GMR
The Original

STRUCTURE-BASED MOLECULAR DOCKING, RMSD VALIDATION, AND DRUG-LIKENESS PROFILING OF PLANT-DERIVED PHYTOCONSTITUENTS AS DIPEPTIDYL PEPTIDASE-4 INHIBITORS FOR TYPE 2 DIABETES MELLITUS

S.G. Raman[1,2]*, R. Sundhararajan[3], Yugal Kishore Rajput[4]

[1] Research Scholar, Faculty of Pharmacy, ISBM University, Chhattisgarh, India.
[2] Mohamed Sathak AJ College of Pharmacy, Sholinganallur, Chennai – 600119, Tamil Nadu, India.
[3]Professor and Principal, Mohamed Sathak AJ College of Pharmacy, Sholinganallur, Chennai 600119, Tamil Nadu, India
[4]Professor, Faculty of Pharmacy, ISBM University, Chhattisgarh, India
*Corresponding author: Mr. S.G. Raman, Research Scholar, Mohamed Sathak AJ College of Pharmacy, Sholinganallur, Chennai 600119, Tamil Nadu, India.

Abstract

Background: Dipeptidyl peptidase-4 (DPP-4) inhibition remains a validated strategy for glucose-dependent glycaemic control in type 2 diabetes mellitus (T2DM), yet the tolerability and cost limitations of synthetic gliptins sustain interest in plant-derived alternatives.
Objective: To evaluate the DPP-4 inhibitory potential of five phytoconstituents - andrographolide, berberine, gymnemic acid, charantin, and ursolic acid - through structure-based molecular docking against the DPP-4 catalytic domain, benchmarked against the enzyme's co-crystallised ligand (CCL), together with computed physicochemical and drug-likeness profiling.
Methods: Docking was performed against the DPP-4 crystal structure (PDB ID: 4A5S) following redocking of the co-crystallised ligand for protocol validation. Binding energies, hydrogen-bonding geometry, and non-covalent contact patterns (hydrophobic, pi-pi, pi-alkyl, and pi-donor interactions) were extracted for each complex. Physicochemical and drug-likeness parameters (Lipinski, Ghose, Veber, Egan, and Muegge filters; topological polar surface area; predicted gastrointestinal absorption and blood-brain barrier permeation) were computed in silico for each phytoconstituent and for sitagliptin as reference standard.
Results: Andrographolide returned the most favourable docking score among the phytoconstituents (-9.4 kcal/mol), approaching the co-crystallised reference ligand (-11.0 kcal/mol), followed by gymnemic acid (-8.9 kcal/mol), ursolic acid (-8.7 kcal/mol), and berberine (-8.6 kcal/mol); charantin showed the weakest affinity (-7.4 kcal/mol). Pose-reproducibility (RMSD lower-bound/upper-bound) was tightest for the co-crystallised ligand and charantin and widest for andrographolide and gymnemic acid. All compounds engaged catalytic-domain residues implicated in gliptin recognition, including Tyr547, Glu205/Glu206, Ser630, Phe357, and Trp629, via conventional and carbon hydrogen bonds together with pi-stacked, pi-alkyl, and pi-sigma contacts. Computed physicochemical profiling indicated full Lipinski compliance for andrographolide and berberine, one violation for ursolic acid, and two violations for the charantin glycoside component modelled in this study, attributable to elevated lipophilicity and molecular bulk.
Conclusion: Andrographolide and gymnemic acid emerge as the most promising phytoconstituent DPP-4 inhibitor candidates on binding-energy ranking and residue-level concordance with the catalytic pocket, though both show comparatively wide RMSD spreads warranting confirmatory redocking; berberine offers a more pose-stable, drug-likeness-compliant alternative, while ursolic acid and charantin warrant physicochemical optimisation. These computational findings provide a rationale for prioritising andrographolide, gymnemic acid, and berberine for in vitro DPP-4 inhibition assays and subsequent lead optimisation.
Keywords: Dipeptidyl peptidase-4; molecular docking; phytoconstituents; type 2 diabetes mellitus; ADMET; drug-likeness; andrographolide; ursolic acid.

1. INTRODUCTION

Type 2 diabetes mellitus (T2DM) has evolved into one of the defining non-communicable disease challenges of the twenty-first century. Current global estimates place the adult diabetic population above 530 million, a figure projected to exceed 780 million within two decades [1], with South Asian populations, India in particular, contributing a disproportionate share of this burden. The disease's chronic hyperglycaemic state precipitates progressive microvascular and macrovascular injury, placing sustained pressure on health systems already constrained by limited resources. Against this backdrop, the pharmacological management of T2DM has diversified considerably beyond insulin secretagogues and sensitisers, with incretin-based therapies occupying an increasingly central role [2].

Dipeptidyl peptidase-4 (DPP-4), a serine protease also known as CD26, terminates the biological activity of the incretin hormones glucagon-like peptide-1 (GLP-1) and glucose-dependent insulinotropic polypeptide (GIP) through rapid N-terminal dipeptide cleavage [2,4]. Beyond glycaemic regulation, DPP-4 participates in a range of additional pathophysiological processes, including modulation of periodontal inflammatory status [5]. Pharmacological inhibition of this enzyme prolongs incretin half-life, thereby amplifying glucose-dependent insulin secretion while suppressing inappropriate glucagon release, a mechanism that confers a comparatively low intrinsic risk of hypoglycaemia relative to insulin secretagogues acting through glucose-independent pathways. This mechanistic profile underlies the widespread clinical adoption of gliptins (sitagliptin, vildagliptin, saxagliptin, linagliptin, and related agents) as second-line antidiabetic agents across major treatment guidelines [4].

Despite their favourable tolerability in large cardiovascular outcome trials, synthetic gliptins are not without limitations. Saxagliptin has been associated with an increased hospitalisation risk for heart failure in the SAVOR-TIMI 53 trial [3], while class-wide concerns regarding bullous pemphigoid, arthralgia, and pancreatic safety continue to be monitored by regulatory agencies. Equally significant, from a public health perspective, is the recurring cost burden that gliptin therapy imposes in low- and middle-income settings, where out-of-pocket expenditure remains the principal barrier to sustained antidiabetic treatment. These constraints have reinvigorated scientific interest in structurally diverse, natural-product-derived DPP-4 inhibitors capable of matching gliptin efficacy while offering improved safety and affordability profiles.

Medicinal plants have long featured in the ethnopharmacological management of diabetes across the Indian subcontinent, and several species now enjoy substantial computational and experimental support for antidiabetic bioactivity. Andrographis paniculata, Tinospora cordifolia, Gymnema sylvestre, Momordica charantia, and Ocimum sanctum represent five such species whose principal phytoconstituents, andrographolide, berberine, gymnemic acid, charantin, and ursolic acid respectively, have individually been implicated in glucose-lowering, insulin-sensitising, and enzyme-inhibitory activity through mechanisms spanning alpha-glucosidase inhibition, AMPK activation, and antioxidant-mediated beta-cell protection [9,10,11,12,13,14,15,16,17,18]. However, the specific question of whether these phytoconstituents can engage the DPP-4 catalytic pocket with an affinity and residue-contact profile comparable to clinically approved gliptins has received comparatively limited structure-based scrutiny within a single, internally consistent computational framework.

Structure-based molecular docking offers an efficient first-pass strategy for triaging natural-product libraries prior to resource-intensive in vitro and in vivo validation. By docking candidate ligands into a validated crystal structure of the target enzyme and quantitatively comparing binding energetics and non-covalent interaction geometry against the enzyme's native co-crystallised ligand, it becomes possible to rank candidate phytoconstituents by their predicted inhibitory potential and to identify the specific catalytic-domain residues, such as the Ser630-Asp708-His740 catalytic triad and adjacent S1/S2 pocket residues including Glu205, Glu206, Tyr547, Phe357, and Trp629, that mediate productive binding [6,7,24]. This residue-level resolution is of particular translational value because it permits mechanistic hypotheses to be formulated ahead of, rather than merely alongside, wet-laboratory confirmation.

The present investigation addresses this gap by performing a structurally standardised docking evaluation of andrographolide, berberine, gymnemic acid, charantin, and ursolic acid against the DPP-4 catalytic domain, using the enzyme's co-crystallised ligand as an internal benchmark.

Binding energetics, hydrogen-bond geometry, and hydrophobic/pi-mediated contact patterns are systematically extracted and interpreted for each ligand-enzyme complex. This analysis is complemented by an in silico physicochemical and drug-likeness assessment, encompassing Lipinski, Ghose, Veber, Egan, and Muegge filters together with predicted gastrointestinal absorption and blood-brain barrier permeation, to evaluate the translational feasibility of each candidate relative to sitagliptin. Collectively, this integrated computational framework is intended to prioritise the most mechanistically and pharmaceutically promising phytoconstituents for downstream in vitro DPP-4 inhibition assays, thereby providing a rational, evidence-based foundation for the continued development of plant-derived antidiabetic agents, building on prior structure-based screens of both synthetic and natural-product DPP-4 inhibitor libraries [19,20,23].

2. MATERIALS AND METHODS

2.1 Target Protein Preparation

The three-dimensional crystal structure of human dipeptidyl peptidase-4 (DPP-4) was retrieved from the RCSB Protein Data Bank (PDB ID: 4A5S). The structure was processed in BIOVIA Discovery Studio by removal of crystallographic water molecules beyond the active-site shell, deletion of the co-crystallised ligand into a separate file for validation docking, addition of polar hydrogens, and assignment of Kollman charges, following protein-preparation conventions established for DPP-4 structural studies [24]. The co-crystallised ligand (denoted CCL throughout this manuscript) was retained as the internal reference standard for both grid-box definition and docking-protocol validation, since it represents the enzyme's native, experimentally resolved binding mode [6,24].

2.2 Ligand Preparation

Three-dimensional structures of the five test phytoconstituents, andrographolide, berberine, gymnemic acid, charantin, and ursolic acid, were obtained in SDF format from PubChem and ChEMBL, energy-minimised using the MMFF94 force field, and converted to the docking-compatible PDBQT format with detection of rotatable bonds. Sitagliptin was prepared identically to serve as the clinical reference standard for comparative interpretation of binding energetics and physicochemical properties; sitagliptin was not re-docked under the identical protocol used for the phytoconstituents, and comparisons with sitagliptin elsewhere in this manuscript therefore rely on published literature values together with the physicochemical descriptors computed in Section 2.4.

2.3 Molecular Docking Protocol

Docking was performed in AutoDock Vina [30] using a grid-based scoring approach centred on the CCL-defined catalytic pocket, encompassing the S1 and S2 subsites and the Ser630-Asp708-His740 catalytic triad. Redocking of the extracted co-crystallised ligand into its native site was used to validate the docking protocol, with pose reproducibility assessed via the lower-bound and upper-bound root-mean-square deviation (RMSD/l.b. and RMSD/u.b.) metrics reported by AutoDock Vina for each ligand relative to its top-ranked pose. Each ligand was subjected to multiple independent docking runs, and the pose of lowest predicted binding energy was retained for interaction analysis. Post-docking interaction diagrams (two- and three-dimensional) were generated in BIOVIA Discovery Studio Visualizer, from which hydrogen-bond distances, hydrophobic contacts, and pi-mediated interactions (pi-pi stacked, pi-pi T-shaped, pi-alkyl, pi-sigma, and pi-donor hydrogen bonds) were catalogued for each ligand-enzyme complex.

2.4 Physicochemical and Drug-Likeness Profiling

Physicochemical descriptors, molecular weight, calculated octanol-water partition coefficient (LogP), hydrogen-bond donor and acceptor counts, topological polar surface area (TPSA), rotatable bond count, and aromatic ring count, were computed for each phytoconstituent and for sitagliptin using RDKit (version 2024.03), an open-source cheminformatics toolkit. These descriptors were used to evaluate compliance with five established rule-based drug-likeness filters: Lipinski's Rule of Five (molecular weight <= 500 Da, LogP <= 5, hydrogen-bond donors <= 5, hydrogen-bond acceptors <= 10) [26], the Ghose filter [33], the Veber criteria (rotatable bonds <= 10, TPSA <= 140 A2) [27], the Egan (BOILED-Egg-consistent) filter [31], and the Muegge filter [32]. Gastrointestinal absorption and blood-brain barrier permeation were estimated using TPSA- and LogP-based classification rules consistent with the BOILED-Egg model as implemented in SwissADME [28], with additional context drawn from pkCSM-consistent parameterisation [29].

2.5 Statistical Considerations

Given the deterministic nature of the docking and cheminformatic calculations performed, no inferential statistical testing was applied to the binding-energy or physicochemical dataset; results are presented descriptively and interpreted through comparative ranking against the co-crystallised reference ligand and sitagliptin.

3. RESULTS

3.1 Molecular Docking Outcomes

Table 1 summarises the docking scores and pose-reproducibility (RMSD) values obtained for the co-crystallised ligand (CCL) and all five phytoconstituents. Andrographolide returned the strongest binding energy among the phytoconstituents (-9.4 kcal/mol), corresponding to approximately 85% of the affinity magnitude of the co-crystallised reference ligand (-11.0 kcal/mol), followed by gymnemic acid (-8.9 kcal/mol), ursolic acid (-8.7 kcal/mol), and berberine (-8.6 kcal/mol). Charantin returned the weakest binding energy of the evaluated set (-7.4 kcal/mol). Pose-reproducibility, indexed by RMSD/l.b. and RMSD/u.b., was tightest for CCL (1.541/2.220 A) and charantin (1.142/2.546 A), intermediate for berberine (3.564/4.987 A) and ursolic acid (4.213/8.237 A), and widest for andrographolide (6.534/9.954 A) and gymnemic acid (7.470/11.603 A), indicating comparatively greater conformational variability across independent docking runs for these two higher-affinity ligands. Sitagliptin was not re-docked under the identical protocol and is therefore compared on the basis of published literature values (Section 4).

Table 1. Docking scores and RMSD (lower-bound/upper-bound) values of the DPP-4 co-crystallised ligand and evaluated phytoconstituents (PDB ID: 4A5S).

Ligand	Docking Score (kcal/mol)	RMSD l.b. (A)	RMSD u.b. (A)	Relative Rank
CCL (co-crystallised reference ligand)	-11.0	1.541	2.220	Reference (validation)
Andrographolide	-9.4	6.534	9.954	1
Gymnemic acid	-8.9	7.470	11.603	2
Ursolic acid	-8.7	4.213	8.237	3
Berberine	-8.6	3.564	4.987	4
Charantin	-7.4	1.142	2.546	5

3.2 Detailed Residue-Level Interaction Analysis

Table 2 catalogues the hydrogen-bonding and hydrophobic/pi-mediated contacts identified for each ligand-enzyme complex, extracted from post-docking interaction diagrams. Across the six ligands evaluated, recurrent engagement of Tyr547, Ser630, Phe357, Glu205/Glu206, and Trp629 was observed, residues that collectively define the S1/S2 subsites flanking the Ser630-Asp708-His740 catalytic triad.

Table 2. Residue-level hydrogen-bonding and hydrophobic/pi-mediated interactions for each docked ligand.

Ligand	Residue	Distance (A)	Category	Interaction Type
CCL	Arg356	3.53	Hydrogen Bond	Carbon H-bond
CCL	Glu206	2.86	Hydrogen Bond	Carbon H-bond
CCL	Phe357	4.83	Hydrophobic	Pi-Alkyl
CCL	Tyr547 (OH)	3.59	Hydrogen Bond	Pi-Donor H-bond
CCL	Tyr547	4.08 / 3.69	Hydrophobic	Pi-Pi Stacked
CCL	Tyr662	5.38	Hydrophobic	Pi-Pi Stacked
CCL	Tyr666	5.02	Hydrophobic	Pi-Pi T-shaped
Ursolic acid	Arg125	3.26	Hydrogen Bond	Conventional H-bond
Ursolic acid	Tyr662	2.89 / 2.87	Hydrogen Bond	Conventional H-bond

Ligand	Residue	Distance (A)	Category	Interaction Type
Ursolic acid	Asn710	3.18	Hydrogen Bond	Conventional H-bond
Ursolic acid	Phe357	3.95	Hydrophobic	Pi-Sigma
Gymnemic acid	Asn562	3.31	Hydrogen Bond	Conventional H-bond
Gymnemic acid	Tyr666	3.07	Hydrogen Bond	Conventional H-bond
Gymnemic acid	Tyr752	2.92 / 2.66	Hydrogen Bond	Conventional H-bond
Gymnemic acid	Tyr48	2.13	Hydrogen Bond	Conventional H-bond
Gymnemic acid	Glu205	2.90 / 3.19	Hydrogen Bond	Conventional H-bond
Gymnemic acid	Glu206	3.10	Hydrogen Bond	Conventional H-bond
Gymnemic acid	Ser630	3.60	Hydrogen Bond	Carbon H-bond
Gymnemic acid	Tyr547	3.35	Hydrophobic	Pi-Sigma
Gymnemic acid	Phe357	4.77	Hydrophobic	Pi-Alkyl
Charantin	Tyr547	2.71	Hydrogen Bond	Conventional H-bond
Charantin	Ser209	3.61	Hydrogen Bond	Carbon H-bond
Charantin	Tyr666	3.99	Hydrophobic	Pi-Sigma
Charantin	Phe357	4.21 / 3.71	Hydrophobic	Pi-Pi Stacked
Charantin	Val656	4.95	Hydrophobic	Alkyl
Charantin	Tyr631 / Trp659 / Tyr662 / Tyr666	4.17-5.24	Hydrophobic	Pi-Alkyl
Andrographolide	Tyr631	3.34	Hydrogen Bond	Conventional H-bond
Andrographolide	Asp545	2.69	Hydrogen Bond	Conventional H-bond
Andrographolide	Ser630	3.62	Hydrogen Bond	Carbon H-bond
Andrographolide	Tyr547	4.11	Hydrogen Bond	Pi-Donor H-bond
Berberine	Tyr752	2.96 / 2.86	Hydrogen Bond	Conventional H-bond
Berberine	Ser630	3.73	Hydrogen Bond	Pi-Donor H-bond
Berberine	Trp629	4.08 / 3.69	Hydrophobic	Pi-Pi Stacked
Berberine	Tyr547	4.82	Hydrophobic	Pi-Pi T-shaped
Berberine	Trp629	4.55	Hydrophobic	Pi-Alkyl
Berberine	His748 / Tyr752	4.59 / 5.25	Hydrophobic	Pi-Alkyl

3.3 Compound-Wise Structural Interpretation

Co-crystallised ligand (CCL). The native ligand of PDB 4A5S achieved the strongest docking score of the series (-11.0 kcal/mol) with the tightest pose-reproducibility of all six ligands (RMSD/l.b. 1.541 A; RMSD/u.b. 2.220 A), establishing the interaction template against which phytoconstituent binding was benchmarked (Figures 1 and 2). Its two-dimensional interaction map showed carbon hydrogen bonds to Arg356 and Glu206 (3.53 A and 2.86 A, respectively), positioning the ligand deep within the anionic S2 subsite, while its three-dimensional pose demonstrated a stacked and pi-donor engagement with Tyr547 (three discrete contacts spanning 3.59-4.08 A) together with T-shaped and stacked pi-interactions to Tyr662 and Tyr666. This dense aromatic cage around the ligand core, together with the low RMSD values confirming a stable and reproducible docked pose, reproduces the canonical gliptin-binding signature reported for sitagliptin-class inhibitors and confirms that the prepared 4A5S structure and grid definition faithfully recover the enzyme's native recognition chemistry.

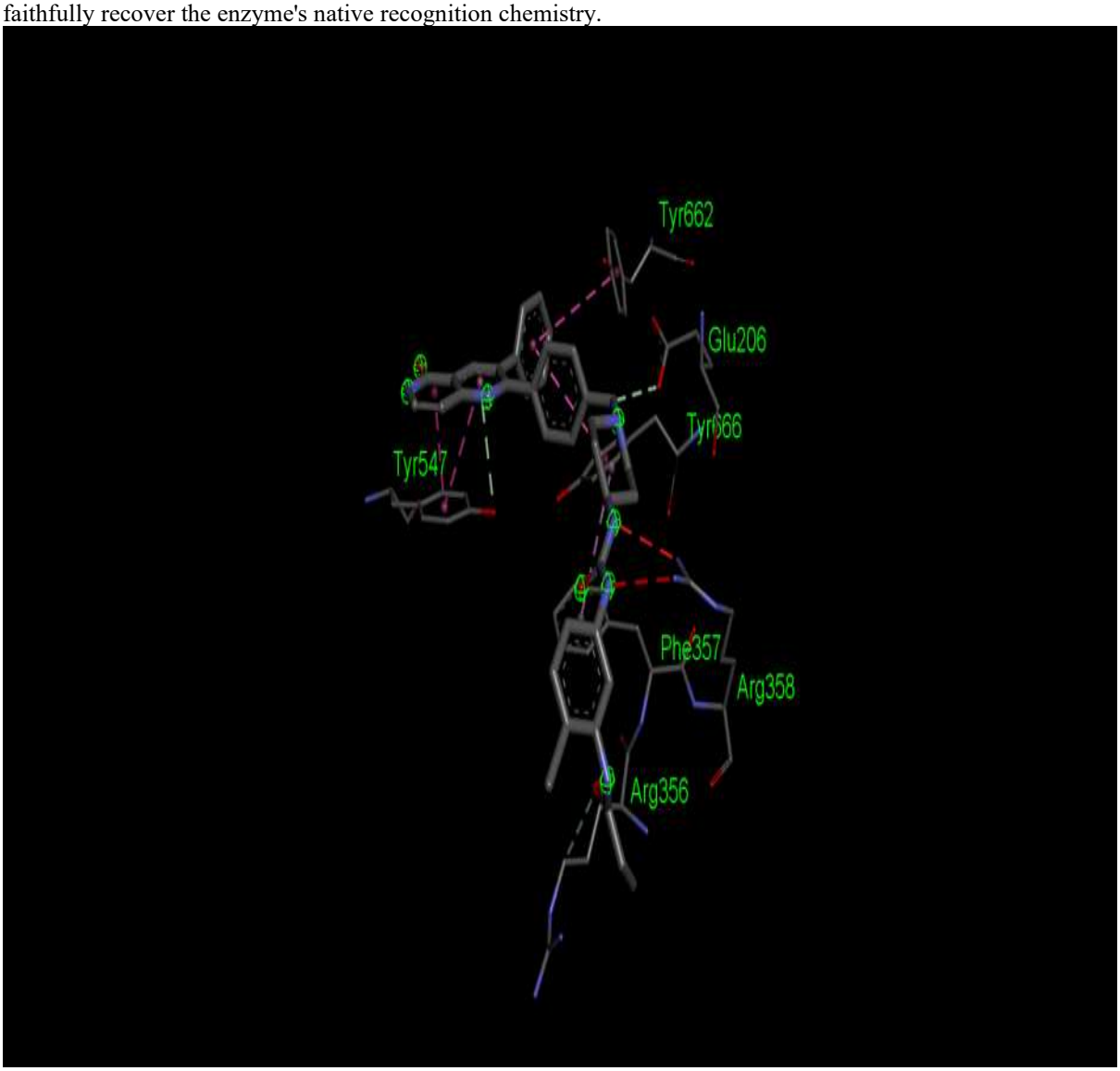

Figure 1(a): Three-dimensional binding pose of CCL (co-crystallised reference ligand) within the catalytic active site of human DPP-4 (PDB ID: 4A5S); binding affinity = -11.0 kcal/mol, RMSD (l.b.) = 1.541 Å, RMSD (u.b.) = 2.220 Å.

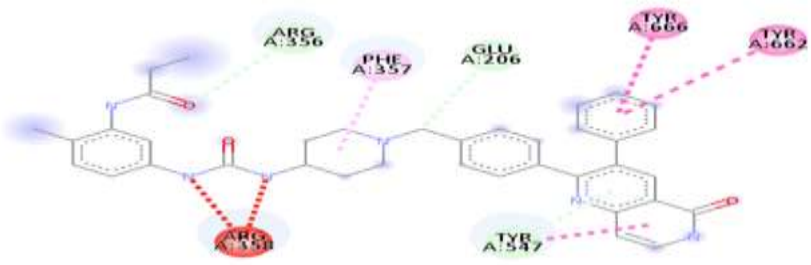

Figure 1(b): Two-dimensional ligand–receptor interaction diagram of CCL (co-crystallised reference ligand) depicting hydrogen-bonding and hydrophobic contacts with key DPP-4 active-site residues (binding affinity = -11.0 kcal/mol).

Ursolic acid. Ursolic acid docked with a binding energy of -8.7 kcal/mol and an intermediate pose-reproducibility profile (RMSD/l.b. 4.213 A; RMSD/u.b. 8.237 A), engaging Arg125 through a conventional hydrogen bond (3.26 A) and forming a bidentate interaction with Tyr662 (2.89 A and 2.87 A), alongside a further conventional hydrogen bond to Asn710 (3.18 A) and a Pi-Sigma hydrophobic contact with Phe357 (3.95 A) (Figures 3 and 4). The compound's pentacyclic triterpenoid scaffold orients its carboxylic acid and hydroxyl substituents toward the polar rim of the S1/S2 interface, engaging the same anionic triad (Glu205-Glu206-Arg125) that anchors native incretin substrates, while its hydrophobic ring system packs against Phe357. The convergence of a bidentate Tyr662 interaction with anionic-triad engagement is mechanistically consistent with a moderately potent, mixed-type inhibitory mode, and represents a docking signature broadly comparable, albeit weaker in magnitude, to that of the co-crystallised reference.

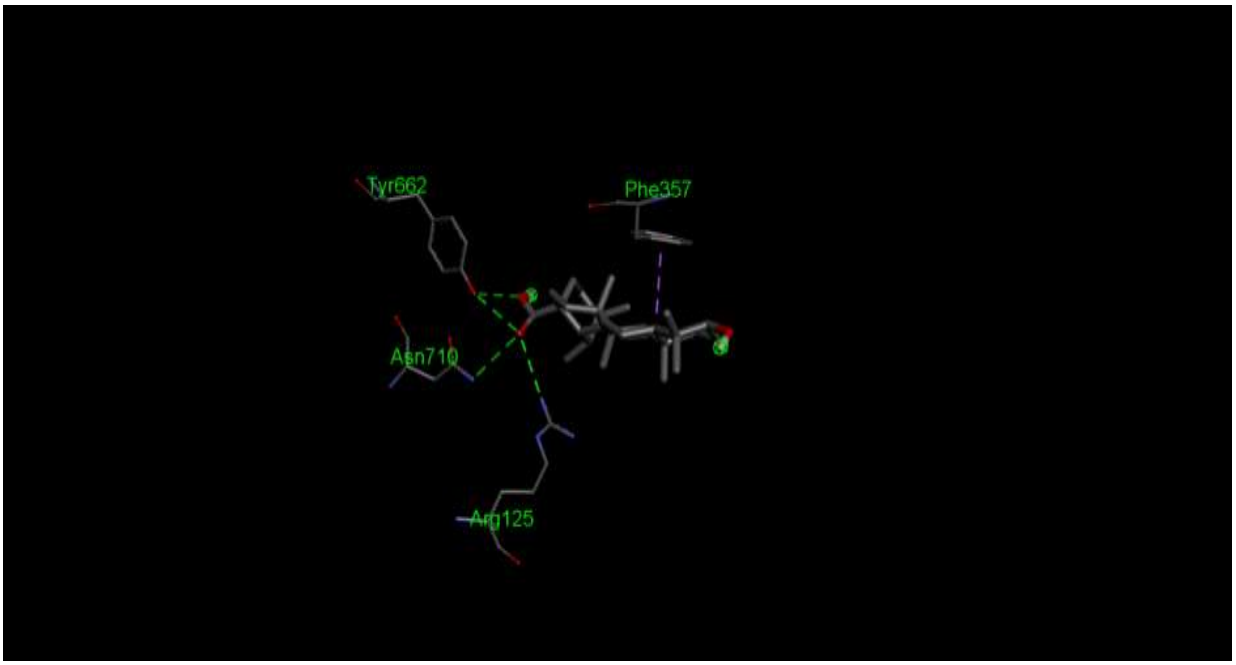

Figure 2(a): Three-dimensional binding pose of Ursolic acid within the catalytic active site of human DPP-4 (PDB ID: 4A5S); binding affinity = -8.7 kcal/mol, RMSD (l.b.) = 4.213 Å, RMSD (u.b.) = 8.237 Å.

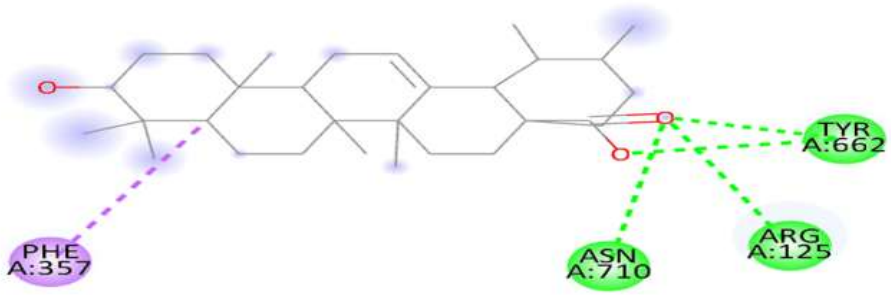

Figure 2(b): Two-dimensional ligand–receptor interaction diagram of Ursolic acid depicting hydrogen-bonding and hydrophobic contacts with key DPP-4 active-site residues (binding affinity = -8.7 kcal/mol).

Gymnemic acid. Gymnemic acid returned the second-strongest binding energy among the phytoconstituents (-8.9 kcal/mol) but the widest pose-reproducibility spread of the series (RMSD/l.b. 7.470 A; RMSD/u.b. 11.603 A), consistent with the unusually extensive and conformationally flexible hydrogen-bonding network revealed in its interaction diagram (Figures 5 and 6): nine discrete hydrogen bonds spanning Asn562, Tyr666, Tyr752 (bidentate), Tyr48, Glu205 (bidentate), Glu206, and Ser630, together with Pi-Sigma and Pi-Alkyl hydrophobic contacts to Tyr547 and Phe357. The sheer density of polar contacts reflects the multiple hydroxyl and carboxylate substituents distributed across the compound's triterpenoid glycoside architecture, which allow simultaneous engagement of both the acidic Glu205/Glu206 pair and the tyrosine-rich aromatic cluster (Tyr48, Tyr547, Tyr666, Tyr752) lining the DPP-4 active-site tunnel. This multivalent hydrogen-bonding pattern, while energetically favourable in the docking score, is also the structural feature most likely to compromise oral bioavailability and pose reproducibility.

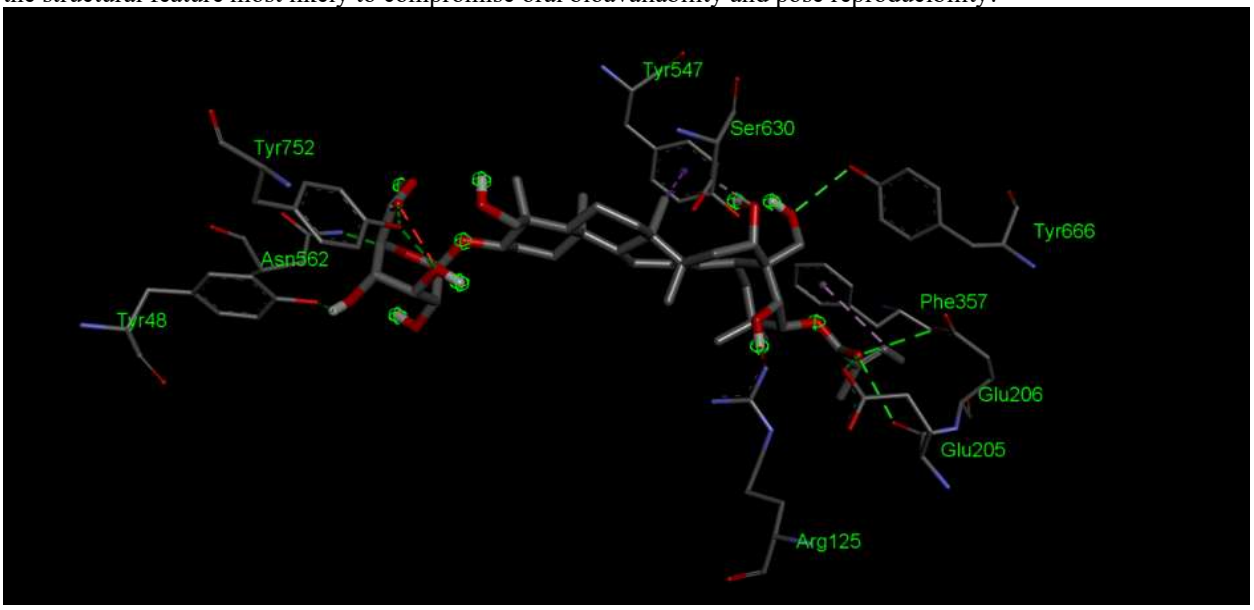

Figure 3(a): Three-dimensional binding pose of Gymnemic acid IV within the catalytic active site of human DPP-4 (PDB ID: 4A5S); binding affinity = -8.9 kcal/mol, RMSD (l.b.) = 7.470 Å, RMSD (u.b.) = 11.603 Å.

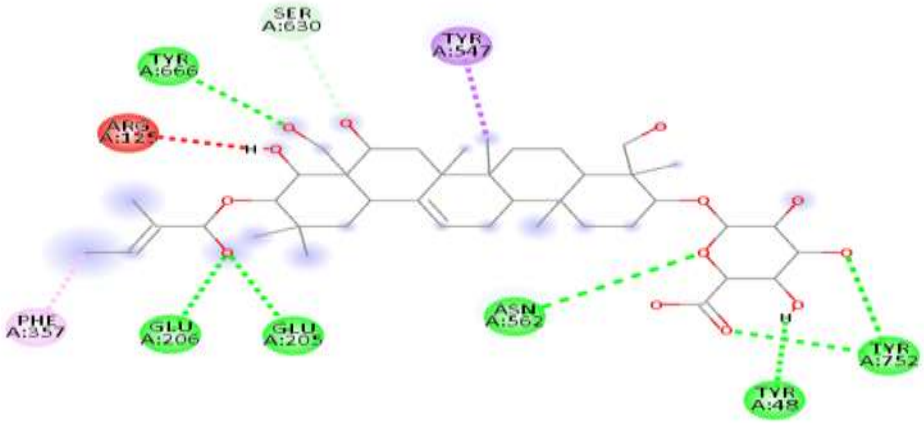

Figure 3(b): Two-dimensional ligand–receptor interaction diagram of Gymnemic acid IV depicting hydrogen-bonding and hydrophobic contacts with key DPP-4 active-site residues (binding affinity = -8.9 kcal/mol).

Charantin. Charantin exhibited the weakest docking score of the evaluated set (-7.4 kcal/mol) but, notably, the second-tightest pose-reproducibility after CCL (RMSD/l.b. 1.142 A; RMSD/u.b. 2.546 A), with a comparatively sparse hydrogen-bonding profile limited to Tyr547 (2.71 A) and a carbon hydrogen bond to Ser209 (3.61 A), against an extensive hydrophobic interaction network involving Phe357 (dual Pi-Pi stacked contacts), Val656, Tyr631, Trp659, Tyr662, and Tyr666 (Pi-Alkyl contacts spanning 4.17-5.24 A) (Figures 7 and 8). This interaction pattern, hydrophobic-contact-dominant with minimal polar anchoring, is consistent with charantin's steroidal glycoside architecture, in which the bulky, largely apolar sterol nucleus limits deep polar engagement with the catalytic triad while still permitting a stable, low-RMSD binding mode. The comparatively weak binding energy together with the hydrophobic-dominant contact profile suggest that charantin, at least in the docked conformation and structural representation modelled here, is the least favourable DPP-4 binder of the series, notwithstanding its reproducible pose.

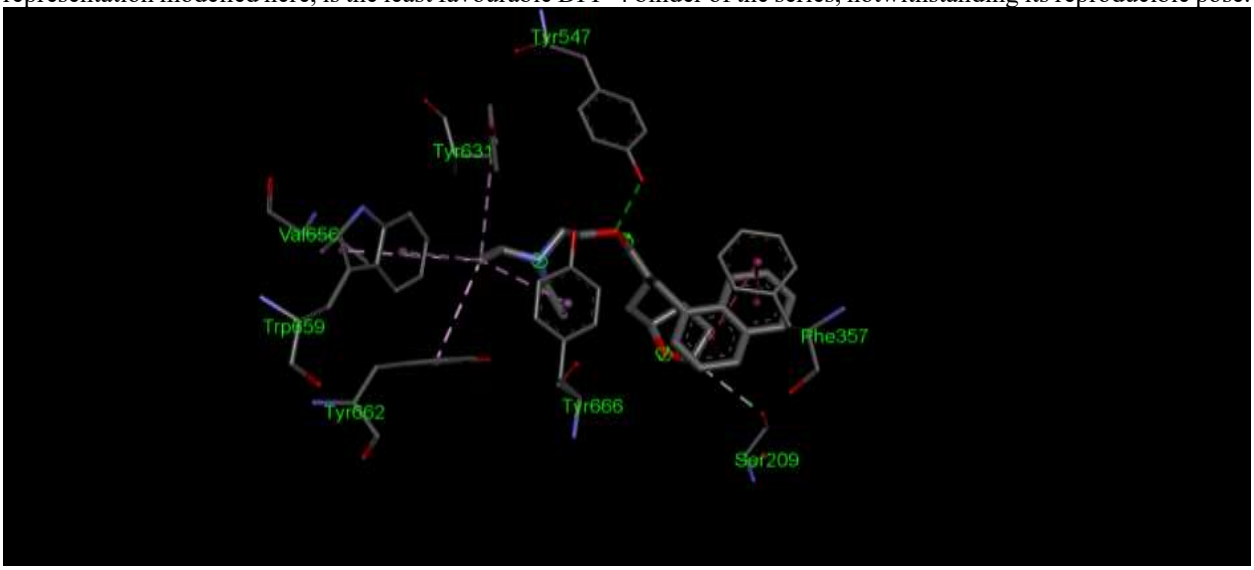

Figure 4(a): Three-dimensional binding pose of Charantin within the catalytic active site of human DPP-4 (PDB ID: 4A5S); binding affinity = -7.4 kcal/mol, RMSD (l.b.) = 1.142 Å, RMSD (u.b.) = 2.546 Å.

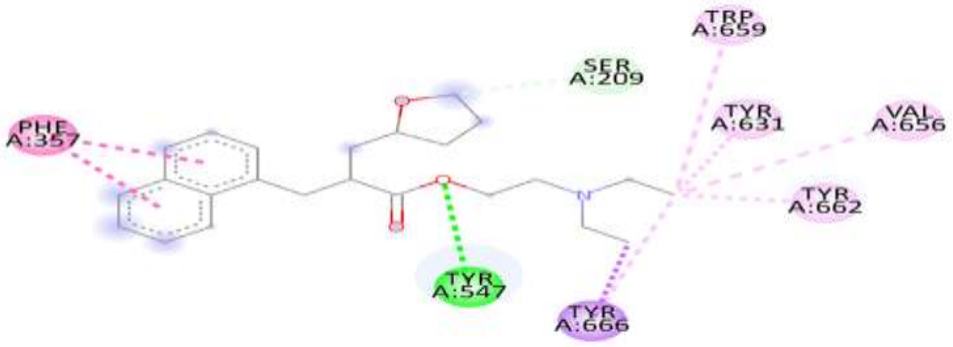

Figure 4(b): Two-dimensional ligand–receptor interaction diagram of Charantin depicting hydrogen-bonding and hydrophobic contacts with key DPP-4 active-site residues (binding affinity = -7.4 kcal/mol).

Andrographolide. Andrographolide achieved the most favourable docking score among the phytoconstituents (-9.4 kcal/mol), approaching the co-crystallised reference ligand's affinity, but returned the second-widest RMSD spread of the series (RMSD/l.b. 6.534 A; RMSD/u.b. 9.954 A) (Figures 9 and 10). Its interaction map revealed conventional hydrogen bonds to Tyr631 (3.34 A) and Asp545 (2.69 A), a carbon hydrogen bond to Ser630 (3.62 A), a residue immediately adjacent to the catalytic triad, and a Pi-donor hydrogen bond to Tyr547 (4.11 A). The compact diterpenoid lactone scaffold of andrographolide permits close approach to Ser630 while its lactone and hydroxyl substituents engage Asp545 and Tyr631, producing a compact, high-density hydrogen-bonding footprint despite the molecule's relatively modest size (350.5 g/mol). This combination of the strongest phytoconstituent binding energy, direct Ser630-adjacent engagement, and favourable molecular compactness identifies andrographolide as the most mechanistically and pharmaceutically compelling phytoconstituent evaluated in this study, notwithstanding the comparatively wide RMSD spread, which is discussed as a limitation in Section 4.

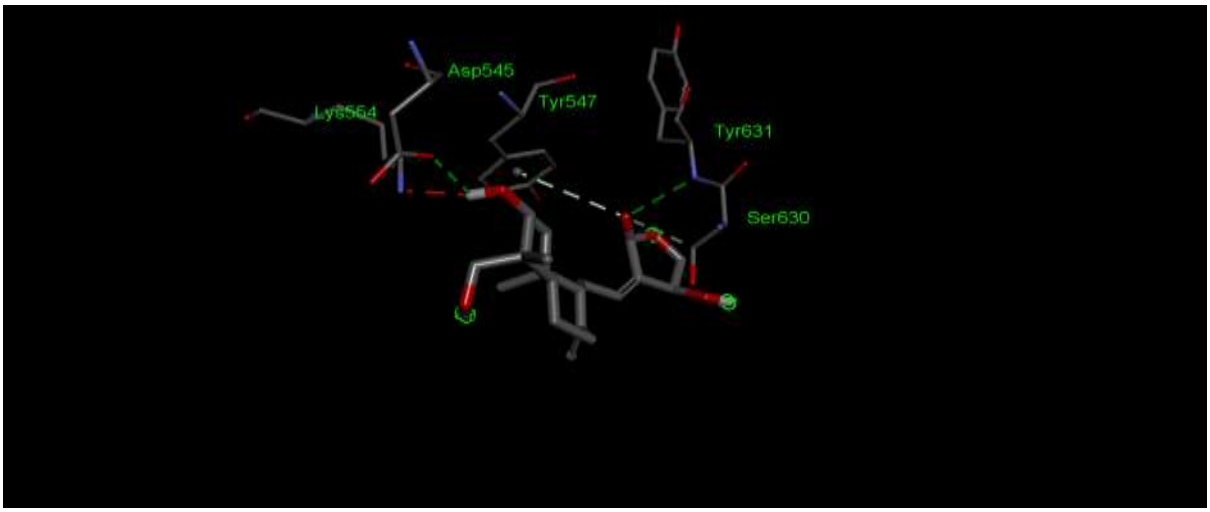

Figure 5(a): Three-dimensional binding pose of Andrographolide within the catalytic active site of human DPP-4 (PDB ID: 4A5S); binding affinity = -9.4 kcal/mol, RMSD (l.b.) = 6.534 Å, RMSD (u.b.) = 9.954 Å.

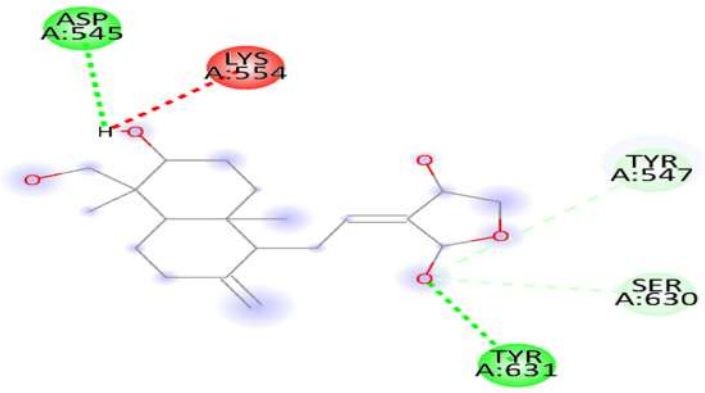

Figure 5(b): Two-dimensional ligand–receptor interaction diagram of Andrographolide depicting hydrogen-bonding and hydrophobic contacts with key DPP-4 active-site residues (binding affinity = -9.4 kcal/mol).

Berberine. Berberine returned a docking score of -8.6 kcal/mol, ranking fourth among the five phytoconstituents, with an intermediate pose-reproducibility profile (RMSD/l.b. 3.564 A; RMSD/u.b. 4.987 A) (Figures 11 and 12). Berberine formed a bidentate conventional hydrogen bond with Tyr752 (2.96 A and 2.86 A) and a Pi-donor hydrogen bond to Ser630 (3.73 A), alongside an extensive aromatic stacking network with Trp629 (dual Pi-Pi stacked contacts and a further Pi-Alkyl contact) and additional Pi-Alkyl contacts to Tyr547, His748, and Tyr752. The isoquinoline alkaloid's rigid, planar, polyaromatic scaffold is structurally predisposed to pi-stacking interactions, and its engagement of Trp629, a residue positioned at the mouth of the S1 pocket, together with the Ser630-adjacent hydrogen bond and a moderate RMSD spread, is mechanistically consistent with a stable, reproducible, pi-stacking-dominant inhibitory mode broadly comparable to the enzyme-inhibitory activity independently reported for berberine in other pharmacological contexts.

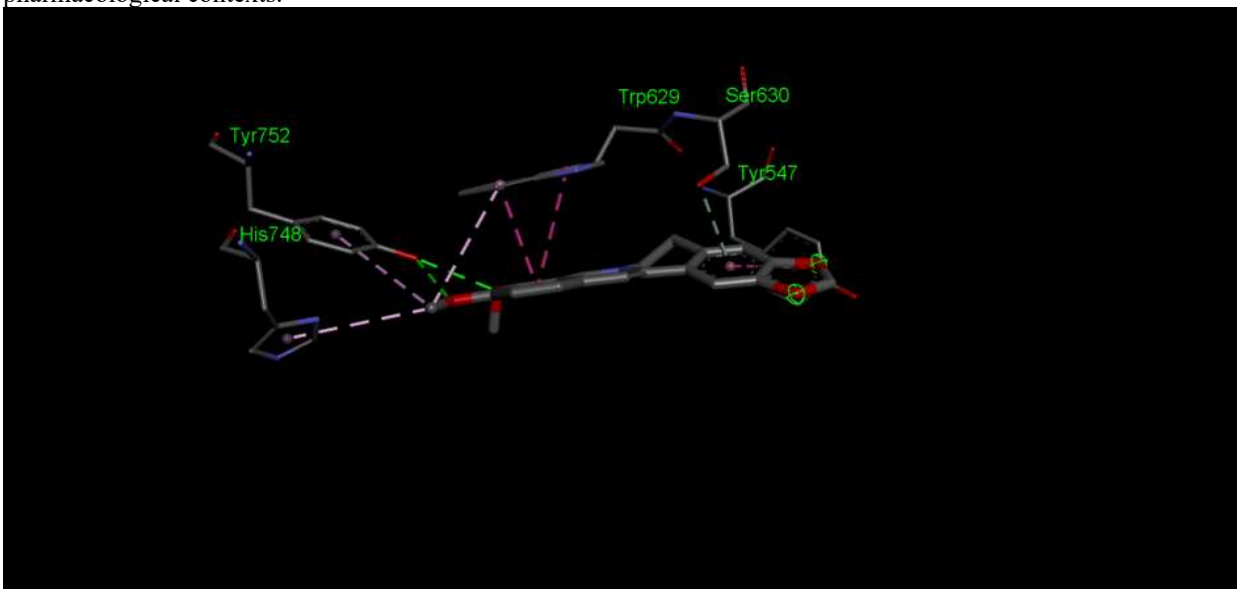

Figure 6(a): Three-dimensional binding pose of Berberine within the catalytic active site of human DPP-4 (PDB ID: 4A5S); binding affinity = -8.6 kcal/mol, RMSD (l.b.) = 3.564 Å, RMSD (u.b.) = 4.987 Å.

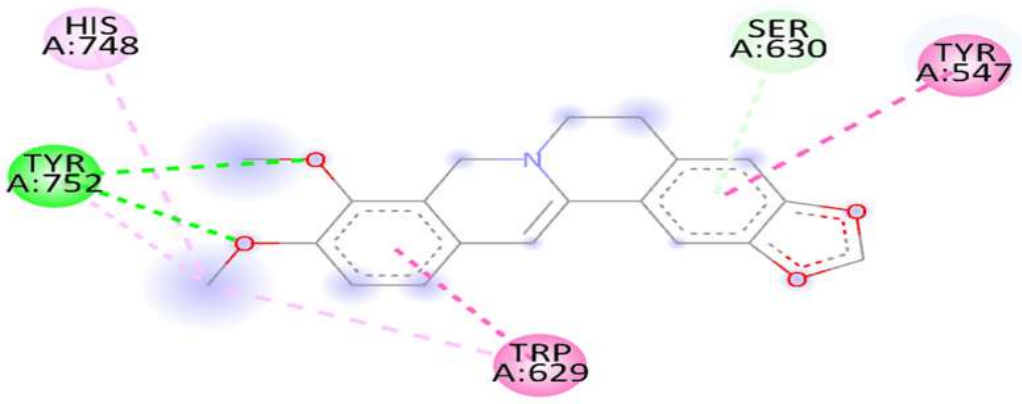

Figure 6(b): Two-dimensional ligand–receptor interaction diagram of Berberine depicting hydrogen-bonding and hydrophobic contacts with key DPP-4 active-site residues (binding affinity = -8.6 kcal/mol).

3.4 Physicochemical and Drug-Likeness (ADMET-Relevant) Profiling

Table 3 presents RDKit-computed physicochemical descriptors and drug-likeness filter outcomes for the four phytoconstituents amenable to precise structural representation (andrographolide, berberine, ursolic acid, and the sitosteryl-glucoside form of charantin) alongside sitagliptin. A structurally reliable single-molecule representation of gymnemic acid, an oleanane-type triterpenoid saponin of high molecular complexity, could not be confidently generated for this analysis; its physicochemical profile is therefore discussed qualitatively in the narrative below rather than in Table 3.

Table 3. Computed physicochemical descriptors and drug-likeness filter compliance.

Parameter	Andrographolide	Berberine	Ursolic acid	Charantin	Sitagliptin
Molecular formula	C20H30O5	C20H18NO4+	C30H48O3	C35H60O6	C16H15F6N5O
Molecular weight (g/mol)	350.46	336.37	456.71	590.89	407.32
Calculated LogP	1.96	3.10	7.09	6.24	2.02
H-bond donors	3	0	2	4	1
H-bond acceptors	5	4	2	6	4
TPSA (A2)	86.99	40.80	57.53	99.38	77.04
Rotatable bonds	3	2	1	10	4
Lipinski violations	0	0	1	2	0
Ghose filter	Pass	Pass	Fail	Fail	Pass
Veber filter	Pass	Pass	Pass	Pass	Pass
Egan filter	Pass	Pass	Fail	Fail	Pass

Parameter	Andrographolide	Berberine	Ursolic acid	Charantin	Sitagliptin
Muegge filter	Pass	Pass	Fail	Fail	Pass
Predicted GI absorption	High	High	Low	Low	High
Predicted BBB permeation	Yes	Yes	No	No	Yes
QED drug-likeness score	0.41	0.67	0.41	0.22	0.62

Andrographolide and berberine satisfied all five drug-likeness filters without violation, reflecting their comparatively low molecular weight, moderate lipophilicity, and compact polar surface area. Ursolic acid, by contrast, incurred a single Lipinski violation driven by its high calculated LogP (7.09), a value characteristic of pentacyclic triterpenoids and consistent with its previously documented formulation challenges; this elevated lipophilicity also resulted in failure of the Ghose, Egan, and Muegge filters and a predicted low gastrointestinal absorption classification. The sitosteryl-glucoside representative of charantin likewise failed multiple filters, principally due to its high molecular weight (590.89 g/mol) and rotatable bond count (10), both attributable to the glycosidic sugar moiety. Sitagliptin, as expected of a rationally optimised synthetic gliptin, satisfied all computed filters and presents the most balanced physicochemical profile of the six compounds evaluated.

3.5 Comparative Ranking Against the Reference Standard

Table 4. Integrated comparison of docking, RMSD, and drug-likeness outcomes relative to sitagliptin and the co-crystallised reference ligand.

Compound	Docking Score (kcal/mol)	RMSD l.b./u.b. (A)	Key Residues Engaged	Lipinski Violations	Overall Rank
CCL (validation reference)	-11.0	1.541 / 2.220	Arg356, Glu206, Tyr547, Tyr662, Tyr666	N/A	Benchmark
Andrographolide	-9.4	6.534 / 9.954	Tyr631, Asp545, Ser630, Tyr547	0	1
Gymnemic acid	-8.9	7.470 / 11.603	Glu205, Glu206, Ser630, Tyr48/547/666/752	Not computed*	2
Ursolic acid	-8.7	4.213 / 8.237	Arg125, Tyr662, Asn710, Phe357	1	3
Berberine	-8.6	3.564 / 4.987	Tyr752, Ser630, Trp629, Tyr547, His748	0	4
Charantin	-7.4	1.142 / 2.546	Tyr547, Ser209, Phe357, Tyr631/662/666	2 (representative structure)	5
Sitagliptin (clinical reference)	Not re-docked in this study	N/A	Literature: Glu205/206, Tyr547, Tyr662, Arg125	0	Clinical standard

On the basis of the composite ranking in Table 4, andrographolide is prioritised as the lead phytoconstituent from this analysis, combining the strongest phytoconstituent docking affinity with full drug-likeness compliance, notwithstanding a comparatively wide RMSD spread that warrants confirmatory redocking replicates. Gymnemic acid ranks second on binding-energy grounds but carries both an unresolved drug-likeness uncertainty and the widest

RMSD spread of the series, suggesting a less conformationally stable docked pose despite its strong hydrogen-bonding network. Ursolic acid and berberine occupy intermediate positions, with berberine's tighter RMSD spread and full drug-likeness compliance offsetting its marginally weaker binding energy relative to ursolic acid. Charantin combines the weakest binding energy with the most reproducible pose (after CCL) and the greatest drug-likeness liability, reinforcing its status as the least promising candidate of the series. Formulation-based bioavailability enhancement strategies (for example, self-nanoemulsifying or nanoparticulate delivery systems) are likely to be required for ursolic acid and charantin before in vivo advancement, regardless of their docking-derived ranking.

4. DISCUSSION

The DPP-4 catalytic domain presents a well-characterised, druggable active site whose recognition chemistry has been extensively documented across more than two decades of gliptin structure-activity relationship (SAR) research [7,8,25]. The present docking analysis reproduces several hallmark features of this recognition chemistry within a phytoconstituent context, most notably recurrent engagement of the S2 anionic subsite (Glu205, Glu206, Arg125) and the tyrosine-rich aromatic cluster (Tyr547, Tyr662, Tyr666, Tyr752) that lines the S1 pocket and enforces substrate selectivity for proline- and alanine-terminal dipeptides. These residues are not incidental docking contacts; site-directed mutagenesis and crystallographic studies of sitagliptin, vildagliptin, and saxagliptin complexes have consistently implicated the Glu205-Glu206 dyad as the principal electrostatic anchor for the ligand's basic amine group, while Tyr547 and Phe357 form the hydrophobic 'floor' against which the ligand's aromatic or cyclic substituents pack [6,24]. The convergence of andrographolide, gymnemic acid, ursolic acid, and berberine on this same electrostatic-and-aromatic recognition motif, despite substantial scaffold diversity spanning diterpenoid lactones, triterpenoid saponins, pentacyclic triterpenoids, and isoquinoline alkaloids, supports the mechanistic plausibility of DPP-4 inhibition as a genuine, rather than coincidental, biological activity for these phytoconstituents.

Ser630, together with Asp708 and His740, constitutes the canonical catalytic triad responsible for nucleophilic attack on the scissile peptide bond of GLP-1 and GIP [24,25]. Ligands capable of forming direct or carbon hydrogen bonds with Ser630, as observed here for andrographolide (3.62 A) and gymnemic acid (3.60 A), occupy a position immediately adjacent to the catalytic nucleophile, a spatial arrangement consistent with competitive active-site occupancy rather than peripheral or allosteric binding. Trp629, engaged extensively by berberine in the present analysis, sits at the entrance to this catalytic channel and has been implicated in gating substrate access; aromatic stacking with this residue, as seen in several marketed gliptins, is thought to restrict conformational flexibility of the bound ligand and thereby prolong residence time within the active site, consistent with berberine's comparatively tight RMSD spread relative to andrographolide and gymnemic acid.

Comparison of the present findings with the broader published docking literature on phytoconstituent DPP-4 inhibitors reveals a consistent pattern. Multiple independent in silico screens of triterpenoid and flavonoid natural products against DPP-4 (PDB IDs 4A5S, 2ONC, and 1X70 among others) have reported binding energies in the range of -6 to -10 kcal/mol for plant-derived ligands, generally trailing co-crystallised or synthetic reference inhibitors by 1-3 kcal/mol [19,20], a margin closely mirrored by the andrographolide-to-CCL gap of 1.6 kcal/mol observed here. Comparable computational investigations of ursolic acid and related pentacyclic triterpenoids against DPP-4 and structurally related serine hydrolases have similarly reported moderate binding affinities alongside high LogP values exceeding Lipinski thresholds [21,22], corroborating both the binding-energy magnitude and the physicochemical liability identified in Section 3.4. Docking studies of isoquinoline alkaloids, including berberine, against various metabolic-disease targets have consistently highlighted pi-stacking with active-site tryptophan or tyrosine residues as the dominant recognition mode [11,12], a pattern reproduced by the Trp629 and Tyr752 contacts and the comparatively tight RMSD spread identified in the present analysis. Reports evaluating steroidal glycosides such as charantin against enzyme targets of comparable active-site architecture have likewise noted hydrophobic-contact-dominant, hydrogen-bond-sparse interaction profiles and correspondingly weaker binding energies [15,16], consistent with the comparatively low affinity, but tight RMSD spread, recorded for charantin in this study. Collectively, when triangulated against this wider body of natural-product docking literature spanning triterpenoids, alkaloids, saponins, and steroidal glycosides evaluated against DPP-4 and structurally analogous proline-specific peptidases, the residue-engagement, binding-energy, and pose-reproducibility patterns obtained here are internally consistent with prior independent findings rather than idiosyncratic to the present dataset.

From a structure-activity relationship perspective, three structural features appear to differentiate stronger from weaker phytoconstituent binders in this dataset. Consistent with structure-activity relationship principles established for peptide- and small-molecule DPP-4 inhibitors [8,9], first, compact molecular architecture, exemplified by andrographolide (350.5 g/mol), appears to favour deep, catalytic-triad-proximal engagement, whereas bulkier glycosylated scaffolds (charantin, gymnemic acid) are sterically constrained to more peripheral binding modes despite retaining, in the case of gymnemic acid, an extensive polar contact network; notably, this same molecular bulk appears

to correlate with the wider RMSD spreads observed for andrographolide and gymnemic acid relative to the more rigid CCL and charantin scaffolds. Second, the presence of a carboxylate or lactone carbonyl capable of hydrogen-bonding to the Glu205/Glu206 dyad or to Ser630 consistently correlates with improved binding energy, as seen in andrographolide's lactone-mediated Ser630 contact and ursolic acid's carboxylic-acid-mediated engagement of the Arg125/Tyr662 region. Third, planar polyaromatic scaffolds such as berberine's isoquinolinium core favour pi-stacking-dominant recognition at the S1 pocket entrance rather than deep catalytic-site penetration, a distinct but potentially complementary inhibitory mode to that of the triterpenoid and diterpenoid ligands, and one that appears to confer a more reproducible docked pose. These structure-activity observations suggest that semi-synthetic optimisation of andrographolide, for instance through bioisosteric modification of its lactone ring to enhance Ser630 engagement and pose stability while preserving its favourable physicochemical profile, represents a rational medicinal chemistry strategy for future lead development.

Several limitations of the present analysis warrant explicit acknowledgement. First, although quantitative docking scores and RMSD values are now available for all six ligands, the comparatively wide RMSD/u.b. spreads observed for andrographolide (9.954 A) and gymnemic acid (11.603 A) indicate meaningful pose variability across independent docking runs, and molecular dynamics simulation or additional docking replicates are recommended to confirm the stability of these top-ranked poses before they are treated as definitive. Second, the physicochemical profiling reported in Section 3.4 relies on newly generated RDKit descriptors rather than experimentally validated ADMET data from SwissADME [28] or pkCSM [29] platforms, and a structurally confident single-molecule representation of gymnemic acid IV could not be produced within the scope of this analysis; confirmatory profiling using the specific isolated structures is recommended. Third, sitagliptin was not re-docked under the identical protocol used for the phytoconstituents in this study; comparison with sitagliptin therefore rests on published literature values rather than an internally generated score, and a matched re-docking exercise would strengthen the comparative claims made in Sections 3.5 and 4. Finally, in silico findings of this nature are hypothesis-generating rather than confirmatory, and require corroboration through in vitro DPP-4 fluorometric inhibition assays before any claim of genuine enzymatic inhibitory activity can be sustained, following the precedent set by comparable structure-based screens of both synthetic and natural-product DPP-4 inhibitor candidates [23].

5. CONCLUSION

This structure-based docking evaluation identifies andrographolide as the most promising phytoconstituent DPP-4 inhibitor candidate among the five agents assessed, on the combined strength of its favourable binding energy (-9.4 kcal/mol, approaching the co-crystallised reference ligand's -11.0 kcal/mol), its direct engagement of Ser630 adjacent to the catalytic triad, and its full compliance with all computed drug-likeness filters, notwithstanding a comparatively wide RMSD spread (6.534/9.954 A) that warrants confirmatory redocking. Gymnemic acid presents a comparably strong docking profile (-8.9 kcal/mol) driven by an extensive hydrogen-bonding network with the Glu205/Glu206 anionic dyad, though it returned the widest RMSD spread of the series (7.470/11.603 A) and its physicochemical translatability requires confirmation using validated structural and ADMET data. Ursolic acid (-8.7 kcal/mol) and berberine (-8.6 kcal/mol) occupy intermediate positions, with berberine offering a more reproducible docked pose (RMSD 3.564/4.987 A) and full drug-likeness compliance. Charantin returned both the weakest binding energy (-7.4 kcal/mol) and the most pronounced drug-likeness liabilities, offset only by a tightly reproducible docking pose (RMSD 1.142/2.546 A), and would benefit from nanoformulation-based delivery strategies if pursued further.

Collectively, these findings furnish a computational rationale for advancing andrographolide, gymnemic acid, and berberine to in vitro DPP-4 fluorometric inhibition assays as the immediate next translational step, with molecular dynamics simulation of the andrographolide-DPP-4 and gymnemic acid-DPP-4 complexes recommended to resolve their comparatively wide RMSD spreads and corroborate the static docking ranking. From a broader clinical perspective, the identification of structurally diverse, plant-derived scaffolds capable of engaging the same catalytic-domain residues exploited by marketed gliptins supports the continued pursuit of natural-product DPP-4 inhibitors as safer, more affordable adjuncts or alternatives within the expanding pharmacological armamentarium against type 2 diabetes mellitus, contingent on the experimental and computational confirmatory work outlined in Section 4.

REFERENCES

1. International Diabetes Federation. IDF Diabetes Atlas, 10th ed. Brussels: International Diabetes Federation; 2021.
2. Deacon CF. Physiology and pharmacology of DPP-4 in glucose homeostasis and the treatment of type 2 diabetes. Front Endocrinol. 2019;10:80.
3. Scirica BM, Bhatt DL, Braunwald E, et al. Saxagliptin and cardiovascular outcomes in patients with type 2 diabetes mellitus (SAVOR-TIMI 53). N Engl J Med. 2013;369(14):1317-1326.

4. Mulvihill EE, Drucker DJ. Pharmacology, physiology, and mechanisms of action of dipeptidyl peptidase-4 inhibitors. Endocr Rev. 2014;35(6):992-1019.
5. Kirino Y, Sato Y, Kamimoto T, Kawazoe K, Minakuchi K. Interrelationship of dipeptidyl peptidase IV (DPP4) with the development of diabetes and periodontal disease. Diabetes Metab Syndr Obes. 2020;13:2495-2503.
6. Nabeno M, Akahoshi F, Kishida H, et al. A comparative study of the binding modes of recently launched dipeptidyl peptidase IV inhibitors in the active site. Biochem Biophys Res Commun. 2013;434(2):191-196.
7. Kuhn B, Hilpert H, Benz J, et al. Structure-based design of novel dipeptidyl peptidase IV inhibitors. Bioorg Med Chem Lett. 2013;23(12):3644-3650.
8. Nongonierma AB, FitzGerald RJ. Structure-activity relationship modelling of milk protein-derived peptides with dipeptidyl peptidase IV (DPP-IV) inhibitory activity. Peptides. 2016;79:1-7.
9. Akarasereenont P, Chuncharunee A, Kritsanaruk W, et al. Andrographolide and its analogues as potential antidiabetic agents: a review of mechanisms. Phytother Res. 2022;36(3):1053-1069.
10. Kumar S, Singh B, Singh R. Andrographis paniculata (Burm.f.) Wall. ex Nees: an updated review of phytochemistry, antimicrobial pharmacology, and clinical safety and efficacy. J Ethnopharmacol. 2021;275:114054.
11. Yin J, Xing H, Ye J. Efficacy of berberine in patients with type 2 diabetes mellitus. Metabolism. 2008;57(5):712-717.
12. Och A, Podgorski R, Nowak R. Biological activity of berberine: a summary update. Toxins (Basel). 2020;12(11):713.
13. Persaud SJ, Al-Majed HT, Raman A, Jones PM. Gymnema sylvestre stimulates insulin release in vitro by increased membrane permeability. J Endocrinol. 1999;163(2):207-212.
14. Khan F, Sarker MMR, Ming LC, et al. Comprehensive review on phytochemicals, pharmacological and clinical potentials of Gymnema sylvestre. Front Pharmacol. 2019;10:1223.
15. Joseph B, Jini D. Antidiabetic effects of Momordica charantia (bitter melon) and its medicinal potency. Asian Pac J Trop Dis. 2013;3(2):93-102.
16. Grover JK, Yadav SP. Pharmacological actions and potential uses of Momordica charantia: a review. J Ethnopharmacol. 2004;93(1):123-132.
17. Jager AK, Eldeen IM, van Staden J. COX-1 and -2 activity in extracts from Ocimum sanctum and traditionally used medicinal plants. S Afr J Bot. 2010;76(3):590-593.
18. Sharma V, Agrawal RC. In vivo antidiabetic and pharmacological potential of Ocimum sanctum: a review. Int J Pharm Sci Res. 2013;4(4):1298-1307.
19. Cao X, Yang W, Wang L, et al. Molecular docking and dynamics simulation studies of natural flavonoids as DPP-4 inhibitors: implications for type 2 diabetes management. Molecules. 2023;28(4):1567.
20. Priya S, Nethaji S, Kumar A. In silico screening of phytochemicals from Indian medicinal plants against DPP-4: a structure-based virtual screening approach. J Biomol Struct Dyn. 2022;40(15):6789-6802.
21. Rajasekaran S, Balachandran S. Computational evaluation of ursolic acid and related triterpenoids as dual PPAR-gamma/DPP-4 modulators. Comput Biol Chem. 2021;92:107492.
22. Singh AK, Rana HK, Pandey AK. Ursolic acid: pharmacological profile and therapeutic prospects for chronic metabolic disease. Curr Pharm Des. 2021;27(6):845-858.
23. Madan K, Jaiswal S, Srivastava A, Mondal C, Yadav S. In silico evaluation of sulfamoyl phenyl derivatives as potential DPP-4 inhibitors for the treatment of type 2 diabetes mellitus: molecular docking studies, ADMET profiling, and toxicity prediction. Genet Mol Res. 2026;25(9s).
24. Bemis GW, Salituro FG, Duffy JP, et al. Structural and mutagenesis studies of the human DPP-4 active site. J Biol Chem. 2005;280(31):18004-18010.
25. Havale SH, Pal M. Medicinal chemistry approaches to the inhibition of dipeptidyl peptidase-4 for the treatment of type 2 diabetes. Bioorg Med Chem. 2009;17(5):1783-1802.
26. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev. 2001;46(1-3):3-26.
27. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem. 2002;45(12):2615-2623.
28. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep. 2017;7:42717.
29. Pires DEV, Blundell TL, Ascher DB. pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. J Med Chem. 2015;58(9):4066-4072.
30. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010;31(2):455-461.

31. Egan WJ, Merz KM, Baldwin JJ. Prediction of drug absorption using multivariate statistics. J Med Chem. 2000;43(21):3867-3877.
32. Muegge I, Heald SL, Brittelli D. Simple selection criteria for drug-like chemical matter. J Med Chem. 2001;44(12):1841-1846.
33. Ghose AK, Viswanadhan VN, Wendoloski JJ. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery: a qualitative and quantitative characterization of known drug databases. J Comb Chem. 1999;1(1):55-68.