Chemistry Africa (2026) 9:140
https://doi.org/10.1007/s42250-025-01544-1

ORIGINAL ARTICLE

Bioprospection for Putative Antidiabetic Leads from *Momordica balsamina* Using Network Pharmacology, Molecular Modelling and Density Functional Theory Calculations

Halimat Yusuf Lukman[1] · **Rukayat Abiola Abdulsalam**[1] · **Saheed Sabiu**[1]

Received: 21 January 2025 / Accepted: 1 November 2025 / Published online: 23 March 2026
© The Author(s) 2026

Abstract

Purpose With the increasing incidence and mortality rate of type 2 diabetes mellitus (T2DM), alternative management options with better efficacy and less adverse effects are continuously being researched. Despite reports validating the antidiabetic property of *Momordica balsamina*, computational bioprospection for its putative antidiabetic leads and mechanism of action remain elusive.

Method The study employed network pharmacology, molecular docking, molecular dynamics simulation, density functional theory, and in silico pharmacokinetics prediction to identify the putative leads and discern the antidiabetic mechanism of action *M. balsamina*.

Result An initial screening of the 74 compounds retrieved from *M. balsamina* showed oral drug-likeness while gene ontology analysis identified EGFR tyrosinase kinase inhibitor as the most significant pathway with *IGF1R* and *GSK3B* being the most prominent genes in the pathway. Analysis of the binding interactions of the top-ranked compounds and the genes revealed that except for myricetin-*GSK3B* complex, other resulting complexes had lower binding free energy values for the genes relative to the standard gene inhibitors. Kaempferol-3-O-glucoside (-51.50 ± 7.97 kcal/mol) and balsaminoside B (-61.79 ± 7.35 kcal/mol), however, presented superior binding affinities with *IGF1R* and *GSK3B*, respectively, suggestive of their better potential to interact and downregulate the genes than other top *M. balsamina* compounds.

Conclusion Evidence from this study is the potential of *M. balsamina* compounds to downregulate *IGF1R* and *GSK3B* expressions and formation of thermodynamically stable and compact interactions with the genes for glucose homeostasis via insulin sensitivity and glycogen metabolism regulation. Thus, the identified leads can be further explored as possible candidates for antidiabetic drug development.

Keywords Bioprospection · *Momordica balsamina* · Molecular dynamic simulation · Network pharmacology · Type 2 diabetes mellitus

Abbreviations

T2DM	Type 2 diabetes mellitus
EGFR	Epidermal Growth Factor Receptor
IGF1R	Insulin-like Growth Factor 1 Receptor
GSK3B	Glycogen Synthase Kinase 3B
DFT	Density Functional Theory
NP	Network Pharmacology
MD	Molecular Dynamics
OMIM	Online Mendelian Inheritance in Man
SEA	Similarity Ensemble Approach
STP	Swiss Target Prediction
PPI	Protein-Protein Interaction
PCT	Pathway Compound Target
GGI	Gene-Gene Interaction
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
RMSD	Root Mean Square Deviation
RMSF	Root Means Square Fluctuation
SASA	Solvent Accessible Surface Area

✉ Saheed Sabiu
sabius@dut.ac.za

[1] Department of Biotechnology and Food Science, Faculty of Applied Sciences, Durban University of Technology, P. O. Box 1334, Durban 4000, South Africa

ROG	Radius of Gyration
MM/GBSA	Molecular Mechanics with Generalized Born Surface Area
LUMO	Lowest Unoccupied Molecular Orbital
HOMO	Highest Occupied Molecular Orbital

1 Introduction

As part of the continuous efforts to reduce the failure rates at clinical phase of drug development and cost of production of more effective therapeutics, computer-aided drug design has proven to be an efficient and reliable tool [1]. Of significance is network pharmacology (NP), an established computational technique employed in investigating the underlying mechanism of drugs relative to the biological targets involved in a particular disease [2]. The technique has been used to elucidate the intricate relationship between the bioactive constituents of medicinal plants and metabolic diseases including type 2 diabetes mellitus (T2DM), a disease requiring safer and more effective treatment alternative due to its rising and predicted incidence rate [3–5]. Besides NP, molecular dynamics (MD) simulation has also been used to analyze the degree of affinity, stability and interactions between hit compounds and biological targets [6, 7].

Studies have employed both NP and MD simulation to assess plant constituents for the signaling pathways and genes they could target for further consideration in the development of potent antidiabetic drug candidates. For example, Akoonjee et al. [4] profiled *Khaya grandifoliola* metabolites and identified *PRKCA and MMP2* as the target genes while Gogoi et al. [8] documented *GSK3B* as most significant target with auraptene, anisessine, and adhatodine from North-East India antidiabetic plants. However, the putative mechanism of action via signaling pathways and genes of several antidiabetic plants including *Momordica balsamina* are limited.

Commonly known as Balsam apple or African pumpkin, *M. balsamina* belongs to the family Cucurbitaceae and is widely distributed across tropical, subtropical regions, and parts of Africa, including South Africa [9, 10]. Traditionally, this plant has been used to treat ailments such as hepatitis, rheumatism, wounds, gastroenteritis, syphilis, fever, and diabetes [11]. Phytochemical studies reveal that *M. balsamina* contains flavonoids, phenolics, triterpenoids, carotenoids, sterols, and cucurbitanes which contribute to its biological activities including diabetes management [12–14]. The methanolic extract of *M. balsamina* reportedly produced hypoglycemic effect in animal model studies [15, 16]. Similarly, betulinic acid from the plant was proposed by Kabir et al. [17] as one of the potential lead compounds contributing to the observed antidiabetic activity in STZ-induced diabetic rats. Despite these studies, the mechanisms of antidiabetic action of *M. balsamina* remain poorly understood. Therefore, this study employed both NP and MD simulation to decipher the possible signaling pathways and genes implicated in *M. balsamina* and T2DM.

2 Methods

2.1 Retrieval and Screening of *M. balsamina* Compounds

Momordica balsamina compounds were retrieved from PubMed server (https://pubmed.ncbi.nlm.nih.gov/) using the search term bioactive + compounds + of + Momordica + balsamina) (accessed on 2 October 2023), and Dr. Duke phytochemical and ethnobotanical database (https://phytochem.nal.usda.gov/) (accessed on 4 October 2023). These compounds were screened for oral drug-likeness properties using the Lipinski rule of five (Ro5) by inputting their canonical Simplified Molecular Input Line Entry System (SMILES) obtained from PubChem (https://pubchem.ncbi.nlm.nih.gov/) into the SwissADME server ((http://swissadme.ch/index.php) [18].

2.2 Retrieval of T2DM and *M. balsamina* Genes

The DisGeNet (https://www.disgenet.org/search) [19] and the Online Mendelian Inheritance in Man (OMIM) (https://www.ncbi.nlm.nih.gov/omim) [20], were accessed for T2DM genes in *Homo sapiens* mode while the genes for *M. balsamina* compounds were retrieved from the SEA (Similarity Ensemble Approach) (http://sea.bkslab.org/) and Swiss Target Prediction (STP) (http://www.swisstargetprediction.ch/) using their SMILES. The common genes between T2DM and *M. balsamina* compounds were obtained using the VENNY 2.1 tool (https://bioinfogp.cnb.csic.es/tools/venny/).

2.3 Protein-protein Interaction and Pathway Compound Target Network Construction of Intersecting Targets

The protein-protein interaction (PPI) of the identified common genes between T2DM and *M. balsamina* compounds was constructed with the STRING database Version 11.0b (https://string-db.org/) at a confidence level of $P > 0.4$. The pathway compound target (PCT) network was constructed using Cytoscape's merge algorithm plugin (Cytoscape 3.8.2) by first obtaining the gene-gene

interaction (GGI) between the common T2DM targets and *M. balsamina* compounds. The network was analyzed to identify the degree of interactions algorithm (Eq. 1) observable through the numbers of nodes (PCT) and edges (interactions) [21]. The network topology was analyzed using the CytoHubba module in Cytoscape (Cytoscape 3.8.2 software) [22].

$$Deg\,(v) = |N\,(v)| \tag{1}$$

Where v is each node neighbor and N(v) is the node's neighbors.

2.4 Gene Ontology (GO) Analysis of the Common Genes

The GO in *Homo sapiens* mode was analyzed using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) database (https://david.ncifcrf.gov/tools.jsp), an official gene symbol identifier [23].

2.5 Kyoto Encyclopedia of Genes and Genomes (KEGG) Analysis of the Common Genes

The KEGG pathways of the common genes were retrieved from the STRING database. The pathways closely related to T2DM were selected and the most significant pathway was identified using the false discovery rate (FDR) at a P-value < 0.05. The KEGG bubble plot and GO enriched bar chart were constructed using the SRC bioinformatics database (https://www.bioinformatics.com.cn/) [23].

2.6 Targets and Compounds Preparation for Molecular Docking

The 3D structures of the most enriched genes [*IGF1R* (PDB ID: 2OJ9) and *GSK3B* (PDB ID: 4ACD)] of the most significant pathway identified from the KEGG enrichment analysis were obtained from the RSCB protein Data Bank (https://www.rcsb.org/). The 3D structures of the compounds that interacted with these genes, reference standard (metformin) and gene antagonists (7-[3-(azetidin-1-ylmethyl) cyclobutyl]−5-(3-phenylmethoxyphenyl)pyrrolo[2,3-d] pyrimidin-4-amine (Nvp-aew541) and 2-hydroxy-3-[5-(4-morpholinylmethyl)−2-pyridinyl]−1 H-indole-5-carbonitrile (AZD-1080) for *IGF1R* and *GSK3B*, respectively, were retrieved from PubChem (https://pubchem.ncbi.nlm.nih.gov/). The target genes were optimized (by removing water molecules, protein residue connectivity, and missing side chains addition) using UCSF Chimera software V1.14 [24] while the compounds were optimized by adding Gasteiger charges and non-polar hydrogen atom [25].

2.7 Molecular Docking and Molecular Dynamics (MD) Simulation of the Targets and Compounds

Docking of the optimized target genes and compounds was done using Autodock Vina Plugin on Chimera with well-defined grid box coordinates, and dimensions [center (X: 22.75; Y: −20.41; Z: 8.284) and dimension (40.08; Y: 35.49; Z: 29.43) for 4ACD) and center (X: −11.79; Y: 15.29; Z: 71.72) and dimensions (X: 37.25; Y: 41.03; Z: 32.96) for 2OJ9]. The docking protocol validation which was done to prevent pseudo-positive binding by ensuring the compounds bind at similar binding pocket as the native ligand gave a Root Mean Square Deviation (RMSD) value of < 1.0 Å (Fig. 1) [26]. The top-ranked compounds that produced the highest negative docking scores with each target were saved in pdb format. The complexes were prepped on UCSF Chimera by saving the compound and gene separately in. mol and.pdb extensions, respectively, for subsequent MD simulation analysis.

The MD simulation was carried out using the graphical processing unit (GPU) version with the AMBER 18 package of the Centre for High Performance Computing (CHPC), Cape Town South Africa [6]. The study employed the AMBER 18 force field with FF18SB variant, and the simulation with a restraint in hydrogen atom bonds using the SHAKE algorithm of the SPFP precision model and simulation step size of 2 fs during the 120 ns simulation [27, 28]. The post-dynamics analysis of the RMSD, root mean square fluctuation (RMSF), solvent accessible surface area (SASA), and radius of gyration (ROG), and number of hydrogen (H) bonds were done with the CPPTRAJ module of AMBER 18 suite, and their plots obtained using Origin software V18 (Origin Lab Corporation, Northampton, MA, USA). The molecular mechanics with the generalized born surface area method [MM/GBSA] was used to obtain energy components [29] while the 2D interaction plots of the resulting complexes were visualized on Discovery Studio version 21.1.0.

2.8 Quantum Chemical Calculations for Top *M. balsamina* Metabolites

The DFT/B3LYP/631G/+ (d, p) basis set was employed for the molecular property prediction of top-ranked *M. balsamina* compounds identified in the MD simulation [30, 31]. The Gaussian 16 program package of CHPC, Cape Town, South Africa and Gauss View V 6.0.16 were employed for the optimization and visualization, respectively [32]. The highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) of the optimized compounds were analyzed using the Parr and Pearson interpretation of DFT and Koopmans theorem. Thereafter,

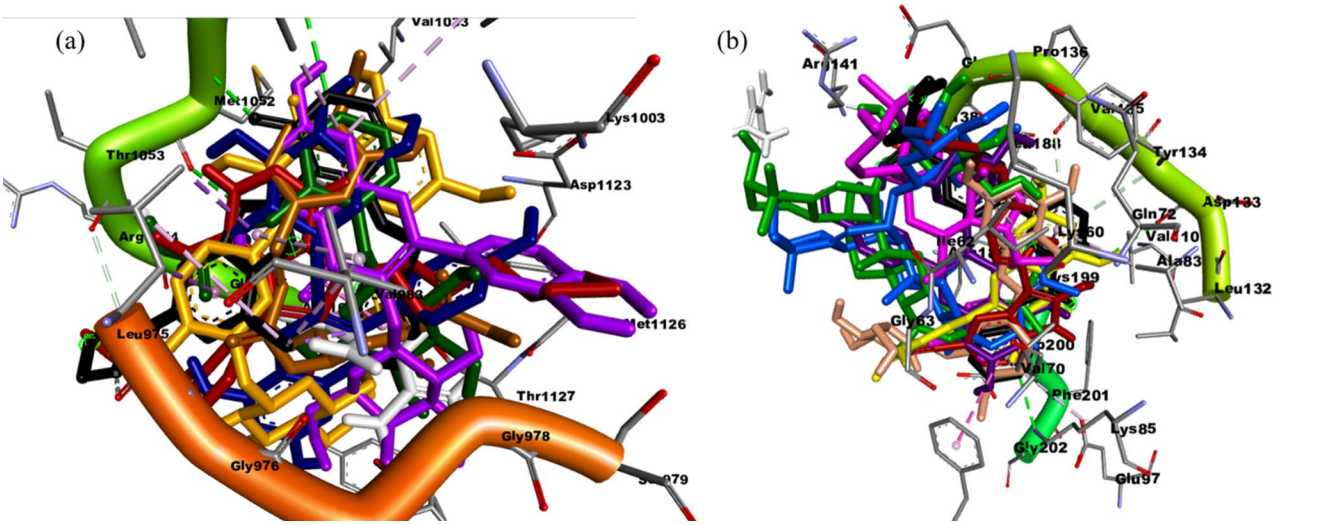


Fig. 1 Superimposition of the top ranked compounds of M. balsamina at the active site of the crystal structure of (a) IGF1R with 3-caffeoylquinic acid (brown), quercetin-3-O-glucoside (orange), quercetin 3-galactoside (purple), melilotoside (green), kaempferol-3-O-glucoside (blue), and standards [Nvp-aew541 (red), metformin (white)] with the native gene inhibitor (black) and (b) GSK3B with balsaminoside B (green), balsaminoside C (blue), cucurbalsaminol A (pink), myricetin (yellow), quercetin-3-O-glucoside (orange), quercetin 3-galactoside (purple), and standards [AZD-1080 (red), metformin (white)] with the native gene inhibitor (black)

other chemical descriptors such as energy gap (ΔE), ionization energy (I), electron affinity (A), chemical hardness (η), softness (δ), electronegativity (χ), chemical potential (Cp), global electrophilicity (ω) were estimated (Supplementary Table 1) [33–35].

2.9 Pharmacokinetics and Toxicity Profiles of the Top-Ranked Compounds of *M. balsamina*

The pharmacokinetics (absorption, distribution, metabolism, excretion) properties of the top-ranked compounds of *M. balsamina* were estimated by inserting their SMILES into the SwissADME server [18] while their toxicities were predicted using the ProTox server (https://tox-new.charite.de/protox_II/) (accessed on 30 April 2024) [36].

3 Results and Discussion

3.1 Screening of *M. balsamina* Compounds

The antidiabetic mechanism of action of *M. balsamina* and identification of its potential lead compounds were explored using NP and MD simulation techniques. Of the seventy-four compounds retrieved from *M. balsamina* data mining, 13 violated three (3) Lipinski's rule of five (Ro5) which states that H-bond donor, H-bond acceptor, molecular weight, lipophilicity, and bioavailability score should be ≤5, 10, 500 g/mol, 4.5, and 0.55, respectively (Table 1), and were hence excluded from the study. This suggests that compounds with ≤two (2) violations possess better oral drug-likeness profiles and suitable for possible development as antidiabetic agents.

3.2 Common Genes between *M. balsamina* and T2DM and PPI

The interactions between T2DM and *M. balsamina* compounds genes are shown in Fig. 2. Of the total 3134 and 15,025 T2DM genes retrieved from DisGeNet and GeneCards, respectively, 2653 were found to overlap (Fig. 2a), while 175 genes overlapped between the 6111 STP and 1847 SEA genes obtained for the screened *M. balsamina* compounds (Fig. 2b). Interaction of these overlapping genes resulted in 57 common genes between T2DM and *M. balsamina* compounds (Fig. 2c). The PPI of the common genes produced 57 nodes, 271 edges, an average node degree of 9.51, average local clustering coefficient of 0.672, and an enrichment P-value of 1.0×10^{-16}. However, *AMY1A* and *GPR35* were outliers, indicating their non-interaction with the whole network (Fig. 3). The observed PPI profile depicts *M. balsamina* relevance to target T2DM at the genetic level. Compounds with pharmacological effects at the molecular/genetic levels could possibly target multiple diseases since a particular gene can be implicated in several pathologies [37]. Flavonoids such as quercetin, kaempferol, and myricetin present in *M. balsamina* offer numerous pharmacological values including antidiabetic properties [38, 39]. Specifically, kaempferol has been reported to regulate lipid metabolism and enhance insulin receptor

Table 1 List of *M. balsamina* compounds used in this study and their oral drug-likeness properties

Compound ID (C#)	Compound name	MW	#H-bond acceptors	#H-bond donors	iLOGP	BS	LV
1	+) -Dehydrovomifoliol	222.28	3	1	1.91	0.55	0
2	3,4-Dicaffeoylquinic acid	516.45	12	7	1.25	0.11	3
3	3,5-Dicaffeoylquinic acid	516.45	12	7	1.05	0.11	3
4	3-Caffeoylquinic acid	354.31	9	6	0.96	0.11	1
5	4,5-Dicaffeoylquinic acid	516.45	12	7	1.25	0.11	3
6	4- Caffeoylquinic acid	354.31	9	6	0.96	0.11	1
7	4-p-Coumaroylquinic acid	338.31	8	5	1.45	0.56	0
8	Balsaminagenin A	474.72	4	4	4.4	0.55	1
9	Balsaminagenin B	488.74	4	3	4.42	0.55	1
10	Balsaminagenin C	470.68	4	2	3.95	0.55	0
11	Balsaminapentaol A	490.72	5	5	3.71	0.55	0
12	Balsaminol A	474.72	4	4	3.99	0.55	1
13	Balsaminol B	488.74	4	3	4.57	0.55	1
14	Balsaminol C	470.68	4	2	3.95	0.55	0
15	Balsaminol D	430.62	4	2	3.13	0.55	0
16	Balsaminol E	456.7	3	2	4.38	0.55	1
17	Balsaminol F	458.72	3	3	4.24	0.55	1
18	Balsaminoside A	634.88	8	5	4.71	0.55	1
19	Balsaminoside B	620.86	8	6	4.57	0.17	2
20	Balsaminoside C	620.86	8	6	4.57	0.17	2
21	Catechin	290.27	6	5	1.47	0.55	0
22	4-Feruloylquinic acid	368.34	9	5	1.52	0.11	1
23	4-Coumaroylquinic acid	338.31	8	5	1.45	0.56	0
24	Cucurbalsaminol A	474.72	4	4	4.51	0.55	1
25	Cucurbalsaminol B	488.74	4	3	4.63	0.55	1
26	Cucurbalsaminol C	488.74	4	3	4.38	0.55	1
27	Cucurbitacin S	498.65	6	2	3.37	0.55	0
28	Cucurbitadienol	426.72	1	1	5.18	0.55	1
29	Fevicordin A	528.63	8	4	3.24	0.55	1
30	Kaempferol-O-rutinoside	594.52	15	9	2.79	0.17	3
31	Isorhamnetin-3-O-glucoside	478.4	12	7	1.21	0.17	2
32	Kaempferol	286.24	6	4	1.7	0.55	0
33	Kaempferol-3- O-rutinoside	594.52	15	9	2.79	0.17	3
34	Kaempferol-3-O-glucoside	448.38	11	7	0.53	0.17	2
35	Karavilagenin A	486.77	3	1	5.05	0.55	1
36	Karavilagenin C	472.74	3	2	4.91	0.55	1
37	Karavilagenin E	456.7	3	2	4.77	0.55	1
38	Karavoate A	514.78	4	1	5.31	0.17	2
39	Karavoate B	556.82	5	0	5.33	0.17	2
40	Karavoate C	528.81	4	1	5.45	0.17	2
41	Karavoate D	584.87	5	0	6.13	0.17	2
42	Karavoate E	542.83	4	1	5.86	0.17	2
43	Karavoate F	612.92	5	0	6.61	0.17	2
44	Karavoate H	680.95	5	0	5.91	0.17	2
45	Karavoate I	621.85	6	1	5.3	0.17	2
46	Karavoate J	770.95	9	0	5.47	0.17	3
47	Karavoate K	611.29	4	1	5.92	0.17	2
48	Karavoate L	749.85	5	0	5.99	0.17	2
49	Karavoate M	606.87	5	1	5.57	0.17	2
50	Karavoate N	741.01	7	0	7.05	0.17	2
51	Karavoate O	602.89	4	1	6.24	0.17	2
52	Karavoate P	733.03	5	0	6.54	0.17	2
53	Kuguacin B	456.7	3	2	4.55	0.55	1
55	Kuguacin C	414.62	3	1	3.79	0.55	1

Table 1 (continued)

Compound ID (C#)	Compound name	MW	#H-bond acceptors	#H-bond donors	iLOGP	BS	LV
56	Kuguaglycoside A	634.88	8	5	3.66	0.55	1
57	Melilotoside	326.3	8	5	0.53	0.56	0
58	Methylquinic acid	368.34	9	5	1.33	0.11	0
59	Myricitin	318.24	8	6	1.08	0.55	1
60	Phenethyl rutinoside	430.45	10	6	1.98	0.55	1
61	Pseudolaroside A	330.29	9	5	1.13	0.56	0
62	Quercetin	302.24	7	5	1.63	0.55	0
63	Quercetin 3-galactoside	464.38	12	8	2.11	0.17	2
64	Quercetin-3,7-dirhamnoside	594.52	15	9	1.21	0.17	3
65	Quercetin-3,7-O-diglucoside	626.52	17	11	2.06	0.17	3
66	Quercetin-3-O-glucoside	464.38	12	8	2.11	0.17	2
67	Quercetin-3-O-robinobioside	610.52	16	10	1.58	0.17	3
68	Quercetin-3-O-robinoside	772.66	21	13	0.86	0.17	3
69	Quercetin-3-O-rutinoside	610.52	16	10	1.58	0.17	3
70	Quercetin-3-O-sambubioside	596.49	16	10	1.41	0.17	3
71	Quercetin-3-rutinoside	610.52	16	10	1.58	0.17	3
72	Verapamil	454.6	6	0	4.5	0.55	0
73	Quinic acid	192.17	6	5	−0.12	0.56	0
74	Rhamnetin-3-O-glucoside	462.4	11	6	2.81	0.17	2

[a]*BS* Bioavailability Score, *LV* Lipinski violation number

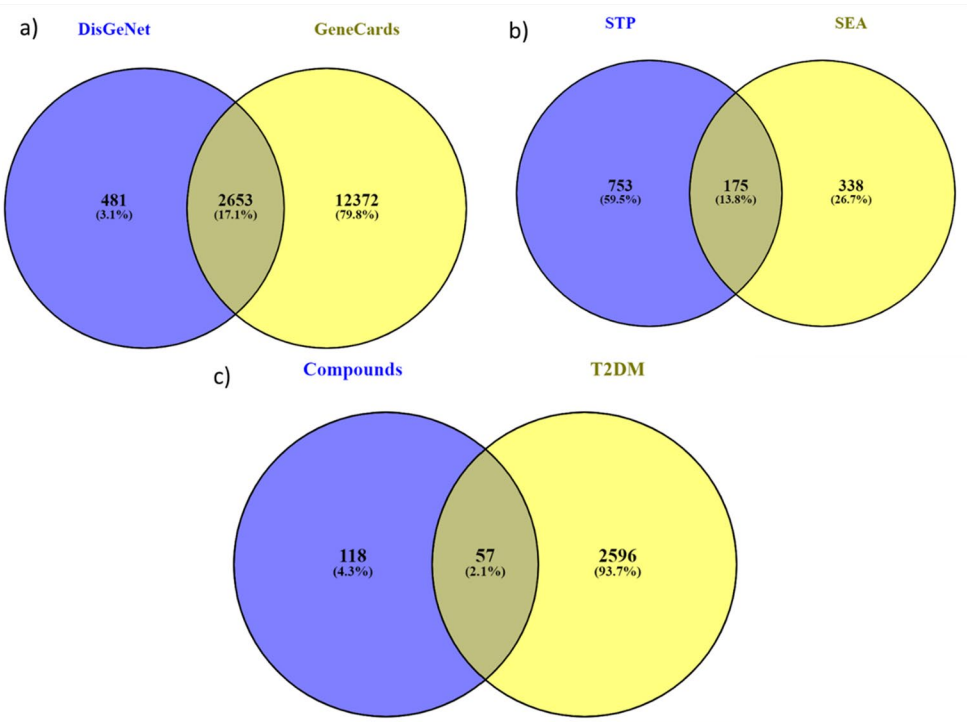


Fig. 2 Overlapping genes between **a** T2DM genes from DisGeNet and GeneCards **b** *M. balsamina* compounds' genes from STP and SEA (c) T2DM and *M. balsamina* compounds

signaling pathways to bring about glucose homeostasis and prevent diabetic complications [40], while myricetin reportedly delayed cellular uptake of carbohydrates and enhance insulin secretion via the GLP-1 receptor agonist [41]. The presence of these compounds in *M. balsamina* supports its potential as an antidiabetic agent.

3.3 Gene Ontology Analysis of the Common Genes between *M. balsamina* Compounds and T2DM

The GO analysis presented a total of 188 biological processes (BP), 51 molecular functions (MF), and 23 cellular components (CC). A bar chart of the top 10 functional

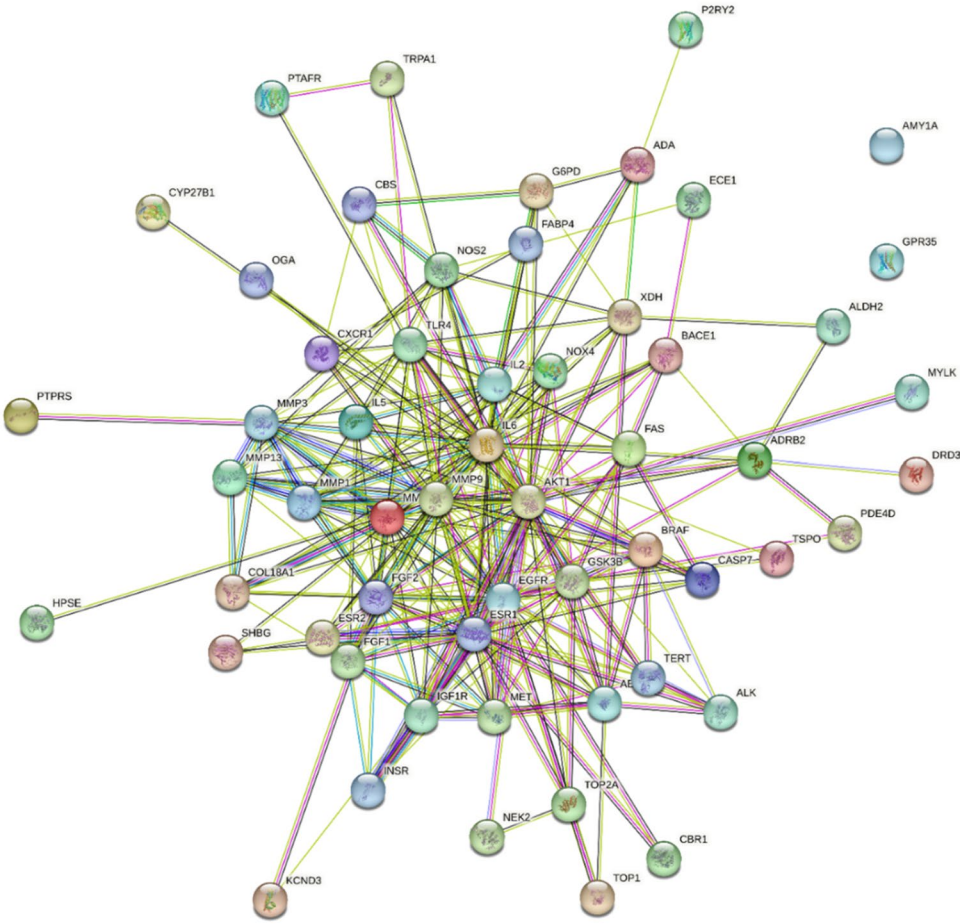


Fig. 3 PPI between the 57 common genes of T2DM and *M. balsamina* compounds

annotations of the common genes between *M. balsamina* compounds and T2DM is shown in Fig. 4. Positive regulation of protein phosphorylation (P value = 9.5×10^{-9}), receptor complex (P value = 1.7×10^{-6}), and endopeptidase activity (P value = 7.5×10^{-8}), formed the top BP, MF, and CC, respectively (Fig. 4). Notably, phosphorylation of signaling pathways either induces or prevents pathological conditions, for example, the phosphorylation of tyrosine and serine residues of insulin receptor substrate 1 (IRS1) stimulates insulin responses positively [42]. Thus, the finding that positive regulation of protein phosphorylation was the most significant BP could suggest that in elevated blood glucose levels, *M. balsamina* compounds could stimulate protein kinases necessary for the phosphorylation of insulin signaling pathways resulting in insulin sensitivity. The finding that extracellular matrix is the second top CC and the endopeptidase activity as the top MF denotes the potential involvement of *M. balsamina* compounds in diabetic foot ulcer (DFU), management since extracellular matrix enzymes such as matrix metalloproteinase-2 (MMP-2) are responsible for degrading and replacing damaged fibrils during wound healing processes while endopeptidase activates MMP-2 [43–45]. Similarly, the observation that MAP kinase, G-protein coupled receptors, and receptor complex reportedly implicated in insulin sensitivity and glucose homeostasis maintenance [46] were among the processes annotated in the GO analysis is indicative of *M. balsamina* compounds potential to enhance insulin sensitivity and maintain glucose homeostasis. Furthermore, other annotated processes indicate that *M. balsamina* could prevent pancreatic B-cell death via prevention of insulin resistance and mitochondria dysfunction and enhancement of glucose transport across the cell [47, 48].

3.4 KEGG Analysis of the Common Genes between *M. balsamina* Compounds and T2DM

The KEGG analysis identified 104 signaling pathways between the common genes of *M. balsamina* compounds and T2DM at a P-value < 0.05 (Supplementary Table 2), with the EGFR tyrosine kinase inhibitor resistance and endocrine resistance being the first and second most significant pathways, respectively (Table 2; Fig. 5). The bubble plot of the top 10 signaling pathway related to T2DM identified in the KEGG enrichment analysis is shown in Fig. 6. The EGFR, a transmembrane receptor

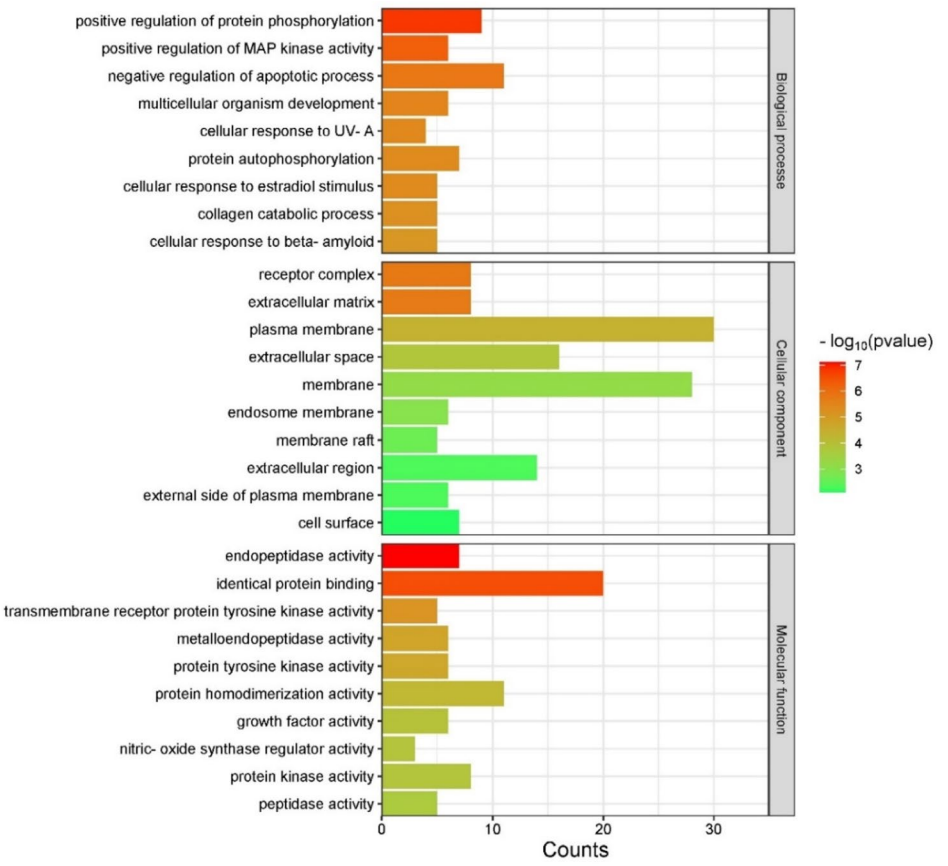


Fig. 4 GO bar plot of the top 10 functional parameters for the common genes between *M. balsamina* compounds and T2DM

protein belonging to the ErbB/HER family, enhances tyrosine kinase receptor activity upon activation via the binding of specific growth factors and small molecules and phosphorylates intracellular tyrosine residues [49]. Mostly expressed in human tissues, the EGFR is implicated in several pathological processes such as cell proliferation, and differentiation, chronic kidney disease, diabetes, and cancer [50]. With respect to diabetes, studies have revealed that EGFR is a promising target for insulin resistance and diabetic kidney disease treatment, with animal studies showing that the inhibition of EGFR activity can help maintain normal blood glucose level and body weight in diabetic rats [51–53].

3.5 Network Analysis of the EGFR Pathway, Genes, and *M. balsamina* Compounds

The interaction of the enriched genes of the most significant pathway presented 8 gene nodes, 26 edges, an average node degree of 6.50, average local clustering coefficient of 0.94, and an enrichment P-value of 1.14×10^{-3} (Fig. 6a). Further interaction analysis and layout of the PCT gave a network where EGFR tyrosine kinase inhibitor resistance interacted with 51 compounds of *M. balsamina* (Fig. 6b). The observation that the EGFR was the significant pathway targeted by *M. balsamina* compounds is in line with the computational study of Romero [54], who identified EGFR has a hub target and its potential inhibitors for diabetes and cancer management. The hub genes of the EGFR pathway identified in this study are presented in Fig. 7. Furthermore, the observation that the endocrine resistance pathway was the second most significant pathway suggests *M. balsamina* compounds' involvement with endocrine genes implicated in insulin resistance or T2DM.

Analysis of the degree of interaction of the enriched genes in the EGFR pathway showed that *IGF1R* (22) and *GSK3B* (19) interacted with the highest number of *M. balsamina* compounds (Fig. 8) and were hence used further in the study. The tyrosine-protein kinase receptor, *IGFIR*, found in mitogen acting cells induces EGFR-TKI resistance. Combined inhibitors of EGFR-TKI resistance and *IGF1R* reportedly improve antidiabetic activity [55], thus supporting *IGFIR* targeting in T2DM therapy. Glycogen synthase kinase (GSK) of protein kinase family affects glucose regulation and insulin signaling by phosphorylating the SER/THR residues in diabetes signaling pathways. The phosphorylation and dephosphorylation of the SER and TYR residues of the *GSK3α* and β subunits

Table 2 Signaling pathways and genes related to *M. balsamina* compounds and T2DM

S/*N*	Pathway ID	Pathway name	False discovery rate	Genes
1	hsa01521	EGFR tyrosine kinase inhibitor resistance	1.31×10^{-08}	*FGF2*, *EGFR*, *MET*, *GSK3B*, *IL6*, *BRAF*, *AKT1*, *IGF1R*
2	hsa01522	Endocrine resistance	4.30×10^{-08}	*MMP2*, *EGFR*, *ESR2*, *MMP9*, *ESR1*, *BRAF*, *AKT1*, *IGF1R*
3	hsa04151	PI3K-Akt signaling pathway	2.64×10^{-07}	*IL2*, *FGF2*, *EGFR*, *INSR*, *MET*, *GSK3B*, *TLR4*, *IL6*, *AKT1*, *FGF1*, *IGF1R*
4	hsa04657	IL-17 signaling pathway	5.39×10^{-07}	*IL5*, *MMP13*, *MMP3*, *MMP1*, *GSK3B*, *MMP9*, *IL6*
5	hsa04066	HIF-1 signaling pathway	8.24×10^{-07}	*EGFR*, *INSR*, *NOS2*, *TLR4*, *IL6*, *AKT1*, *IGF1R*
6	hsa04926	Relaxin signaling pathway	2.76×10^{-07}	*MMP2*, *MMP13*, *EGFR*, *MMP1*, *NOS2*, *MMP9*, *AKT1*
7	hsa04015	Rap1 signaling pathway	3.49×10^{-06}	*FGF2*, *EGFR*, *INSR*, *MET*, *BRAF*, *AKT1*, *FGF1*, *IGF1R*
8	hsa04010	MAPK signaling pathway	3.51×10^{-06}	*FGF2*, *EGFR*, *INSR*, *MET*, *BRAF*, *AKT1*, *FGF1*, *IGF1R*, *FAS*
9	hsa04668	TNF signaling pathway	1.90×10^{-05}	*MMP3*, *CASP7*, *MMP9*, *IL6*, *AKT1*, *FAS*
10	hsa04510	Focal adhesion	3.03×10^{-05}	*EGFR*, *MET*, *GSK3B*, *MYLK*, *BRAF*, *AKT1*, *IGF1R*
11	hsa04068	FoxO signaling pathway	3.46×10^{-05}	*EGFR*, *INSR*, *IL6*, *BRAF*, *AKT1*, *IGF1R*
12	hsa04915	Estrogen signaling pathway	4.30×10^{-05}	*MMP2*, *EGFR*, *ESR2*, *MMP9*, *ESR1*, *AKT1*
13	hsa04014	Ras signaling pathway	6.07×10^{-05}	*FGF2*, *EGFR*, *INSR*, *MET*, *AKT1*, *FGF1*, *IGF1R*
14	hsa04932	Non-alcoholic fatty liver disease	6.37×10^{-05}	*INSR*, *GSK3B*, *CASP7*, *IL6*, *AKT1*, *FAS*
15	hsa04931	Insulin resistance	2.0×10^{-04}	*INSR*, *GSK3B*, *OGA*, *IL6*, *AKT1*
16	hsa04923	Regulation of lipolysis in adipocytes	2.6×10^{-04}	*FABP4*, *INSR*, *ADRB2*, *AKT1*
17	hsa05230	Central carbon metabolism in cancer	5.5×10^{-04}	*EGFR*, *MET*, *G6PD*, *AKT1*
18	hsa04933	AGE-RAGE signaling pathway in diabetic complications	1.3×10^{-03}	*MMP2*, *NOX4*, *IL6*, *AKT1*
19	hsa04660	T cell receptor signaling pathway	1.5×10^{-03}	*IL2*, *IL5*, *GSK3B*, *AKT1*
20	hsa04020	Calcium signaling pathway	1.6×10^{-03}	*EGFR*, *ADRB2*, *NOS2*, *MYLK*, *PTAFR*

results in alternating activation and inactivation of the gene [56, 57]. Moreso, inhibition of GSK is crucial in glycogen metabolism regulation [58]. The finding that *M. balsamina* compounds interacted with *IGF1R* and *GSK3B* genes indicates their downregulatory potential on the genes' activities towards T2DM management, which could probably be via insulin signaling and sensitivity and glycogen metabolism regulation.

3.6 Molecular Docking and Dynamics Simulation of *M. balsamina* Compounds with EGFR Enriched Genes

The docking scores and the type of interactions of *M. balsamina* compounds with *IGF1R* and *GSK3B* are shown on Supplementary Tables 3 and 4 while the docking scores of the top-ranked *M. balsamina* compounds and their energy component are presented in Table 3. With *IGF1R*, all the compounds exhibited higher negative docking scores than metformin (−4.5 kcal/mol) while only quercetin-3-O-glucoside (−8.9 kcal/mol) had higher negative score than Nvp-aew541 (−8.5 kcal/mol) (Table 3). Similarly, for *GSK3B*, only quercetin-3-O-glucoside (−9.4 kcal/mol) had higher negative docking score than AZD-1080 (−8.7 kcal/mol) while metformin (−4.5 kcal/mol) had the least negative docking scores compared to all the compounds of *M. balsamina* (Table 3).

Since docking score is a preliminary binding interaction analysis [6], MD simulation was performed to gain better insights into the binding stability and compactness of the resulting complex in each case [59]. Nvp-aew541 (−28.20 ± 7.72) and metformin (−11.66 ± 6.65) exhibited low negative binding free energies (ΔG_{bind}) than the compounds,

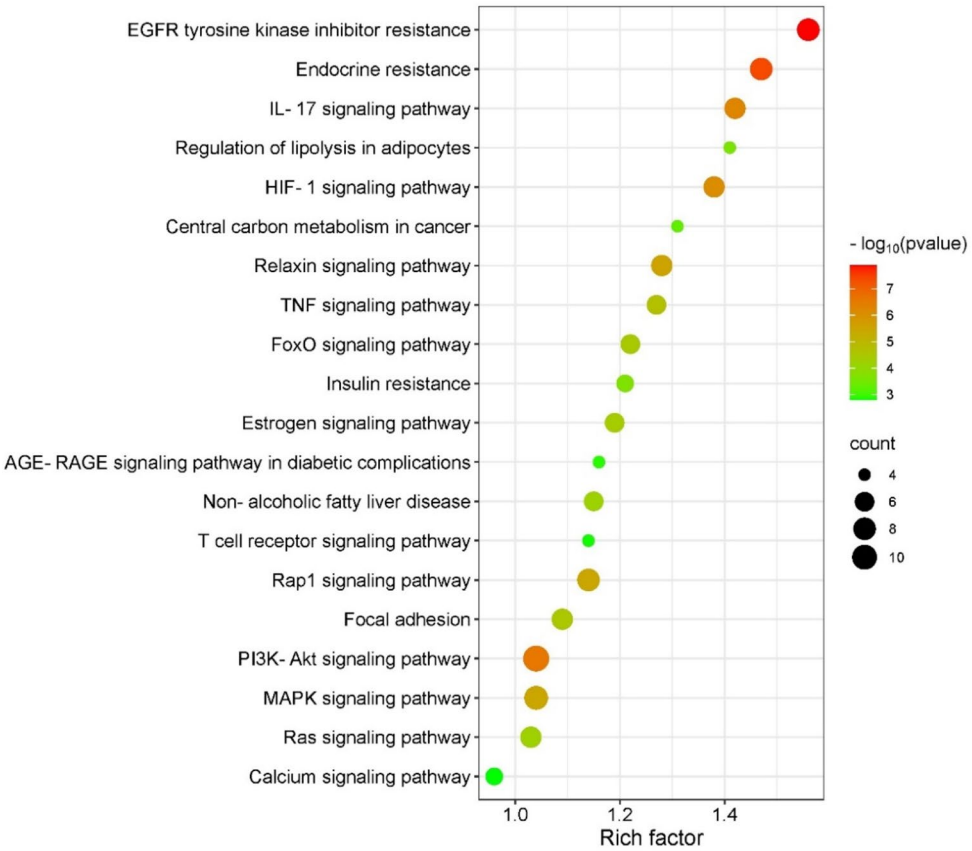


Fig. 5 KEGG bubble plot of the top 20 signaling pathways associated with the *M. balsamina* compounds and T2DM

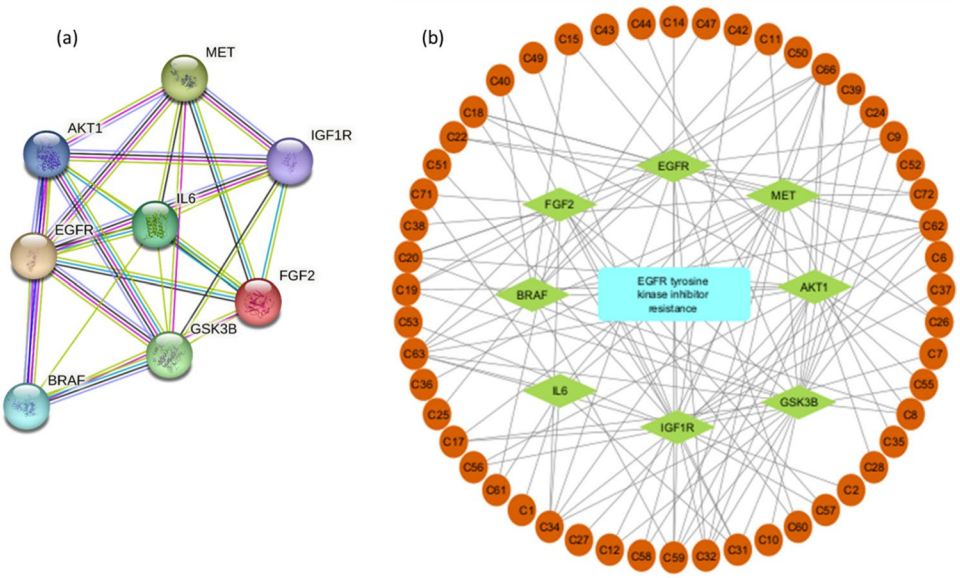


Fig. 6 EGFR tyrosine kinase inhibitor **a** gene-gene interaction and **b** PCT layout with *M. balsamina* compounds

with kaempferol-3-O-glucoside (−51.50 ± 7.97) having the highest negative ΔG_{bind} despite not having the best docking score (Table 3). This connotes the superior binding affinity potential of kaempferol-3-O-glucoside over the other investigated compounds. Despite balsaminoside B, C, and cucurbalsaminol A having lower negative docking scores than quercetin-3-O-glucoside (−43.63 ± 6.58 kcal/mol), they displayed higher negative ΔG_{bind} values of −61.79 ± 7.35, −59.87 ± 10.47, −59.53 ± 7.54 kcal/mol, respectively (Table 3), suggestive of their better binding affinities and modulatory potential against the *GSK3B* in the management of T2DM.

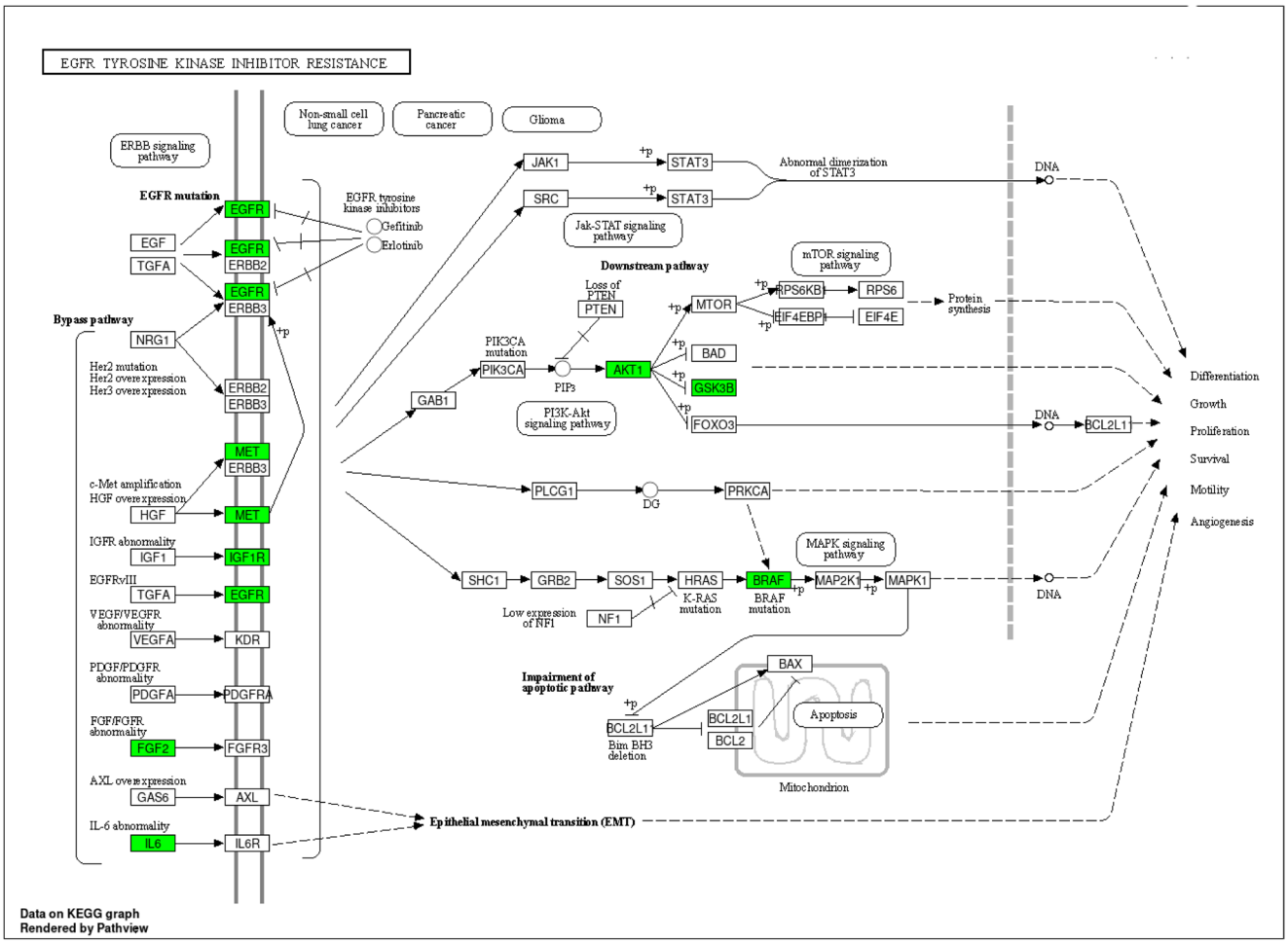


Fig. 7 Enriched genes (Green) of EGFR-TKI signaling pathway identified in T2DM in this study

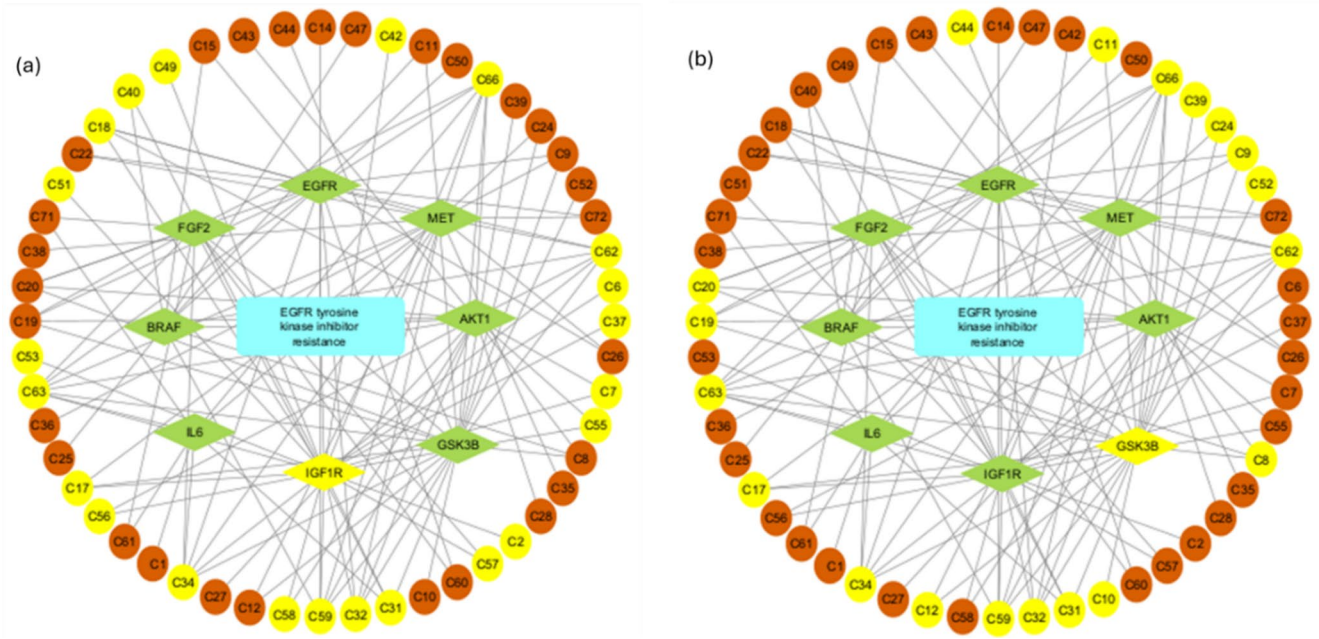


Fig. 8 Network construction of *M. balsamina* compounds (yellow) with genes (yellow) related to the EGFR tyrosine kinase inhibitor resistance signaling pathway (turquoise blue) **a** 22 compounds interacted with *IGF1R***b** 19 compounds interacted *GSK3B*

Table 3 Docking scores and energy components of top-ranked compounds of *M. balsamina* with *IGF1R* and *GSK3B* genes

Complexes	kcal/mol						
	Docking scores	ΔE_{vdW}	ΔE_{elec}	ΔG_{gas}	ΔG_{solv}	ΔE_{sur}	ΔG_{bind}
IGF1R							
Nvp-aew541	−8.5	−31.63±6.86	−208.31±21.69	−239.94±25.99	211.74±21.08	−4.08±0.85	−28.20±7.72
Metformin	−4.5	−5.83±3.95	−515.35±93.82	−521.19±95.29	509.52±89.84	−1.76±0.68	−11.66±6.65
3-caffeoylquinic acid	−8.3	−42.45±3.08	−24.74±6.31	−67.18±5.78	31.15±3.61	−6.02±0.20	−36.04±3.65
kaempferol-3-O-glucoside	−7.9	−48.88±6.58	−51.25±11.96	−100.13±10.27	48.63±9.38	−6.75±0.27	−51.50±7.97
Melilotoside	−7.9	−35.52±4.24	−38.28±17.41	−73.79±15.35	34.81±8.71	−5.42±0.26	−38.98±7.62
Quercetin 3-galactoside	−8.2	−44.32±5.25	−51.21±12.61	−95.53±10.52	51.06±8.08	−6.65±0.29	−44.47±5.62
Quercetin-3-O-glucoside	−8.9	−44.93±5.91	−53.53±15.39	−98.46±13.08	48.74±8.63	−6.84±0.47	−49.72±6.36
GSK3B							
AZD-1080	−8.7	−35.39±3.37	147.45±33.24	112.06±32.99	−147.42±29.95	−4.83±0.36	−35.36±6.21
Metformin	−5.0	−2.16±3.62	49.68±78.78	47.51±80.43	−49.69±78.07	−0.53±0.77	−2.18±4.23
Balsaminoside B	−8.4	−56.89±5.41	−31.47±10.57	−95.19±10.85	33.39±6.38	−7.78±0.55	−61.79±7.35
Balsaminoside C	−8.5	−50.76±5.09	−41.90±14.55	−99.96±16.08	40.09±8.41	−7.24±0.63	−59.87±10.47
Cucurbalsaminol A	−8.5	−61.56±6.19	−9.60±4.41	−80.84±9.37	21.31±3.72	−8.32±0.90	−59.53±7.54
Myricetin	−8.2	−35.58±3.95	−30.49±10.92	−66.07±9.71	35.53±7.21	−5.16±0.28	−30.54±3.79
Quercetin 3-galactoside	−8.3	−47.39±3.36	−27.54±8.93	−74.93±8.96	36.45±6.38	−6.62±0.28	−38.48±4.21
Quercetin-3-O-glucoside	−9.4	−44.49±4.90	−57.82±23.61	−102.31±20.64	58.68±15.79	−6.53±0.34	−43.63±6.58

[b]ΔE_{VdW}: Vander Waals energy; ΔE_{elec}: electrostatic energy; ΔE_{gas}: gas-phase free energy; ΔG_{solv}: solvation free energy; ΔG_{bind}: total binding free energy.

3.6.1 Post-dynamic Parameters and Hydrogen Bonds of Top-ranked Compounds of *M. balsamina* with IGF1R and GSK3B Genes

The post-dynamic parameters and hydrogen bonds of top-ranked compounds of *M. balsamina* with *IGF1R* and *GSK3B* are shown in Figs. 9 and 10. While the stability of a compound-target complex in term of convergence and divergence is measured as RMSD [60], RMSF describes the flexibility in the amino acid residues present at the active site of the target in the complex [61]. With ROG and SASA, the degree of compactness/folding of the complex and

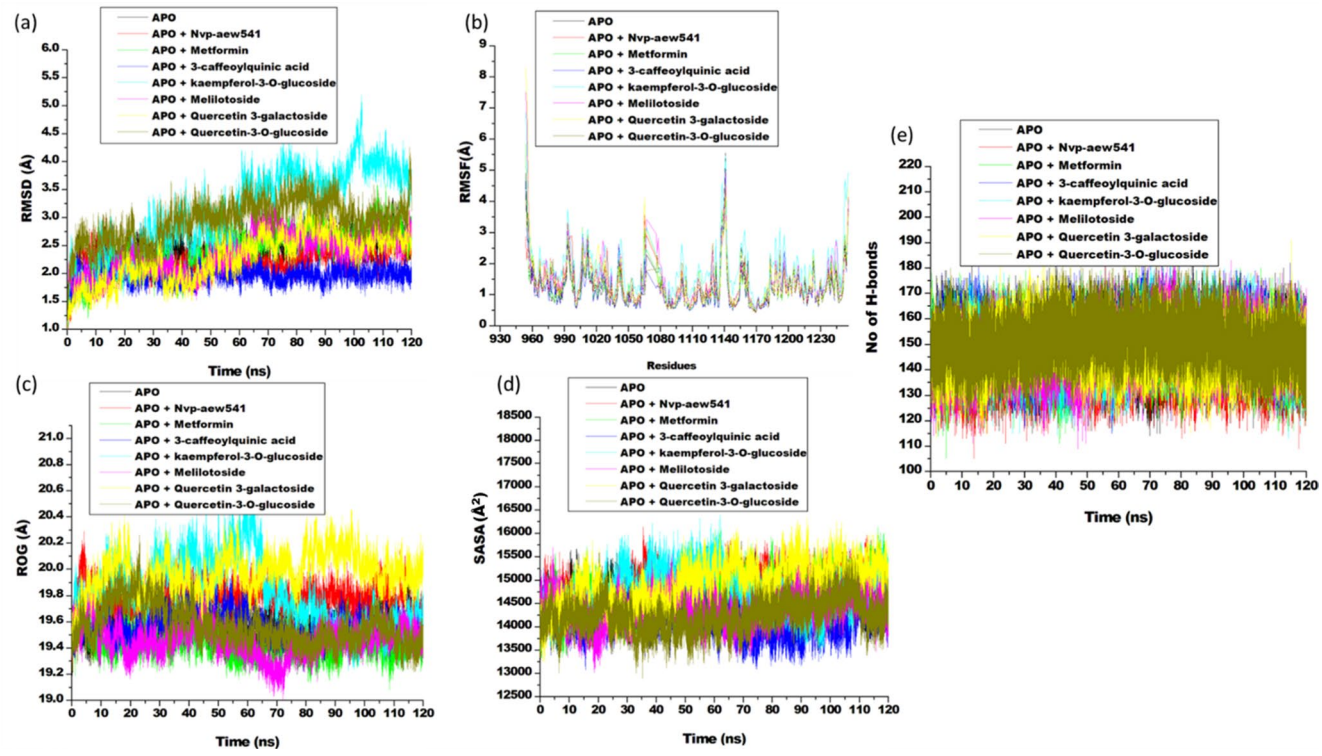


Fig. 9 Comparative post-dynamic data of the **a** RMSD, **b** RMSF, **c** ROG, **d** SASA, **e** intermolecular H-bond number of Nvp-aew541, metformin, and 3-caffeoxylquinic acid, kaempferol-3-O-glucoside, melliotoside, quercetin-3-galactoside, and quercetin-3-O-glucoside of *M. balsamina* and *IGF1R* complexes

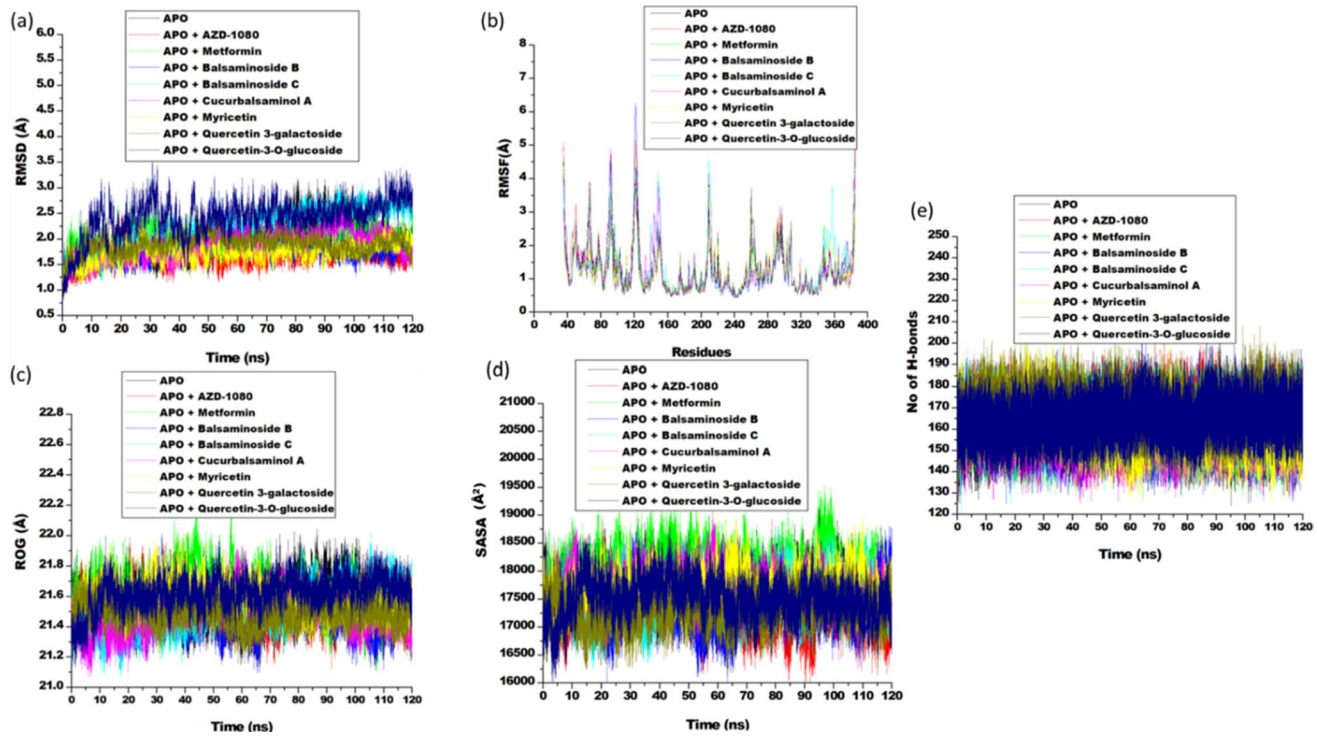


Fig. 10 Comparative post-dynamic data of the **a** RMSD, **b** RMSF, **c** ROG, **d** SASA, **e** intermolecular H-bond number of AZD-1080, metformin, and balsaminoside B, balsaminoside C, cucurbalsaminol A, myricetin, quercetin-3-galactoside, and quercetin-3-O-glucoside of *M. balsamina* and *GSK3B* complexes

interaction of the complex with solvent are inferred, respectively [62], intermolecular H-bonds number determines the stability of complexes via intermolecular attraction between the two components [26].

A stable RMSD pattern was observed in the *IGF1R* bound complex (Fig. 9a). Kaemferol-3-glucoside which presented slight fluctuation between the 100–110 ns, eventually stabilizes at the end of the 120 ns simulation. The average RMSD values were all < 3.5 Å with kaempferol-3-O-glucoside (3.17 ± 0.69 Å) and 3-caffeoylquinic acid (1.96 ± 0.15 Å) having the highest and least average RMSD values, respectively (Table 4). The reduced fluctuations in the RMSD plots throughout the simulation depict greater convergence and better stability [63, 64]. To corroborate this claim, the bound systems had average RMSD values < 3.5 Å, which is reported to be the limit for a stable complex [65]. The RMSF pattern showed several peaks which were prominent from 1050 to 1080 and 1140 amino acid residues (Fig. 9b). The average RMSF values of the complex was again highest for kaempferol-3-O-glucoside (1.85 ± 0.84) and least for 3-caffeoylquinic acid (1.21 ± 0.65) (Table 4). The higher the fluctuations of RMSF peaks, the less the compactness and stability [66]. Notably, the reduced fluctuations, particularly in the active site amino acid residues of *IGF1R* is suggestive of stable complex formation of *M. balsamina* compound with the gene. Additionally, the non-significant differences in the average RMSF values and the high ΔG_{bind} values corroborates the formation of the stable complexes.

Table 4 Average post-dynamic data and hydrogen bond number of top-ranked *M. balsamina* compounds with *IGF1R* and *GSK3B* genes

Complexes	RMSD (Å)	RMSF(Å)	ROG (Å)	SASA (Å^2)	Intermolecular H-bonds number
IGF1R					
Apo-gene	2.42±0.24	1.20±0.68	19.59±0.11	14644.50±306.57	147.52±8.11
Nvp-aew541	2.25±0.22	1.33±0.67	19.78±0.11	14751.57±329.94	142.34±7.74
Metformin	2.47±0.36	1.32±1.06	19.43±0.09	14696.94±334.35	147.55±8.23
3-caffeoylquinic acid	1.96±0.15	1.21±0.65	19.58±0.09	14194.68±316.35	150.14±8.01
kaempferol-3-O-glucoside	3.17±0.69	1.85±0.84	19.80±0.24	14759.16±416.17	149.49±8.14
Melilotoside	2.42±0.39	1.48±0.82	19.45±0.11	14386.13±320.86	147.95±8.38
Quercetin 3-galactoside	2.32±0.40	1.42±0.89	19.99±0.14	14953.59±434.15	147.86±8.21
Quercetin-3-O-glucoside	2.99±0.37	1.39±0.61	19.55±0.15	14241.55±321.74	151.05±8.35

An equilibrated ROG pattern was observed at the beginning of the simulation (Fig. 9c). However, kaempferol-3-O-glucoside which slightly deviated around the 60 ns finally equilibrated towards the end of the simulation while quercetin-3-O-galactoside showed greater fluctuation than the other compounds at the 120 ns with an average ROG value of 19.99 ± 0.14 Å (Table 4). High ROG values are directly proportional to stability [66]. Thus, quercetin-3-galactoside and kaempferol-3-O-glucoside possess greater compactness and thermodynamic organization than the other compounds. The SASA plot was stable throughout the simulation with quercetin-3-O-galactoside again displaying higher fluctuations (Fig. 9d) and higher average SASA value (14953.59 ± 434.15 Å) than the other compound-*IGFIR* complex. The other compound-bound complexes had no remarkable differences with the apo-gene (Table 4) suggestive of better stability since higher SASA value connotes reduced stability [67]. This could have also been influenced by their higher negative ΔG_{bind} values. Notably, a stable pattern was observed in the number of hydrogen (H) bonds with quercetin-3-O-glucoside displaying greater stability in the bound system (Fig. 9e). The non-significant differences in the average H-bond numbers in all the bound complexes (Table 4) depict the formation of stable complexes and indicate the presence of strong intermolecular attractions between the bound systems [68].

Regarding GSK3B, the RMSD plot shows an initial stable complex, however, balsaminoside C and quercetin-3-O-glucoside exhibited slightly greater fluctuations (Fig. 10a) than the other compounds with average RMSD values of 2.25 ± 0.38 and 2.47 ± 0.38 Å, respectively (Table 5). However, the gene inhibitor had the least RMSD value (1.66 ± 0.18 Å) depicting greater convergence and better stability. Prominent peak vibrations were observed in the RMSF plot (Fig. 10b). The vibrations were minimal between 160 and 200, 240, and 320–340 residues. Except for balsaminoside B with slightly lower mean RMSF value (1.19 ± 0.78 Å), other compounds have values not significantly different from the apo-gene and the standards (Table 5). The high peak fluctuations and the average RMSF values variations compared to the apo-gene might not significantly impact the compounds' binding and stability to the active site amino acids of the gene. The recorded ΔG_{bind} values could have also greatly influenced the complexes stability. Additionally, a stable ROG pattern was observed throughout the simulation period (Fig. 10c) with non-significant differences between the complexes formed (Table 5), suggesting better compactness and folding of the complexes. The SASA plot was stable (Fig. 10d) with metformin showing higher mean SASA value (18082.76 ± 330.73 Å) (Table 5). This observation is indicative of metformin-*GSK3B* complex to be more likely less stable, which is further supported by its lower negative docking, and ΔG_{bind} values. However, *M. balsamina*-*GSK3B* complexes have the potential to interact with biological solvent as they displayed SASA values close to the apo-genes. The intermolecular H-bond numbers plot formed a stable bound complex (Fig. 10e) with no significant differences in all the complexes (Table 5), indicative of the formation of stable complexes between the bound systems.

3.6.2 2D Interaction Plots of the Top-ranked Compounds of *M. balsamina* with IGF1R and GSK3B Genes Binding Following 120 ns Simulation

Several interacting bonds including van der Waals, alkyls, sigma, and H-bonds were observed in the 2D interaction views of the simulated *IGF1R* and *GSK3B* genes and top-ranked *M. balsamina* compounds complexes. Notably, H-bonds contribute to the formation of strong and stable complexes [68]. Different types of amino acids were observed in the snapshots of the complexes, while some were maintained throughout the simulation, new ones were seen at the end of the 120 ns simulation. With *IGF1R*, Nvp-aew541 presented 13 and 12 interactions at the 0 ns and 120 ns, respectively, and retained the 2 H-bonds (MET99; THR100) formed at the 0ns (Fig. 11a). Metformin presented

Table 5 Average post-dynamic data and hydrogen bond number of top-ranked *M. balsamina* compounds with *GSK3B* gene

Complexes	RMSD (Å)	RMSF(Å)	ROG (Å)	SASA (Å^2)	Intermolecular H-bond number
GSK3B					
Apo-gene	1.94±0,32	1.33±0.75	21.54±0.14	17678.04±320.28	160.68±8.39
AZD-1080	1.66±0.18	1.24±0.78	21.51±0.11	17325.88±378.86	164.94±8.49
Metformin	1.99±0.21	1.35±1.88	21.61±0.14	18082.76±330.73	161.60±8.25
Balsaminoside B	1.82±0.22	1.19±0.78	21.48±0.10	17427.93±361.96	160.93±8.35
Balsaminoside C	2.25±0.38	1.37±0.81	21.54±0.14	17568.83±345.43	162.52±8.32
Cucurbalsaminol A	1.86±0.27	1.33±0.83	21.48±0.12	17652.7±304.41	159.71±8.33
Myricetin	1.74±0.19	1.27±0.73	21.59±0.10	17672.14±324.26	162.04±8.99
Quercetin 3-galactoside	1.89±0.18	1.20±0.69	21.49±0.08	17301.31±291.65	168.30±8.46
Quercetin-3-O-glucoside	2.47±0.38	1.27±0.70	21.62±0.10	17490.02±331.87	164.69±8.61

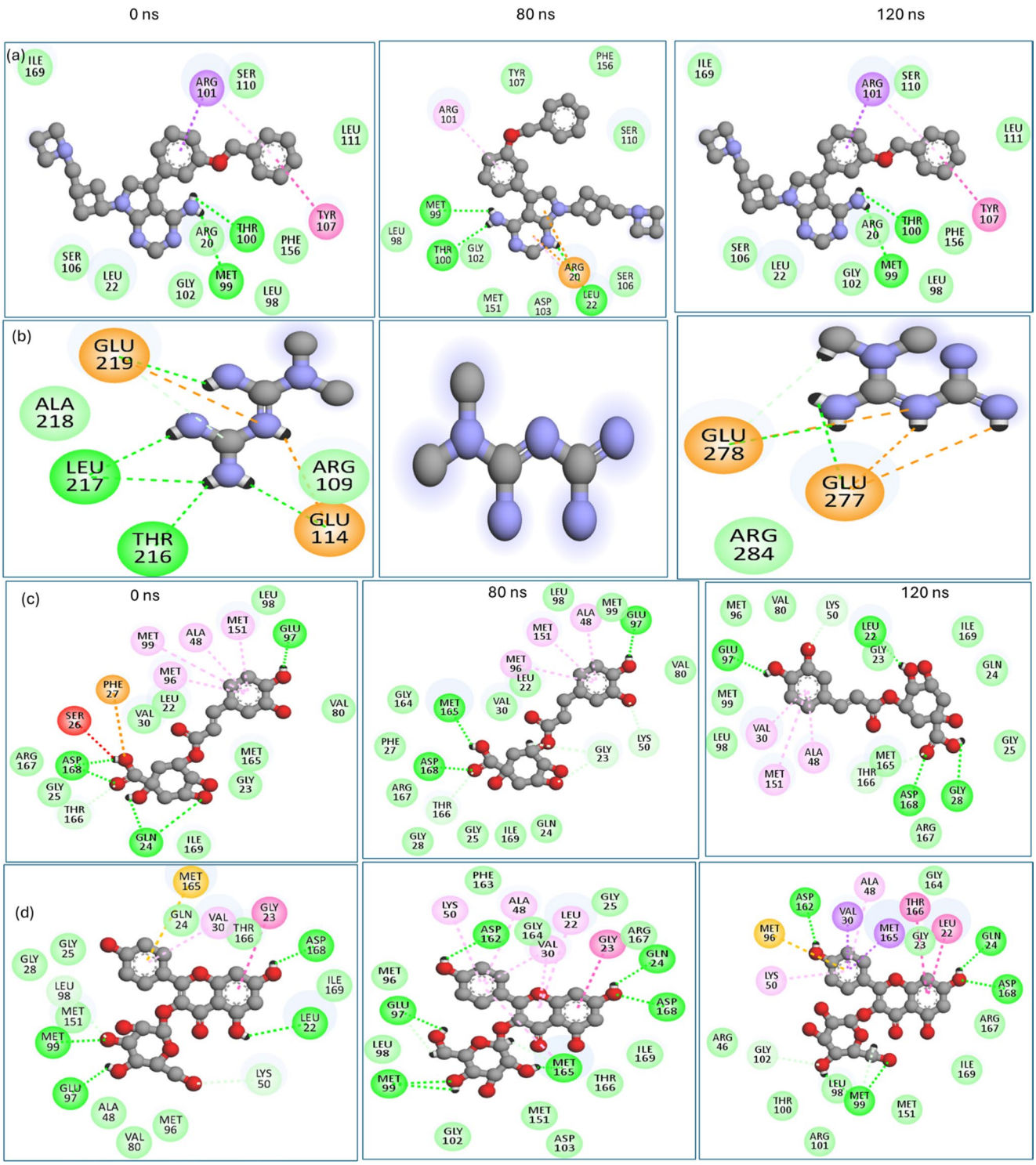


Fig. 11 2D interaction plots of *IGF1R* with **a** Nvp-aew541, **b** metformin, **c** 3-caffeoylquinic acid, **d** kaempferol-3-O-glucoside, **e** melilotoside, **f** quercetin 3-galactoside and **g** quercetin-3-O-glucoside over 120 ns simulation

fewer interactions, 6 bonds (comprising 2 H-bonds: LEU217; THR216) at the 0 ns and 3 bonds at 120 ns with no H-bonds (Fig. 11b). 3-caffeoylquinic acid showed a total of 19 interactions comprising 4 H-bonds (GLU97; ASP168; GLY28; LEU22) at the end of the simulation (Fig. 11c). Kaempferol-3-O-glucoside had more interactions with a total of 21 interactions including 4 H-bonds at the end of the simulation (GLN24; ASP162,168; GLY28; MET99) (Fig. 11d) while melilotoside showed 18 and 20 interactions at the 0 and 120 ns and maintained the same number of

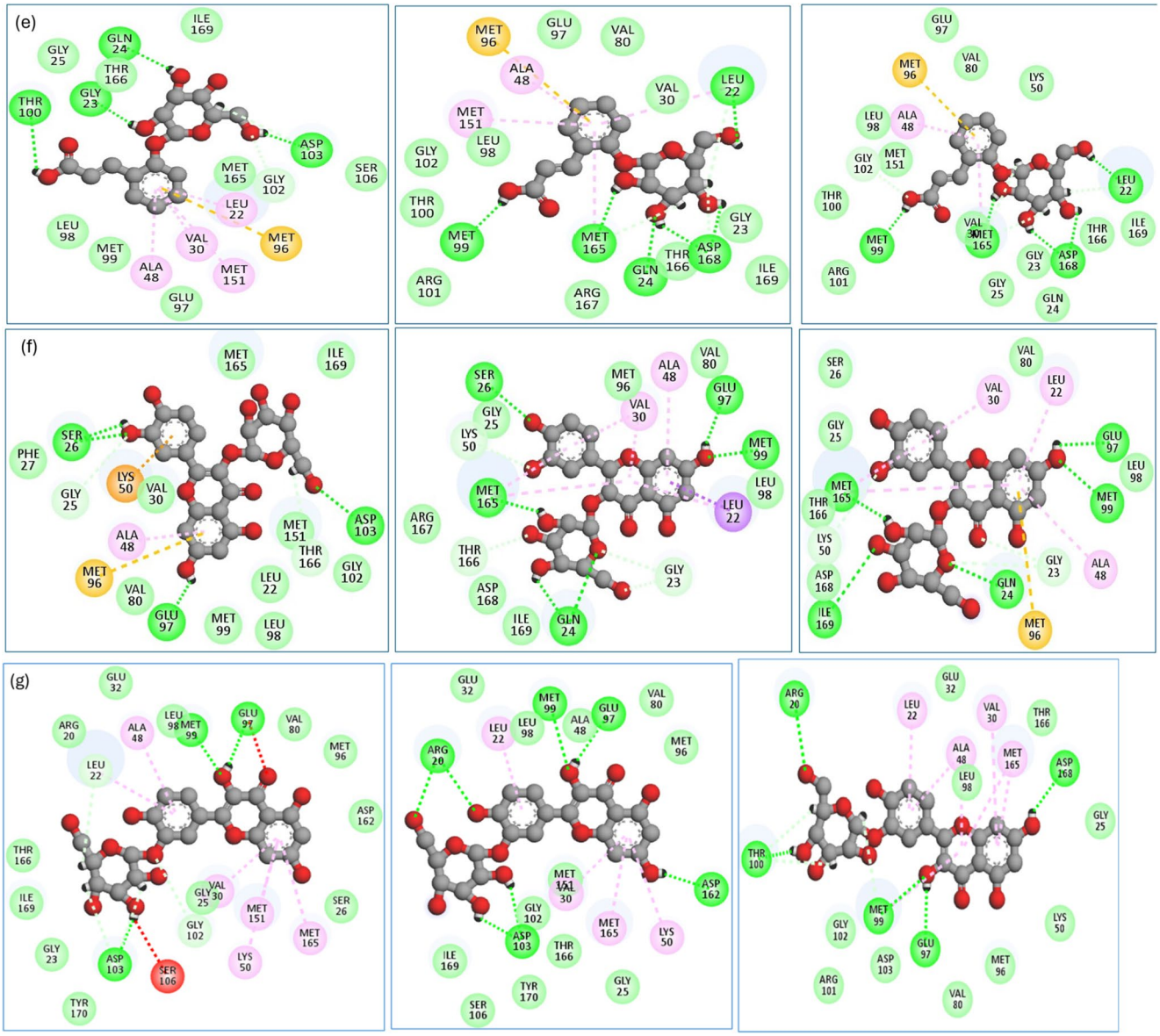


Fig. 11 (continued)

H-bonds (4) (Fig. 11e). Quercetin 3-galactoside displayed lesser interaction at the end (17) than at the beginning (18) of the simulation. However, it showed more H-bond interactions (MET165; ILE169; GLN24; MET99; GLU97) (Fig. 11f). 23 interactions were observed in quercetin-3-O-glucoside complex at the 0 ns and were reduced to 19 at 120 ns simulation with 5 H-bonds (ARG20; THR100; MET99; GLU97; ASP168) (Fig. 11g).

With *GSK3B*, 17 interactions observed at the 0 ns were reduced to 13 interactions comprising 3 H-bonds (VAL101; GLN38; ASP99) at 120 ns simulation (Fig. 12a). Metformin showed poor interaction with only 7 interactions and no H-bond at the end of the simulation (Fig. 12b). The balsaminoside B-*GSK3B* complex showed 23 and 18 interactions at 0 and 120 ns, respectively, with increased H-bond interaction at 120 ns (Fig. 12c). Balsaminoside C-*GSK3B* complex, which presented 21 interactions at the 0 ns had 18 interactions including 1 H-bond (GLU103) at 120 ns (Fig. 12d). The existing 21 interactions between cucurbalsaminol A and *GSK3B* were maintained at the end of the simulation producing 1 H-bond (TYR106) interaction (Fig. 12e). With myricetin, 15 interactions existed, the interaction increased to 20 and maintained the same number of H-bonds (4) (Fig. 12f). Quercetin 3-galactoside had 17 interactions with 2 H-bonds (THR104; VAL101) at the 0 ns and 20 interactions comprising 3 H-bonds (VAL101;

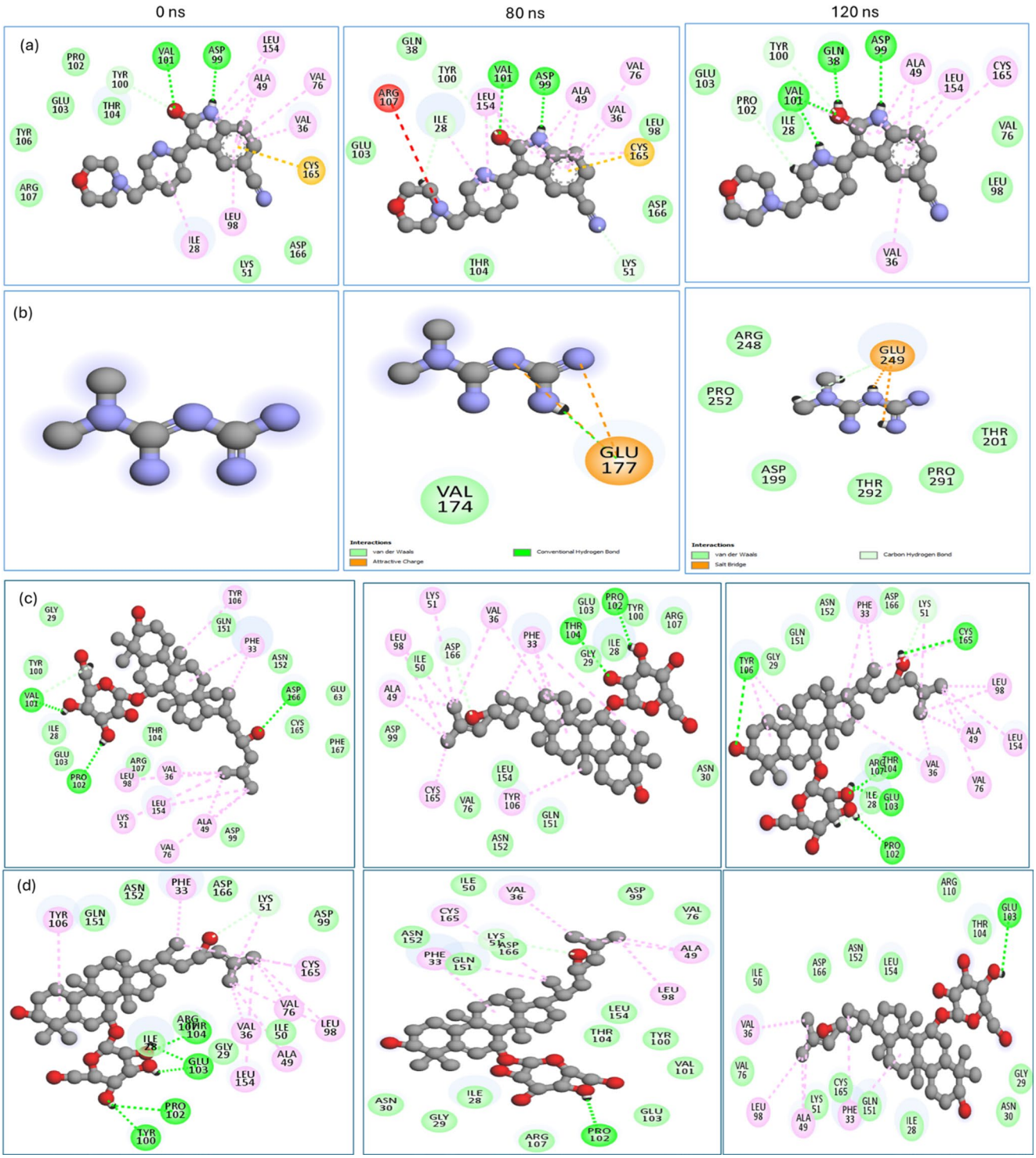


Fig. 12 2D interactions of *GSK3B* with **a** AZD-1080, **b** metformin, **c** balsaminoside B, **d** balsaminoside C, **e** cucurbalsaminol A, **f** myricetin **g** quercetin 3-galactoside, and **h** quercetin-3-O-glucoside over 120 ns simulation

PRO102; TYR37) at the 120 ns (Fig. 12g), while quercetin-3-O-glucoside had 20 interactions at 0 ns, and presented more interactions and H-bonds (GLY34; ASN152; ASP99,166; LYS51; CYS165; VAL101) at 120 ns simulation (Fig. 12h).

The finding that the *M. balsamina* compounds exhibited greater interactions and H-bonds numbers than the reference standards signify their potential to interact with the amino acid at the active site of the targets and stable inhibitory

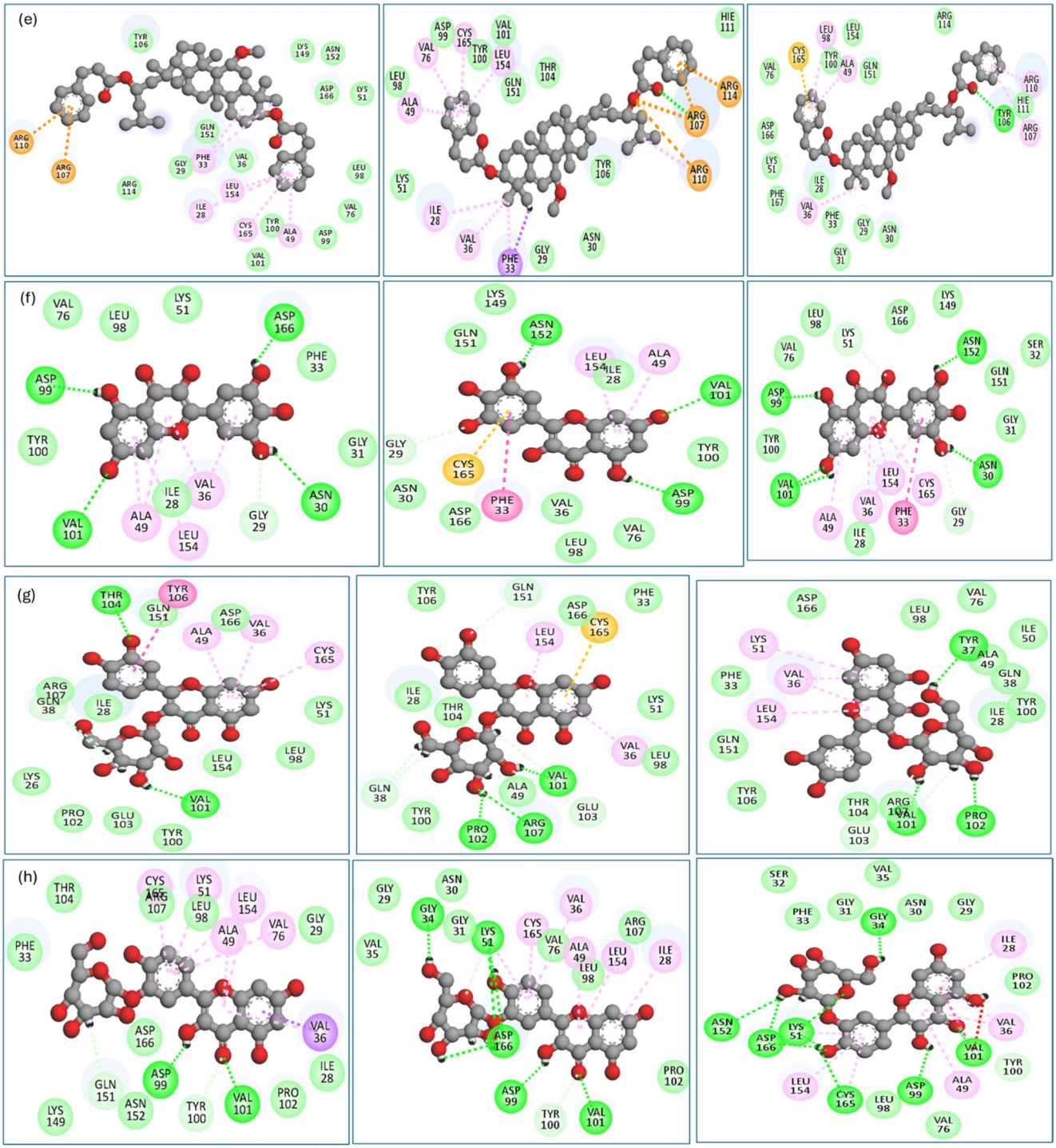


Fig. 12 (continued)

effects on the gene. This is in line with the observation of Abdomaleki et al. [69] who also reported better interaction in profiled compounds than the standard. Overall, the high number of H-bonds, stability, number of interactions, and compactness observed in the top-ranked *M. balsamina* compounds with the genes could be responsible for the higher binding free energies recorded.

3.7 The Conceptual DFT Analysis of Top Ranked *M. balsamina* Compounds Post Simulation

The DFT analysis depicts the chemical properties and potential therapeutic effects of a compound where properties such as the energy gap derived from HOMO and LUMO are critical in establishing the potential chemical reactivity of the

compound [70]. The lower the energy gap, the lower the stability and hence higher possibility of interaction of the compound with a target receptor [33, 34]. The conceptual DFT parameters for the top ranked *M. balsamina* compounds are displayed in Table 6. For *IGF1R* top interacting compounds, LUMO values ranging from − 2.86375 to −0.86968 eV and HOMO values ranging from − 6.96426 to −5.62627 eV were recorded. Kaemferol-3-glucoside and melilotoside presented the highest negative LUMO and HOMO values, respectively, while metformin and Nvp-aew541 presented the lowest negative LUMO and HOMO values, respectively relative to the compounds (Table 6; Fig. 13). Consequently, the standards had higher energy gaps, hardness, and chemical potential while displaying lower electron affinity, ionization energy, electronegativity, global electrophilicity, and softness than the compounds (Table 6; Fig. 13).

With regards *GSK3B*, the compounds exhibited LUMO values between − 0.47566 and − 2.60932 eV and HOMO values between − 5.66845 and − 6.55636 eV. Except for myricetin (3.91031 eV), quercetin 3-galactoside (3.86378 eV), and quercetin-3-O-glucoside (4.09290 eV), all the top ranked compounds had higher energy gaps than the standards

Table 6 cDFT parameters of top-ranked compounds against *IGF1R* and *GSK3B*

Compound	cDFT parameters (eV)									
	LUMO	HOMO	EG	EA	IE	HN	SN	EN	CP	GE
IGF1R										
Nvp-aew541	−0.92138	−5.62627	4.70489	0.92139	5.62627	2.35244	0.42509	3.27383	−3.27383	2.27805
Metformin	−0.86968	−6.33513	5.46545	0.86968	6.335133	2.73273	0.36594	3.60240	−3.60241	2.37443
3-caffeoylquinic acid	−2.43952	−6.46112	4.02160	2.43952	6.461122	2.01080	0.49731	4.45032	−4.45032	4.92474
Kaempferol-3-O-glucoside	−2.86375	−6.73460	3.87085	2.86375	6.73460	1.93543	0.51668	4.79917	−4.79917	5.95013
Melilotoside	−2.60007	−6.96426	4.36420	2.60007	6.96427	2.18210	0.45828	4.78217	−4.78217	5.24017
Quercetin 3-galactoside	−2.60932	−6.47310	3.86378	2.60932	6.47310	1.93189	0.51763	4.54121	−4.54121	5.33741
Quercetin-3-O-glucoside	−1.96005	−6.05295	4.09290	1.96005	6.05295	2.04645	0.48865	4.00650	−4.00650	3.92193
GSK3B										
AZD-1080	−1.77882	−6.40153	4.62271	1.77882	6.40153	2.31135	0.43265	4.09018	−4.09018	3.61899
Metformin	−0.47566	−5.66845	5.19279	0.47566	5.66845	2.59640	0.38515	3.07205	−3.07205	1.81743
Balsaminoside B	−0.84927	−6.50602	5.65675	0.84927	6.50602	2.82837	0.35356	3.67748	−3.67765	2.39097
Balsaminoside C	−0.75839	−6.43799	5.67961	0.75839	6.43799	2.83980	0.35214	3.59819	−3.59819	2.27956
Cucurbalsaminol A	−0.62995	−6.55636	5.92641	0.62995	6.55636	2.96321	0.33747	3.59316	−3.59316	2.17851
Myricetin	−2.49231	−6.40262	3.91031	2.49231	6.40262	1.95515	0.51147	4.44746	−4.44746	5.05841
Quercetin 3-galactoside	−2.60932	−6.47310	3.86378	2.60932	6.47310	1.93189	0.51763	4.541208	−4.54121	5.33741
Quercetin-3-O-glucoside	−1.96005	−6.05295	4.09290	1.96005	6.05295	2.04645	0.48865	4.00650	−4.00650	3.92193

[c]*LUMO* Lowest unoccupied molecular orbital, *HUMO* Highest occupied molecular orbital, *EG* Energy gap, *EA* Electron affinity, *IE* Ionization energy, *HN* Hardness, *SN* Softness, *EN* Electronegativity, *CP* Chemical potential, *GE* Global electrophilicity

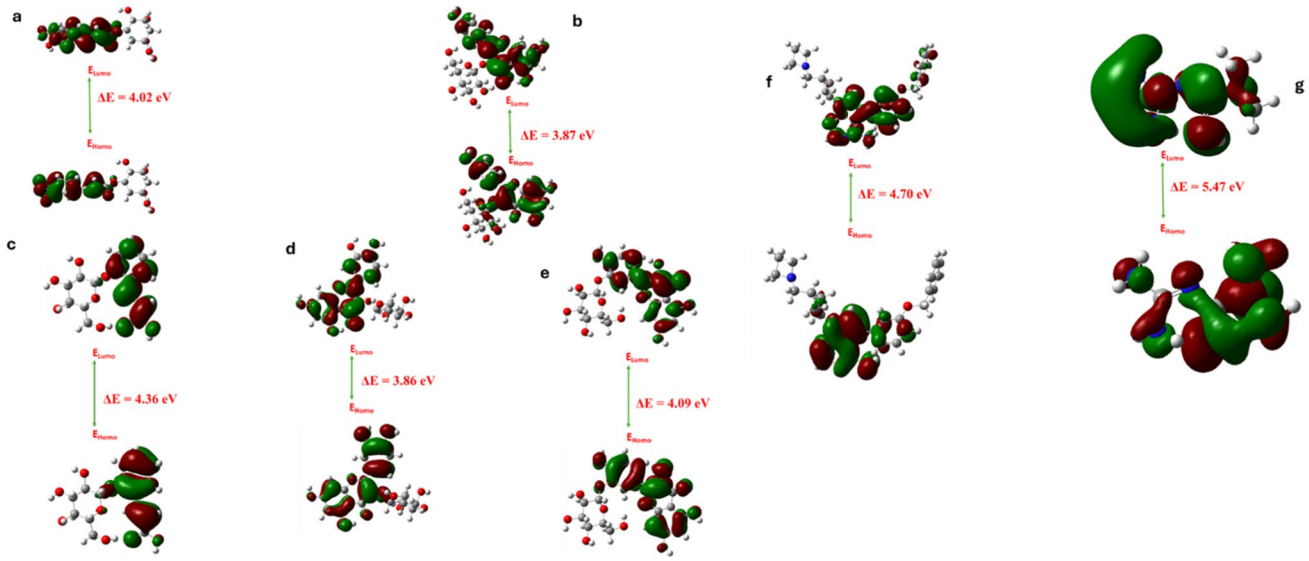


Fig. 13 Frontier molecular orbitals of top ranked *M. balsamina* compounds [**a** 3-caffeoylquinic acid **b** kaempferol-3-o-glucoside **c** melilotoside **d** quercetin 3-galactoside **e** quercetin-3-o-glucoside], **f** gene antagonist (NVP-aew541), and **g** reference standard (metformin) that interacted with *IGF1R*

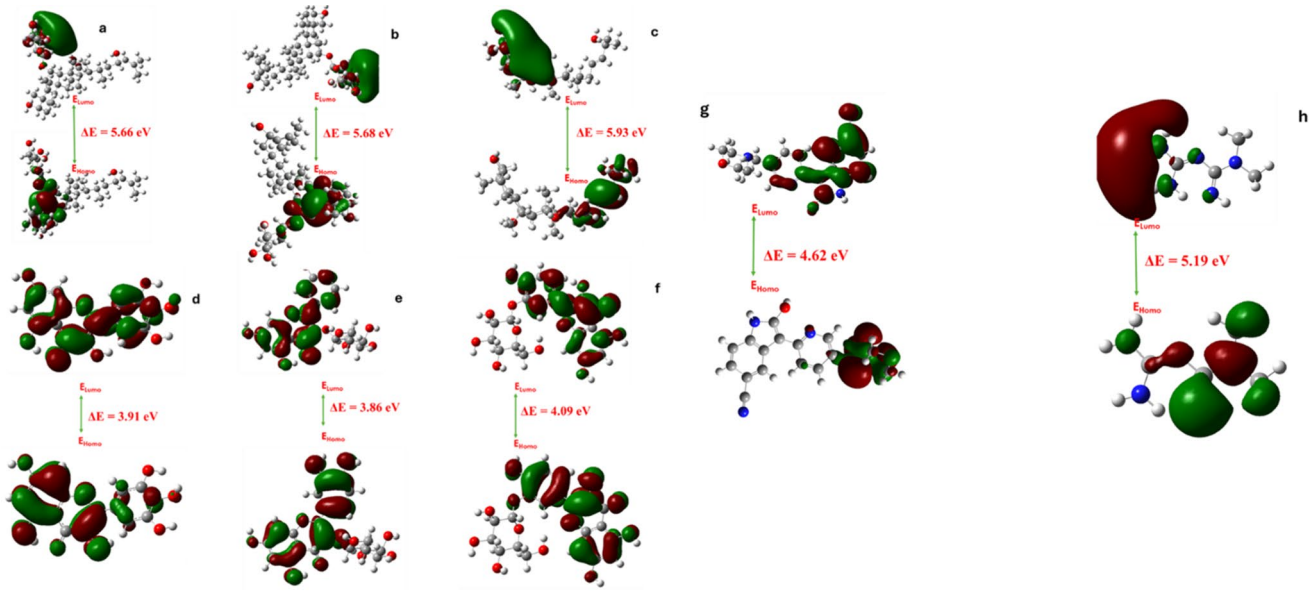


Fig. 14 Frontier molecular orbitals of top ranked *M. balsamina* compounds **a** balsaminoside B **b** balsaminoside C **c** cucurbalsaminol A **d** myricetin **e** quercetin 3-galactoside **f** quercetin-3-O-glucoside, **g** gene antagonist (AZD-1080), and (f) reference standard (metformin) that interacted with *GSK3B*

(Table 6; Fig. 14). Similarly, these three compounds displayed higher electron affinity, softness, global electrophilicity, and electronegativity and lower hardness and chemical potential relative to other compounds and standards (Table 6). However, among the top ranked compounds for both genes, quercetin 3-galactoside had the least energy gap. The observation that the standards exhibited higher energy gaps, hardness, and lower values for some of the parameters analyzed relative to the top ranked compounds indicate lesser reactivity. Meanwhile, higher reactivity potential of the compounds could have further enhanced the binding affinities for the genes. Myricetin, quercetin 3-galactoside, and quercetin-3-O-glucoside which exhibited lower energy gaps could produce better stability and binding potential with *GSK3B* relative to the other compounds [31]. Except for AZD-108, all the top compounds had higher electronegativities and electrophilicities, indicative of their high probability to attract or donate electrons, respectively. However, the observation that quercetin 3-galactoside which demonstrated better chemical properties had lower negative ΔG_{bind} might reduce its potential to serve dual purpose in downregulating the expression of both genes relative to most of the top-ranked compounds.

3.8 ADMET Properties of *M. balsamina* Compounds

The elimination of toxic and non-druggable compounds has been easier with the advent of computation study [71]. The ADMET properties of the top-ranked compounds of *M. balsamina* are presented in Tables 7 and 8. Except for cucurbalsaminol A, the profiled *M. balsamina* compounds had low

Table 7 ADME properties of top-ranked compounds of *M. balsamina*

Compounds	GIA	BBB-*P*	Pgp-S	CYP1A2-IB	CYP2C19-IB	CYP2C9-IB	CYP2D6 -IB	CYP3A4-IB
3-caffeoylquinic acid	Low	No	No	No	No	No	No	No
Kaempferol-3-O-glucoside	Low	No	No	No	No	No	No	No
Melilotoside	Low	No	No	No	No	No	No	No
Quercetin 3-galactoside	Low	No	No	No	No	No	No	No
Quercetin 3-glucoside	Low	No	No	No	No	No	No	No
Balsaminoside B	Low	No	Yes	No	No	No	No	No
Balsaminoside C	Low	No	Yes	No	No	No	No	No
Cucurbalsaminol A	High	No	Yes	No	No	No	No	No
Myricetin	Low	No	No	Yes	No	No	No	Yes
Metformin	High	No	No	No	No	No	No	No
Nvp-aew541	High	Yes	Yes	Yes	Yes	Yes	Yes	Yes
AZD-1080	High	No	Yes	Yes	No	Yes	Yes	Yes

[d]*GIA* Gastrointestinal absorption, *BBB-P* Blood brain barrier permeability *Pgp-S*- Permeability glycoprotein substrate, *CYP* Cytochrome P450, *IB* Inhibitor

Table 8 Toxicity profile of the top-ranked compounds of *M. balsamina*

Compounds	Hepatotoxicity	Neurotoxicity	Nephrotoxicity	Respiratory toxicity	Cardiotoxicity	Carcinogenicity	Immunotoxicity	Mutagenicity	Cytotoxicity	LD50 (mg/kg)/toxicity class
3-caffeoylquinic acid	I (0.72)	I (0.89)	A (0.56)	A (0.57)	I (0.99)	I (0.68)	A (0.99)	I (0.93)	I (0.80)	5000/5
Kaempferol-3-O-glucoside	I (0.82)	I (0.88)	A (0.76)	A (0.61)	I (0.67)	I (0.85)	I (0.64)	I (0.76)	I (0.69)	5000/5
Meliotoside	I (0.82)	I (0.88)	A (0.76)	I (0.67)	A (0.50)	I (0.76)	A (0.56)	I (0.78)	I (0.87)	1500/4
Quercetin 3-galactoside	I (0.82)	I (0.88)	A (0.76)	A (0.61)	I (0.67)	I (0.85)	A (0.66)	I (0.76)	I (0.69)	5000/5
Quercetin 3-glucoside	I (0.82)	I (0.88)	A (0.76)	A (0.61)	I (0.67)	I (0.85)	A (0.66)	I (0.76)	I (0.69)	5000/5
Balsaminoside B	I (0.96)	I (0.92)	A (0.61)	A (0.51)	A (0.99)	I (0.68)	A (0.99)	I (0.87)	I (0.92)	6000/6
Balsaminoside C	I (0.96)	I (0.92)	A (0.61)	A (0.51)	A (0.99)	I (0.68)	A (0.99)	I (0.87)	I (0.92)	6000/6
Cucurbalsaminol A	I (0.79)	I (0.87)	I (0.67)	A (0.96)	A (0.56)	A (0.61)	A (0.99)	I (0.74)	I (0.86)	2000/4
Myricetin	I (0.69)	I (0.89)	A (0.62)	A (0.83)	I (0.99)	A (0.68)	I (0.86)	A(0.51)	I (0.99)	159/3
Metformin	I (0.74)	I (0.84)	I (0.79)	I (0.60)	I (0.56)	I (0.68)	I (0.99)	I (0.59)	I (0.69)	680/4
Nvp-aew541	I (0.74)	A (0.81)	I (0.74)	A (0.97)	I (0.83)	A (0.60)	A (0.82)	I (0.50)	I (0.60)	416/4
AZD-1080	I (0.77)	A (0.83)	I (0.54)	A (0.79)	I (0.84)	I (0.63)	I (0.98)	I (0.61)	I (0.60)	200/3

[c]/ Inactive *A* Active

gastrointestinal absorption while only cucurbalsaminol A, balsaminoside B and C were glycoprotein substrate permeant (Table 7). Interestingly, all the compounds were non-blood brain barrier permeant and non-cytochrome (CYP) P450 isoenzyme inhibitors (Table 7). The compounds displayed varying probabilities for organ and end-point toxicities; however, they were potentially non-hepatotoxic, non-mutagenic, and non-cytotoxic (Table 8). Apart from myricetin with a lethal dose 50 (LD_{50}) of 159 mg/kg and toxicity class of 3, other compounds exhibited $LD_{50} \geq 1500$ mg/kg and toxicity class ≥ 3 (Table 8). Virtual screening showed that the top *M. balsamina* compounds can be administered via oral route, which is the most common and convenient mode of drug administration. Equally, the compounds showed better bioavailability, biotransformation, and safety profiles than the standard drugs. Evidence from this study highlights the putative lead compounds of *M. balsamina* for possible exploration as antidiabetic drug candidates. To improve the delivery, bioavailability, and activity of the promising compounds, specific transporter proteins or nanocarriers can be employed [72–74]. Furthermore, pre-clinical studies on the degree of modulation of the hub genes by the identified putative leads can be explored.

4 Conclusion

Findings from this work show that *M. balsamina* is a hub for antidiabetic drug candidates with remarkable pharmacokinetics and low toxicity profiles. Similarly, the EGFR tyrosine kinase resistance inhibitor was the top significant pathway identified with *M. balsamina* denoting its possible antidiabetic mechanism. The *IGF1R* and *GSK3B* of EGFR tyrosine kinase resistance inhibitors, which were the most enriched due to their interactions with *M. balsamina* compounds interacted better with kaempferol-3-O-glucoside and balsaminoside B, respectively. Additionally, all the compounds exhibited remarkable chemical properties as well as significant binding free energies, stability, compactness, and thermodynamic organization with the genes than the standard drugs. Thus, *M. balsamina* compounds demonstrate potential antidiabetic activities by downregulating *IGF1R* and *GSK3B* genes for glucose homeostasis. However, further preclinical and clinical studies to validate antidiabetic potential of these compounds for drug development purposes are recommended.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s42250-025-01544-1.

Acknowledgements The authors acknowledge the Directorate of Research and Postgraduate Support, Durban University of Technology, Durban, South Africa and the Centre for High Performance Computing (CHPC), Cape Town, South Africa for providing the platform for the advanced computational analysis.

Authors' Contributions Halimat Yusuf Lukman: Formal analysis, Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing. Rukayat Abiola Abdulsalam: Formal analysis, Methodology, Validation, Writing – review & editing. Saheed Sabiu: Conceptualization, Funding acquisition, Methodology, Supervision, Software, Writing – review & editing. All authors have read and agreed to publish this version of the manuscript.

Funding Open access funding provided by Durban University of Technology. The authors thank the South African National Research Foundation (NRF) Competitive Programme for Rated Researchers Support (SRUG2204193723) for providing funding to Prof. Saheed Sabiu.

Declarations

Ethical Approval Not applicable.

Informed Consent Not applicable.

Conflict of interest The authors declare that they have no competing interest that could influence the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Surabhi, Singh BK (2018) Computer aided drug design: an overview. J Drug Deliv Ther 8(5):504–509. https://doi.org/10.22270/jddt.v8i5.1894
2. Guo W, Huang J, Wang N, Tan HY, Cheung F, Chen F, Feng Y (2019) Integrating network Pharmacology and Pharmacological evaluation for Deciphering the action mechanism of herbal formula Zuojin pill in suppressing hepatocellular carcinoma. Front Pharmacol 10:1185
3. Adnan M, Jeon B-B, Chowdhury MHU, Oh K-K, Das T, Chy MNU, Cho D-H (2022) Network pharmacology study to reveal the potentiality of a methanol extract of Caesalpinia sappan L wood against type-2 diabetes mellitus. Life 12(2):277
4. Akoonjee A, Lukman HY, Ajao AA, Uthman TO, Sabiu S (2025) A network pharmacology- and molecular dynamics simulation-based bioprospection of *Khaya grandifoliola* C. DC. for diabetes care. J Biomol Struc Dyn 43(13):7155–7174. https://doi.org/10.1080/07391102.2024.2446672
5. International Diabetes Federation Atlas (2021) IDF Diabetes Atlas 10th ed. https://diabetesatlas.org/atlas/tenth-edition/. Accessed 22 May 2023
6. Lukman HY, Aribisala JO, Akoonjee A, Sulyman AO, Wudil A, Sabiu S (2024) Modulation of dipeptidyl peptidase by Rooibos tea metabolites towards type 2 diabetes care: Evidence from molecular dynamics simulation and density functional theory. Sci Afr e02173
7. Mpofana N, Peter C, Lukman HY, Makgobole MU, Dlova NC, Gqaleni N, Hussein A, Sabiu S (2024) Mechanisms of selected Cassipourea metabolites for Melasma treatment: network pharmacology and molecular dynamics study. F1000Res 13:952. https://doi.org/10.12688/f1000research.153996.1
8. Gogoi B, Gogoi D, Gogoi N, Mahanta S, Buragohain AK (2022) Network pharmacology based high throughput screening for identification of multi targeted anti-diabetic compound from traditionally used plants. J Biomol Struct Dyn 40(17):8004–8017
9. Thakur G, Bag M, Sanodiya B, Bhadauriya P, Debnath M, Prasad G, Bisen P (2009) Momordica balsamina: a medicinal and neutraceutical plant for health care management. Curr Pharmaceut Biotechnol 10:667–682. https://doi.org/10.2174/138920109789542066
10. Akinniyi JA, Sultanbawa MUA, Mnawaku D (1986) Ethnobotany and ethnopharmacology of Nigerian medicinal plants. The state of medicinal plant research in Nigeria. University Press; pp 154 – 65, Ibadan, Nigeria
11. Ramalhete C, Mansoor TA, Mulhovo S, Molnár J, Ferreira MJU (2009) Cucurbitane-type triterpenoids from the African plant *Momordica balsamina*. J Nat Prod 72(11):2009–2013. http://dx.doi.org/10.1021/np900457u
12. Faujdar S, Paliwal SK, Mehta S, Kalia AN (2013) Pharmacognostical and phytochemical evaluation of *Momordica balsamina* fruits. Int J Pharm Sci Rev Res 19:77–79
13. Ramalhete C, Gonçalves BMF, Barbosa F, Duarte N, Ferreira MU (2022) *Momordica balsamina*: phytochemistry and pharmacological potential of a gifted species. Phytochemistry reviews: proceedings of the Phytochemical Society of Europe 21(2):617–646
14. Sabiu S, Aruwa CE, Mohanlall V, Baijnath H (2020) *Momordica balsamina* L: an appraisal on morphology, ecological diversity, phytochemistry, pharmacological and biotechnological application. Curr Tradit Med 6:1–00
15. Sani AA, Idrisi SB, Yakubu S (2019) Anti-Diabetic activity of methanolic extract of momordica balsamina in rabbits. SCIRJ 7(6). https://doi.org/10.31364/SCIRJ/v7.i6.2019.P0619657
16. Siboto A, Sibiya A, Khathi A, Ngubane P (2018) The effects of momordica balsamina methanolic extract on kidney function in STZ-Induced diabetic rats: effects on selected metabolic markers. JDR 7341242. https://doiorg/101155/2018/7341242
17. Kabir N, Umar IA, Dama HA, James DB, Inuwa HM (2021) Isolation and structural Elucidation of novel antidiabetic compounds from leaves of *Momordica balsamina* Linn and *Leptadenia hastata* (Pers) decne. IJPR 20(2):390–402
18. Daina A, Michielin O, Zoete V (2017) SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness, and medicinal chemistry friendliness of small. Molecules Sci Rep 7:42717
19. Piñero J, Ramírez-Anguita JM, Saüch-Pitarch J, Ronzano F, Centeno E, Sanz F, Furlong LI (2020) The disgenet knowledge platform for disease genomics. Nucleic Acids Res 48:D845–D855
20. Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA (2005) Online Mendelian inheritance in man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Res 33:D514
21. Huang Z, Shi X, Li X, Zhang L, Wu P, Mao J, Xing R, Zhang N, Wang P (2020) Network Pharmacology approach to uncover the mechanism governing the effect of Simiao powder on knee osteoarthritis. BioMed Res Int :6971503
22. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504
23. Oh KK, Adnan M, Cho DH (2020) Network pharmacology of bioactives from *Sorghum bicolor* with targets related to diabetes mellitus. PLoS One 15(12):e0240873

24. Branden G, Sjogren T, Schnecke V, Xue Y (2014) Structure-based ligand design to overcome CYP inhibition in drug discovery projects. Drug Discov Today 19:905–911
25. Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker BA, Wang J, Yu B, Zhang J, Bryant SH (2016) PubChem substance and compound databases. Nucleic Acids Res 44(D1):D1202-1213
26. Aribisala JO, Abdulsalam RA, Dweba Y, Madonsela K, Sabiu S (2022) Identification of secondary metabolites from *Crescentia cujete* as promising antibacterial therapeutics targeting type 2A topoisomerases through molecular dynamics simulation. Comput Biol Med 145:105432
27. Nair PC, Miners JO (2014) Molecular dynamics simulations: from structure function relationships to drug discovery. Silico Pharmacol 2:1–4
28. Gonnet P (2007) P-SHAKE: a quadratically convergent SHAKE in O (n2). J Comput Phys 220(2):740–750
29. Seifert E (2014) OriginPro 91: scientific data analysis and graphing software-software review. J Chem Inf Model 54(5):1552
30. Pearson RG (1986) Absolute electronegativity and hardness correlated with molecular orbital theory. Proc Natl Acad Sci 83(22):8440–8441
31. Pearson RG (1995) The HSAB principle—more quantitative aspects. Inorg Chim Acta 240(1–2):93–98
32. Lanrewaju AA, Enitan-Folami AM, Nyaga MM, Sabiu S, Swalaha FM (2023) Cheminformatics bioprospection of selected medicinal plants metabolites against trypsin cleaved VP4 (spike protein) of rotavirus A. J Biomol Struc Dyn :1–20
33. Aihara J (1999) Reduced HOMO-LUMO gap as an index of kinetic stability for polycyclic aromatic hydrocarbons. J Phys Chem A 103:7487–7495
34. Ayers PW, Parr RG, Pearson RG (2006) Elucidating the hard/soft acid/ base principle: A perspective based on half-reactions. J Chem Phys 124(194107):1–8
35. Luo J, Xue ZQ, Liu WM, Wu JL, Yang ZQ (2006) Koopmans' theorem for large molecular systems within density functional theory. J Phys Chem 110(43):12005–12009
36. Drwal1 MN, Banerjee P, Dunkel M, Wettig MR, Preissner R (2014) ProTox: a web server for the *in Silico* prediction of rodent oral toxicity. Nucleic acids Res-42, web server issue W53–W58. https://doi.org/10.1093/nar/gku401
37. Casas AI, Hassan AA, Larsen SJ, Gomez Rangel V, Elbatreek M, Kleikers PW, Guney E, Egea J, López MG, Baumbach JJ (2019) From single drug targets to synergistic network pharmacology in ischemic stroke. Proc Natl Acad Sci USA 116:7129–7136
38. Wang TY, Li Q, Bi KS (2018) Bioactive flavonoids in medicinal plants: structure, activity and biological fate. Asian Pharm Sci 13:12–23. https://doi.org/10.1016/j.ajps.2017.08.004
39. Marunaka Y, Niisato N, Miyazaki H, Nakajima KI, Taruno A, Sun H et al (2018) Quercetin is a useful medicinal compound showing various actions including control of blood pressure, Neurite elongation and epithelial ion transport. Curr Med Chem 25:4876–87. https://doiorg/102174/0929867323666160919095043
40. Yang Y, Chen Z, Zhao X, Xie H, Du L, Gao H, Xie C (2022) Mechanisms of kaempferol in the treatment of diabetes: a comprehensive and latest review. Front Endocrinol 13:990299. https://doi.org/10.3389/fendo2022990299
41. Niisato N, Marunaka Y (2023) Therapeutic potential of multifunctional myricetin for treatment of type 2 diabetes mellitus. Front Nutr 10:1175660. https://doi.org/10.3389/fnut.2023.1175660
42. Martínez Báez A, Ayala G, Pedroza-Saavedra A, González-Sánchez HM, Chihu Amparan L (2024) Phosphorylation codes in IRS-1 and IRS-2 are associated with the activation/inhibition of insulin canonical signaling pathways. Curr Issues Mol Biol 46(1):634–649. https://doi.org/10.3390/cimb46010041
43. Huang Y, Kyriakides TR (2020) The role of extracellular matrix in the pathophysiology of diabetic wounds matrix. Biology Plus 6:100037. https://doi.org/10.1016/j.mbplus2020100037
44. Arreguin-Cano JA, Ayerdi-Nájera B, Tacuba-Saavedra A, Navarro-Tito N, Dávalos-Martínez A, Emigdio-Vargas A, Barrera-Rodríguez E, Blanco-García N, Gutiérrez-Venegas G, Ventura-Molina E, León-Dorantes G (2019) MMP-2 salivary activity in type 2 diabetes mellitus patients. Diabetol Metab Syndr 11:113. https://doi.org/10.1186/s13098-019-0510-2
45. Rubio JA, Jiménez S, Lázaro-Martínez JL (2020) Mortality in patients with diabetic foot ulcers: causes, risk factors, and their association with evolution and severity of ulcer. J Clin Med 9(9):3009. https://doi.org/10.3390/jcm9093009
46. Ng GYQ, Loh ZW, Fann DY, Mallilankaraman K, Arumugam TV, Hande MP (2024) Role of mitogen-activated protein (MAP) kinase pathways in metabolic diseases. Genome Integr 15:1. https://doi.org/10.14293/genint141004
47. Tomita T (2016) Apoptosis in pancreatic β-islet cells in type 2 diabetes. Bosn J Basic Med Sci 16(3):162–179. https://doi.org/10.17305/bjbms.2016.919
48. Gupta P, Taiyab A, Hassan MI (2020) Emerging role of protein kinases in diabetes mellitus: From mechanism to therapy. Adv Protein Chem Struct Biol 124:47–85. https://doi.org/10.1016/bs.apcsb.2020.11.001
49. Forrester SJ, Kawai T, O'Brien S, Thomas W, Harris RC, Eguchi S (2016) Epidermal growth factor receptor transactivation: mechanisms, pathophysiology, and potential therapies in the cardiovascular system. Annu Rev Pharmacol Toxicol 56:627–653. https://doi.org/10.1146/annurev-pharmtox-070115-095427
50. Harskamp LR, Gansevoort RT, van Goor H, Meijer E (2016) The epidermal growth factor receptor pathway in chronic kidney diseases. Nat Rev Nephrol 12(8):496–506. https://doi.org/10.1038/nrneph201691
51. Balogun TA, Ipinloju N, Abdullateef OT, Moses SI, Omoboyowa DA, James AC et al (2021) Computational evaluation of bioactive compounds from *Colocasia affinis* Schott as a novel EGFR inhibitor for cancer treatment. Cancer Inform. https://doi.org/10.1177/11769351211049244
52. Li Z, Li Y, Overstreet JM, Chung S, Niu A, Fan X et al (2018) Inhibition of epidermal growth factor receptor activation is associated with improved diabetic nephropathy and insulin resistance in type 2 diabetes. Diabetes 67(9):1847–1857. https://doi.org/10.2337/db17-1513
53. Zhang MZ, Wang Y, Paueksakon P, Harris RC (2014) Epidermal growth factor receptor inhibition slows progression of diabetic nephropathy in association with a decrease in endoplasmic reticulum stress and an increase in autophagy. Diabetes 63(6):2063–2072
54. Romero R (2023) Identification of potential EGFR inhibitors for type 2 diabetes and pancreatic cancer treatment: A computational approach. BioRxiv Preprint https://doi.org/10.1101/2023.09.05.556125
55. Hevener AL, Zhou Z, Moore TM, Drew BG, Ribas V (2018) The impact of ER alpha action on muscle metabolism and insulin sensitivity—strong enough for a man made for a woman. Mol Met 15:20–34
56. Meyer MR, Clegg DJ, Prossnitz ER, Barton M (2011) Obesity insulin resistance and diabetes: sex differences and role of oestrogen receptors. Acta Physiol 203:259–269
57. Wang H, Yang X, Sun Y, Li Y, Dong Y, Shan D, Yu Z (2020) Synergistic effect of metformin and EGFR-TKI in the treatment of non-small cell lung cancer. Transl Cancer Res 9(1):372–381
58. Teli DM, Gajjar AK (2023) Glycogen synthase kinase-3: a potential target for diabetes. Bioorg Med Chem 92:117406
59. Karmakar S, Basak HK, Paswan U, Pramanik AK, Chatterjee A (2022) Designing of next-generation dihydropyridine-based

calcium channel blockers: an *in Silico* study. J Appl Pharm Sci 12(4):127–35
60. Kazeem MI, Abdulsalam RA, Mellem JJ, Sabiu S (2025) Lemongrass (*Cymbopogon citratus*) infusions exhibit neuroprotective properties: evidence from in vitro and in Silico studies. Food Bioscience 68:106355
61. Martinez L (2015) Automatic identification of mobile and rigid substructures in molecular dynamics simulations and fractional structural fluctuation analysis. PLoS One 10:e0119264
62. Hess B (2002) Convergence of sampling in protein simulations. Phys Rev 65:031910
63. Cerón-Carrasco JP (2022) When virtual screening yields inactive drugs: dealing with false theoretical friends. Chem Med Chem 17:e202200278. https://doi.org/10.1002/cmdc.202200278
64. Childers MC, Daggett V (2017) Insights from molecular dynamics simulations for computational protein design. Mol Syst Des Eng 12:9–33. https://doi.org/10.1039/C6ME00083E
65. Kumi RO, Soremekun OS, Issahaku AR, Fisayo CA, Olotu A, Soliman MES (2020) Exploring the ring potential of 2,4-diaminopyrimidine derivatives towards the identification of novel caspase-1 inhibitors in Alzheimer's disease therapy. J Mol Model 26:1–17
66. Adedirin O, Abdulsalam RA, Nasir-Naeem KO, Oke AA, Jubril AO, Sabiu S (2025) Density functional theory and molecular dynamics simulation-based bioprospection of *Agathosma betulina* essential oil metabolites against protein tyrosine phosphatase 1B for interventive antidiabetic therapy. Heliyon. https://doi.org/10.1016/j.heliyon.2025.e42239
67. Sulyman AO, Aliyu OM, Ajani EO, Abdulkareem YF, Afe IA, Abdulyakeen FO, Lukman HY, Sabiu S, Singh K, Naidoo K (2025) Mechanisms of L-citrulline on phosphodiesterase 5 in erectile dysfunction intervention. Sci Afr 27:e02572. https://doi.org/10.1016/j.sciaf.2025.e02572
68. Hubbard RE, Haider MK (2010) Hydrogen bonds in proteins: role and strength. eLS John Wiley & Sons, Ltd, Hoboken, NJ, USA
69. Abdolmaleki A, Shiri F, Ghasemi JB (2021) Part IV applications and case studies of docking. Chapter 11; Use of molecular docking as a decision-making tool in drug discovery. Molecular docking for Computer-Aided Drug Design. Academic Press, pp 229–243. https://doi.org/10.1016/b978-0-12-822312-3.00010-2
70. Basak HK, Paswan U, Ghosh J, Saha S, Chatterjee A (2021) In silico evaluation of savirin derivatives as inhibitors of the agr quorum sensing system of Staphylococcus aureus. Anal Chem Lett 11(5):661–683
71. Ibrahim SO, Lukman HY, Zubair MF, Amusan OT, Abdulkadri FR, Lawal B, Abdulfatah LB, Atolani O (2024) An insight into the physicochemical, drug-likeness, pharmacokinetics and toxicity profile of Kigelia africana (Lam) bioactive compounds. BJEPS 4(1):4. https://doi.org/10.55810/2313-0083.1050
72. An G, Gallegos J, Morris ME (2011) The bioflavonoid kaempferol is an Abcg2 substrate and inhibits Abcg2-mediated quercetin efflux. Drug Metab Dispos 39(3):426–432. https://doi.org/10.1124/dmd.110.035212
73. Luo H, Jiang B, Li B, Li Z, Jiang BH, Chen YC (2012) Kaempferol nanoparticles achieve strong and selective inhibition of ovarian cancer cell viability. Int J Nanomed 7:3951–9. https://doi.org/10.2147/IJN.S33670
74. Lukman HY, Adedirin O, Sabiu S (2025) Antidiabetic potential of citrus essential oil nanoparticles: A systematic review of current evidence and future directions. J Essent Oil-Bear Plants 28(5):886–901

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.