Biointerface Research in Applied Chemistry
Open-Access Journal (ISSN: 2069-5837)

Article
Volume 13, Issue 5, 2023, 454
https://doi.org/10.33263/BRIAC135.454

Computer-aided Identification of Bioactive Compounds from *Brachystegia eurycoma* with Therapeutic Potential against Drug Targets of Type 2 Diabetes mellitus

Ayodeji Benjamin Akawa [1,2], Babatunji Emmanuel Oyinloye [2,3,4,*], Basiru Olaitan Ajiboye [4,5]

1 Medical Biochemistry Unit, College of Medicine and Health Sciences, Afe Babalola University, PMB 5454, Ado-Ekiti 360001, Nigeria; akawaab@abuad.edu.ng (A.B.A.);
2 Phytomedicine, Biochemical Toxicology and Biotechnology Research Laboratories, Department of Biochemistry, College of Sciences, Afe Babalola University, PMB 5454, Ado-Ekiti 360001, Nigeria; babatunjioe@abuad.edu.ng (B.E.O.);
3 Biotechnology and Structural Biology (BSB) Group, Department of Biochemistry and Microbiology, University of Zululand, KwaDlangezwa 3886, South Africa
4 Institute of Drug Research and Development, Elias Bogoro Center, Afe Babalola University, PMB 5454, Ado-Ekiti 360001, Nigeria
5 Phytomedicine and Molecular Toxicology Research Laboratory, Department of Biochemistry, Federal University Oye-Ekiti, Ekiti State, Nigeria; bash1428@yahoo.co.uk (B.O.A.);
* Correspondence: babatunjioe@abuad.edu.ng (B.E.O);

Scopus Author ID 54395911800
Received: 16.12.2021; Accepted: 19.10.2022; Published: 21.12.2022

Abstract: Diabetes mellitus (DM) is by far the most common metabolic disease impacting human health, and Type II diabetes (T2DM) accounts for almost all occurrences of diabetes. This work examined the anti-diabetic efficacy of *Brachystegia eurycoma* compounds against druggable proteins associated with T2DM and its complications. Fourteen proteins were identified in the literature as T2DM treatment targets and downloaded from the protein data bank. Preliminary screening of the compounds with the protein targets via molecular docking studies showed that the compounds, notably quercetin, kaempferol, and catechin, had high selectivity for GLUT1, aldose reductase, and GLP-1 receptor. Eleven compounds from the plants were chosen as hits based on their favorable binding energies with the proteins. Following molecular docking studies, binding free energy, DFT calculation, ADMET predictions, and QSAR were used to examine further the drug-likeness, efficacy, toxicity, stability, and inhibitory/agonizing prowess of these compounds. The findings in this study showed that these eleven bioactive compounds, which belong to the group of flavonoids and phenolic acids that formed stable complexes with the three proteins, had moderate/low toxicity, are bio-orally available and non-inhibitors of some/all of the CYP450 isozymes. Using trustworthy correlation coefficients (R2), the predicted QSAR models demonstrated the potency of the compounds to function as inhibitors (pIC50) of aldose reductase and GLUT1 and as agonists (pEC50) of GLP-1R. According to DFT calculation of frontier molecular orbitals (FMOs) and global descriptive parameters, it was shown that Quercitrin is the most chemically inert molecule, whereas chlorogenic acid is the most reactive compound. This experimental approach may be utilized to develop drugs that can modulate proteins associated with T2DM without causing off-target effects, as shown in this research.

Keywords: *Brachystegia eurycoma*; flavonoids and phenolic acids; molecular docking; GLUT1; GLP-1 receptor.

1. Introduction

Medical progress in identifying cures and therapies for multi-factorial chronic health disorders like diabetes mellitus (DM) has been made for many years. However, diabetes mellitus still affects many people and contributes significantly to the decline in life expectancy worldwide [1,2]. 8.5% of individuals aged 18 and up had diabetes in 2014. Diabetic complications were directly responsible for the death of 1.5 million people in 2019 [3,4]. There was a 5% rise in diabetes-related premature death between 2000 and 2016 [5]. From 2000 to 2010, diabetes-related mortality rates in high-income nations declined but rose from 2010 to 2016. Both eras saw a rise in the early death rate attributable to diabetes in low- and middle-income countries (LMIC) [6].

Insulin deficiency (Type 1 diabetes mellitus (T1DM)) or insulin resistance (Type 2 diabetes mellitus (T2DM)) are two common causes of diabetes [7]. Insulin is a hormone that controls blood sugar levels. Increased blood sugar levels are a typical complication of uncontrolled diabetes, and they can cause significant harm to the body's systems, including neurons and blood vessels, over time [8,9]. The health impact of diabetes is overwhelming. When diabetes is left untreated, it can cause damage to the heart, blood vessels, kidneys, eyes, and nerves over time. Heart attacks and strokes are two to three times more likely in adults with diabetes [9, 10]. Neuropathy (nerve damage) in the feet, when combined with decreased blood flow, raises the risk of foot ulcers, infection, and eventual limb amputation [9-11].

There is no cure for DM at the moment. Various treatments exist to help manage and control DM, such as decreasing blood sugar levels or delaying the development of problems [12]. The FDA has authorized several medications as first-line treatments for T2DM. To reduce blood glucose levels and modulate tissue insulin sensitivity, the most popular anti-diabetic medication is metformin, which may be taken orally [13]. As with other anti-diabetic medications, metformin, too, is commonly accompanied by mild to severe adverse effects [13,14]. Therefore, treatments combining a combination of various anti-diabetic medicines which act on protein targets in molecular pathways implicated in diabetes development have given more satisfying outcomes than drug monotherapy alone [15]. Problems with this technique include adverse effects, drug-drug interactions, and hepatotoxicity, among others [16,17]. Compounds combining the active ingredients can be safer and more effective than single-target medicines since they can selectively affect different diabetic target proteins and pathways. By prioritizing specific disease targets, the multi-target ligand drug design method maximizes the relative efficacy of a molecule toward each receptor [17,18].

Modern drug discovery research relies heavily on computational approaches, which have shown to be invaluable [19]. It is possible to anticipate ligand-receptor interactions at the atomic level with these techniques and save costly and time-consuming experiments by making use of these technologies [18-20]. For complicated illnesses, a multi-targeted drug design strategy can enable a paradigm change in the use of existing medicines. A combination of disease networks, chemical and physical properties of medicines, and biological receptors are used in these techniques. Various techniques have been created to assist in identifying new and numerous targets for known or new medicines with the availability of powerful computational resources [21].

Among the inhabitants of the southern region of Nigeria, *Brachystegia eurycoma* is a popular leguminous plant for its ethnomedicinal and nutritional benefits. Even though this legume has great promise in terms of food and medicinal research, it has been woefully

underused to date. *B. eurycoma* has nutritional and medicinal advantages. Bioactive compounds isolated from *B. eurycoma* have been shown to have anti-diabetes, analgesics and antimicrobial properties and liver enzyme modulating properties [21, 22]. In addition to their aforementioned therapeutics, they are also claimed to have anti-inflammatory, antioxidant, and anti-tumor properties [20, 21].

This study aims to identify lead phytochemicals from *B. eurycoma* that could interact with one or more druggable proteins examined in this study which could be developed as adjuvant therapy with minimal or no side effect. For instance, current GLP-1 receptor agonists have several potentially beneficial effects on type 2 diabetic patients; known negative side effects such as gastrointestinal disorders, pancreatic inflammation, and pancreatitis, as well as a small increase in heart rate, musculoskeletal disorders, development of acute kidney injury and injection site reactions, may prevent their use, especially at larger doses. Therefore, this study employs different computational techniques to find multi-target ligands from *B. eurycoma* compounds against selected targets for treating T2DM. As therapeutic targets, the present study selected the most researched T2DM-associated proteins [22, 23].

2. Materials and Methods

2.1. Protein identification and its preparation.

This article identified a total of fourteen proteins that are drug targets for diabetes mellitus. The following protein crystal structure with their identification code was retrieved from the protein databank repository (http://www.rcsb.org/): pancreatic alpha-amylase (PDB ID: 1B2Y), aldose reductase (PDB ID; 2R24), PPAR- gamma (PDB ID:5U5L), insulin receptor kinase (PDB ID: 1GAG), glucagon-like peptide-1 (GLP-1) receptor (PDB ID: 6X1A),alpha-glucosidase (PDB ID: 5ZCB), human glutamine-fructose-6-phosphate transaminase 1 (PDB ID: 2V4M), Maltase-Glucoamylase (PDB ID: 2QMJ), Glucose transporter 1 (PDB ID: 5EQG), lysosomal acid-alpha-glucosidase (PDB ID: 5NN5), Adiponectin (PDB ID: 4CFE), and protein tyrosine phosphatase 1B(PDB ID: 4Y14). However, the crystal structures of sodium-glucose cotransporter-2 (SGLT-2) were not available on the protein data bank because their three-dimensional structures have not been characterized. Hence, their 3D structures were modeled using SWISS-MODEL.

The downloaded and modeled protein crystal structures were all prepared by the Schrödinger protein preparation wizard. The reasons for preparing the proteins are to add missing hydrogen atoms, optimize hydrogen bonds, delete water molecules, create disulfide bridges if necessary, fill the missing side chain via Prime refinement, and minimize the structures using the OPLS3 force field [22-24].

The grid files for defining the protein binding pockets were generated by picking the co-crystal ligand within the protein binding pockets. However, for proteins without bound ligands, such as the modeled proteins (GPR120) and (SGLT-2), the active sites of such proteins were predicted by the Sitemap module of the Schrodinger suite.

2.2. Ligand identification and preparation

The phytochemicals of the whole of *Brachystegia eurycoma*, including its seed, stem, and bark, were identified from the literature [22-24]. The 2D structures of the compounds were downloaded from the PubChem database in SDF format. These compounds were sorted into a folder to create a library of compounds from *Brachystegia eurycoma*. This library of

compounds was exported to Maestro graphical interface for compound minimization. The compound minimization was carried out by ligprep module of Schrödinger. The following options were assigned before preparing the ligands on Ligprep: desalt, use Maestro as output format, generate tautomers, use Epik to generate possible states at pH of 7.0 ± 2.0, and utilize OPLS3e force field for minimization.

2.3. Molecular docking studies.

The 2D structures of *Brachystegia eurycoma* phytochemicals were subjected to a molecular docking study with fourteen diabetes-relevant targets. The 78 prepared compounds were initially docked with the fourteen proteins using extra precision (XP) glide docking as the scoring algorithm. This method is the most preferred choice in scoring the compounds according to their order of binding affinity [23,24].

2.4. Calculation of free energy of binding via Prime MM-GBSA.

The free energy of binding, also known as binding free energy, was determined by Prime MM-GBSA, which takes into account the ligand binding energies and ligand strain energies for a set of ligands and a single receptor.

2.5. Density functional theory (DFT) calculation.

Using the Schrödinger Materials science (version 3.9) that accommodates Jaguar fast engine, the reactivity of the compound, which includes frontier molecular orbitals (FMOs) and molecular electrostatic potential (MEP), was estimated using 6-31G** as a basic set and Becke's three-parameter exchange potential and Lee-Yang-Parr correlation (B3LYP) as the level of functional theory. The global descriptors, including ionization potential (IP), electron affinity (EA), chemical hardness (η), chemical softness (ζ), electronegativity (χ), chemical potential (μ), and electrophilicity (ω were calculated from FMOs. The following equations are used to define the global descriptors:

2.6. ADMET prediction.

The selected hit compounds from *Brachystegia eurycoma* were screened for toxicity, drug-likeness, and pharmacokinetics. The ADME parameters, including Lipinski's rule, Ghose rule, Veber rule, blood-brain barrier permeant, and Pgp substrate, were predicted by the SWISSADME server. Other parameters relating to toxicity, such as hepatotoxicity, carcinogenicity, eye irritation and corrosion, and acute oral toxicity, were predicted by ADMETsar.

2.7 Predictive QSAR modeling.

The top criterion for building a predictive quantitative structure-activity relationship (QSAR) model is retrieving experimental compounds with known biological activities against the desired drug target. Therefore, the chemical and bioactivity (IC50) profile of a given set of experimental data sets of GLUT1 and Aldose reductase from the ChEMBL database. Also, the agonists of the GLP-1 receptor alongside their EC50 were retrieved from the ChEMBL database by blasting the PDB fast a sequence of GLP-1R. The predictive models were

generated by AutoQSAR modeling, which calculated the topological and physicochemical descriptors in addition to the binary fingerprints.

3. Results and Discussion

Today's drug discovery research relies heavily on computational approaches. Using these techniques, you can make accurate predictions about ligand-receptor interaction events down to the atomic level without having to build up a complex experimental setup in advance. One of these computational approaches is molecular docking, which determines the binding affinity and suitable conformations of drug molecules in the bound form within a receptor binding site [25, 26].

3.1. Molecular docking of ligands against phytochemicals.

In this study, 78 previously characterized bioactive compounds identified in literature from *Brachystegia eurycoma* were docked against fourteen (14) selected therapeutically important diabetic targets. The results of the compound's docking scores against different proteins implicated in different diabetes molecular pathways are shown in Table S1. Data from this study showed that the bioactive compounds from *Brachystegia eurycoma* had favorable binding/interaction with some of the selected proteins relevant to diabetes. This study, however, scrutinized the available data from molecular docking results and identified the three proteins with the most favorable ligand-receptor interaction; they are glucose transporter 1 (GLUT1), aldose reductase (ALD), and glucagon-like peptide (GLP-1) receptor. Consequently, eleven (11) phytochemicals from *Brachystegia eurycoma* with considerable docking scores with three diabetes targets were further investigated for their binding free energy and electronic behaviors, among others. These compounds are Quercetin, Kaempferol, Catechin, Luteolin, Flavone, 4, 5-dihydroxy-7-methoxy, 3-ethyl-2-hydroxy, Ellagic acid, Rosmarinic acid, Resveratrol, Chlorogenic acid, and Quercitrin. Categorically, the compounds belong to the group of flavonoids and phenolic acids. Broadly, flavonoids and phenolic acids have rich health benefits, which include preventive activity against inflammation and allergies via antioxidants, antimicrobials, and diabetes and its complications [26, 27].

Table 1 shows the findings of the molecular docking simulations. Concerning the aldose reductase protein crystal structure, the compounds, Quercetin, Kaempferol, Catechin, and Luteolin had the lowest binding energy with docking scores of -12.156 kcal/mol, -11.250 kcal/mol, -11.161 kcal/mol and -11.133 kcal/mol respectively. Most phytochemicals recognized as hit compounds exhibited a strong affinity for GLP-1 receptor, with docking scores of -12.862 kcal/mol for Quercitrin, -11.431kcal/mol for 3-ethyl-2-hydroxy, and -11.139 kcal/mol for Ellagic acid. The docking scores of the GLUT1-docked compounds varied from -11.421 to -7.034 (kcal/mol), as shown in Table 1. Further observation denoted that, among the selected hits, Quercitrin had the most favorable energy of binding with GLP-1 receptor and GLUT1, while it also demonstrated strong/favorable binding with aldose reductase. Quercitrin has been shown to interact with many diabetes-relevant targets in different tissues and organs with the ability to potentiate glucose homeostasis throughout the body, emphasizing improving glucose consumption in peripheral tissues through reducing glucose absorption in the intestine, modulation of insulin secretion and sensitivity [26,27]. The tendency of quercetin to modulate several drug targets implicated in the pathogenesis of diabetes may be responsible for its high binding energy with the GLP-1 receptor and GLUT1. Other than quercetin, several

investigations [26-28] have demonstrated the anti-diabetic effects of the compounds with a high docking score (Table 1). By regulating several molecular pathways, these compounds have demonstrated their potential in treating diabetes.

3.2. Interactions of the ligands with critical residues.

The interacting profiles of the compounds with the proteins are shown in Figures 1-3. The results showed that the selected compounds formed different non-covalent interactions with the residues within the active sites of the proteins, and these interactions include hydrogen bond (H-bond) networking and Pi-pi stacking. Glucose transport is carried out by the GLUT1 protein, which is lodged in the cell's outer membrane, where it conveys sugar from the blood and/or other cells to be used as fuel [28]. The GLUT1 protein is responsible for glucose transport across the blood-brain barrier, which serves as the brain's primary energy source [28]. Suppression of GLUT1 has been suggested as therapeutic to treat complications of diabetes such as diabetic retinopathy and diabetes-induced microvasculopathy [29]. Interacting profiles of the hit compounds with GLUT1 highlight that most of the compounds formed pi-pi interaction with TRP111. Chlorogenic acid and Rosmarinic acid formed the most hydrogen bond network with GLUT1. The rich H-bond interactions exhibited by these compounds may be responsible for their favorable binding affinity with GLUT1. It is interesting to know that these intermolecular interactions between the ligands and the residues of GLUT1 have been previously reported [30].

GLP-1 is of special relevance due to its glucose-lowering properties [29, 30], including its ability to delay stomach emptying and inhibit glucagon production [29, 30]. People with type 2 diabetes have diminished incretin action. The most current understanding of this deficiency indicates that it is related to a worsening of the GLP-1 action, with decreased insulin secretion, increased insulin resistance, and hyperglycemia, perhaps resulting in a reduction in GLP-1 receptor expression and subsequent GLP-1 resistance [31]. GLP-1 receptor agonists activate GLP-1 receptors, resulting in comparable increases in insulin production in response to oral and intravenous glucose. By establishing Pi-pi stacking with the compounds' phenyl rings, the residue PHE381 established significant intermolecular interactions. Additionally, THR207, SER31, LYS197, GLN234, and GLN221 were responsible for the majority of hydrogen interactions with the compounds. The compounds' high docking scores may be attributed to their abundant hydrogen bond donor and acceptor as a mechanism of interaction with GLP-1R. Similar amino acid interactions are seen in putative GLP-R receptor agonists derived from *Phyllanthus emblica* phytocompounds that stimulate insulin secretion [31]. Aldose reductase (AR) activity is believed to significantly promote diabetes complications such as nephropathy, neuropathy, and possibly stroke [30, 31]. The polyol pathway is one of the best-researched metabolic pathways linked with hyperglycemia. This is a two-step process in which AR plays a significant role as the rate-limiting enzyme by using NADPH as a cofactor to convert glucose to the alcohol sorbitol [31, 32]. Numerous structurally varied compounds, either naturally occurring or chemically manufactured, be effective against aldose reductase in wet experiments [33]. The interaction profiles of the phytocompounds with aldose reductase are shown in Figure 1 and Table 2. Previous studies have shown that aromatic molecules are the ideal substrate for this enzyme due to the hydrophobic nature of the active site. In accordance with this, aside from the significant lipophilic and van der Waals contributions, pi-pi interactions also add to the binding energy of these compounds. These previous findings correspond to the data obtained from this study – the compounds majorly exhibited non-covalent interaction with amino acid residues of aldose reductase using its aromatic rings. As listed in Table 1, the compounds formed mainly pi-pi interactions with TRP219, TRP111, or TRP48. In interacting profiles of the compound (Quercetin) with the highest docking score, two hydrogen bonds, five π–π interactions, and numerous hydrophobic interactions were observed.

Table 1. Docking score and interacting profiles of hit compounds from *Brachystegia eurycoma* against selected drug targets of diabetes

	Docking score			Interacting residues			Binding free energy		
Entry name	**Aldose Reductase**	**GLP-1 Receptor**	**GLUT1**	**Aldose Reductase**	**GLP-1 Receptor**	**GLUT1**	**Aldose Reductase**	**GLP-1 Receptor**	**GLUT1**
Quercetin	-12.156	-11.335	-8.107	TRP111, TRP79, TR48, TRP20, TRP48, [2 Hbond]	TRP33, PHE381, THR20, GLN221 [3 HBOND]	TRP412, PHE26, GLU380, ASN288 [3 HBOND]	-53.28	-51.77	-37.10
Kaempferol	-11.250	-10.315	-8.099	TRP111, TRP20	LYS197, PHE381, TRP33, GLN221, THR207 [2 HBOND]	TRP412, ASN411, GLN283, ASN288 [3 HBOND]	-53.53	-41.82	-33.61
Catechin	-11.161	-11.075	-9.297	LEU300, TRP111 [1 HBOND]	TRP33, THR207 [2 HBOND]	ASN411, GLU380 [3 HBOND]	-48.35	-55.84	-35.89
Luteolin	-11.133	-10.648	-7.728	TRP111, TRP79	PHE381, TRP33, GLN221, THR207 [3 HBOND]	ASN411, TRP388, HIS160, GLN161 [3 HBOND]	-52.41	-52.08	-32.83
Flavone,4,5-dihydroxy-7-methoxy	-10.734	-8.326	-8.530	TRP111, VAL47 [1 HBOND]	PHE381	GLN283, ASN288 [3 HBOND]	-53.49	-47.60	-40.47
3-ethyl-2-hydroxy	-10.305	-11.431	-7.034	TRP111, TYR48, TRP20 [2 HBOND]	MET233, PHE230, TRP33, PHE381, TRP203, ARG380, GLN221 [3 HBOND]	SER80, ASN411, GLN283, ASN288 [4 HBOND]	-40.03	-72.66	-37.82
Ellagic acid	-10.023	-11.139	-9.087	HIP110, TRP20, CYS298, TYR48	TRP33, TRP203, THR207	GLN283, ASN288, GLU380 [3 HBOND]	-44.59	-70.99	-45.83
Rosmarinic acid	-9.857	-10.133	-8.515	THR113, TRP111, TRP111, TYR48, HIP110, GLN183, TYR209 [5 HBOND]	PHE381, TRP33, PHE230, ARG299, LYS197, TRP297 [3 HBOND]	GLU380, ASN288, GLU283 [5 HBOND]	-62.32	-63.48	-31.37
Resveratrol	-9.441	-9.411	-7.148	TRP111, LEU300 [1 HBOND]	PHE381, SER31, TRP33, THR207 [3 HBOND]	TRP412, THR30 [1 HBOND]	-53.22	-63.48	
Chlorogenic acid	-8.346	-10.166	-7.311	ALA299, THR113, TRP111, HIP110, TYR48, TRP111 [5 HBOND]	SER31, THR207, GLN221 [3 HBOND]	SER80 [1 HBOND]	-56.23	-45.80	-33.31
Quercitrin	-9.289	-12.862	-11.421	LYS221, TRP219, ALA299, LEU301 [3 HBOND]	PHE381, LYS197, TRP33, SER31 [2 HBOND]	ASN288, ASN411, GLU380 [4 HBOND]	-33.33	-64.84	-59.30

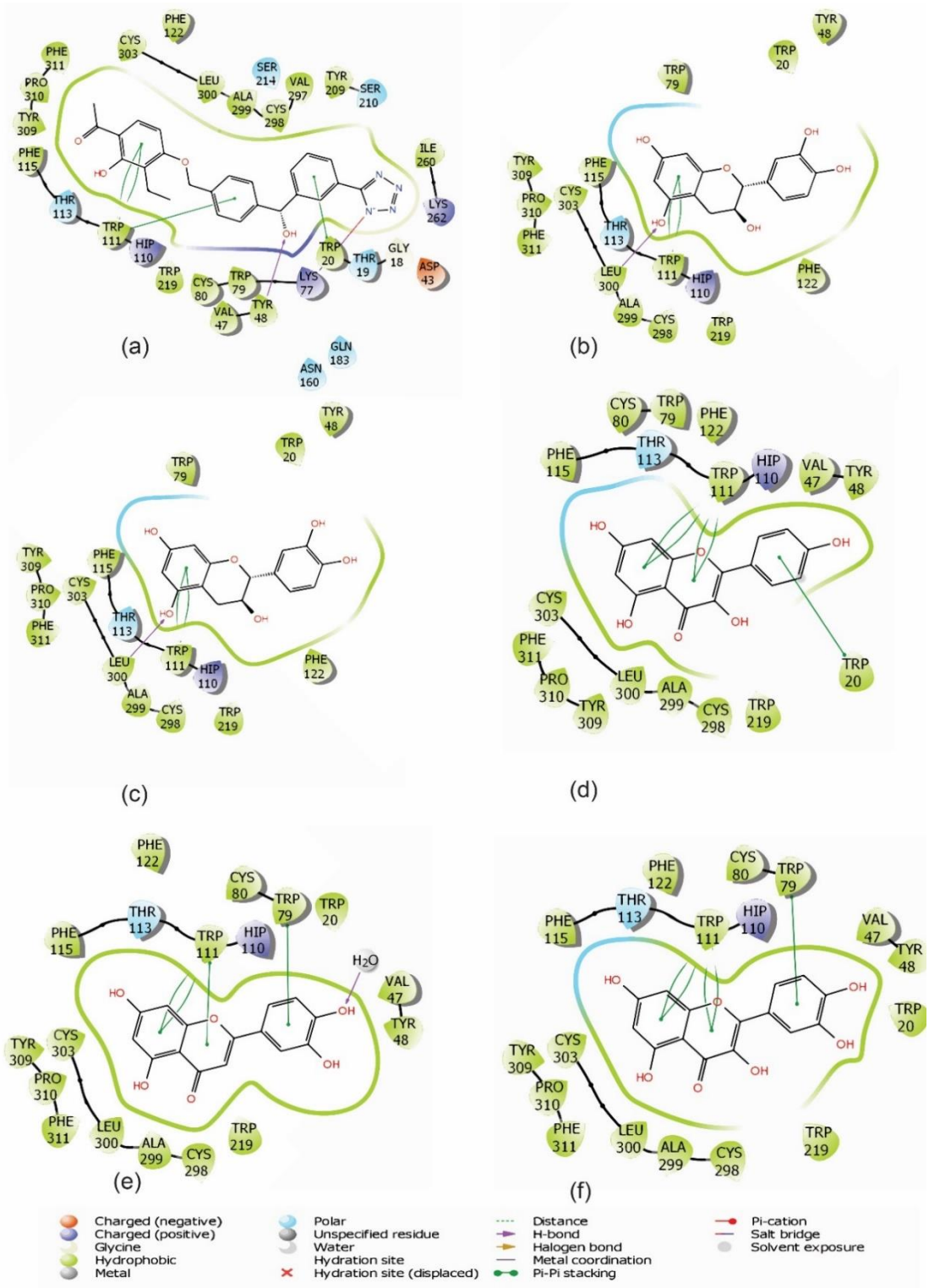

Figure 1. Two-dimensional interactions of (**a)** Quercetin; (**b)** Quercitrin; (**c)** Kampferol; (**d)** Catechin; (**e**) Luteolin; (**f)** Flavone,4,5-dihydroxy-7-methoxy with residues at the binding pocket of Aldose reductase.

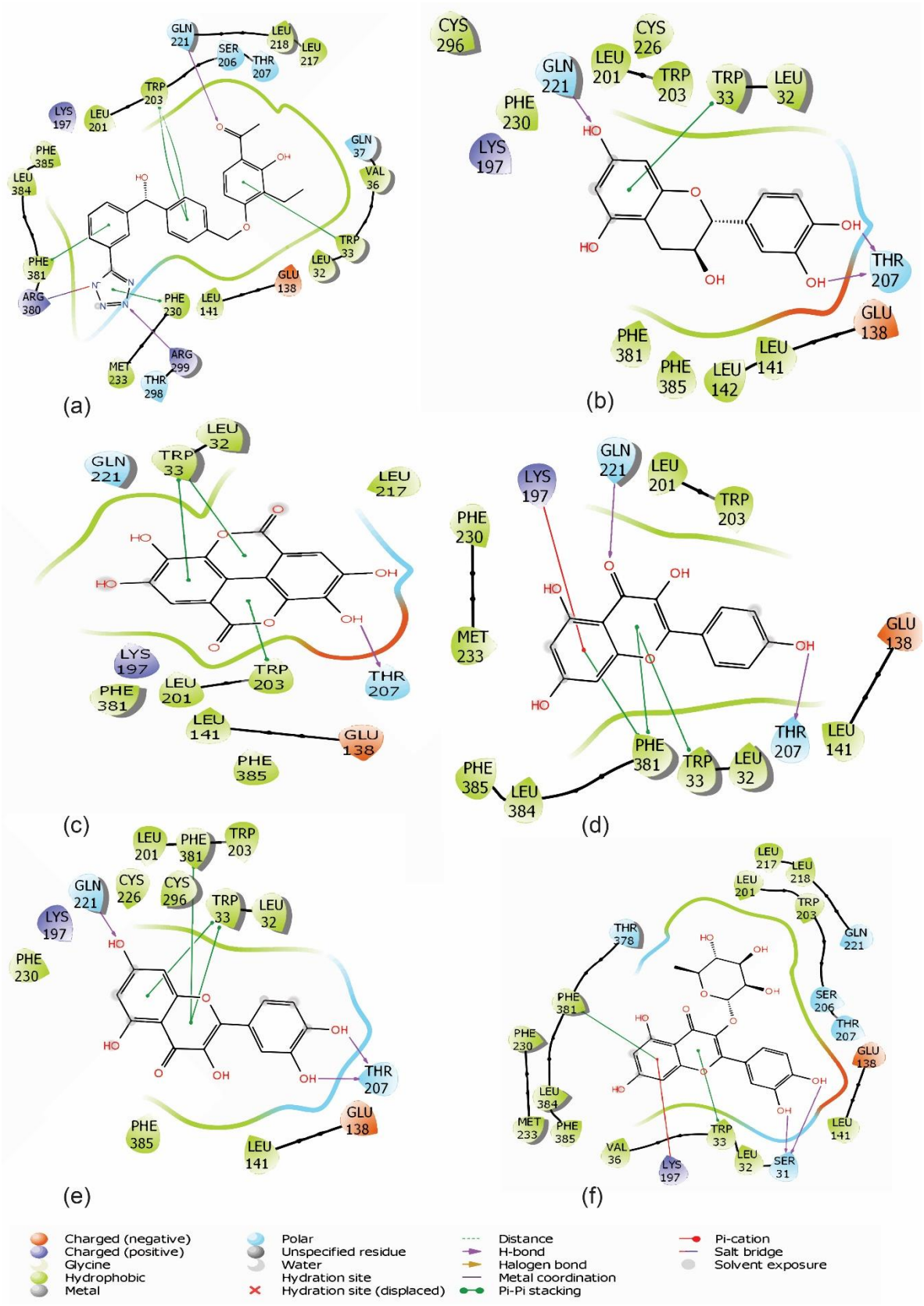

Figure 2. Two-dimensional interactions of (**a)** Quercitrin; (**b)** Quercetin; (**c)** 3-ethyl-2-hydroxy; (**d)** Kaempferol; (**e)** Ellagic acid; (**f)** Catechin with residues at the binding pocket of GLP-1 receptor.

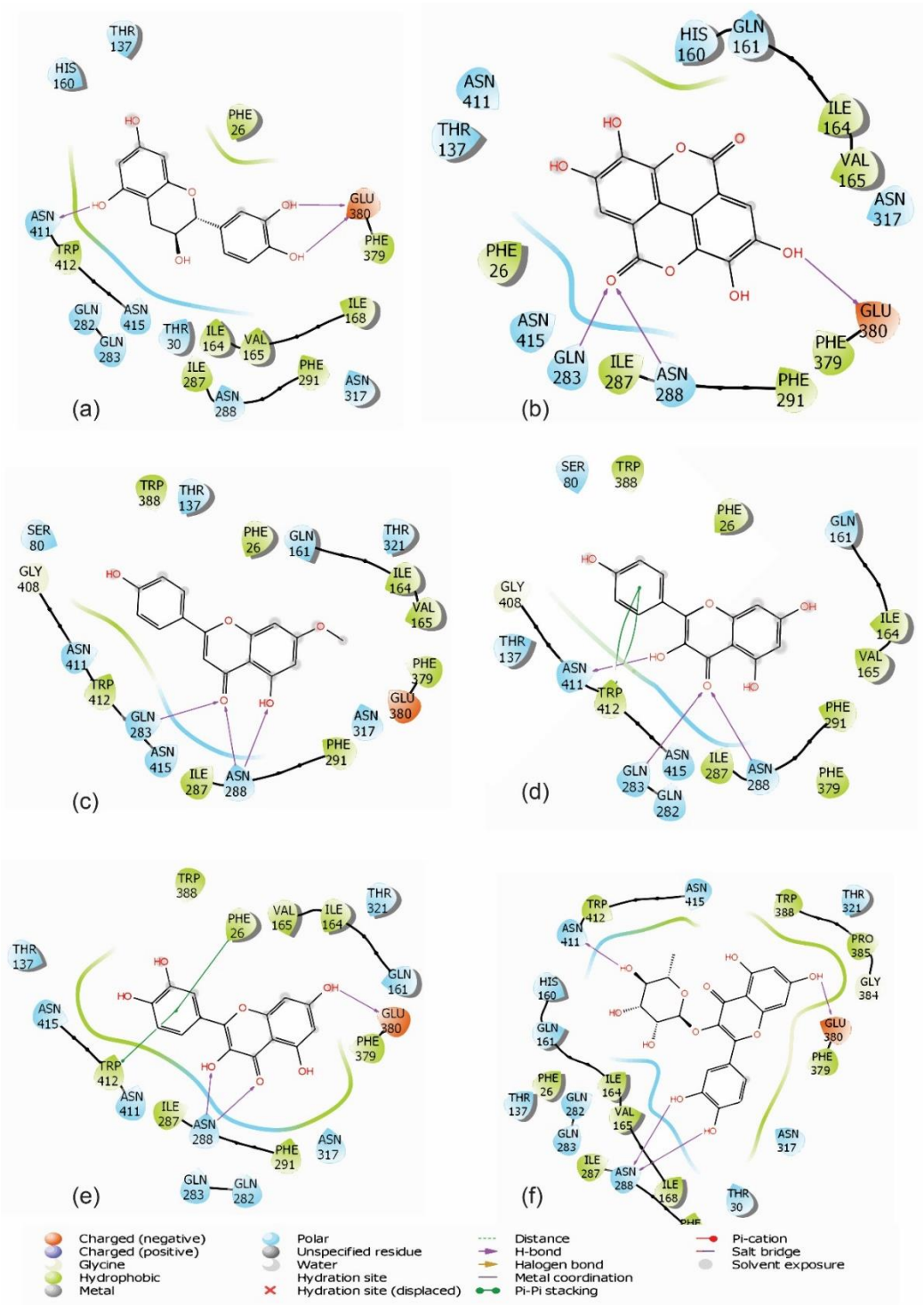

Figure 3. Two-dimensional interactions of (**a)** Quercitrin; (**b)** Catechin; (**c)** Flavone,4,5-dihydroxy-7-methoxy; (**d)** Resveratrol; (**e)** Ellagic acid (**f)** Kaempferol with residues at the binding pocket of GLUT1.

Although quercetin, which has already been reported as a potent inhibitor of aldose reductase [33], demonstrated the most favorable binding with aldose reductase, other compounds selected as hits equably had high docking scores. It can be concluded that since these compounds shared a similar mode of interactions with aldose reductase, it is likely they possess a similar inhibitory profile as quercetin.

Table 2: DFT calculations showing the FMOs and global descriptive parameters of the hit compounds from *Brachystegia eurycoma.*

Parameters	A	B	C	D	E	F	G	H	I	J	K
HOMO	-0.21649	-0.20227	-0.20449	-0.21797	-0.21779	-0.06920	-0.22444	-0.06270	-0.20620	-0.01921	-0.21728
LUMO	-0.05097	-0.04739	-0.01409	-0.03628	-0.05254	-0.00741	-0.07364	-0.02893	-0.01928	-0.00864	-0.06538
Band gap	0.16552	0.15488	0.19404	0.18169	0.16525	0.06179	0.1508	0.03377	0.18692	0.01057	0.28266
Ionization potential (I)	0.21649	0.20227	0.20449	0.21797	0.21779	0.06920	0.22444	0.06270	0.20620	0.01921	0.21728
Electron affinity (A)	0.05097	0.04739	0.01409	0.03628	0.05254	0.00741	0.07364	0.02893	0.01928	0.00864	0.06538
Chemical hardness (η)	0.08276	0.07744	0.09702	0.09085	0.08263	0.030895	0.0754	0.01689	0.09346	0.00528	0.14133
Chemical softness (ζ)	6.0425	6.4567	5.1536	5.5036	6.0511	16.1838	6.6667	29.6031	5.3498	94.6970	3.5378
Electronegativity (χ)	0.24195	0.22597	0.21154	0.23611	0.24406	0.07291	0.29808	0.09163	0.22548	0.02353	0.24997
Chemical potential (μ)	-0.08276	-0.07744	-0.09702	-0.09085	-0.08263	-0.03089	-0.0754	-0.01689	-0.09346	-0.00528	-0.14133

A = Quercetin; B = Kaempferol; C = Catechin; D = Luteolin; E = Flavone,4,5-dihydroxy-7-methoxy; F = 3-ethyl-2-hydroxy; G = Ellagic ; H = Rosmarinic acid; I = Resveratrol; J = Chlorogenic acid; K = Quercitrin

Calculating the binding free energy of ligand-bound protein stability complexes is one of the most accurate post-docking analytic methods for validating docking score findings [33, 34]. A negative number indicates that the outcomes were favorable. The binding free energy results (Table 1) showed that the compounds all formed favorable stability with the protein targets. Chlorogenic acid formed the most stable complex with aldose reductase; Quercitrin, Resveratrol, and Rosmarinic acid are compounds with the most stable molecules with GLP-1 receptor; and finally, Ellagic acid formed the most favorable stability with GLUT1.

Chemical compound stability and reactivity may be predicted using quantum mechanics, which uses the most up-to-date mathematical models for electronic behavior in water, benzene, dichloromethane, and chlorobenzene, among others [34]. With the use of the frontier molecular orbitals (HOMO and LUMO orbitals), scientists may learn about a chemical compound's ability to donate and absorb electrons. The "band gap" is the difference between the HOMO and LUMO orbitals. For a given chemical, the bandgap defines its stability; a greater bandgap indicates a more stable molecule, whereas a lower bandgap indicates a reactive or less stable compound [33, 34]. In this study, the FMOs and the global descriptive parameters of the understudied compounds were calculated in the gas phase. The comprehensive results are shown in Table 2. From the data enlisted in the aforementioned table, and going by the description of HOMO/LUMO, Luteolin had the highest electron-donating capacity with a HOMO score of -0.21797 eV, and Ellagic acid had the highest electron receiving capacity with LUMO score of -0.07364 eV. The metabolic stability of drug candidates is often determined in both liver microsome and hepatocyte assays. However, the behavior of the compounds electronically can give insight into the level of reactivity of a molecular compound in the gas, aqueous or liquid phase.

Therefore, based on the energy gap value and its implications, chlorogenic acid is the most reactive compound among the hits, contributing to the compound's proclivity to participate in the spontaneous reaction. However, Quercitrin is the most chemically inert molecule and is unlikely to initiate any spontaneous reaction in the gas phase.

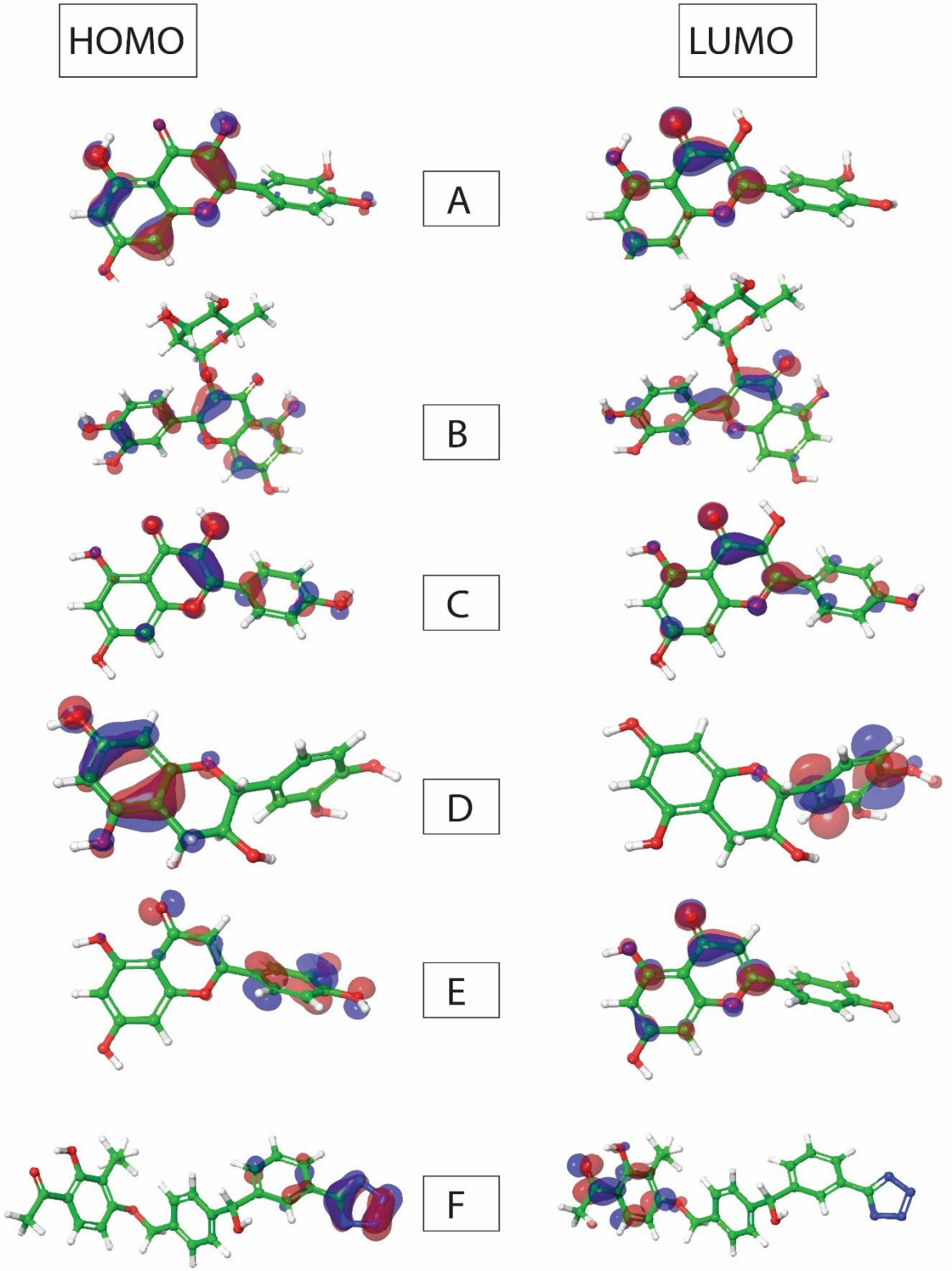

Figure 4. Frontier molecular orbitals showing HOMO/LUMO of (**a)** Quercetin; (**b)** Kaempferol; (**c)** Catechin; (**d)** Luteolin; (**e)** Flavone,4,5-dihydroxy-7-methoxy; (**f)** 3-ethyl-2-hydroxy.

The global descriptors, which include the ionization potential (IP), electron affinity (EA), softness, and hardness, determine the reactivity and stability of the compounds. The values of these descriptors are derived from the HOMO/LUMO and bandgap using defined equations, which are then used to calculate the reactivity and stability of the compounds. Based on IP and EA, Luteolin is the most stable, and chlorogenic acid is the most reactive of the drug candidates.

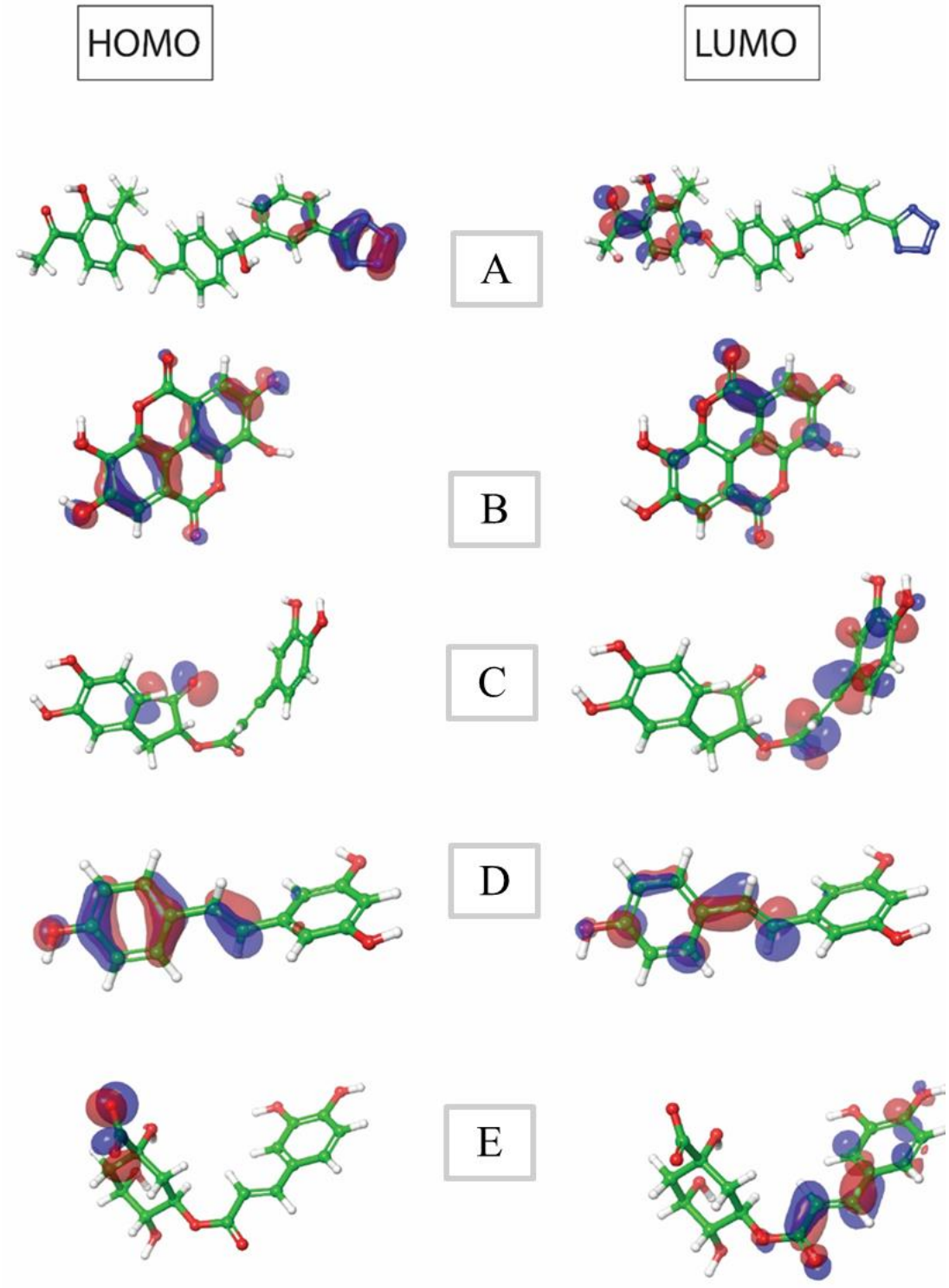

Figure 5. Frontier molecular orbitals showing HOMO/LUMO of **(a)** Ellagic (b**)** Rosmarinic acid (c**)** Resveratrol (d**)** Chlorogenic acid (e**)** Quercitrin.

According to the hardness and softness metrics used to assess a chemical compound's reactivity and stability, the stability of a compound rises with hardness but decreases with softness, whereas the reactivity of a compound reduces with hardness but increases with softness. According to this description, the reactive chemicals are designated as chlorogenic acids. Other descriptive values calculated by DFT are shown in Table 2. The 3D diagram of the compounds showing the atomic sites prone to donating and receiving electrons is shown in Figures 4 and 5.

3.3. Druglikeness prediction.

All eleven compounds identified as hits were tested for pharmacokinetics, drug-likeness, and toxicity. This prediction is critical in drug discovery and development because it allows for the elimination of medicines or prospective drug candidates that have undesirable side effects or off-target effects. Lipinski's rule, Ghose's hypothesis, and Veber's rule are some rules formulated to determine the drug-likeness of a chemical compound. The Lipinski rule of five is useful in differentiating between compounds similar to drugs and molecules not similar to drugs. It forecasts a high likelihood of success or failure due to drug resemblance for compounds that meet two or more of the following criteria: Less than 500 Dalton in molecular mass, high lipophilicity (defined as LogP less than 5), fewer than 5 hydrogen bonds donors, fewer than 10 hydrogen bond acceptors, and a molar refractivity ranging from 40 to 130 [35]. Following the Veber criteria, active drugs with total hydrogen bonds less than 12, rotatable bonds less than 10, and polar surface area (PSA) less than 140 are likely to have oral bioavailability $\geq 20\%$ percent. Log P (0.45.6), MR (molar refractivity 40150), MW 160480, number of atoms 2070, and PSA 140 are all values that meet the Ghose criteria for an active pharmaceutical ingredient [36]. The results of the ADME/tox predictions, as shown in Table 3, demonstrated that all the compounds except Chlorogenic acid and Quercitrin were in accordance with Lipinski and Ghose rule. Additionally, only three compounds violated the Veber rule. Due to these compounds' drug-likeness or oral bioavailability, they tend to possess high gastrointestinal absorption (GI). The profiles of the compounds' drug-drug interactions provided further information. From a computational standpoint, additional information regarding the drugs' pharmacokinetic characteristics in conjunction with five of the major cytochrome p450 isoenzymes involved in drug transformation (CYP1A2, CYP2C19, CYP2D6, CYP3A4, and CYP2C9) was considered. Some of the compounds, which are Catechin, Rosmarinic acid, Chlorogenic acid, and Quercitrin, are not a substrate of any of the five isoenzymes. However, other compounds are inhibitors of one or more principal isoenzymes of cytochrome p450, which denote their tendency to interact with the metabolites of another bio-transformed drug, which may increase toxicity or reduce the efficacy of another drug. The toxicity profiles of the compounds in Table 3 attest that the compounds had varying toxicity levels. Despite the differences in toxicities among the compounds, Chlorogenic acid showed the cleanest toxicity profiles. Results showed that this compound is neither carcinogenic nor hepatotoxic nor can it induce eye corrosion or irritation.

Table 3. ADMET/tox of the hit compounds.

Parameters	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**
Lipinski's rule violation	0	0	0	0	0	0	0	0	0	0	2
Ghose #violations	0	0	0	0	0	0	0	0	0	1	0
Veber #violations	0	0	0	0	0	0	1	1	0	1	1
GI absorption	High	High	High	High	High	High	High	Low	High	Low	Low
BBB permeant	No	No	No	No	No	No	No	No	Yes	No	No
Pgp substrate	No	No	Yes	No	No	Yes	No	No	No	No	No
CYP1A2 inhibitor	Yes	Yes	No	Yes	Yes	No	Yes	No	Yes	No	No
CYP2C19 inhibitor	No	No	No	No	No	No	No	No	No	No	No
CYP2C9 inhibitor	No	No	No	No	Yes	Yes	No	No	No	No	No

Parameters	A	B	C	D	E	F	G	H	I	J	K
CYP2D6 inhibitor	Yes	Yes	No	Yes	Yes	No	No	No	No	No	No
Carcinogenicity	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative
Eye corrosion	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative
Eye irritation	Positive	Positive	Positive	Positive	Positive	Negative	Positive	Positive	Positive	Negative	Negative
Hepatotoxicity	Positive	Positive	Negative	Positive	Positive	Positive	Positive	Positive	Positive	Negative	Positive
Acute Oral Toxicity	III	III	IV	III	III	III	II	III	III	III	III
Aromatase binding	Positive	Positive	Positive	Positive	Positive	Positive	Positive	Negative	Positive	Positive	Negative

A = Quercetin; B = Kaempferol; C = Catechin; D = Luteolin; E = Flavone,4,5-dihydroxy-7-methoxy; F = 3-ethyl-2-hydroxy; G = Ellagic ; H = Rosmarinic acid; I = Resveratrol; J = Chlorogenic acid; K = Quercitrin.

Table 4 summarizes the prediction models used to calculate the bioactivities of the compounds against functional proteins. Kernel partial least squares regression with binary fingerprint was used to create the models (radial, molprint2d as descriptors). For GLUT1, aldose reductase, and GLP-1R, reliable correlation coefficients (R2) of 0.8067, 0.8600, and 0.7537 were observed [37].

Table 4. Best model generated for GLUT1, Aldose reductase, and GLP-1 receptor.

Best model generated for GLUT1						
S/n	**Model code**	**Score**	**S.D**	**R^2**	**RMSE**	**Q^2**
1	kpls_molprint2D_38	0.8099	0.4361	0.8067	0.4027	0.8166
Best model generated for Aldose reductase						
S/n	**Model code**	**Score**	**S.D**	**R^2**	**RMSE**	**Q^2**
	Kpls_radial_4	0.8448	0.4768	0.8600	0.4660	0.8466
Best model generated for GLP-1 receptor						
S/n	**Model code**	**Score**	**S.D**	**R^2**	**RMSE**	**Q^2**
1	Kpls_molprint2d_20	0.7592	0.3092	0.7537	0.2763	0.7481

The compounds' pIC50 (inhibitory activities) against GLUT1 and aldose reductase were substantially anticipated, ranging between 6.762 and 5.661. (Table 5). The pEC50 values of the compounds also indicate their efficacy as GLP-1R agonists. By and large, the plant compounds chosen as hits can function as inhibitors of GLUT1 and aldose reductase while having a positive effect on GLP-1 receptors, therefore increasing insulin release in beta-pancreatic cells.

Table 5. Predicted activities of the hit compounds against the three protein targets.

Entry name	GLUT1	Aldose reductase	GLP-1 receptor
Quercetin	5.32	6.59	6.63
Kaempferol	5.27	6.14	6.87
Catechin	5.48	5.43	6.12
Luteolin	5.34	6.63	6.99
Flavone,4,5-dihydroxy-7-methoxy	5.38	7.01	6.73
3-ethyl-2-hydroxy	5.41	6.17	6.01
Ellagic acid	5.47	5.63	6.64
Rosmarinic acid	5.39	6.60	6.91
Resveratrol	5.41	5.47	6.17
Chlorogenic acid	5.40	6.07	6.77
Quercitrin	5.32	5.22	6.60

Figure 6. illustrates the scatter plots comprising both the test and training sets used to create prediction models. Table S2-S4 compares the AutoQSAR predicted activities to the observed activities for the studied proteins.

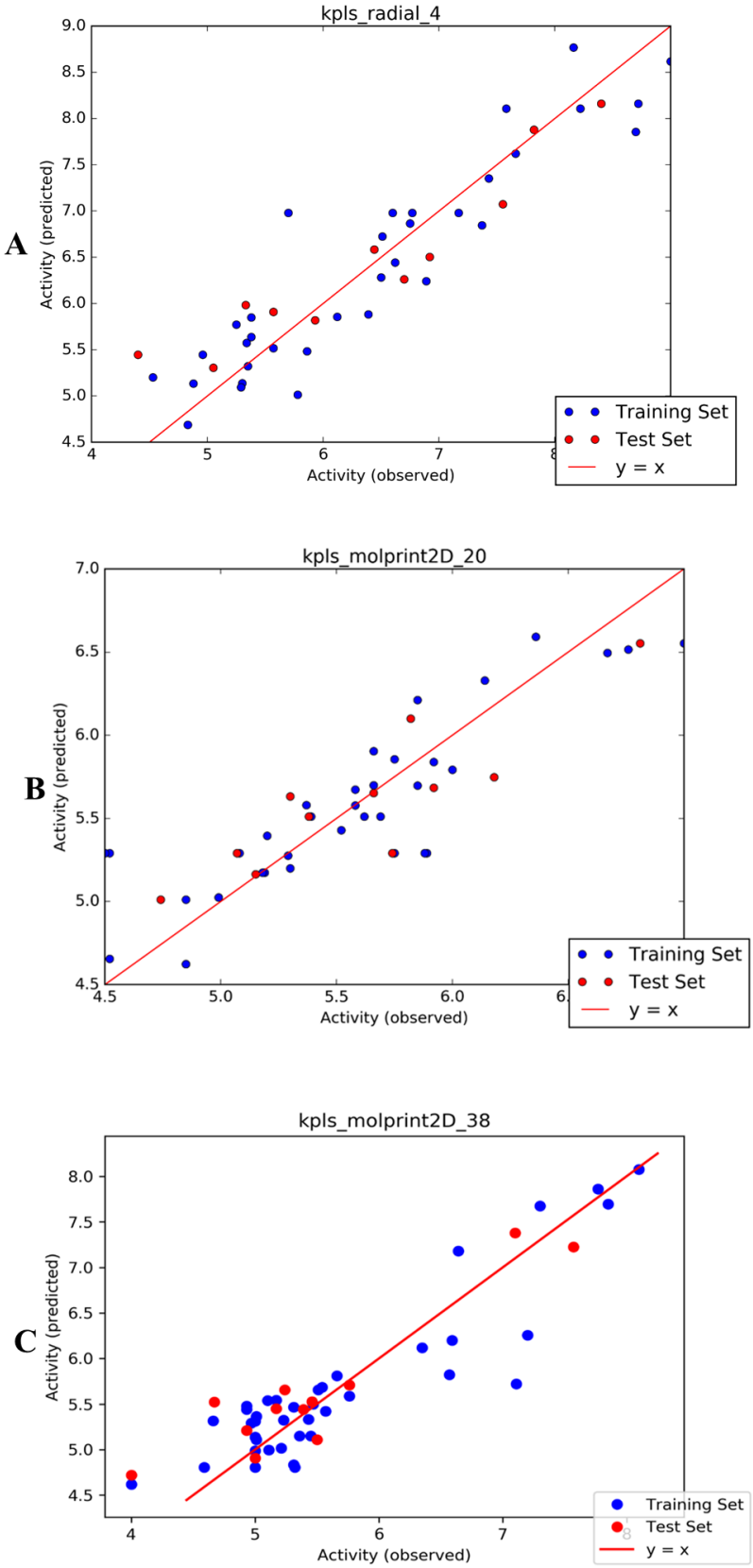

Figure 6. Scatter plot of (**a)** best QSAR model for Aldose reductase; (**b)** best QSAR model for GLP-1 receptor; (**c)** best QSAR model for GLUT1.

4. Conclusions

In this study, 78 compounds identified in *B. eurycoma* were docked with fourteen proteins linked to different molecular pathways involved in the pathogenesis of type 2 diabetes. After extensive scrutinization, the compounds made the most favorable interactions with GLUT1, aldose reductase, and the GLP-1 receptor. Docking scores indicate that the compounds have low binding energy and a high affinity for the proteins. Eleven compounds were selected as hits and were further brought forward for other computational analyses. The compounds with the highest docking scores interacted with critical amino acid residues for protein inhibition or activation and formed stable complexes with the proteins, exhibiting moderation for ADMET parameters. This work shows that quercetin, kaempferol, and catechin are unique bioactive compounds from *B. eurycoma* leaf that could interact with one or more druggable proteins examined in this study which could be developed as adjuvant therapy with minimal or no side effect. These compounds are worth subjecting to further analysis.

Funding

This research received no external funding. authors

Acknowledgments

The authors are grateful to the management of Afe Babalola University for their continuous support.

Conflicts of Interest

The authors declare that there is no conflict of interest.

References

1. Asoiro, F.U.; Ezeoha, S.L.; Anyanwu, C.N.; Aneke, N.N. Physical properties of Irvingia gabonensis, Detarium microcapum, Mucuna pruriens and *Brachystegia eurycoma* seeds. *Heliyon* **2020** 19;6, e04885, https://doi.org/10.1016/j.heliyon.2020.e04885.
2. Palit, S.P.; Patel, R.; Jadeja, S.D.; Rathwa, N,; Mahajan, A.; Ramachandran, A.V.; Dhar, M.K.; Sharma, S.; Begum, R. A. genetic analysis identifies a haplotype at adiponectin locus: Association with obesity and type 2 diabetes. *Sci Rep.* **2020,** *10*, 2904, https://doi.org/10.1038/s41598-020-59845-z.
3. Ajala, L.O.; Okafor, M.C.; Ndukwe, M.K.; Okoro, O.E.; Ogundele, G.J.; Ogundele, R.B. In Vitro Element Bioavailability Studies of Some Underutilized Seeds in Southeast Nigeria. *Biol Trace Elem Res*. **2021,** *199*, 3977-3986, https://doi.org/10.1007/s12011-020-02507-y.
4. Skyler, J.S.; Bakris, G.L.; Bonifacio, E.; Darsow, T.; Eckel, R.H.; Groop, L. Differentiation of diabetes by pathophysiology, natural history, and prognosis. *Diabetes care.* **2017** , *45*, 241–255, https://doi.org/10.2337/db16-0806.
5. Kottaisamy, C.; Raj, D. S.; Prasanth, H.; Kumar, V.; Sankaran, U. Experimental animal models for diabetes and its related complications-a review. *Laboratory animal research* **2021** , *37*, 23, https://doi.org/10.1186/s42826-021-00101-4.
6. Santos, M.; Cavalcante, A.; Amaral, L.; Souza, G.; Santos, B.; Portugal, L. C.; Bittencourt Junior, F. F.; Troquez, T.; Rafacho, B.; Hiane, P. A.; Santos, E. Combination of cafeteria diet with intraperitoneally streptozotocin in rats. A type-2 diabetes model. *Acta cirurgicabrasileira* **2021**, *36*, e360702, https://doi.org/10.1590/ACB360702.
7. Uki, Kawai.; Kazushi, Uneda.; Takayuki, Yamada.; Sho, Kinguchi.; Kazuo, Kobayashi.;Kengo, Azushima.; Tomohiko, Kanaoka. Comparison of effects of SGLT-2 inhibitors and GLP-1 receptor agonists on cardiovascular and renal outcomes in type 2 diabetes mellitus patients with/without albuminuria: A systematic

review and network meta-analysis. *Diabetes Research and Clinical Practice* **2021**, 109146, https://doi.org/10.1016/j.diabres.2021.109146.
8. Zaccardi, F.; Webb, D.R.; Yates, T.; Davies, M.J. Pathophysiology of type 1 and type 2 diabetes mellitus: a 90-year perspective. *Postgrad Med J*. **2016**, *92*, 63-9, https://doi.org/10.1136/postgradmedj-2015-133281.
9. Kasole, R.; Martins, H.D.; Kimiywe, J. Traditional Medicine and Its Role in the Management of Diabetes Mellitus: "Patients' and Herbalists' Perspectives". *Evid Based Complement Alternat Med* **2019** 2835691, https://doi.org/10.1155/2019/2835691.
10. Rena, G.; Hardie, D.G.; Pearson, E.R. the mechanisms of action of metformin. *Diabetologia* **2017** *60*, 1577-1585, https://doi.org/10.1007/s00125-017-4342-z.
11. McGovern, A.; Tippu, Z.; Hinton, W.; Munro, N.; Whyte, M.; De Lusignan, S. Comparison of medication adherence and persistence in type 2 diabetes: A systematic review and meta-analysis. *Diabetes ObesMetab* **2018**, *20*, 1040–1043. https://doi.org/10.1111/dom.13160.
12. Lavernia, F.; Adkins, S.E.; Shubrook, J.H. Use of oral combination therapy for type 2 diabetes in primary care: Meeting individualized patient goals. *Postgrad Med.* **2015**, *127*, 808-17, https://doi.org/10.1080/00325481.2015.1085293.
13. Debrah, A.; Godfrey, O. M., Rital, K. Prevalence and risk factors associated with type 2 diabetes in elderly patients aged 45-80 years at Kanungu District. *Journal of Diabetes research* **2020**, *2020*, 5, https://doi.org/10.1155/2020/5152146.
14. Kumar, A.; Tiwari, A.; Sharma, A. Changing Paradigm from one Target one Ligand towards Multi-target Directed Ligand Design for Key Drug Targets of Alzheimer Disease: An Important Role of In Silico Methods in Multi-target Directed Ligands Design. *CurrNeuropharmacol.* **2018**, *16*, 726-739, https://doi.org/10.2174/1570159X16666180315141643.
15. Artasensi, A.; Pedretti, A.; Vistoli, G.; Fumagalli, L. Type 2 Diabetes Mellitus: A Review of Multi-Target *Drugs. Molecules.* **2020**, *25*, 1987, https://doi.org/10.3390/molecules25081987.
16. Macalino, S.J.; Gosu, V.; Hong, S.; Choi, S. Role of computer-aided drug design in modern drug discovery. *Arch Pharm Res.* **2015**, 38, 1686-701, https://doi.org/10.1007/s12272-015-0640-5.
17. Uyoyoghene, O.U.; Michael, O.M.; Lawrence, C.C. Proximate analysis and phytochemical profile of *Brachystegia eurycoma* leaves. *Asian Journal of research in biochemistry* **2019**, *4*, http://dx.doi.org/10.9734/ajrb/2019/v4i230064.
18. Ogunbinu, A.O.; Ogunwande, I.A.; Walker, T.M.; Setzer, W.N. Identification of the volatile constituents of *Brachystegia eurycoma* Harms. *International Journal of Aromatherapy* **2016**, *16*, 155-158, https://doi.org/10.1016/j.ijat.2006.09.002
19. Irondi, E.A.; Oboh, G.; Akindahunsi, A.A.; Boligon, A.A.; Athade, L.M. Phenolics composition and anti-diabetic property of *Brachystegia eurycoma* seed flour in high –fat diet, low dose streptozotocin-induced type-2 diabetes in rats. *Journal of tropical diseases* **2015**, S159-S165, https://doi.org/10.1016/S2222-1808(15)60880-5.
20. Bhowmick, A.; Banu, S. Therapeutic targets of type 2 diabetes: an overview. *MOJ Drug Des develop therapy* **2017**, *1*, 60–64, https://doi.org/10.15406/mojddt.2017.01.00011.
21. Opeyemi, Iwaloye.; Olusola, Olalekan.; Elekofehinti, Babatomiwa.; Kikiowo, Emmanuel.; Ayo, Oluwarotimi.; Toyin, Mary Fadipe. Machine Learning-based Virtual Screening Strategy Reveals Some Natural Compounds as Potential PAK4 Inhibitors in Triple Negative Breast Cancer". *Current Proteomics* **2021**, *18*,. https://doi.org/10.2174/1570164618999201223092209.
22. Gupta, N.; Pandya, P.; Verma, S. Computational Predictions for Multi-Target Drug Design. In: Roy K. (eds) Multi-Target Drug Design Using Chem-Bioinformatic Approaches. Methods in Pharmacology and Toxicology. *Humana Press, New York, NY* **2018**, https://doi.org/10.1007/7653_2018_26.
23. Tahia, F.; Dagogo-Jack S. Role of ceramides in the pathogenesis of diabetes mellitus and its complications. *Journal of Diabetes Complications.* **2021,** *35*, 107734, https://doi.org/10.1016/j.jdiacomp.2020.107734.
24. Eid, H.M.; Haddad, P.S. The Antidiabetic Potential of Quercetin: Underlying Mechanisms. *Curr Med Chem.* **2017**; *24*, 355-364, https://doi.org/10.2174/0929867323666160909153707.
25. Xincheng, Y.; Ling, Z.; Yuxin, C.; Jun, T.; Youwei, W. In vivo and in vitro antioxidant activity and α-glucosidase, α-amylase inhibitory effects of flavonoids from Cichorium glandulosum seeds. *Journal of food chemistry* **2019**, *139*, 59-66, https://doi.org/10.1016/j.foodchem.2012.12.045.
26. Vinayagam, R.; Xu, B. Antidiabetic properties of dietary flavonoids: a cellular mechanism review. *NutrMetab (Lond)* **2015**, *12*, 60, https://doi.org/10.1186/s12986-015-0057-7.

27. Lu, L.; Seidel, C.P.; Iwase, T.; Stevens, R.K.; Gong, Y.Y.; Wang, X.; Hackett, S.F.; Campochiaro, P.A. Suppression of GLUT1; a new strategy to prevent diabetic complications. *J Cell Physiol*. **2013,** *228*, 251-7, https://doi.org/10.1002/jcp.24133. PMID: 22717959.
28. Kikiowo, B.; Ogunleye, A. J.; Iwaloye. O.; Ijatuyi, T. T.; Adelakun, N. S.; Alashe, W. O. Induced Fit Docking and Automated QSAR Studies Reveal the ER-α Inhibitory Activity of Cannabis sativa in Breast Cancer". *Recent Patents on Anti-Cancer Drug Discovery* **2021**, *16*, https://doi.org/10.2174/1574892816666210201115359
29. Elekofehinti, O.O.; Iwaloye, O.; Olawale, F. Newly designed compounds from scaffolds of known actives as inhibitors of survivin: computational analysis from the perspective of fragment-based drug design. *In Silico Pharmacol*. **2021**, *47*, https://doi.org/10.1007/s40203-021-00108-8.
30. Vikas, Kumar.; Kalicharan,Sharma.; Bahar, Ahmed, F. A.; Al-Abbasi, Firoz Anwar, Amita. Deconvoluting the dual hypoglycemic effect of wedelolactone isolated from Wedeliacalendulacea: investigation via experimental validation and molecular docking. *RSC Advances* **2018**, *54*, 18180-18196, https://doi.org/10.1039/C7RA12568B.
31. Sharma, P.; Joshi, T.; Joshi, T.; Chandra, S.; Tamta, S. In silico screeningof potential anti-diabetic phytochemicals from Phyllanthus emblica against therapeutic targets of type 2diabetes. *Journal of Ethnopharmacology* **2019**, https://doi.org/10.1016/j.jep.2019.112268.
32. Reddy, T.N.; Ravinder, M.; Bagul, P.; Ravikanti, K.; Bagul, C.; Nanubolu,J.B. *Synthesis and biological evaluation of new epalrestat analogues as aldose reductase inhibitors (ARIs). Eur J Med Chem* **2014,** *71*, 53–66, https://doi.org/10.1016/j.ejmech.2013.10.043.
33. Elekofehinti, O.O.; Iwaloye, O.; Molehin, O.R. Identification of lead compounds from large natural product library targeting 3C-like protease of SARS-CoV-2 using E-pharmacophore modelling, QSAR and molecular dynamics simulation. *In Silico Pharmacol* **2021**, *49*, https://doi.org/10.1007/s40203-021-00109-7.
34. American Diabetes Association; 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes—2020. *Diabetes Care* **2020**, *43*, S14–S31, https://doi.org/10.2337/dc20-S002.
35. Oyinloye, B. E.; Iwaloye, O.; Ajiboye, B. O. Polypharmacology of Gongronemalatifolium leaf secondary metabolites against protein kinases implicated in Parkinson's disease and Alzheimer's disease. *Scientific African* **2021** , e00826, https://doi.org/10.1016/j.sciafi.2021.e00862.
36. Elekofehinti, O.O.; Iwaloye, O.; Josiah, S.S. Molecular docking studies, molecular dynamics and ADME/tox reveal therapeutic potentials of STOCK1N-69160 against papain-like protease of SARS-CoV-2. *Mol Divers* **2021** , 1761–1777, https://doi.org/10.1007/s11030-020-10151-w.

Supplementary materials

Table S1. Preliminary study showing the docking scores of best compounds against selected diabetes targets.

PROTEIN ID: 2R24			**PROTEIN ID: 5U5L**		
S/N	**COMPOUNDS NAME**	**DOCKING SCORE**	**S/N**	**COMPOUNDS NAME**	**DOCKING SCORE**
1	Quercetin	-11.813	1	3-ethyl-2-hydroxy	-9.261
2	Catechin	-11.339	2	Coumarin	-8.303
3	Kaempferol	-11.250	3	9-[2-Deoxy-beta-d-ribohexopyranosyl] Purin-6 (1H) one	-7.655
4	Luteolin	-11.132	4	Bis (2-ethylhexyl) Phthalate	-7.491
5	Flavone,4,5-dihydroxy-7-methoxy	-10.734	5	Flavone,4,5-dihydroxy-7-methoxy	-7.048
6	Ellagic acid	-10.023	6	Gallic acid	-6.953
7	Rosmarinic acid	-9.857	7	4-H-Pyran-4-one, 2,3-dihydro-3,5-dihydroxy-6-methyl	-6.925
8	Resveratrol	-9.441	8	Xanthoprotein	-6.924
9	Gallic acid	-8.520	9	2- Cyclopenten-1-one	-6.305
10	Anthraquinone	-8.401	10	Ferulic acid	-6.230
11	Coumarin	-7.910	11	Ethyl-2- hydroxybenzyl sulfone	-6.087
12	1,3-Dimethyl naphthalene	-7.894	12	2,3-dihydro-Benzofuran Coumaran	-5.596
13	Geranyl acetone	-7.865	13	Resveratrol	-5.496
14	Ferulic acid	-7.696	14	1,3-Dimethyl naphthalene	-5.419
15	Piperidine, 1-chloroacetyl-	-7.426	15	2-hydroxy-5-methylisophthalaldehyde	-5.344
PROTEIN ID: 4Y14			PROTEIN ID: **5ZCB**		
S/N	**COMPOUNDS NAME**	**DOCKING SCORE**	**S/N**	**COMPOUNDS NAME**	**DOCKING SCORE**
1	4-H-Pyran-4-one, 2,3-dihydro-3,5-dihydroxy-6-methyl	-9.491	1	Rosmarinic acid	-8.143
2	Gallic acid	-7.509	2	rutin	-7.984
3	Xanthoprotein	-7.149	3	Quercitrin	-6.489
4	caffeic acid	-7.002	4	Catechin	-5.470
5	Chlorogenic acid	-6.848	5	9-[2-Deoxy-beta-d-ribohexopyranosyl] Purin-6 (1H) one	-5.393
6	Ferulic acid	-6.617	6	Gallic acid	-5.026
7	Coumarin	-6.353	7	Luteolin	-4.376
8	2- Cyclopenten-1-one	-5.646	8	1,2,3-Trihydroxybenzene	-4.376
9	Luteolin	-4.716	9	Xanthoprotein	-4.175
10	2-hydroxy-5-methylisophthalaldehyde	-4.694	10	Kaempferol	-3.879
11	Resveratrol	-4.539	11	catehol	-3.836
12	3-ethyl-2-hydroxy	-4.040	12	Quinines	-3.753
13	Furan-3-Carboxaldehyde-2-methoxy-2,3-dihydro	-3.873	13	1,6-Anhydro-beta-D-glucopyranose	-3.739
14	Hydroquinone	-3.273	14	Piperidine, 1-chloroacetyl-	-3.336
15	4-H-Pyran-4-one, 2,3-dihydro-3,5-dihydroxy-6-methyl	-9.491	15	4-H-Pyran-4-one, 2,3-dihydro-3,5-dihydroxy-6-methyl	-3.291
PROTEIN ID: 1B2Y			PROTEIN ID: 6X1A		
S/N	**COMPOUNDS NAME**	**DOCKING SCORE**	**S/N**	**COMPOUNDS NAME**	**DOCKING SCORE**
1	rutin	-9.881	1	Quercitrin	-12.983
2	Chlorogenic acid	-8.813	2	Chlorogenic acid	-12.277

3	Quercetin	-7.365	3	3-ethyl-2-hydroxy	-11.538
4	Luteolin	-7.240	4	Ellagic acid	-11.038
5	9-[2-Deoxy-beta-d-ribohexopyranosyl] Purin-6 (1H) one	-7.119	5	Luteolin	-10.971
6	Flavone,4,5-dihydroxy-7-methoxy	-6.486	6	Diterpene	-10.612
7	Ellagic acid	-5.538	7	Bis (2-ethylhexyl) Phthalate	-10.325
8	Catechin	-5.357	8	Kaempferol	-10.315
9	2-O-methyl-D-mannopyranosa	-5.256	9	Catechin	-10.035
10	3-ethyl-2-hydrox	-4.901	10	Resveratrol	-9.411
11	2-hydroxy-5-methylisophthalaldehyde	-4.830	11	Flavone,4,5-dihydroxy-7-methoxy	-8.326
12	caffeic acid	-4.647	12	9-[2-Deoxy-beta-d-ribohexopyranosyl] Purin-6 (1H) one	-7.695
13	Gallic acid	-4.597	13	Anthraquinone	-7.658
14	Piperidine, 1-chloroacetyl-	-4.463	14	Alloaromadendrene	-7.583
15	1,6-Anhydro-beta-D-glucopyranose	-4.173	15	b-Caryophyllene	-7.503
PROTEIN ID: **1GAG**			PROTEIN ID: GPR120		
S/N	**COMPOUNDS NAME**	**DOCKING SCORE**	**S/N**	**COMPOUNDS NAME**	**DOCKING SCORE**
1	Chlorogenic acid	-11.807	1	rutin	-12.812
2	4-H-Pyran-4-one, 2,3-dihydro-3,5-dihydroxy-6-methyl	-9.858	2	Chlorogenic acid	-8.480
3	9-[2-Deoxy-beta-d-ribohexopyranosyl] Purin-6 (1H) one	-9.554	3	3-ethyl-2-hydroxy	-8.117
4	rutin	-9.173	4	Luteolin	-7.816
5	Gallic acid	-8.926	5	Kaempferol	-7.705
6	Eicosanoic acid	-8.810	6	Rosmarinic acid	-6.912
7	Gallic acid	-8.804	7	Flavone,4,5-dihydroxy-7-methoxy	-6.815
8	Tetracosanoic acid	-8.788	8	1,3-Dimethyl naphthalene	-6.525
9	Quercitrin	-8.194	9	p-Cymene	-6.494
10	3-ethyl-2-hydroxy	-7.870	10	Alloaromadendrene	-6.378
11	Ferulic acid	-7.777	11	Catechin	-6.363
12	2-hydroxy-5-methylisophthalaldehyde	-7.347	12	Resveratrol	-6.363
13	Xanthoprotein	-6.682	13	Anthraquinone	-5.707
14	Rosmarinic acid	-6.418	14	Alloaromadendrene	-5.251
15	Catechin	-6.321	15	caffeic acid	-5.040
PROTEIN ID: **ADIPONECTIN**			PROTEIN ID: **5NN5**		
S/N	**COMPOUNDS NAME**	**DOCKING SCORE**	**S/N**	**COMPOUNDS NAME**	**DOCKING SCORE**
1	rutin	-12.581	1	rutin	-9.274
2	Procyanidin B2	-10.316	2	Quercitrin	-7.366
3	Quercitrin	-9.921	3	Ellagic acid	-7.134
4	Ellagic acid	-8.184	4	Luteolin	-6.125
5	Catechin	-7.729	5	Catechin	-5.673
6	9-[2-Deoxy-beta-d-ribohexopyranosyl] Purin-6 (1H) one	-7.286	6	9-[2-Deoxy-beta-d-ribohexopyranosyl] Purin-6 (1H) one	-5.666
7	Luteolin	-7.241	7	Kaempferol	-5.451
8	Flavone,4,5-dihydroxy-7-methoxy	-6.410	8	3-ethyl-2-hydroxy	-5.431
9	1,6-Anhydro-beta-D-glucopyranose	-5.327	9	2-O-methyl-D-mannopyranosa	-5.197

10	Gallic acid	-5.318	10	Flavone,4,5-dihydroxy-7-methoxy	-4.975
11	1,2,3-Trihydroxybenzene	-5.130	11	1,6-Anhydro-beta-D-glucopyranose	-4.528
12	Piperidine, 1-chloroacetyl-	-4.610	12	Resveratrol	-4.522
13	Furan-3-Carboxaldehyde-2-methoxy-2,3-dihydro	-4.424	13	1,2,3-Trihydroxybenzene	-4.315
14	Xanthoprotein	-4.302	14	Gallic acid	-4.141
15	2-hydroxy-5-methylisophthalaldehyde	-4.215	15	Anthraquinone	-3.902
PROTEIN ID: **2QMJ**			PROTEIN ID: **5EQG**		
S/N	**COMPOUNDS NAME**	**DOCKING SCORE**	**S/N**	**COMPOUNDS NAME**	**DOCKING SCORE**
1	rutin	-8.222	1	Ellagic acid	-9.709
2	Desulphosinigrin	-7.798	2	rutin	-15.457
3	2-O-methyl-D-mannopyranosa	-6.881	3	Quercitrin	-11.421
4	Quercetin	-6.480	4	9-[2-Deoxy-beta-d-ribohexopyranosyl] Purin-6 (1H) one	-9.465
5	3-O-Methyl-d-glucose	-6.332	5	Catechin	-9.297
6	Cis-9-Hexadecenal,9-hexadecenal,(Z)	-6.332	6	Flavone,4,5-dihydroxy-7-methoxy	-8.860
7	Catechin	-6.322	7	3-ethyl-2-hydroxy	-8.475
8	1,6-Anhydro-beta-D-glucopyranose	-6.115	8	Kaempferol	-8.042
9	Kaempferol	-5.712	9	Luteolin	-7.740
10	Quinines	-5.644	10	Chlorogenic acid	-7.311
11	1,2,3-Trihydroxybenzene	-5.490	11	Beta-Sitosterol	-7.095
12	a-Pinene	-5.398	12	Quinines	-6.393
13	9-[2-Deoxy-beta-d-ribohexopyranosyl] Purin-6 (1H) one	-5.303	13	Resveratrol	-6.148
14	Catechol	-5.250	14	Anthraquinone	-6.050
15	Terpinen-4-ol	-5.151	15	Piperidine, 1-chloroacetyl-	-5.872
PROTEIN ID: 2V4M				PROTEIN ID: 4CFE	
S/N	**COMPOUNDS NAME**	**DOCKING SCORE**	**S/N**	**COMPOUNDS NAME**	**DOCKING SCORE**
1	9-[2-Deoxy-beta-d-ribohexopyranosyl] Purin-6 (1H) one.1	-7.849	1	Chlorogenic acid.1	-7.188
2	Gallic acid.1	-6.337	2	4-H-Pyran-4-one, 2,3-dihydro-3,5-dihydroxy-6-methyl.1	-6.675
3	Catechin.1	-6.124	3	Gallic acid.1	-5.902
4	caffeic acid.1	-5.994	4	Kaempferol.1	-5.566
5	4-H-Pyran-4-one, 2,3-dihydro-3,5-dihydroxy-6-methyl.1	-5.839	5	Xanthoprotein.1	-5.516
6	Furan-3-Carboxaldehyde-2-methoxy-2,3-dihydro.1	-5.230	6	Catechin.1	-5.456
7	Rosmarinic acid.1	-4.410	7	2-hydroxy-5-methylisophthalaldehyde.1	-5.358
8	Xanthoprotein.1	-4.305	8	Quercetin.1	-5.192
9	Ferulic acid.1	-3.973	9	1,2,3-Trihydroxybenzene.1	-4.983
10	1,6-Anhydro-beta-D-glucopyranose.1	-3.699	10	Rosmarinic acid.1	-4.337
11	Piperidine, 1-chloroacetyl-.1	-3.652	11	Terpinen-4-ol.1	-4.230
12	Chlorogenic acid.1	-2.208	12	Luteolin.1	-4.216
13	3-ethyl-2-hydroxy.1	-0.624	13	1,2,4-Trimethyl-3-nitrobicyclo (3,3,1) nonan-9-one.1	-4.203

14	2-hydroxy-5-methylisophthalaldehyde.1	-0.135	14	caffeic acid.1	-4.144
15	-	-	15	Furan-3-Carboxaldehyde-2-methoxy-2,3-dihydro.1	-3.981

Table S2. Details of AutoQSAR predicted activities compared with the observed activities for GLUT1.

s/n	CHEMBL CID	Set	Observed pIC_{50}	Predicted pIC_{50}	Residue Error
1	CHEMBL3780235	train	7.3000	7.6729	0.3729
2	CHEMBL4632395	train	6.5900	6.2002	-0.3898
3	CHEMBL579920	train	5.1000	5.5375	0.4375
4	CHEMBL4165190	train	4.6600	5.3162	0.6562
5	CHEMBL546792	train	5.7600	5.5890	-0.1710
6	CHEMBL531938	train	5.2100	5.0150	-0.1950
7	CHEMBL4283327	train	5.0000	5.3120	0.3120
8	CHEMBL4289139	train	7.1100	5.7208	-1.3892
9	CHEMBL4290111	train	5.0000	5.1369	0.1369
10	CHEMBL3780460	train	8.1000	8.0780	-0.0220
11	CHEMBL3781308	train	7.8500	7.6945	-0.1555
12	CHEMBL591362	train	5.0100	5.3634	0.3534
13	CHEMBL602586	train	4.9300	5.4438	0.5138
14	CHEMBL4064730	train	6.3500	6.1194	-0.2306
15	CHEMBL4061928	train	6.5700	5.8227	-0.7473
16	CHEMBL4069907	train	4.9300	5.4801	0.5501
17	CHEMBL590933	test	5.1700	5.4482	0.2782
18	CHEMBL533252	test	5.2400	5.6568	0.4168
19	CHEMBL547470	train	7.2000	6.2553	-0.9447
20	CHEMBL533298	train	5.5400	5.6844	0.1444
21	CHEMBL528147	train	5.5700	5.4197	-0.1503
22	CHEMBL580275	train	5.4500	5.1510	-0.2990
23	CHEMBL547127	train	5.3100	5.4664	0.1564
24	CHEMBL526944	test	5.5000	5.1088	-0.3912
25	CHEMBL546770	train	5.5100	5.6563	0.1463
26	CHEMBL3780785	test	7.5700	7.2239	-0.3461
27	CHEMBL3781151	train	7.7700	7.8592	0.0892
28	CHEMBL3781535	test	7.1000	7.3796	0.2796
29	CHEMBL3781847	train	6.6400	7.1791	0.5391
30	CHEMBL479638	train	4.9700	5.2908	0.3208
31	CHEMBL45068	test	4.6700	5.5244	0.8544
32	CHEMBL4294011	test	4.0000	4.7185	0.7185
33	CHEMBL4292959	train	4.0000	4.6195	0.6195
34	CHEMBL4283795	test	5.0000	4.9072	-0.0928
35	CHEMBL4542975	train	4.5900	4.8060	0.2160
36	CHEMBL4542417	train	5.3200	4.8060	-0.5140
37	CHEMBL548777	train	5.2300	5.3256	0.0956
38	CHEMBL529822	test	5.7600	5.7093	-0.0507
39	CHEMBL547715	train	5.6600	5.8106	0.1506
40	CHEMBL588752	train	5.3600	5.1492	-0.2108
41	CHEMBL581191	train	5.4700	5.4999	0.0299
42	CHEMBL546544	train	5.1100	4.9956	-0.1144
43	CHEMBL527924	train	5.3100	4.8330	-0.4770
44	CHEMBL536892	test	5.4600	5.5263	0.0663

s/n	CHEMBL CID	Set	Observed pIC_{50}	Predicted pIC_{50}	Residue Error
45	CHEMBL532775	test	5.3900	5.4400	0.0500
46	CHEMBL548111	train	5.4300	5.3314	-0.0986
47	CHEMBL587923	train	5.0100	5.1114	0.1014
48	CHEMBL547266	test	4.9300	5.2118	0.2818
49	CHEMBL589920	train	5.0000	4.9863	-0.0137
50	CHEMBL584015	train	5.1700	5.5450	0.3750
51	CHEMBL4542975	train	5.0000	4.8060	-0.1940

Table S3. Details of AutoQSAR predicted activities compared with the observed activities for aldose reductase.

s/n	CHEMBL CID	Set	Observed pIC_{50}	Predicted pIC_{50}	Residue Error
1	CHEMBL399370	train	5.2000	5.3960	0.1960
2	CHEMBL401292	train	4.5200	4.6557	0.1357
3	CHEMBL250747	test	5.3000	5.6323	0.3323
4	CHEMBL250310	test	5.9200	5.6845	-0.2355
5	CHEMBL402918	train	6.1400	6.3288	0.1888
6	CHEMBL250928	test	5.6600	5.6535	-0.0065
7	CHEMBL249091	train	6.6700	6.4950	-0.1750
8	CHEMBL250325	train	5.5800	5.6724	0.0924
9	CHEMBL250091	train	5.5800	5.5779	-0.0021
10	CHEMBL401183	train	5.6600	5.6992	0.0392
11	CHEMBL251128	train	5.7500	5.8554	0.1054
12	CHEMBL399713	train	6.7600	6.5155	-0.2445
13	CHEMBL398714	train	7.0000	6.5526	-0.4474
14	CHEMBL398714	test	6.8100	6.5526	-0.2574
15	CHEMBL250529	train	4.8500	4.6239	-0.2261
16	CHEMBL399369	test	6.1800	5.7475	-0.4325
17	CHEMBL400800	train	5.3700	5.5797	0.2097
18	CHEMBL400500	train	4.8500	4.6239	-0.2261
19	CHEMBL401184	test	5.1500	5.1638	0.0138
20	CHEMBL249510	test	5.8200	6.1000	0.2800
21	CHEMBL399720	train	5.6600	5.9043	0.2443
22	CHEMBL248912	train	5.2900	5.2753	-0.0147
23	CHEMBL261639	train	5.8500	5.6971	-0.1529
24	CHEMBL399841	train	6.0000	5.7918	-0.2082
25	CHEMBL250748	train	5.5200	5.4290	-0.0910
26	CHEMBL400700	train	5.8500	6.2112	0.3612
27	CHEMBL398962	train	5.3000	5.2007	-0.0993
28	CHEMBL3359269	test	5.3800	5.5110	0.1310
29	CHEMBL3359270	train	5.6200	5.5110	-0.1090
30	CHEMBL3359270	train	5.6900	5.5110	-0.1790
31	CHEMBL3359269	train	5.3900	5.5110	0.1210
32	CHEMBL250112	train	6.3600	6.5917	0.2317
33	CHEMBL3359272	test	5.7400	5.2906	-0.4494
34	CHEMBL3359272	train	5.0800	5.2906	0.2106
35	3CHEMBL3359272	train	5.8900	5.2906	-0.5994
36	CHEMBL3359272	train	5.8800	5.2906	-0.5894
37	CHEMBL3359271	train	5.1900	5.1730	-0.0170
38	CHEMBL3359271	train	5.1800	5.1730	-0.0070

s/n	CHEMBL CID	Set	Observed pIC_{50}	Predicted pIC_{50}	Residue Error
39	CHEMBL3359272	test	5.0700	5.2906	0.2206
40	CHEMBL3359272	train	5.7500	5.2906	-0.4594
41	CHEMBL3359272	train	4.5200	5.2906	0.7706
42	CHEMBL3359272	train	4.5000	5.2906	0.7906
43	CHEMBL251760	train	5.9200	5.8387	-0.0813
44	CHEMBL3359274	train	4.9900	5.0245	0.0345
45	CHEMBL3359274	train	4.9900	5.0245	0.0345
46	CHEMBL3359275	train	4.8500	5.0122	0.1622
47	CHEMBL3359275	test	4.7400	5.0122	0.2722

Table S4. Details of AutoQSAR predicted activities compared with the observed activities for GLP-1R.

s/n	CHEMBL CID	Set	Observed pIC_{50}	Predicted pIC_{50}	Residue Error
1	CHEMBL418008	train	6.5000	6.2839	-0.2161
2	CHEMBL266497	train	5.7000	6.9803	1.2803
3	CHEMBL142599	test	7.5500	7.0742	-0.4758
4	CHEMBL20884	test	6.9200	6.5060	-0.4140
5	CHEMBL282017	train	7.3700	6.8484	-0.5216
6	CHEMBL304257	train	6.6200	6.4439	-0.1761
7	CHEMBL20024	train	8.7000	7.8582	-0.8418
8	CHEMBL7183	train	6.5100	6.7269	0.2169
9	CHEMBL422920	test	5.3300	5.9836	0.6536
10	CHEMBL168614	train	5.3500	5.3222	-0.0278
11	CHEMBL171315	test	5.0500	5.3073	0.2573
12	CHEMBL82749	train	5.3000	5.1374	-0.1626
13	CHEMBL441254	train	6.8900	6.2426	-0.6474
14	CHEMBL363040	test	4.4000	5.4469	1.0469
15	CHEMBL20796	test	7.8200	7.8820	0.0620
16	CHEMBL10372	train	8.7200	8.1636	-0.5564
17	CHEMBL266497	train	6.7700	6.9803	0.2103
18	CHEMBL69675	train	5.2900	5.0950	-0.1950
19	CHEMBL266773	test	6.7000	6.2623	-0.4377
20	CHEMBL139666	train	7.6600	7.6217	-0.0383
21	CHEMBL142934	train	5.3800	5.8518	0.4718
22	CHEMBL69277	train	4.8300	4.6877	-0.1423
23	CHEMBL240725	train	9.0000	8.6183	-0.3817
24	CHEMBL241577	train	8.2200	8.1087	-0.1113
25	CHEMBL344526	train	7.4300	7.3552	-0.0748
26	CHEMBL75630	test	5.5700	5.9120	0.3420
27	CHEMBL71594	train	4.8800	5.1345	0.2545
28	CHEMBL304945	train	5.7800	5.0135	-0.7665
29	CHEMBL304797	train	5.5700	5.5180	-0.0520
30	CHEMBL470376	train	6.1200	5.8569	-0.2631
31	CHEMBL283996	train	8.1600	8.7709	0.6109
32	CHEMBL307121	train	6.3900	5.8830	-0.5070
33	CHEMBL66326	test	5.9300	5.8184	-0.1116
34	CHEMBL1744091	train	5.8600	5.4830	-0.3770
35	CHEMBL372861	test	6.4400	6.5842	0.1442
36	CHEMBL73415	train	6.7500	6.8683	0.1183

s/n	CHEMBL CID	Set	Observed pIC_{50}	Predicted pIC_{50}	Residue Error
37	CHEMBL335376	train	5.3400	5.5757	0.2357
38	CHEMBL20009	train	7.5800	8.1093	0.5293
39	CHEMBL283984	train	5.3800	5.6387	0.2587
40	CHEMBL69818	train	4.5300	5.2021	0.6721
41	CHEMBL70014	train	5.2500	5.7729	0.5229
42	CHEMBL266497	train	7.1700	6.9803	-0.1897
43	CHEMBL266497	train	6.6000	6.9803	0.3803
44	CHEMBL189464	train	4.9600	5.4469	0.4869
45	CHEMBL189464	test	8.4000	8.1636	-0.2364