Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Target Fishing and Molecular Docking of Terpenes: A Systematic Evaluation of Method Concordance and Ligand Efficiency

Manasa Hegde
University of Alabama in Huntsville
Armin Ahmadi
University of Alabama in Huntsville
Daniel McShan
Terpedia, LLC
Susan Trapp
Terpedia, LLC
Jerome Baudry
jerome.baudry@uah.edu

University of Alabama in Huntsville

Research Article

Keywords: Target fishing, Molecular docking, Terpenes, Proteins

Posted Date: November 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11236711/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: Competing interest reported. Susan Trapp and Daniel McShan are affiliated with Terpedia, LLC, which develops terpene-focused scientific information resources. The other authors declare no competing interests.

Abstract

Terpenes are naturally occurring small molecules with diverse chemical structures and potential therapeutic applications. Despite their wide-ranging reported bioactivities including antimicrobial, anti-inflammatory, analgesic, and anticancer properties, the molecular targets underlying these effects remain poorly characterized for most terpenes. Here, we apply a combined *in silico* pipeline, ligand-based target fishing using four complementary tools, followed by molecular docking, to 23 terpenes spanning monoterpene and sesquiterpene classes, with the dual goals of prioritizing candidate protein targets and benchmarking the agreement between the two computational methods. Consensus target fishing across four tools identified five relatively high-frequency protein targets: cannabinoid receptor 2 (CB2), androgen receptor (AR), 11β-hydroxysteroid dehydrogenase 1 (11β-HSD1), cytochrome P450 19A1 (CYP19A1), and estrogen receptor β (ERβ). Molecular docking against these targets revealed that known high-affinity ligands consistently outscored terpenes for four of five targets, while several terpenes exceeded known-ligand scores for CYP19A1. Critically, the correspondence between target fishing predictions and docking performance was inconsistent: target-fishing-positive terpenes did not reliably rank higher in docking than target-fishing-negative ones, and enrichment analysis confirmed that docking scores carried limited discriminatory power in this chemical series. These findings describe both the promise and the current limitations of integrated in silico approaches for terpene pharmacology and provide a ranked set of experimentally testable terpene-target interactions. We conclude that it is imperative to develop, curate, and distribute more data on terpenes to enable the pharmacological development of that rich class of natural compounds.

Introduction

Terpenes represent the most chemically and structurally diverse family of natural products [1, 2]. They are synthesized as secondary metabolites predominantly in plants, where they constitute the major fraction of essential oils, and their structural diversity, spanning monoterpenes, sesquiterpenes, diterpenes, and beyond, arises from the combinatorial cyclization and rearrangement of isoprene-derived precursors [3]. The term “terpenoid” formally refers to oxygenated or otherwise functionalized terpene derivatives, though the two terms are used interchangeably in much of the literature [2, 4, 5]; in this paper, “terpene” is used in the broader sense, encompassing both.

Terpenes have attracted sustained pharmacological interest due to their reported antimicrobial, anti-inflammatory, antioxidant, analgesic, and anticancer activities [6, 7]. Well-known examples include α-pinene, with documented antitumor and analgesic properties [8, 9], and paclitaxel (Taxol), a diterpene-derived chemotherapeutic agent originally isolated from the bark of the Pacific yew (*Taxus brevifolia*), that remains a frontline treatment for breast, ovarian, and lung cancers [10, 11]. More broadly, terpenes have been investigated as modulators of inflammatory pathways, acting on pro-inflammatory cytokines, membrane receptors, and transcription factors, and as potential leads for a wide range of therapeutic indications [4, 12, 15]. Yet despite this extensive body of evidence, a fundamental problem persists for most terpenes, it is not known which specific protein targets are responsible for the observed effects,

nor how terpenes interact with those targets at the molecular level [13, 14]. This gap between reported bioactivity and mechanistic understanding severely limits the ability to rationally develop terpene-based therapeutics or to predict their polypharmacological profiles and side-effect liabilities.

Two computational approaches are commonly applied to the problem of target identification for small molecules. Ligand-based target fishing (TF) identifies candidate protein targets by comparing a query molecule against databases of small molecules with known binding partners, exploiting the principle that structurally similar ligands tend to share targets [16–18]. Molecular docking evaluates the geometric and energetic compatibility of a ligand within a protein's binding site and returns a score intended to rank binding likelihood [19–21]. Each method has well-documented limitations even in conventional drug discovery settings. TF predictions depend entirely on the coverage and composition of the underlying reference databases and are unreliable for compound classes that are poorly represented therein; natural products, including terpenes, fall squarely into this category [17]. Docking scores correlate imperfectly with experimental binding affinities, are sensitive to protein structure preparation, and are known to be influenced by molecular size independently of binding quality [19, 20]. Whether applying TF and docking in combination improves target identification over either method alone, and whether such a pipeline is useful at all for terpenes, which are small, hydrophobic, and structurally distinct from typical drug-like molecules, remains an open question.

Here, we address this gap using a set of 23 terpenes, comprising of 18 monoterpenes and 5 sesquiterpenes, selected based on their prevalence in plants and their frequent occurrence in terpene profiles reported for commercial cannabis samples. We apply four complementary ligand-based TF tools (SwissTargetPrediction, SEA, STarFish, and PPB2), adopt a consensus criterion requiring prediction by at least two tools, and then perform GBVI/WSA molecular docking of all 23 terpenes against the five most frequently predicted consensus targets using the Molecular Operating Environment (MOE) platform. Docking is benchmarked against known high- and low-affinity ligands for each target drawn from the IUPHAR/BPS Guide to PHARMACOLOGY database. We then explicitly compare TF predictions and docking rankings using enrichment analysis to ask: does docking preferentially recover TF positive terpenes? The results reveal characteristic agreements and disagreements between the two methods, shed light on the structural and algorithmic factors that drive these discrepancies, and yield a curated set of terpene-target interactions prioritized for experimental follow-up. Two prior studies have applied analogous TF–plus-docking pipelines to individual bioactive molecules, glyphosate [48] and capsaicin [49], but a systematic evaluation of the method's coherence and limitations across a chemically diverse natural product class has not been reported. This work provides such an evaluation for terpenes.

Methodology

Target fishing

Four publicly available ligand-based target fishing (TF) tools were used: SwissTargetPrediction (STP) [23], Similarity Ensemble Approach (SEA) [24], Stacked Target Fishing (STarFish) [25], and

Polypharmacology Browser (PPB2) [26]. All share the same operating principle: the query compound is compared against a database of known bioactive ligands using molecular fingerprints, and proteins whose known ligands are most similar to the query are returned as predicted targets. The tools differ, however, in the fingerprint types used, the statistical framework applied, and the reference database queried (Table 1), so their predictions are partially independent. To reduce the influence of any single tool's database biases, we retained only consensus targets, defined as proteins predicted by at least two of the four tools for a given terpene. When a tool returned no predictions for a given terpene (see Section 3.1), that tool was excluded from the denominator for that compound, and the ≥ 2 threshold was applied to the tools that did return results. This means that consensus strength is not uniform across the dataset: a prediction supported by 2 out of 4 tools (2/4, 50% agreement across methodologically diverse approaches) is not equivalent to one supported by 2 out of 2 tools (2/2, 100% agreement but from fewer and potentially less independent sources). Predictions resting on only two tools, particularly where both are machine-learning classifiers, should be interpreted as lower-confidence consensus. These consensus targets were then pooled across a set of 23 terpenes, comprising 18 monoterpenes and 5 sesquiterpenes, and ranked by prediction frequency; the five most frequently predicted proteins were selected for molecular docking (Table 2). The terpene structures were provided as SMILES strings [22] to each tool.

The four tools differ in the fingerprints and statistical frameworks they employ. STP was queried with Homo sapiens as the target species. It represents each molecule with both a 2D Tanimoto similarity over atom-connectivity path fingerprints and a 3D similarity computed as the Manhattan distance between ElectroShape 5D vectors; a multiple logistic regression model combines the two measures to score each candidate target. SEA aggregates pairwise similarities across Morgan, atom-pair, topological, MACCS keys, and pharmacophore fingerprints into a raw score per target, converts it to a z-score against a background distribution of chemically diverse ligands, and reports a p-value whose significance indicates likely interaction. STarFish is a machine-learning framework designed for natural products: Morgan functional-class fingerprints serve as input features for an ensemble of k-nearest neighbor, random forest, and multilayer perceptron classifiers, operated via command-line scripts. Because STarFish ranks all 1,907 targets in its database for every query compound, a score threshold of ≥ 0.03 was applied to define a positive.

To identify common targets across tools, ChEMBL IDs were used as a shared identifier. STP and PPB2 report ChEMBL IDs directly in their output. For SEA and STarFish, which report target keys and UniProt IDs respectively, ChEMBL IDs were retrieved via the UniProt ID mapping tool (UniProtKB AC/ID → ChEMBL). Targets predicted by at least two tools for a given terpene were then identified by matching ChEMBL ID across all four outputs.

PubChem and the target fishing tools employed in this study were accessed in September 2025, and all analyses were conducted during that time.

Table 1
Details of the fingerprints, methods, and URL of the target fishing tools

TF Tool	Fingerprint	Method	URL
STP	Atom connectivity paths (2D) and Electro Shape 5D based on physicochemical properties (3D)	Uses 2D and 3D fingerprints combined through multiple logistic regression to search for similar molecules	http://www.swisstargetprediction.ch/
SEA	Morgan, atom pair, topological, MACCS keys, and pharmacophore-based fingerprints	Uses fingerprints and ranks predicted proteins based on a statistically significant p-value	https://sea16.docking.org/
PPB2	Xfp, MQN, and ECfp4 fingerprints	Uses molecular fingerprints through user-selected NN similarity searches, NB models, or a combined NN + NB approach	https://ppb2.gdb.tools/
STarFish	Morgan functional-class connectivity fingerprint	Uses a combination of k-nearest neighbors, random forest, and multilayer perceptron	https://github.com/ntcockroft/STarFish

Molecular docking

Molecular docking was performed using the Molecular Operating Environment (MOE, Chemical Computing Group) [47].

Ligand preparation. Terpene and known-ligand structures were retrieved from PubChem [29] as three-dimensional conformers in SDF format. Structures were imported into MOE and processed using the Compute / Wash function, which standardizes protonation states, adds explicit hydrogens, and removes counterions, using default parameters.

Protein preparation. For each of the five consensus target proteins, a crystal structure was retrieved from the Protein Data Bank (PDB) [46], selecting among available structures based on resolution, structural completeness, and the presence of a co-crystallized ligand (Table 2). When the PDB entry contained multiple copies of the target chain or unrelated protein chains, a single appropriate copy was selected, and all extraneous chains and non-essential components were removed. Protein structures were then prepared using the MOE QuickPrep module with default parameters, which assigns protonation states, adds missing hydrogens, and performs energy minimization to a minimum gradient

cutoff. The binding site was identified using the MOE SiteFinder function, and the resulting dummy atoms were used to define the docking site.

Docking protocol. Docking was performed in two sequential stages. In the first stage, placement used the Triangle Matcher algorithm with the London dG scoring function, retaining 15 poses per ligand. In the second stage, the top poses were subjected to induced-fit refinement using the GBVI/WSA ΔG scoring function, retaining 5 refined poses per ligand. The resulting docked complexes were further processed using the MOE SVL script dock_refine_advanced.svl, with the GBVI/WSA dG scoring function selected and all output options enabled (GBVI/PBSA, strain, solvation, affinity, and PDB export); all other parameters were kept at default settings. The pose with the most negative MM/PBSA scores among the five refined poses was retained as the representative binding pose for each compound-protein pair, and this score was used for all subsequent analyses. Known ligands for each target were selected as described in the following section and were docked using the identical protocol; their scores serve as a benchmark for interpreting terpene docking results.

Identification of known ligands for target proteins

To provide a reference frame for the terpene docking scores, known ligands were selected for each target protein from two sources. First, the co-crystallized ligand present in the PDB structure used for docking was included. Second, the highest-affinity known ligand (HAKL) and lowest-affinity known ligand (LAKL) for each protein were retrieved from the IUPHAR/BPS Guide to PHARMACOLOGY (GtoPdb) [52], a curated database of quantitative ligand-target interactions. Binding affinity was quantified as pK_i (-log10 K_i) for four of the five proteins; pIC50 (-$\log_{10}$ IC50) was used for the remaining protein, as K_i data were not available. Together, the co-crystallized ligand, HAKL, and LAKL bracket the affinity range of well-characterized binders and provide the upper and lower scoring benchmarks against which the terpene docking results are evaluated.

Results and Discussion

Target fishing of selected terpenes

Target fishing was carried out on a set of 23 terpenes selected for their widespread occurrence in plants and reported presence in cannabis, representing a subset of the more than 120 terpenes [27, 28]. The chemical structures of these compounds were retrieved from PubChem [29] and are shown in Fig. 1. The dataset comprised 18 monoterpenes and 5 sesquiterpenes, representing two structurally distinct terpene classes.

The complete target fishing results for each terpene across all target fishing tools are provided in Online Resource 1. Target proteins predicted by two or more target fishing tools were subsequently selected based on consensus and are listed in Online Resource 2. However, tool participation was uneven across the dataset: SEA returned no predictions for 12 of 23 terpenes (52%), and STP returned no predictions

for 2 terpenes (sabinene and β-pinene); for these two compounds, consensus rested entirely on PPB2 and STarFish, both machine-learning based tools. Only 11 of 23 terpenes (48%) received predictions from all four tools. As described in Section 2.1, the ≥ 2 threshold was applied to tools that returned results, so for sabinene and β-pinene the effective consensus is 2/2 (complete agreement among available tools, but from only two methodologically similar ML-based sources) rather than the 2/4 or better support available for other terpenes. A 2/2 consensus from two tools of the same class provides less independent evidence than a 2/4 consensus across mechanistically diverse methods. This uneven participation means that the consensus criterion does not represent equal evidence across terpenes and should be interpreted with corresponding caution.

Selection of target proteins for docking

From the consensus derived target set, the top five target proteins predicted across the terpenes were selected for docking analysis, as summarized in Table 2. The ranking by prediction frequency identified CB2 (18 terpenes), AR (15), CYP19A1 (13), ERβ (11), and 11β-HSD1 (10) as the top five. The cutoff at five is pragmatic: the gap between the fifth-ranked protein (11β-HSD1, 10 terpenes) and the sixth (serotonin transporter, 9 terpenes) is a single terpene, so no natural break in the data justifies this threshold. The five targets were selected to make the study tractable and to focus on proteins with well-characterized ligand sets suitable for benchmarking. During the compilation of target proteins, different TF tools occasionally provided distinct PDB IDs for the same protein; in those cases, the structure with higher resolution, greater completeness, and a co-crystallized ligand was preferred.

Table 2
Consensus derived target proteins selected for docking analysis, ranked by the number of terpenes predicting each target protein.

Predicted target	PDB ID	UniProt ID	Number of terpenes predicting
Cannabinoid receptor 2 (CB2)	6KPF	P34972	18
Androgen Receptor (AR)	3V49	P10275	15
Cytochrome P450 19A1 (CYP19A1)	3EQM	P11511	13
Estrogen receptor β (ERβ)	1L2J	Q92731	11
11-β-hydroxysteroid dehydrogenase 1 (11β-HSD1)	1XU7	P28845	10

Known ligands

Table 3 lists the ligands known to bind to target proteins selected for docking based on the criteria described in the Methods. The set consisted of a co-crystallized ligand, the highest affinity known ligand, and the lowest affinity known ligand for each protein.

Table 3
Known ligands selected for the target proteins, including ligand type, affinity, PubChem CID, and PubChem name.

Protein	Ligand type	PubChem name	PubChem CID
CB2	Co-crystallized ligand	E3R	145915889
	Highest affinity known ligand (pKi = 10.12)	AM10257	137321161
	Lowest affinity known ligand (pKi = 5.1)	Ibipinabant	9826744
AR	Co-crystallized ligand	CHEMBL2177238	24811980
	Highest affinity known ligand (pKi = 9.6)	Testosterone propionate	5995
	Lowest affinity known ligand (pKi = 5.43)	Flutamide	3397
11β-HSD1	Co-crystallized ligand	CHAPS	107670
	Highest affinity known ligand (pIC50 = 8.7)	AZD4017	24946280
	Lowest affinity known ligand (pIC50 = 7.54)	Enoxolone	10114
CYP19A1	Co-crystallized ligand	Androstenedione	6128
	Highest affinity known ligand (pKi = 10.7)	Letrozole	3902
	Lowest affinity known ligand (pKi = 4.46)	Testolactone	13769
ERβ	Co-crystallized ligand	R, R-THC	446849
	Highest affinity known ligand (pKi = 10.6)	4-hydroxytamoxifen	449459
	Lowest affinity known ligand (pKi = 6.39)	Ospemifene	3036505

Molecular docking

Molecular docking was performed for the selected five target proteins listed in Table 2, with 23 terpenes shown in Fig. 1 and known ligands listed in Table 3. The terpenes for which a target protein was predicted by target fishing were classified as TF+ terpenes, and those for which none of the target fishing tools predicted a target protein were classified as TF− terpenes. Both TF − and TF+ terpenes were used for docking to analyze their interactions in comparison to known ligands.

Cannabinoid receptor 2

Among the consensus targets for the terpenes, cannabinoid receptor 2 (CB2) emerged as the most frequently predicted, with 18 out of 23 terpenes predicting it as a potential target protein with high

consensus. CB2 is a G protein-coupled receptor that belongs to the cannabinoid receptor family, predominantly expressed in the immune system and peripheral tissues. CB2 is found in immune cells, including B cells, T cells, and macrophages, and plays an essential role in modulating immune responses, inflammation, and pain perception. CB2 is a promising therapeutic target for various diseases, including autoimmune diseases, chronic pain, and cancer [30, 31]. The docking scores of known ligands and monoterpenes with CB2 are listed in Table 4. AM10257 demonstrated the best docking score of -75.01 kcal/mol, followed by -74.57 kcal/mol in Ibipinabant and − 72.67 kcal/mol in E3R, as shown in Table 4. In radioligand binding studies of HEK cells by Finlay D B et al. (2020) [35], terpenes from cannabis, such as limonene, myrcene, α-pinene, and β-pinene, did not interact with CB2 receptors. The docking scores of these same terpenes in the present study are among the lower values in the monoterpene series, which is consistent with the experimental result, although agreement on this small dataset cannot be taken as validation: raw docking scores at this scale cannot define a reliable threshold separating binders from non-binders.

Table 4
Docking scores of monoterpenes and known ligands for CB2, ranked by docking score.

Ligand	Docking score (kcal/mol)
AM10257[a]	-75.01
Ibipinabant[a]	-74.57
E3R[a]	-72.67
Linalool	-40.75
Bornyl acetate	-39.12
Terpineol	-37.73
Eucalyptol	-36.82
Perillyl alcohol	-36.58
Fenchol	-32.93
γ-Terpinene	-32.8
β-Myrcene	-32.72
β-Ocimene	-31.95
(S)-Carvone	-31.95
(R)-Carvone	-31.46
p-Cymene	-30.65
Borneol	-30.61
Terpinolene	-30.29
Limonene	-30.19
Sabinene	-29.79
α-Pinene	-29.34
β-Pinene	-27.58

Unmarked: TF+ terpene; [a]known ligand.

Table 5 shows the docking scores of known ligands along with sesquiterpenes for CB2. Known ligands showed the best docking scores, followed by nerolidol at -50.41 kcal/mol, α-bisabolol at -49.06 kcal/mol, and humulene at -44 kcal/mol. β-caryophyllene showed the lowest docking score of -42.01 kcal/mol among sesquiterpenes. Gertsch J et al. (2008) found β-caryophyllene to be a full agonist at CB2 [32], and

its anti-inflammatory and analgesic effects in vivo have been confirmed in mice [33, 34]. This experimental evidence supports β-caryophyllene at CB2 specifically and independently of the docking result; however, the literature finding cannot be used to set a score threshold for the remaining sesquiterpenes, as the relationship between docking score magnitude and binding probability is not calibrated at this dataset size.

Table 5
Docking scores of sesquiterpenes and known ligands for CB2, ranked by docking score.

Ligand	Docking score (kcal/mol)
AM10257[a]	-75.01
Ibipinabant[a]	-74.57
E3R[a]	-72.67
Nerolidol	-50.41
α-Bisabolol	-49.06
Humulene	-44
Caryophyllene oxide	-42.5
β-Caryophyllene	-42.01

Unmarked: TF+ terpene; [a]known ligand.

Androgen receptor

Androgen receptor (AR) emerged as one of the frequently predicted targets, with 15 out of 23 terpenes predicting it as a potential target protein with high consensus. AR plays a key role in modulating the expression of male phenotype and is crucial in maintaining male reproductive health. It is significant in bone and muscle development [36]. The docking scores of known ligands and monoterpenes with AR are listed in Table 6. Known ligands binding to AR showed significantly higher docking scores compared to monoterpenes. The docking score of the co-crystallized ligand was − 80.20 kcal/mol. Flutamide, a LAKL, showed a docking score of -66.36 kcal/mol, which is better than that of HAKL, testosterone propionate, with a docking score of -62.14 kcal/mol.

Perillyl alcohol showed a docking score of -41.27 kcal/mol, while previous studies showed that perillyl alcohol inhibits the activity and expression of AR in prostate cancer cells [38]. Perillyl alcohol has shown activity in key regulatory and metabolic pathways in other types of cancer cells [39]. Terpineol showed a docking score of -35.99 kcal/mol, and Magangcha et al. (2023) [40], reported about terpineol binding to AR in several cancers, including prostate cancer. However, bornyl acetate, a TF− terpene, showed better

docking scores compared to TF+ terpenes. Similarly, other TF− terpenes interspersed between TF+ terpenes.

Table 6
Docking scores of monoterpenes and known ligands for AR, ranked by docking score.

Ligand	Docking score (kcal/mol)
CHEMBL2177238[a]	-80.2
Flutamide[a]	-66.36
Testosterone propionate[a]	-62.14
Bornyl acetate[b]	-47.84
Perillyl alcohol	-41.27
Terpineol	-35.99
Linalool	-35.75
(R)-Carvone	-33.05
Fenchol	-32.01
β-Myrcene[b]	-32
Terpinolene	-30.65
γ-Terpinene	-30.45
Limonene	-30.28
β-Ocimene[b]	-30.11
(S)-Carvone	-30.06
Sabinene	-29.2
Eucalyptol[b]	-29.15
p-Cymene[b]	-28.52
β-Pinene	-28.28
Borneol	-26.99
α-Pinene	-26.59

Unmarked: TF+ terpene; [a]known ligand; [b]TF− terpene.

Table 7 shows the docking scores of sesquiterpenes and known ligands with AR. The known ligands exhibited better docking scores than the sesquiterpenes. Among the sesquiterpenes, nerolidol showed a score of − 49.36 kcal/mol, followed by α-bisabolol of − 47.21 kcal/mol. The interaction of α-bisabolol with AR is consistent with the virtual-screening results reported by Han et al. (2020), although the direct AR binding and functional activity were not experimentally confirmed in that study [37].

Table 7
Docking scores of sesquiterpenes and known ligands for AR, ranked by docking score.

Ligand	Docking score (kcal/mol)
CHEMBL2177238[a]	-80.2
Flutamide[a]	-66.36
Testosterone propionate[a]	-62.14
Nerolidol	-49.36
α-Bisabolol	-47.21
Humulene	-37.31
β-Caryophyllene	-35.98
Caryophyllene oxide	-35.95

Unmarked: TF+ terpene; [a]known ligand.

11- β-hydroxysteroid dehydrogenase 1

11-β-hydroxysteroid dehydrogenase 1 (11β-HSD1) is an enzyme involved in converting inactive cortisone to cortisol, located within the brain, liver, and adipose tissues. It plays a key role in the functioning of the cardiovascular and central nervous systems [41]. The docking results of known ligands and monoterpenes are listed in Table 8. Docking scores of known ligands were higher compared to terpenes. Among them, CHAPS, a co-crystallized ligand, showed the highest docking score of -151.43 kcal/mol, followed by AZD-4017 and enoxolone. Borneol showed a docking score of -40.44 kcal/mol, followed by perillyl alcohol and linalool, which were TF− terpenes. However, the rest of the TF− terpenes showed low docking scores and remained at the bottom of the list. Experimental evidence of monoterpenes interacting with 11β-HSD1 was not found in the literature.

Table 8
Docking scores of monoterpenes and known ligands for 11β-HSD1, ranked by docking score.

Ligand	Docking score (kcal/mol)
CHAPS[a]	-151.43
AZD-4017[a]	-96.09
Enoxolone[a]	-71.89
Borneol	-40.44
Perillyl alcohol[b]	-38.10
Linalool[b]	-35.12
Terpineol	-34.83
Bornyl acetate	-34.38
Fenchol	-30.81
Eucalyptol	-30.63
(R)-Carvone	-29.66
(S)-Carvone	-28.74
Terpinolene[b]	-26.17
β-Ocimene[b]	-25.78
β-Myrcene[b]	-24.96
Limonene[b]	-24.06
β-Pinene[b]	-23.17
Sabinene[b]	-23.16
α-Pinene[b]	-22.98
γ-Terpinene[b]	-21.74
p-Cymene[b]	-20.63

Unmarked: TF+ terpene; [a]known ligand; [b]TF− terpene.

All the sesquiterpenes predicted for 11β-HSD1 were TF− terpenes. Table 9 shows the docking results of known ligands and sesquiterpenes for 11β-HSD1. Known ligands demonstrated better docking scores

compared to sesquiterpenes. Among sesquiterpenes, nerolidol showed the highest docking score, followed by α-bisabolol, caryophyllene oxide, β-Caryophyllene and humulene. However, experimental evidence of sesquiterpenes interacting with 11β-HSD1 was not found in the literature.

Table 9
Docking scores of sesquiterpenes and known ligands for 11β-HSD1, ranked by docking score.

Ligand	Docking score (kcal/mol)
CHAPS[a]	-151.43
AZD-4017[a]	-96.09
Enoxolone[a]	-71.89
Nerolidol[b]	-44.85
α-Bisabolol[b]	-40.99
Caryophyllene oxide[b]	-37.97
β-Caryophyllene[b]	-33.61
Humulene[b]	-30.52

[a]known ligand; [b]TF− terpene.

Cytochrome P450 19A1

Cytochrome P450 19A1 (CYP19A1) is an enzyme responsible for the conversion of androgens to estrogen. It is located within the endoplasmic reticulum and plays a central role in hormonal balance and tissue growth. The docking scores of known ligands and monoterpenes with CYP19A1 are listed in Table 10. On raw scores, bornyl acetate ranks first overall at -56.48 kcal/mol, above all three known ligands (letrozole − 48.49, androstenedione − 44.73, testolactone − 42.43 kcal/mol). When scores are normalized by heavy atom count to give ligand efficiency (LE, kcal/mol per heavy atom), the finding becomes even more striking: every monoterpene in the dataset outscores every known ligand at CYP19A1, with LE values ranging from 2.50 to 4.35 for terpenes compared to 2.02 to 2.20 for known ligands. This class-wide advantage is not a size artefact; it survives normalization. Sharma et al. (2024) reported binding and inhibitory effects of limonene, terpineol, and α-pinene with CYP19A1 [61]. In the present study, terpineol and limonene were TF+ terpenes with docking scores of -40.37 and − 28.84 kcal/mol and LE values of 3.67 and 2.88, respectively. α-pinene, classified as TF−, scored − 30.28 kcal/mol (LE 3.03), illustrating the TF false-negative problem discussed in Section 3.6.3: experimental evidence of binding exists, yet the tool failed to predict it. Linalool, also TF−, scored − 41.69 kcal/mol (LE 3.79), the second highest LE in the monoterpene series at this target.

Table 10
Docking scores of monoterpenes and known ligands for CYP19A1, ranked by docking score.

Ligand	Docking score (kcal/mol)
Bornyl acetate	-56.48
Letrozole[a]	-48.49
Androstenedione[a]	-44.73
Testolactone[a]	-42.43
Perillyl alcohol	-42.07
Linalool[b]	-41.69
Borneol	-40.91
Terpineol	-40.37
Fenchol	-40.00
(R)-Carvone	-36.93
(S)-Carvone	-35.21
β-Ocimene[b]	-34.77
γ-Terpinene[b]	-32.94
Eucalyptol[b]	-30.78
p-Cymene[b]	-30.32
α-Pinene[b]	-30.28
Sabinene[b]	-29.89
β-Myrcene[b]	-29.36
Limonene	-28.84
β-Pinene[b]	-27.17
Terpinolene[b]	-25.04

Unmarked: TF+ terpene; [a]known ligand; [b]TF− terpene.

The docking scores of known ligands and sesquiterpenes with CYP19A1 are shown in Table 11. Nerolidol, classified TF − for CYP19A1, achieves the highest raw docking score among sesquiterpenes

and exceeds all three known ligands on both raw score and ligand efficiency (LE 3.44 kcal/mol per heavy atom, vs known ligand range 2.02–2.20). This is the most pronounced example in the sesquiterpene series of a TF− compound outperforming TF+ compounds computationally. α-bisabolol, also TF − for CYP19A1, scored − 53.13 kcal/mol (LE 3.32); Sharma et al. (2024) reported bisabolol as an effective inhibitor of CYP19A1 [61], providing experimental support for that compound independently of the docking result. The docking result and the literature evidence are consistent but remain separate observations: neither validates the other, and both require dedicated binding assays to confirm.

Table 11
Docking scores of sesquiterpenes and known ligands for CYP19A1, ranked by docking score.

Ligand	Docking score (kcal/mol)
Nerolidol[b]	-55.01
α-Bisabolol	-53.13
Letrozole[a]	-48.49
Androstenedione[a]	-44.73
Testolactone[a]	-42.43
Caryophyllene oxide[b]	-39.38
β-Caryophyllene[b]	-38.37
Humulene[b]	-35.75

Unmarked: TF+ terpene; [a]known ligand; [b]TF− terpene.

Estrogen receptor β

Estrogen receptor β (ERβ) is a nuclear receptor that binds to estrogen and regulates the transcription of target genes. It plays important roles in various cellular processes such as growth, development, and apoptosis, and contributes to the functioning of the bone, reproductive system, and brain [42]. The docking results of known ligands and monoterpenes with ERβ are listed in Table 12. Known ligands demonstrated a better docking score compared to monoterpenes. Perillyl alcohol showed a docking score of -37.66 kcal/mol, followed by terpineol with − 36.88 kcal/mol and eucalyptol with − 36 kcal/mol. TF− terpenes remain interspersed with TF+ terpenes in the docking results.

Table 12
Docking scores of monoterpenes and known ligands for ERβ, ranked by docking score.

Ligand	Docking score (kcal/mol)
4-Hydroxytamoxifen[a]	-69.63
Ospemifene[a]	-63.89
R, R-THC[a]	-57.14
Perillyl alcohol	-37.66
Terpineol	-36.88
Eucalyptol	-36.00
Bornyl acetate[b]	-33.26
β-Ocimene[b]	-31.18
Linalool	-31.08
Borneol	-30.22
Terpinolene[b]	-30.06
β-Myrcene[b]	-29.98
Sabinene	-29.11
γ-Terpinene	-29.03
p-Cymene[b]	-27.96
Limonene	-26.90
Fenchol	-26.57
α-Pinene	-26.33
β-Pinene	-26.26
(S)-Carvone	-25.82
(R)-Carvone	-23.79

Unmarked: TF+ terpene; [a]known ligand; [b]TF− terpene.

Table 13 shows the docking scores of known ligands and sesquiterpenes with ERβ. Known ligands scored higher than all sesquiterpenes on raw scores, consistent with the size effect described under ‘effect of molecular size on docking scores’. Among sesquiterpenes, nerolidol, α-bisabolol, β-

caryophyllene, and humulene (all TF + at ERβ) scored above caryophyllene oxide, which is TF−. This is one of the few targets where TF+ status and docking rank are broadly aligned within the sesquiterpene series, though with only five sesquiterpenes the pattern cannot be given statistical weight. No published experimental evidence of sesquiterpene binding to ERβ was identified to independently evaluate these computational results.

Table 13
Docking scores of sesquiterpenes and known ligands for ERβ, ranked by docking score.

Ligand	Docking score (kcal/mol)
4-Hydroxytamoxifen[a]	-69.63
Ospemifene[a]	-63.89
R, R-THC[a]	-57.14
Nerolidol	-45.88
α-Bisabolol	-44.22
β-Caryophyllene	-36.38
Humulene	-34.80
Caryophyllene oxide[b]	-33.82

Unmarked: TF+ terpene; [a]known ligand; [b]TF− terpene.

Across all five targets, known ligands occupied the top of the raw docking score range for four of five proteins, a pattern that ‘effect of molecular size on docking scores’ shows and substantially explained by molecular size rather than by superior binding quality per atom. When scores are normalized by heavy atom count, terpenes are competitive with, or superior to, known ligands at CB2, CYP19A1, and ERβ, and are within a narrow margin at AR and 11β-HSD1. The more pressing question is whether docking scores distinguish TF+ from TF− terpenes within each series. They largely did not. TF− terpenes were interspersed with or outranked TF+ terpenes across most targets: bornyl acetate (TF − at AR) outscored most TF+ monoterpenes at that target; nerolidol and α-bisabolol (TF − at 11β-HSD1 and CYP19A1) ranked first among sesquiterpenes at those targets; and linalool (TF − at 11β-HSD1 and CYP19A1) scored comparably to TF+ terpenes. Where literature evidence supports a particular interaction, for example β-caryophyllene at CB2, perillyl alcohol at AR, and α-bisabolol at CYP19A1, partial reassurance exists, but several of these supported interactions ranked near the bottom of the terpene score range, illustrating that raw docking scores alone are unreliable predictors of experimentally observed binding.

Evaluation of target fishing and docking across target proteins

Comparing TF and docking results across all five target proteins reveals a consistent pattern: the two methods do not reliably identify the same set of terpene-target interactions as plausible. An example of such disagreement involves sesquiterpenes at steroidogenic targets. Nerolidol, α-bisabolol, and humulene received TF− predictions for both 11β-HSD1 and CYP19A1, meaning no TF tool found structural similarity to known ligands of those targets, yet all three ranked among the highest-scoring terpenes in docking against those same proteins. Conversely, α-bisabolol is TF + for four of five targets, but its docking scores at 11β-HSD1 and CYP19A1 do not distinguish it from TF− terpenes. Linalool is TF + for three proteins, TF − for 11β-HSD1 and CYP19A1, yet its docking score at CYP19A1 (− 41.69 kcal/mol) exceeds that of most TF+ monoterpenes at that target. Bornyl acetate is TF − at AR but outscores all TF+ monoterpenes in docking against that receptor. These examples, which are representative, not exceptional, illustrate that TF+ status does not predict high docking scores, and high docking scores do not imply TF+ status. The two methods are probing different and largely non-overlapping aspects of the molecule-target relationship.

Factors affecting target fishing and docking

Working principles

TF tools predict potential protein targets by identifying structural similarity to known bioactive ligands in reference databases, they are fundamentally ligand-centric and database-dependent. Docking, by contrast, evaluates geometric and electrostatic complementarity between a ligand and a specific binding site, it is structure-centric and agnostic to whether a compound resembles known ligands. These two methods therefore operate on entirely different information, The results of this study suggest that, with the current level of information available for each of these two approaches, these two approaches do not converge into mutually agreeing results. That does not necessarily mean that an approach is “wrong”: a TF− prediction means that a terpene is not structurally similar to any known ligand in the database for a given target, but it does not mean the terpene cannot bind. Nerolidol’s TF− status at 11β-HSD1 and CYP19A1 may simply reflect the absence of sesquiterpene-like ligands in those target’s reference databases, rather than a true absence of binding potential, the comparatively high docking scores at those same targets suggest the binding site geometry is not incompatible. Conversely, a TF+ prediction reflects database-mediated structural similarity to known ligands but does not guarantee that the binding site of the protein can accommodate the terpene.

False positives in docking

The sesquiterpenes, nerolidol, α-bisabolol, and humulene, ranked consistently at the top of the terpene docking score distribution across all five proteins, regardless of their TF status. Nerolidol scored − 50.41 kcal/mol at CB2, − 49.36 kcal/mol at AR, and achieved the highest terpene score at CYP19A1, despite being TF − at both 11β-HSD1 and CYP19A1. α-bisabolol similarly appeared among the top-scoring terpenes at every target. Humulene, which was TF − at 11β-HSD1 and CYP19A1, nonetheless outscored most TF+ monoterpenes in docking at those targets. This pattern of non-selective, broad high-scoring across structurally dissimilar targets is inconsistent with the expectation of specific binding and is more

consistent with scoring artifacts driven by the hydrophobic, cavity-filling properties of these larger sesquiterpenes [43]. Broad multi-target pharmacological activity has been reported experimentally for nerolidol [44] and α-bisabolol [45], suggesting some degree of genuine promiscuity. However, promiscuity and scoring artifacts are not mutually exclusive: Wu et al. (2024) demonstrated that certain chemical features can systematically inflate docking scores across targets through non-specific interactions with binding site residues [43]. For hydrophobic, conformationally flexible sesquiterpenes, both effects may contribute simultaneously, making it difficult to distinguish genuine broad-spectrum binding potential from algorithmic bias. This ambiguity represents a fundamental interpretive limitation for this compound class and underscores the need for experimental follow-up before any high-scoring terpene-target pair is treated as a validated interaction.

Limitations in target fishing tools

The clearest evidence that TF tools produce false negatives in this study comes from cases where a terpene received a TF− prediction despite published experimental evidence of interaction with that target. α-pinene was classified TF − at CYP19A1 in our study, yet Sharma et al. (2024) demonstrated inhibitory effects of α-pinene on CYP19A1 activity [61]. Similarly, the entire sesquiterpene class was predicted TF − for 11β-HSD1, despite the structural similarity between certain sesquiterpenes and the steroidal cortisol/cortisone substrates of that enzyme. Nerolidol illustrates the same problem in purely computational terms: it was classified TF + at CB2 and ERβ across all five targets, yet it achieves the highest ligand efficiency among sesquiterpenes at both 11β-HSD1 and CYP19A1, two proteins for which TF returned no prediction at all. The scale of this problem across the dataset is substantial. Analyzing the raw target fishing output (Online Resource 1) reveals that SEA returned no predictions for 12 of the 23 terpenes (52%), and STP returned nothing for sabinene and β-pinene; for those two compounds, consensus rested exclusively on PPB2 and STarFish. At the other extreme, p-cymene received TF− predictions for all five docked targets and had the lowest total consensus target count in the dataset, making it impossible to distinguish whether this reflects genuine absence of binding potential or simply the most severe database coverage gap in the set. In contrast, borneol, fenchol, and (S)-carvone were each TF + for all five targets, illustrating that the range of TF prediction breadth within the same chemical class spans from zero to complete coverage of the docked target set. These disparities arise from a structural mismatch between terpene chemistry and the composition of the TF reference databases. Terpene pharmacology is underrepresented in the curated bioactivity databases underlying SwissTargetPrediction, SEA, and PPB2, which are weighted heavily toward drug-like synthetic compounds. STarFish was designed for natural products and thus has broader coverage in this chemical space [25], but is itself limited by the amount of published quantitative terpene-protein interaction data, which remains sparse. As Mathai et al. (2020) note, false negatives in TF are intrinsic to all ligand-based methods when the query compound class is absent or underrepresented in training data, a condition that applies directly to terpenes [62]. The implication is that TF− predictions in this study should not be interpreted as evidence that a terpene cannot bind a target; they may simply reflect a database coverage gap.

Effect of molecular size on docking scores

Known ligands in this study range from 19 to 42 heavy atoms; sesquiterpenes range from 15 to 16; and monoterpenes range from 10 to 13. Molecular size correlates with GBVI/WSA docking scores because larger molecules can form more simultaneous contacts within a binding pocket, accumulating more favorable interaction terms regardless of binding selectivity [64]. To quantify the contribution of this size effect, we plotted heavy atom count against docking score for all five target proteins (Online Resource 3), computed coefficient of determinations (Table 14), and calculated ligand efficiency (LE, defined as |docking score| / heavy atom count) for every compound at every target. LE normalization places molecules of different sizes on a common scale and is standard practice in medicinal chemistry for cross-series comparisons [56]. Figure 2 shows raw docking score (top row) and LE (bottom row) plotted against heavy atom count (HA) for all five proteins, with known ligands, monoterpenes, and sesquiterpenes distinguished by color. The top row illustrates the size-driven scoring gap between known ligands and terpenes. The bottom row shows that after LE normalization terpenes are competitive with or superior to known ligands at CB2, CYP19A1, and ERβ. At CYP19A1 (fourth column), every terpene outscores every known ligand on a per-atom basis.

Table 14
Coefficient of determination between docking scores and heavy atom counts for terpenes and known ligands, calculated for each target protein.

Target protein	Monoterpenes + Known ligands (MK)	Monoterpenes	Sesquiterpenes + Known ligands (SK)	Sesquiterpenes
CB2	0.945	0.112	0.926	0.245
AR	0.899	0.643	0.787	0.394
11β-HSD1	0.901	0.401	0.876	0.781
CYP19A1	0.307	0.604	0.112	0.154
ERβ	0.925	0.117	0.946	0.312

When known ligands are included alongside terpenes, coefficient of determinations between heavy atom count and docking score are positive for CB2, AR, 11β-HSD1, and ERβ, reflecting the dominance of the larger known ligands at the top of the raw score range. When only terpenes are considered, these correlations drop and are not statistically significant for any protein, confirming that within the terpene series, other structural features, such as shape complementarity and the presence of polar groups, govern docking rank rather than size alone. The LE results sharpen this picture considerably. For CB2, ERβ, and most notably CYP19A1, mean terpene LE exceeds mean known-ligand LE: at CYP19A1, terpene LE values span 2.38 to 4.35 kcal/mol per heavy atom, while known-ligand LE values cluster between 2.02 and 2.20. This reversal is specific to CYP19A1 and is observed for every terpene in the dataset, suggesting that the CYP19A1 active site accommodates terpenoid-scale scaffolds with unusual per-atom efficiency. At AR and 11β-HSD1, known ligands retain a modest LE advantage (mean 3.10 and 3.14, respectively, versus terpene means of approximately 2.85 and 2.65), indicating that these sites extract more interaction energy per atom from larger drug-like molecules. Across all five targets, the conclusion

is the same: raw docking score comparisons between terpenes and drug-like known ligands are not on a meaningful common scale. LE normalization is necessary for any cross-class comparison, and it substantially changes the interpretation, particularly for CYP19A1.

Enrichment factor

To quantify how well docking scores discriminate TF+ from TF− terpenes, we calculated Enrichment Factor (EF) values as defined by Halgren et al. [57]: the ratio of the fraction of true positives (TF+ terpenes and known ligands) in the top-ranked subset to their fraction in the full dataset. EF = 1 corresponds to random ranking; EF > 1 means true positives are overrepresented in the top-ranked fraction; EF < 1 means they are underrepresented. Before interpreting the values in Table 15 and Figs. 3 and 4, it is essential to acknowledge a fundamental constraint: all EF calculations in this study are performed on datasets containing at most 23 terpenes per target, with the TF + and TF− subsets varying in size from target to target, and for some targets, all sesquiterpenes are TF−, leaving no inactives in the dataset at all. As Evangelista et al. [60] showed using hundreds of ligands and decoys per GPCR, meaningful EF statistics require substantial dataset size; at the scale of our study, even EF values substantially above 1 may not be statistically distinguishable from random variation.

$$EF_p = \frac{(TN/N_p)}{(T_{tot}/L_{tot})}$$

Where p is the cutoff percentage, *TN* is the number of true positives within the cutoff, $N_p$is the number of ligands included at that cutoff, $T_{\mathrm{tot}}$is the total number of true positives (TF+ terpenes and known ligands) within the full set, and $L_{\mathrm{tot}}$is the total number of ligands within the full set.

Table 15
Enrichment factors for proteins at different cutoff percentages

Percentage of ligands	Monoterpenes					Sesquiterpenes				
	CB2	AR	11β-HSD1	CYP19A1	ERβ	CB2	AR	11β-HSD1	CYP19A1	ERβ
10	1	1.31	2.10	1.91	1.31	1	1	1	2	1.14
20	1	0.98	2.10	1.91	1.31	1	1	1	2	1.14
30	1	1.09	1.40	1.59	1.31	1	1	1	2	1.14
40	1	1.15	1.58	1.67	0.98	1	1	1	2	1.14
50	1	1.07	1.72	1.74	0.95	1	1	0.75	1.5	1.14
60	1	1.11	1.62	1.47	0.91	1	1	0.60	1.6	1.14
70	1	1.05	1.40	1.27	0.88	1	1	0.50	1.33	1.14
80	1	1.00	1.24	1.12	0.93	1	1	0.50	1.33	1.14
90	1	0.97	1.11	1.11	0.97	1	1	0.43	1.14	1.14
100	1	1.00	1	1	1.00	1	1	0.38	1	1

Table 15 reports EF values at every 10% cutoff. For monoterpenes, CB2 returns EF = 1 throughout, reflecting the absence of any TF− monoterpenes in the CB2 dataset, making enrichment structurally undefined rather than genuinely informative. AR and 11β-HSD1 show EF slightly above 1 at some cutoffs; CYP19A1 shows the highest monoterpene EF values. For sesquiterpenes, 11β-HSD1 returns EF < 1 at several cutoffs, meaning TF− sesquiterpenes were actually preferentially ranked by docking, the opposite of what a useful screening tool would do. CB2 and AR show EF = 1 for sesquiterpenes because all sesquiterpenes in those datasets are TF+, again making enrichment undefined. CYP19A1 and ERβ show modest EF > 1 for sesquiterpenes. In all cases, the numerical EF values must be treated as descriptive summaries of the data, not as validated evidence of discriminatory power, for the dataset size reasons stated above. For monoterpenes, EF values exceeding 1 were observed for AR, 11β-HSD1, and CYP19A1; CB2 exhibited EF equal to 1, indicating no discrimination between ligand classes. For sesquiterpenes, EF below 1 was consistently observed for 11β-HSD1 (preferential ranking of TF− terpenes), while EF equal to 1 was obtained for CB2 and AR. EF values modestly greater than 1 were observed for CYP19A1 and ERβ. These numerical patterns are noted, but as discussed above, their statistical significance and practical meaning are severely constrained by dataset size.

Figure 3 plots the cumulative percentage of known ligands and TF+ terpenes recovered as a function of the fraction of monoterpenes selected by docking rank. Several interpretive issues must be noted upfront. The CB2 line is uninformative as a “baseline” because all monoterpenes in the CB2 dataset are TF+, meaning there are no TF− inactives to distinguish against, any ranking of those compounds trivially recovers 100% of TF+ compounds eventually. For 11β-HSD1 and CYP19A1, recovery of approximately

90% of TF+ compounds and known ligands within the top 50% of the ranked dataset may appear favorable, but this result must be contextualized: the absolute numbers involved are small (a handful of TF+ terpenes and three known ligands), and with such small N, a small number of fortuitous score rankings can produce large apparent enrichment. Furthermore, the practical question is not whether 90% recovery at 50% selection, equivalent to a twofold enrichment, is useful: at this scale, selecting the top 11 of 23 compounds rather than all 23 provides a marginal experimental benefit that may not justify the *in-silico* effort. The ERβ curve falls below CB2, consistent with docking providing no useful discrimination for that target.

Figure 4 shows the equivalent plot for sesquiterpenes. For CB2 and AR, all sesquiterpenes are TF+, again making these curves uninformative baselines rather than evidence of enrichment. The 11β-HSD1 sesquiterpene curve is particularly notable: it falls below the CB2/AR reference at several cut points, meaning that docking preferentially ranked TF− sesquiterpenes over TF+ ones at 11β-HSD1, the opposite of what a useful screening tool would do. For CYP19A1, the sesquiterpene enrichment curve performs better, recovering approximately 50% of TF+ compounds in the top 20% of the ranked dataset; this is the most encouraging result in the sesquiterpene analysis. However, with only five sesquiterpenes in the dataset and varying TF+ fractions per target, even this result involves single-digit compound counts, and no statistical significance can be attributed to it.

Taken together, the enrichment analysis leads to a clear conclusion: docking scores, as computed in this study, do not reliably discriminate TF+ from TF terpenes across the five targets examined. For three of five targets (CB2, AR, ERβ), enrichment is either structurally undefined or essentially absent. For 11β-HSD1, docking actively anti-enriches the TF+ sesquiterpenes. Only for CYP19A1 is there a consistent trend toward preferential recovery of TF+ compounds and known ligands, and even there the trend is based on small absolute numbers and cannot be statistically validated at this dataset size. Critically, an “apparently better than random” result at the scale of 23 terpenes does not constitute evidence that the pipeline is useful for experimental prioritization: the practical benefit of selecting 10 compounds instead of 20 for experimental testing is negligible compared to the time, cost, and expertise required to execute the in silico screening, particularly when the scoring function used (GBVI/WSA) was calibrated on drug-like compounds substantially larger than terpenes. [60]

Conclusion

This study applied a combined *in silico* pipeline, consensus target fishing across four ligand-based tools followed by molecular docking, to 23 terpenes with the goal of identifying candidate protein targets and evaluating whether the two computational approaches agree. Five proteins emerged as high-frequency consensus targets: CB2, AR, 11β-HSD1, CYP19A1, and ERβ. On raw docking scores, known ligands ranked above terpenes at four of the five targets, but ligand efficiency (LE) analysis reveals this advantage to be largely a molecular size artefact: once scores are normalized by heavy atom count, terpenes are competitive with known ligands for CB2 and ERβ, and CYP19A1. Several other terpene-

target interactions were consistent with published experimental evidence, including β-caryophyllene at CB2, perillyl alcohol at AR, and α-bisabolol at CYP19A1, lending partial support to the overall pipeline.

The central methodological question, whether TF and docking reinforce each other usefully when applied to terpenes, must be answered cautiously. TF+ status did not reliably predict higher docking scores: TF− terpenes were frequently interspersed with or outranked TF+ terpenes across multiple targets, and enrichment analysis confirmed that docking provided little discriminatory power for most proteins. LE normalization does not rescue this finding: the failure to discriminate TF+ from TF− terpenes by docking rank is not a size artefact and persists after normalization. Critically, this failure of enrichment cannot be attributed solely to methodological weakness. As Evangelista et al. (2019) showed, meaningful enrichment statistics require large datasets of known actives and inactives, on the order of hundreds of ligands per target [60]. With at most 23 terpenes, of which only a subset is TF + for any given protein, this study operates far below the data density needed to make statistically robust enrichment claims. This is not a limitation unique to our work; it reflects a fundamental challenge for computational approaches applied to natural products, where the pharmacological knowledge base is far sparser than in conventional drug discovery. The most honest answer to the question of pipeline utility is therefore: it generates a prioritized list of hypotheses, but cannot currently validate them computationally. Whether TF or docking alone would perform better than their combination also remains unresolved, as the dataset is too small to distinguish a synergistic benefit from noise introduced by combining two imperfect methods. What is clear is that the two approaches are not redundant: TF and docking each flagged interactions the other missed, suggesting they sample different aspects of molecular recognition, with TF reflecting database-derived chemical similarity and docking reflecting active-site geometry.

What would make this approach more useful? The most pressing need is availability of thorough, organized data on terpenes. The terpene pharmacology literature remains sparse relative to conventional drug targets: most terpenes lack well-characterized protein binding partners, and the reference databases underlying TF tools are biased toward pharmaceutical compounds rather than natural products. Building curated, quantitative terpene-protein interaction datasets, ideally with binding affinities and selectivity data, would directly enable more reliable TF predictions and provide the actives-plus-decoys datasets needed for statistically valid enrichment analysis. Second, the Ligand Efficiency (LE) analysis presented here demonstrates that scoring function recalibration may be necessary: standard GBVI/WSA scoring was developed and validated on drug-like molecules substantially larger than terpenes, and the present results suggest that raw scores are not a meaningful basis for cross-class comparison. Reporting LE alongside raw scores could be standard practice for studies of this kind. In the meantime, this integrated TF−plus-docking approach is best understood as a hypothesis-generation tool: it surveys a large target space at low cost and suggests interactions worth testing, rather than making confident predictions of binding. Used with that expectation, and with awareness of its statistical limitations at small dataset sizes, it represents a possible first-pass screen for terpene-based pharmacological development. It is also our hope that the knowledge-based space of terpenes can be extended by integrating together literature data on this very important class of molecules. With

known experimental targets and binding affinities, together with possibly clinical effects, the use of drug-parametrized tools will become more and more accurate. Terpene-specific databases such as TeroKit [65], TPCN [66], and TERPEDIA [67] provide curated information on terpenes and their reported biological activities. Combining curated data from such resources with the present pipeline may improve the generation and prioritization of testable terpene-protein interaction hypotheses for future experimental studies.

Declarations

Acknowledgements The authors would like to thank the Department of Biological Sciences at The University of Alabama in Huntsville for its support.

Author contributions M.H. and J.B. designed the study, formal analysis, and wrote of the manuscript. A.A. contributed to the initial development of the methodology. S.T. contributed to the selection of terpenes. All authors read and approved the final manuscript.

Data availability The data supporting the findings of this study are provided in the Supplementary Information accompanying this article.

Competing interests *Susan Trapp and Daniel McShan are affiliated with Terpedia, LLC, which develops terpene-focused scientific information resources. The other authors declare no competing interests.*

Funding No funding was received for conducting this study.

Ethics declaration Not applicable.

References

1. Masyita A, Mustika Sari R, Dwi Astuti A et al (2022) Terpenes and terpenoids as main bioactive compounds of essential oils, their roles in human health and potential application as natural food preservatives. Food Chem X 13:100217. https://doi.org/10.1016/J.FOCHX.2022.100217
2. Hu Y, Yuan Z, Wang Q et al (2025) Expanding the Terpene Universe: Synthetic Biology and Non-Natural Chemistry in Engineered Microorganisms. Molecules 2025, Vol 30, Page 4065 30:4065. https://doi.org/10.3390/MOLECULES30204065
3. Christianson DW (2017) Structural and Chemical Biology of Terpenoid Cyclases. Chem Rev 117:11570–11648. https://doi.org/10.1021/ACS.CHEMREV.7B00287
4. Cox-Georgian D, Ramadoss N, Dona C, Basu C (2019) Therapeutic and medicinal uses of terpenes. Med Plants: Farm Pharm 333–359. https://doi.org/10.1007/978-3-030-31269-5_15
5. Paduch R, Kandefer-Szerszeń M, Trytek M, Fiedurek J (2007) Terpenes: Substances useful in human healthcare. Arch Immunol Ther Exp (Warsz) 55:315–327. https://doi.org/10.1007/S00005-007-0039-1

6. Tetali SD (2019) Terpenes and isoprenoids: a wealth of compounds for global use. Planta 249:1–8. https://doi.org/10.1007/S00425-018-3056-X
7. Jaeger R, Cuny E (2016) Terpenoids with Special Pharmacological Significance: A Review. Nat Prod Commun 11:1373–1390. https://doi.org/10.1177/1934578X1601100946
8. Park BB, An JY, Park SU (2021) Recent studies on pinene and its biological and pharmacological activities. EXCLI J 20:812–818. https://doi.org/10.17179/EXCLI2021-3714
9. Salehi B, Upadhyay S, Orhan IE et al (2019) Therapeutic Potential of α- and β-Pinene: A Miracle Gift of Nature. Biomolecules 2019, Vol 9, Page 738 9:738. https://doi.org/10.3390/BIOM9110738
10. Martin V (1993) Overview of Paclitaxel (TAXOL®). Semin Oncol Nurs 9:2–5. https://doi.org/10.1016/S0749-2081(16)30035-3
11. Yu H, Lan F, Zhuang Y et al (2025) Paclitaxel anti-cancer therapeutics: from discovery to clinical use. Chin J Nat Med 23:769–789. https://doi.org/10.1016/S1875-5364(25)60833-8
12. Siddiqui T, Khan MU, Sharma V, Gupta K (2024) Terpenoids in essential oils: Chemistry, classification, and potential impact on human health and industry. Phytomedicine Plus 4:100549. https://doi.org/10.1016/J.PHYPLU.2024.100549
13. Câmara JS, Perestrelo R, Ferreira R et al (2024) Plant-Derived Terpenoids: A Plethora of Bioactive Compounds with Several Health Functions and Industrial Applications-A Comprehensive Overview. Molecules 29:3861. https://doi.org/10.3390/molecules29163861
14. Lowe H, Ali A, Steele B et al (2024) The potential therapeutic value of terpenes. INNOSC Theranostics Pharmacol Sci 7:0332. https://doi.org/10.36922/ITPS.0332
15. Del Prado-Audelo ML, Cortés H, Caballero-Florán IH et al (2021) Therapeutic Applications of Terpenes on Inflammatory Diseases. Front Pharmacol 12. https://doi.org/10.3389/FPHAR.2021.704197
16. Galati S, Di Stefano M, Martinelli E et al (2021) Recent Advances in In Silico Target Fishing. Molecules 2021, Vol 26, Page 5124 26:5124. https://doi.org/10.3390/MOLECULES26175124
17. Cereto-Massagué A, Ojeda MJ, Valls C et al (2015) Tools for in silico target fishing. Methods 71:98–103. https://doi.org/10.1016/J.YMETH.2014.09.006
18. Jenkins JL, Bender A, Davies JW (2006) In silico target fishing: Predicting biological targets from chemical structure. Drug Discov Today Technol 3:413–421. https://doi.org/10.1016/J.DDTEC.2006.12.008
19. Ferreira LG, Dos Santos RN, Oliva G, Andricopulo AD (2015) Molecular Docking and Structure-Based Drug Design Strategies. Molecules 2015 20:13384–13421. https://doi.org/10.3390/MOLECULES200713384
20. Sahoo RN, Pattanaik S, Pattnaik G et al (2022) Review on the use of Molecular Docking as the First Line Tool in Drug Discovery and Development. Indian J Pharm Sci 84:1334–1337. https://doi.org/10.36468/PHARMACEUTICAL-SCIENCES.1031

21. Agu PC, Afiukwa CA, Orji OU et al (2023) Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management. Scientific Reports 2023 13:1 13:13398-. https://doi.org/10.1038/s41598-023-40160-2
22. Weininger D (1988) SMILES, a Chemical Language and Information System: 1: Introduction to Methodology and Encoding Rules. J Chem Inf Comput Sci 28:31–36. https://doi.org/10.1021/CI00057A005
23. Keiser MJ, Roth BL, Armbruster BN et al (2007) Relating protein pharmacology by ligand chemistry. Nat Biotechnol 25:197–206. https://doi.org/10.1038/NBT1284
24. Gallo K, Goede A, Preissner R, Gohlke BO (2022) SuperPred 3.0: drug classification and target prediction-a machine learning approach. Nucleic Acids Res 50:W726–W731. https://doi.org/10.1093/NAR/GKAC297
25. Cockroft NT, Cheng X, Fuchs JR (2019) STarFish: A Stacked Ensemble Target Fishing Approach and its Application to Natural Products. J Chem Inf Model 59:4906–4920. https://doi.org/10.1021/ACS.JCIM.9B00489
26. Awale M, Reymond JL (2019) Polypharmacology Browser PPB2: Target Prediction Combining Nearest Neighbors with Machine Learning. J Chem Inf Model 59:10–17. https://doi.org/10.1021/ACS.JCIM.8B00524
27. Radwan MM, Chandra S, Gul S, Elsohly MA (2021) Cannabinoids, Phenolics, Terpenes and Alkaloids of Cannabis. Molecules 2021, Vol 26, Page 2774 26:2774. https://doi.org/10.3390/MOLECULES26092774
28. Nazzaro F, Fratianni F, Coppola R, De Feo V (2017) Essential Oils and Antifungal Activity. Pharmaceuticals (Basel) 10. https://doi.org/10.3390/PH10040086
29. Kim S, Chen J, Cheng T et al (2025) PubChem 2025 update. Nucleic Acids Res 53:D1516–D1525. https://doi.org/10.1093/NAR/GKAE1059
30. Bie B, Wu J, Foss JF, Naguib M (2018) An overview of the cannabinoid type 2 receptor system and its therapeutic potential. Curr Opin Anaesthesiol 31:407–414. https://doi.org/10.1097/ACO.0000000000000616
31. Turcotte C, Blanchet MR, Laviolette M, Flamand N (2016) The CB2 receptor and its role as a regulator of inflammation. Cell Mol Life Sci 73:4449–4470. https://doi.org/10.1007/S00018-016-2300-4
32. Gertsch J, Leonti M, Raduner S et al (2008) Beta-caryophyllene is a dietary cannabinoid. Proc Natl Acad Sci U S A 105:9099–9104. https://doi.org/10.1073/PNAS.0803601105
33. Ceccarelli I, Fiorenzani P, Pessina F et al (2020) The CB2 Agonist β-Caryophyllene in Male and Female Rats Exposed to a Model of Persistent Inflammatory Pain. Front Neurosci 14:555178. https://doi.org/10.3389/FNINS.2020.00850
34. Bento AF, Marcon R, Dutra RC et al (2011) β-Caryophyllene inhibits dextran sulfate sodium-induced colitis in mice through CB2 receptor activation and PPARγ pathway. Am J Pathol 178:1153–1166. https://doi.org/10.1016/J.AJPATH.2010.11.052

35. Finlay DB, Sircombe KJ, Nimick M et al (2020) Terpenoids From Cannabis Do Not Mediate an Entourage Effect by Acting at Cannabinoid Receptors. Front Pharmacol 11:510858. https://doi.org/10.3389/FPHAR.2020.00359
36. Davey RA, Grossmann M (2016) Androgen Receptor Structure, Function and Biology: From Bench to Bedside. Clin Biochem Rev 37:3
37. Han W, Shi Y, Su J et al (2020) Virtual Screening and Bioactivity Evaluation of Novel Androgen Receptor Antagonists From Anti-PCa Traditional Chinese Medicine Prescriptions. Front Chem 8:582861. https://doi.org/10.3389/FCHEM.2020.582861
38. Chung BH, Lee H young, Lee JS, Young CYF (2006) Perillyl alcohol inhibits the expression and function of the androgen receptor in human prostate cancer cells. Cancer Lett 236:222–228. https://doi.org/10.1016/J.CANLET.2005.05.023
39. Batista Brochado AC, de Moraes JA, de Oliveira R B, et al (2024) Metabolic and Regulatory Pathways Involved in the Anticancer Activity of Perillyl Alcohol: A Scoping Review of In Vitro Studies. Cancers (Basel) 16:4003. https://doi.org/10.3390/CANCERS16234003
40. Mangangcha IR, Shankar V, Evelin H (2023) Elucidation of drug-like properties in metabolites of Curcuma angustifolia Roxb. Indian J Nat Prod Resour (IJNPR) [Formerly Nat Prod Radiance (NPR)] 14:37–49. https://doi.org/10.56042/IJNPR.V14I1.1126
41. Holmes MC, Seckl JR (2006) The role of 11β-hydroxysteroid dehydrogenases in the brain. Mol Cell Endocrinol 248:9–14. https://doi.org/10.1016/J.MCE.2005.12.002
42. Imamov O, Shim GJ, Warner M, Gustafsson JÅ (2005) Estrogen Receptor beta in Health and Disease. Biol Reprod 73:866–871. https://doi.org/10.1095/BIOLREPROD.105.043497
43. Wu Y, Liu F, Glenn I et al (2024) Identifying Artifacts from Large Library Docking. J Med Chem 67:16796–16806. https://doi.org/10.1021/ACS.JMEDCHEM.4C01632
44. Chan WK, Tan LTH, Chan KG et al (2016) Nerolidol: A Sesquiterpene Alcohol with Multi-Faceted Pharmacological and Biological Activities. Molecules 2016, Vol 21, Page 529 21:529. https://doi.org/10.3390/MOLECULES21050529
45. Romagnoli C, Baldisserotto A, Malisardi G et al (2015) A Multi-Target Approach toward the Development of Novel Candidates for Antidermatophytic Activity: Ultrastructural Evidence on α-Bisabolol-Treated Microsporum gypseum. Molecules 20:11765. https://doi.org/10.3390/MOLECULES200711765
46. Berman HM, Westbrook J, Feng Z et al (2000) The Protein Data Bank. Nucleic Acids Res 28:235–242. https://doi.org/10.1093/NAR/28.1.235
47. Chemical Computing Group ULC (2025) Molecular Operating Environment (MOE)
48. Gómez A, Alarcón A, Acosta W, Malagón A (2024) Identification of potential human targets of glyphosate using in silico target fishing. Comput Toxicol 30:100306. https://doi.org/10.1016/J.COMTOX.2024.100306
49. Ye XY, Ling QZ, Chen SJ (2015) Identification of a Potential Target of Capsaicin by Computational Target Fishing. Evidence-Based Complement Altern Med 2015:983951.

https://doi.org/10.1155/2015/983951

50. Corbeil CR, Williams CI, Labute P (2012) Variability in docking success rates due to dataset preparation. J Comput Aided Mol Des 26:775–786. https://doi.org/10.1007/S10822-012-9570-1
51. Naïm M, Bhat S, Rankin KN et al (2007) Solvated Interaction Energy (SIE) for scoring protein-ligand binding affinities. 1. Exploring the parameter space. J Chem Inf Model 47:122–133. https://doi.org/10.1021/CI600406V
52. Harding SD, Armstrong JF, Faccenda E et al (2024) The IUPHAR/BPS Guide to PHARMACOLOGY in 2024. Nucleic Acids Res 52:D1438–D1449. https://doi.org/10.1093/NAR/GKAD944
53. Bofill A, Jalencas X, Oprea TI, Mestres J (2019) The human endogenous metabolome as a pharmacology baseline for drug discovery. Drug Discov Today 24:1806–1820. https://doi.org/10.1016/J.DRUDIS.2019.06.007
54. Bernal FA, Coy-Barrera E (2014) In-Silico Analyses of Sesquiterpene-Related Compounds on Selected Leishmania Enzyme-Based Targets. Molecules 2014, Vol 19, Pages 5550–5569 19:5550–5569. https://doi.org/10.3390/MOLECULES19055550
55. Wojciechowski M, Lesyng B (2004) Generalized Born model: Analysis, refinement, and applications to proteins. J Phys Chem B 108:18368–18376. https://doi.org/10.1021/JP046748B
56. Hopkins AL, Groom CR, Alex A (2004) Ligand efficiency: a useful metric for lead selection. Drug Discov Today 9:430–431. https://doi.org/10.1016/S1359-6446(04)03069-7
57. Halgren TA, Murphy RB, Friesner RA et al (2004) Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. J Med Chem 47:1750–1759. https://doi.org/10.1021/JM030644S
58. Truchon JF, Bayly CI (2007) Evaluating virtual screening methods: Good and bad metrics for the early recognition problem. J Chem Inf Model 47:488–508. https://doi.org/10.1021/CI600426E
59. Bender A, Glen RC (2005) A discussion of measures of enrichment in virtual screening: Comparing the information content of descriptors with increasing levels of sophistication. J Chem Inf Model 45:1369–1375. https://doi.org/10.1021/CI0500177
60. Evangelista Falcon W, Ellingson SR, Smith JC, Baudry J (2019) Ensemble Docking in Drug Discovery: How Many Protein Configurations from Molecular Dynamics Simulations are Needed to Reproduce Known Ligand Binding? J Phys Chem B 123:5189–5195. https://doi.org/10.1021/ACS.JPCB.8B11491
61. Sharma K, Lanzilotto A, Yakubu J et al (2024) Effect of Essential Oil Components on the Activity of Steroidogenic Cytochrome P450. Biomolecules 14:203. https://doi.org/10.3390/BIOM14020203
62. Mathai N, Chen Y, Kirchmair J (2020) Validation strategies for target prediction methods. Brief Bioinform 21:791–802. https://doi.org/10.1093/BIB/BBZ026
63. Peón A, Dang CC, Ballester PJ (2016) How reliable are ligand-centric methods for target fishing? Front Chem 4:194706. https://doi.org/10.3389/FCHEM.2016.00015

64. Siraj MA, Rahman MS, Tan GT, Seidel V (2021) Molecular docking and molecular dynamics simulation studies of triterpenes from vernonia patula with the cannabinoid type 1 receptor. Int J Mol Sci 22:3595. https://doi.org/10.3390/IJMS22073595/
65. Zeng T, Liu Z, Zhuang J et al (2020) TeroKit: A Database-Driven Web Server for Terpenome Research. J Chem Inf Model 60:2082–2090. https://doi.org/10.1021/ACS.JCIM.0C00141
66. Li W, Chen Y, Yang R et al (2024) A terpenoids database with the chemical content as a novel agronomic trait. Database 2024:. https://doi.org/10.1093/DATABASE/BAAE027
67. Terpedia.com https://terpedia.com/. Accessed 28 Jul 2026

Figures

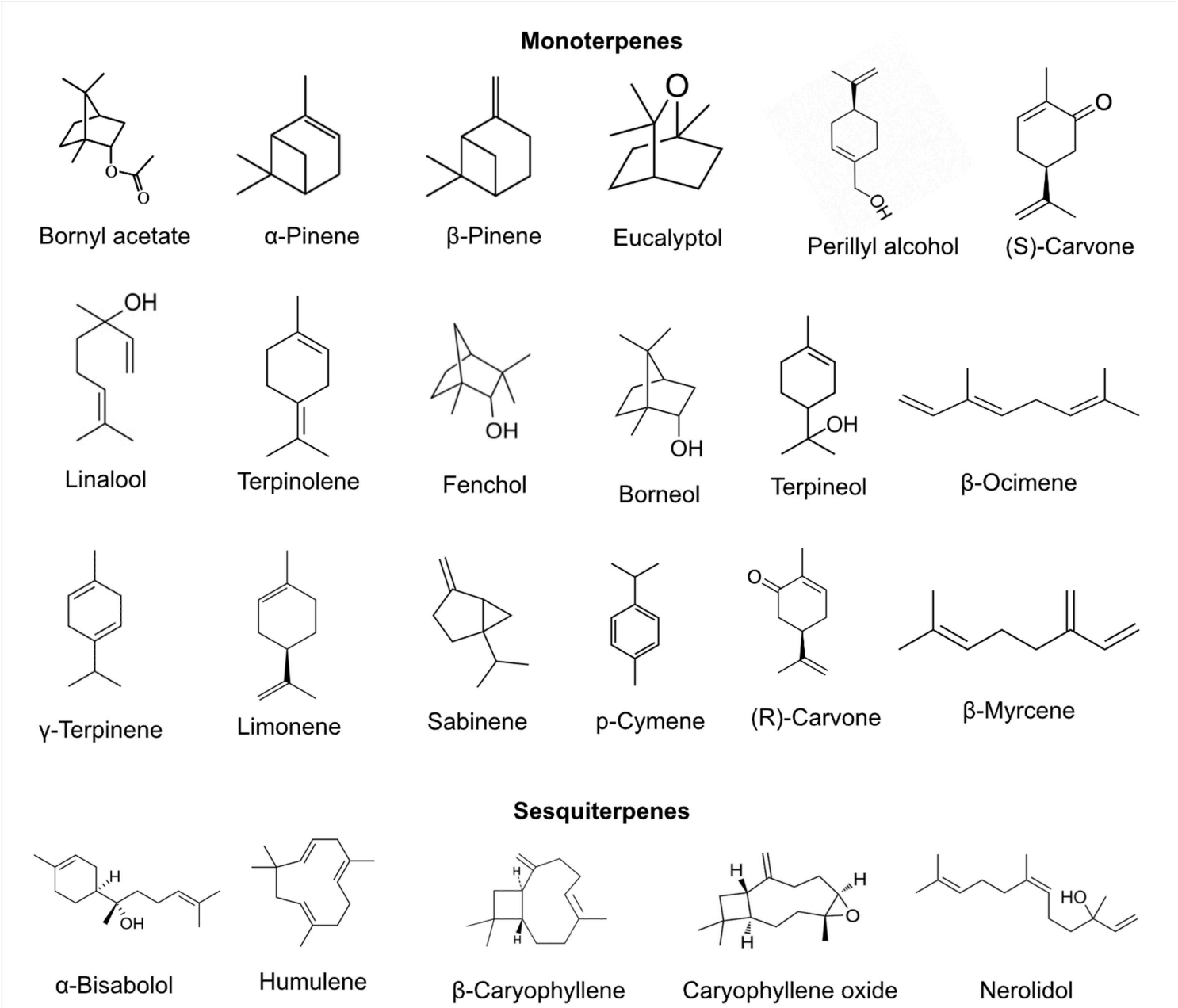


Figure 1

Structure of monoterpenes and sesquiterpenes selected for the study.

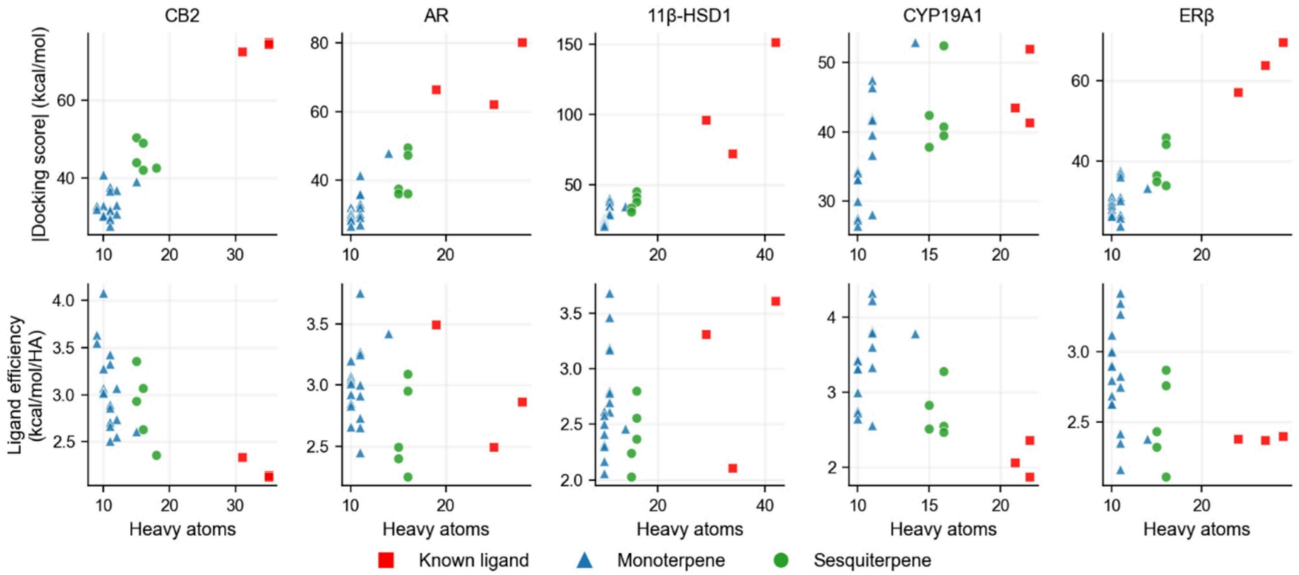


Figure 2

Raw docking scores (top row) and ligand efficiency (LE, |score| / heavy atom count, bottom row) plotted against heavy atom count for each of the five target proteins. Known ligands are shown in red, monoterpenes in blue, and sesquiterpenes in green.

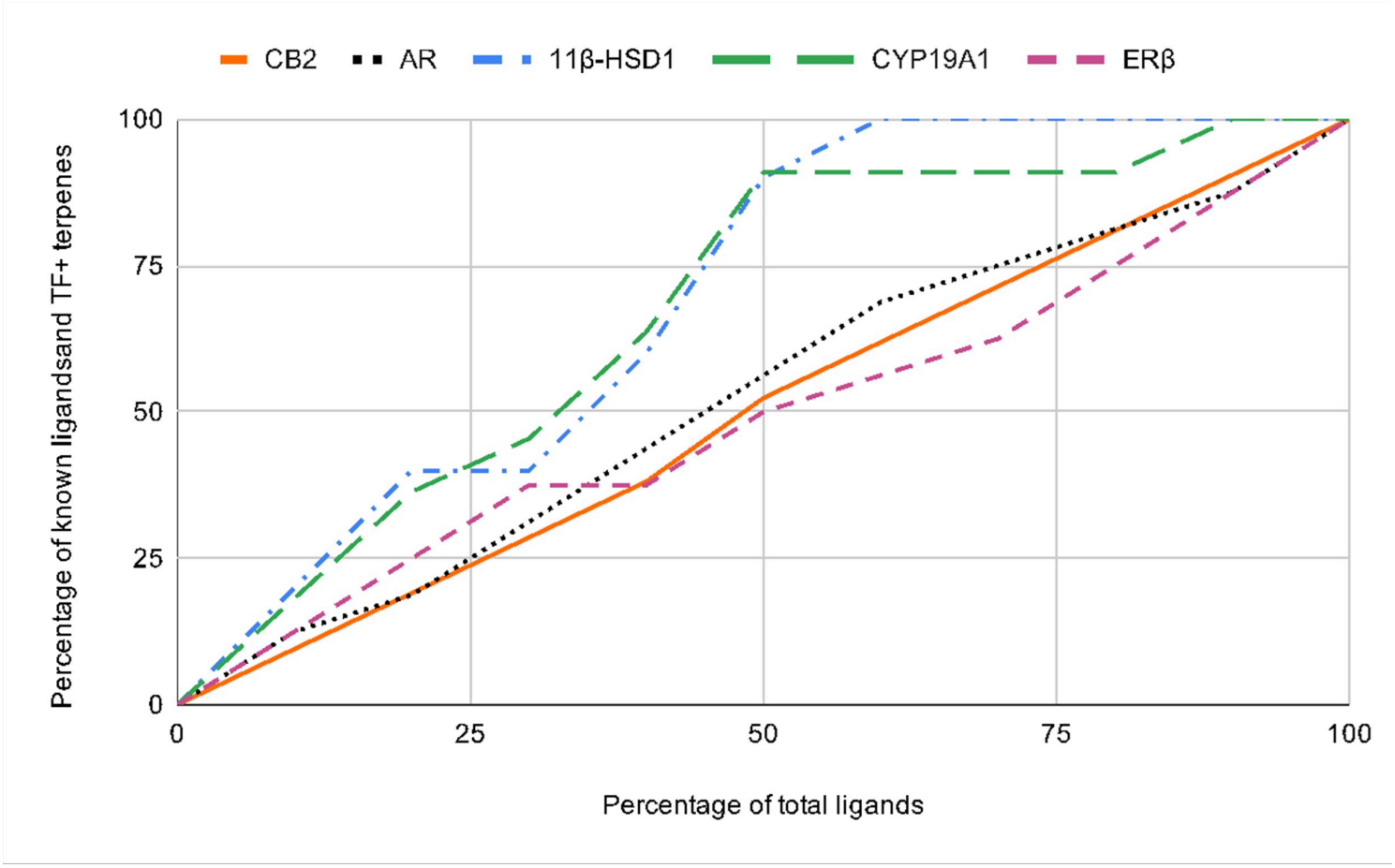


Figure 3

Docking enrichment calculation for monoterpenes

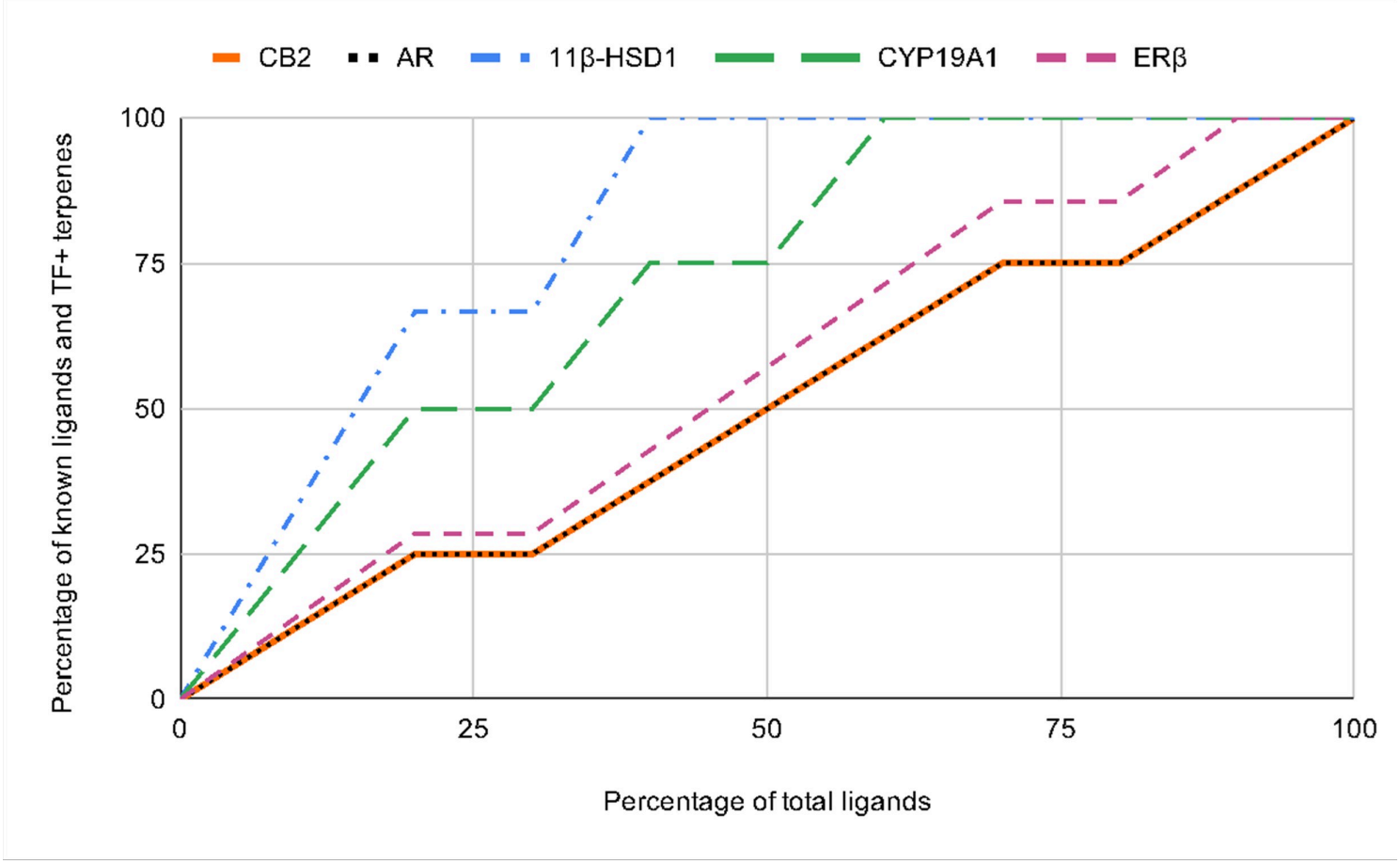


Figure 4

Docking enrichment calculation for sesquiterpenes

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ESM1.xlsx
- ESM2.xlsx
- ESM3.docx