1\text{-}4\ E_{vdW}$	0.00	1.12	0	0.25	0
$\Delta 1\text{-}4\ E_{ele}$	−0.00	4.06	0	0.91	0
ΔE_{PB}	30.07	0.21	7.68	0.05	1.72
$\Delta E_{non\text{-}polar}$	−3.63	0.07	0.27	0.01	0.06
ΔE_{disper}	0.00	0	0	0	0
ΔG_{gas}	−41.25	3.86	8.32	0.86	1.86
ΔG_{sol}	26.44	0.22	7.71	0.05	1.72
ΔE_{total}	−14.81	3.87	3.81	0.87	0.85

including the RMSD, the R_g, the RMSF, the SASA, and a number of hydrogen bonds from simulation trajectories confirmed their stability. We also observed that OTH_UD_XX06_1, GB19, and BMC_000104 interact with key amino acids in the S1, S2, S1′, S2′, and the extensive S2 pockets through conventional hydrogen bonds and hydrophobic contacts. They were also predicted to present novel chemical scaffolds among the existing DPP4, which could make them promising inhibitors of DPP4.

Data availability statement

The authors confirm that the data supporting the findings of this study are available within the article.

Ethics statement

All authors have been personally and actively involved in substantial work leading to the paper, and will take public responsibility

for its content.

CRediT authorship contribution statement

Hans Merlin Tsahnang Fofack: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Maraf Mbah Bake:** Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Simon Petry:** Writing – review & editing, Validation, Software, Resources, Methodology. **Baruch A. Ateba:** Writing – review & editing. **Pascal Amoa Onguéné:** Writing – review & editing. **Haydar Mohammad-Salim:** Resources, Validation, Writing – review & editing. **Fidele Ntie-Kang:** Writing – review & editing, Validation, Resources. **Luc M. Mbaze:** Writing – review & editing, Validation, Supervision, Project administration, Conceptualization. **Serhii Vakal:** Writing – review & editing, Visualization, Validation, Supervision, Software, Investigation, Formal analysis, Data curation. **Cyril A. Kenfack:** Writing – review & editing, Validation, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e35191.

References

[1] World health organization (WHO). https://www.who.int/health-topics/diabetes, 2021. Accessed on: September 24.
[2] P. Saeedi, I. Petersohn, P. Salpea, B. Malanda, S. Karuranga, N. Unwin, S. Colagiuri, L. Guariguata, A.A. Motala, K. Ogurtsova, J.E. Shaw, D. Bright, R. Williams, Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: results from the international diabetes federation diabetes atlas, 9th edition, Diabetes Res. Clin. Pract. 157 (2019) 107843, https://doi.org/10.1016/j.diabres.2019.107843.
[3] J. Reed, S.C. Bain, V. Kanamarlapudi, A review of current trends with type 2 diabetes epidemiology, aetiology, pathogenesis, treatments and future perspectives, diabetes, Metab. Syndr. Obes. Targets Ther. 14 (2021) 3567–3602. Available: https://api.semanticscholar.org/CorpusID:237213965.
[4] S. Chatterjee, K. Khunti, M.J. Davies, Type 2 diabetes, Lancet 389 (10085) (2017) 2239–2251, https://doi.org/10.1016/S0140-6736(17)30058-2.
[5] S. Dodds, The how-to for type 2: an overview of diagnosis and management of type 2 diabetes mellitus, Nurs. Clin. 52 (2017) 513–522, https://doi.org/10.1016/j.cnur.2017.07.002.
[6] S. Park, H.J. Kang, J.-H. Jeon, M.-J. Kim, I.-K. Lee, Recent advances in the pathogenesis of microvascular complications in diabetes, Arch Pharm. Res. (Seoul) 42 (2019) 252–262, https://doi.org/10.1007/s12272-019-01130-3.
[7] S.Q. Pantaleão, et al., Virtual screening and in vitro assays of novel hits as promising DPP-4 inhibitors, Biochim. an Int. J. Biochem. Mol. Biol. (2022), https://doi.org/10.1016/j.biochi.2021.12.007.
[8] J. Huang, et al., Identification of novel uracil derivatives incorporating benzoic acid moieties as highly potent Dipeptidyl Peptidase-IV inhibitors, Bioorg. Med. Chem. 27 (4) (2019) 644–654, https://doi.org/10.1016/j.bmc.2019.01.001.
[9] M.A. Nauck, J.J. Meier, Incretin hormones: their role in health and disease, Diabetes, obesity & metabolism 20 (Suppl 1) (2018) 5–21, https://doi.org/10.1111/dom.13129.
[10] A. Mascolo, C. Rafaniello, L. Sportiello, M. Sessa, D. Cimmaruta, F. Rossi, A. Capuano, Dipeptidyl peptidase (DPP)-4 inhibitor-induced arthritis/arthralgia: a review of clinical cases, Drug Saf. 39 (5) (2016) 401–407, https://doi.org/10.1007/s40264-016-0399-88.
[11] M. Packer, Worsening heart failure during the use of DPP-4 inhibitors, JACC Heart Fail 6 (2018) 445–451, https://doi.org/10.1016/j.jchf.2017.12.016.
[12] S. Padron, E. Rogers, M. Demory Beckler, M. Kesselman, Republished: DPP-4 inhibitor (sitagliptin)-induced seronegative rheumatoid arthritis, Dtb 58 (2020) 12–15, https://doi.org/10.1136/dtb.2019.228981rep.
[13] DIREÇAO NACIONAL DE SAÚDE, Serviço para prevençao e reduçao dos fatores de risco - programa de Prevençao de la Diabetes Mellitus e outros distúrbios metabolicos - manual de controlo e seguimento da Diabetes mellitus, Praia (August 2015) 1–120.

[14] MEDICINA MNEMOTECNIAS - Diabetes, Farmacos hipoglucemiantes, available in: http://medicinamnemotecnias.blogspot.com/2016/04/diabetesfarmacoshipoglucemiantes.html, 2019. Accessed on: January 27.

[15] T. Katsila, G.A. Spyroulias, G.P. Patrinos, M.T. Matsoukas, Computational approaches in target identification and drug discovery, Comput. Struct. Biotechnol. J. 14 (2016) 177–184, https://doi.org/10.1016/j.csbj.2016.04.004.

[16] S. Ali, K. Ahmad, S. Shaikh, J.H. Lim, H.J. Chun, S.S. Ahmad, et al., Identification and evaluation of traditional Chinese medicine natural compounds as potential myostatin inhibitors: an in silico approach, Molecules 27 (2022) 34303, https://doi.org/10.3390/molecules27134303.

[17] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The protein Data Bank, Nucleic Acids Res. 28 (1) (2000) 235–242, https://doi.org/10.1093/nar/28.1.235.

[18] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera-A visualization system for exploratory research and analysis, J. Comput. Chem. 25 (13) (2004) 1605–1612, https://doi.org/10.1002/jcc.20084.

[19] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading, J. Comput. Chem. 31 (2010) 455–461, https://doi.org/10.1002/jcc.21334.

[20] J. Wang, W. Wang, P.A. Kollman, D.A. Case, Automatic atom type and bond type perception in molecular mechanical calculations, Journal of Molecular Graphics and Modeling 25 (2) (2006) 247–260, https://doi.org/10.1016/j.jmgm.2005.12.005.

[21] H. You, Y. Zhang, T. Wu, J. Li, L. Wang, Z. Yu, J. Liu, X. Liu, L. Ding, Identification of dipeptidyl peptidase IV inhibitory peptides from rapeseed proteins, LWT 160 (2022) 113255, https://doi.org/10.3390/foods11101406.

[22] S. Forli, R. Huey, M.E. Pique, M. Sanner, D.S. Goodsell, A.J. Olson, Computational protein-ligand docking and virtual drug screening with the AutoDock suite, Nat. Protoc. 11 (5) (2016) 905–919, https://doi.org/10.1038/nprot.2016.051.

[23] F. Ntie-Kang, P.A. Onguéné, M. Scharfe, L.C.O. Owono, E. Megnassan, L.M. Mbaze, W. Sippl, S.M.N. Efange, ConMedNP: a natural product library from Central African medicinal plants for drug discovery, RSC Adv. 4 (1) (2014) 409–419.

[24] N.M. O'Boyle, M. Banck, C.A. James, C. Morley, T. Vandermeersch, G.R. Hutchison, Open Babel: an open chemical toolbox, J. Cheminf. 3 (1) (2011) 33, https://doi.org/10.1186/1758-2946-3-33.

[25] S. Cosconati, S. Forli, A.L. Perryman, R. Harris, D.S. Goodsell, A.J. Olson, Virtual screening with AutoDock: theory and practice, Expet Opin. Drug Discov. 5 (6) (2010) 597–607, https://doi.org/10.1517/17460441.2010.484460.

[26] S.M. Ivanov, A.A. Lagunin, A.V. Rudik, D.A. Filimonov, V.V. Poroikov, ADVERPred – web service for prediction of adverse effects of drugs, J. Chem. Inf. Model. 58 (1) (2018) 8–11, https://doi.org/10.1021/acs.jcim.7b00568.

[27] G. Xiong, Z. Wu, J. Yi, L. Fu, Z. Yang, C. Hsieh, M. Yin, X. Zeng, C. Wu, A. Lu, X. Chen, T. Hou, D. Cao, ADMETlab 2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties, Nucleic Acids Res. 49 (2021) W5–W14, https://doi.org/10.1093/nar/gkab255.

[28] D.A. Filimonov, A.A. Lagunin, T.A. Gloriozova, A.V. Rudik, D.S. Druzhilovskii, P.V. Pogodin, V.V. Poroikov, Prediction of the biological activity spectra of organic compounds using the PASS online web resource, Chem. Heterocycl. Compd. 50 (3) (2014) 444–457, https://doi.org/10.1007/s10593-014-1496-1.

[29] BIOVIA; Dassault Systèmes, Discovery Studio Visualizer, v20.1.0.19295, Dassault Systèmes, San Diego, CA, USA, 2021.

[30] R.A. Laskowski, M.B. Swindells, LigPlot+: multiple ligand-protein interaction diagrams for drug discovery, J. Chem. Inf. Model. 51 (2011) 2778–2786, https://doi.org/10.1021/ci200227u.

[31] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1–2 (2015) 19–25, https://doi.org/10.1016/j.softx.2015.06.001.

[32] A.W.S.D. Silva, W.F. Vranken, Acpype - AnteChamber PYthon parser interfacE, BMC Res. Notes 5 (2012) 367, https://doi.org/10.1186/1756-0500-5-367.

[33] M.S. Valdés-Tresanco, M.E. Valdés-Tresanco, P.A. Valiente, E. Moreno, gmx_MMPBSA: a new tool to perform end-state free energy calculations with GROMACS, J. Chem. Theor. Comput. 17 (2021) 6281–6291, https://doi.org/10.1021/acs.jctc.1c00645.

[34] Canvas, Schrödinger, LLC, New York, NY, 2019.

[35] L. Schrodinger, W. Delano. http://www.pymol.org/pymol, 2020.

[36] G. Xiong, Z. Wu, J. Yi, L. Fu, Z. Yang, C. Hsieh, M. Yin, X. Zeng, C. Wu, A. Lu, X. Chen, T. Hou, D. Cao, ADMETlab 2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties, Nucleic Acids Res. 49 (2021) W5–W14, https://doi.org/10.1093/nar/gkab255.

[37] P.G. Yap, C.Y. Gan, In vivo challenges of anti-diabetic peptide therapeutics: gastrointestinal stability, toxicity and allergenicity, Trends Food Sci, Technol. 105 (2020) 161–175, https://doi.org/10.3390/nu14204275.

[38] (a) G. Scapin, Structural Chemistry and molecular modeling in the design of DPP4 inhibitors, NATO Science for Peace and Security Series A: Chem. Biol. (2015) 53–67;
(b) G. Scapin, D. Patel, E. Arnold, Multifaceted Roles of Crystallography in Modern Drug Discovery, NATO Science for Peace and Security series, 2015, https://doi.org/10.1007/978-94-017-9719-1.

[39] J.D. Lambert, C.S. Yang, Cancer chemopreventive activity and bioavailability of tea and tea polyphenols, Mutat. Res. 523 (2003) 201–208, https://doi.org/10.1016/s0027-5107(02)00336-6.

[40] J. Liu, J.F. Lu, J. Kan, X.Y. Wen, C.H. Jin, Synthesis, characterization and in vitro anti- diabetic activity of catechin grafted inulin, Int. J. Biol. Macromol. 64 (2014) 76–83, https://doi.org/10.1016/j.ijbiomac.22013.028.

[41] P.K. Kempegowda, F. Zameer, S.K. Murari, Delineating antidiabetic proficiency of catechin from withania somnifera and its inhibitory action on dipeptidyl peptidase-4 (Dpp- 4), Biomed. Res. 29 (16) (2018) 3192–3200, https://doi.org/10.4066/biomedicalresearch.29-18-922.

[42] N. Haron, N. Nadia, P.N. Farahin, D. Susanti, N. Hasniza, K.B.A. Halim, Molecular docking of polyphenol compounds from anacardium occidentale with alpha-glucosidase and dipeptidyl-peptidase-4 enzymes, Malaysian Journal of Fundamental and Applied Sciences 17 (2) (2021) 202–216, https://doi.org/10.11113/mjfas.v17n2.2059.

[43] A. Andrade-Cetto, H. Wiedenfeld, Hypoglycemic effect of Acosmium panamense bark on streptozotocin diabetic rats, J. Ethnopharmacol. 90 (2–3) (2004) 217–220, https://doi.org/10.1016/j.jep.2003.09.049.

[44] W. Ma, L. Xiao, H. Liu, X. Hao, Hypoglycemic natural products with in vivo activities and their mechanisms: a review, Food Sci. Hum. Wellness 11 (2022) 1087–1100, https://doi.org/10.1016/j.fshw.2022.04.001.