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.

Comparative Network Pharmacology Analysis of Ethyl Acetate, n-Butanol and Residual Fractions of *Stachytarpheta jamaicensis* Against Lung Cancer

Muinat O. Lambe
Al- Hikmah University
Tajudeen A. Owolarafe
Nigeria Police Academy, Wudil
Simbiat Omolara Fatai
240603bch003@alhikmah.edu.ng

Al-Hikmah University https://orcid.org/0009-0001-6088-5037

Research Article

Keywords: Stachytarpheta jamaicensis, network pharmacology, molecular docking, lung cancer, solvent fraction comparison

Posted Date: October 22nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11447450/v1

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

Additional Declarations: The authors declare no competing interests.

Abstract

Stachytarpheta jamaicensis (L.) Vahl is a pantropical medicinal plant with documented antioxidant, anti-arthritic, and antimicrobial properties, but its anticancer potential has been evaluated only against breast and cervical cancer cell lines, and no network pharmacology framework has been applied to it. This study comparatively evaluated its ethyl acetate, n-butanol, and residue leaf-extract fractions against lung cancer. Eleven phytochemicals were prioritized across the three fractions based on drug-likeness and reported bioactivity. Predicted targets overlapped lung-cancer-associated genes to a similar extent in every fraction (83.06%, 82.54%, and 82.23% for ethyl acetate, n-butanol, and residue respectively). Protein-protein interaction analysis identified HIF1A as the only hub gene common to all three fractions, with AKT1, SRC, EGFR, and BCL2 each recurring in two; these five were prioritized for docking. Pathway enrichment converged on kinase-mediated cancer pathways, with Pathways in Cancer, apoptosis, and EGFR tyrosine kinase inhibitor resistance enriched in all three fractions. Molecular docking, with every receptor grid validated by redocking of its co-crystallized ligand, identified Silychristin (n-butanol) as the leading compound, returning scores comparable to the matched reference inhibitors at two validated kinase targets: −9.43 kcal/mol at SRC (dasatinib − 9.51) and − 9.30 kcal/mol at EGFR (erlotinib − 9.25). Secoisolariciresinol (residue) followed at − 8.91 kcal/mol at SRC. At HIF1A and BCL2 the reference inhibitors did not separate from the compound set, and these targets were not used to compare fractions. Anticancer-relevant target engagement is therefore distributed across all three fractions rather than concentrated in one.

1. Introduction

Lung cancer is still the most frequently diagnosed malignancy and leading cause of cancer mortality worldwide, accounting for over 2.4 million new cases and 1.8 million deaths annually (Bray et al. 2024). Five-year survival in metastatic non-small-cell lung cancer remains below 10%, driven by acquired resistance and genetic heterogeneity (Merecz-Sadowska et al. 2026), sustaining interest in plant-derived multitarget agents that engage multiple oncogenic nodes simultaneously through polypharmacological mechanisms (Hopkins 2008; Merecz-Sadowska et al. 2026).

Stachytarpheta jamaicensis (L.) Vahl (Verbenaceae) is a pantropical medicinal plant used traditionally for inflammatory, infectious, and metabolic conditions, with documented antioxidant, anti-arthritic, and antimicrobial properties (Liew and Yong 2016; Ololade et al. 2017). Its anticancer potential has been evaluated only against breast (T47D) and cervical (HeLa) cell lines (Widiyastuti et al. 2019); no study has assessed lung cancer activity, and none has applied a network pharmacology framework to this species. A systematic review of 101 network-pharmacology studies across lung, breast, colorectal, and prostate cancer, screened from 7,153 records, did not identify a single study on *S. jamaicensis* (Merecz-Sadowska et al. 2026), confirming this as a genuine gap.

A further gap concerns solvent selection: most phytochemical studies of these species characterize a single crude extract or fraction, leaving open the question of which polarity range concentrates the

compounds most relevant to a given disease target. This study addresses both gaps by comparatively evaluating the drug-likeness, predicted target profiles, network pharmacology convergence, and molecular docking behaviour of phytochemicals across three solvent fractions (ethyl acetate, n-butanol, and residue) of *S. jamaicensis* leaf extract against lung cancer. We ask three specific questions: (1) which fraction yields the greatest proportion of drug-like (2) which fraction shows the strongest and most consistent target convergence with lung-cancer-associated genes; and (3) which fraction's compounds show the most favourable binding against the resulting hub gene targets.

2. Methodology

2.1 Plant Material and Extraction

Fresh leaves of *Stachytarpheta jamaicensis* (L.) Vahl were collected from Oja Tuntun, Ilorin, Kwara State, Nigeria, and authenticated at the Department of Plant Biology, University of Ilorin (Voucher No. UILH/001/1866/2026). Air-dried, powdered leaf material was macerated in absolute ethanol (72 h) and successively partitioned by increasing polarity into n-hexane, ethyl acetate, n-butanol, and residue fractions (Abubakar and Haque 2020); this study reports comparative findings for the ethyl acetate, n-butanol, and residue fractions.

2.2 Phytochemical Profiling

Qualitative phytochemical screening followed established protocols (Finar 1986; Kokate 1999; Odebiyi and Sofowora 1978; Yasuma and Ichikawa 1953). Compound fingerprinting used LC-MS (Waters e2695/W2998/ACQ-QDA) (Garba et al. 2023). The full phytochemical characterisation of these fractions, including the complete LC-MS compound inventory, is reported in a companion manuscript currently under review; the present study uses that profile as its starting point.

2.3 Drug-Likeness Screening

Physicochemical and pharmacokinetic properties were predicted using SwissADME (Daina et al. 2017). Compounds were filtered using five complementary rule-based criteria: Lipinski (Lipinski et al. 2001), Ghose (Ghose et al. 1999), Veber (Veber et al. 2002), Egan (Egan et al. 2000), and Muegge (Muegge et al. 2001). Violations were summed across the five rules, and compounds with no more than one violation in total were considered drug-like (Fig. 1); of these, compounds with reported anticancer or related bioactivity, for the compound itself or its structural class, were carried forward. Silychristin and Marchantin J were retained despite exceeding this threshold because of documented anticancer-relevant activity of Silychristin and of the macrocyclic bisbibenzyl class to which Marchantin J belongs (Viktorová et al. 2019; Xu et al. 2010). Predicted gastrointestinal absorption and bioavailability score were recorded as complementary information and were not used as exclusion criteria.

2.4 Target Prediction and Network Construction

Putative protein targets for prioritized compounds were predicted via SwissTargetPrediction (Daina et al. 2019) and intersected with disease-associated gene sets for five cancer types retrieved from GeneCards (Stelzer et al. 2016) and Open Targets (Ochoa et al. 2023). Statistical significance of target-gene set overlaps for each cancer type was assessed using a hypergeometric enrichment test against a background of 20,000 human protein-coding genes, since disease gene set sizes varied substantially across the five cancers screened. Compound-target networks were built in Cytoscape (Shannon et al. 2003); protein-protein interaction networks used STRING at confidence > = 0.7 (Szklarczyk et al. 2019; Szklarczyk et al. 2021). Hub genes were identified via Maximal Clique Centrality in CytoHubba (Chin et al. 2014). Hub genes were defined as those ranking among the top 10 by MCC in at least two of the three fractions. Of the genes meeting this criterion, five were prioritized for molecular docking on the basis of (i) established roles in lung cancer, (ii) an experimentally resolved three-dimensional structure with a well-defined ligand-binding site suitable for docking, and (iii) a known reference inhibitor against which phytochemical binding could be benchmarked. GO and KEGG enrichment used ShinyGO v0.85.1 (Ge et al. 2020). For each fraction, the top 50 KEGG pathways ranked by false discovery rate (FDR) were retained, and cancer-relevant pathways were selected for cross-fraction comparison.

2.5 Molecular Docking

Molecular docking was performed using the Schrödinger Suite (Release 2024-1; Maestro, LigPrep, Protein Preparation Wizard, and Glide). Three-dimensional structures of the prioritized phytochemicals and of the reference inhibitors acriflavine (HIF1A), dasatinib (SRC), ipatasertib (AKT1), erlotinib (EGFR) and venetoclax (BCL2) were retrieved from PubChem and prepared with LigPrep: ionization and tautomeric states were generated with Epik at pH 7.4 ± 2.0, specified chiralities were retained, and structures were energy-minimized with the OPLS4 force field. Where stereochemistry was undefined in the source record, stereoisomers were generated during ligand preparation and the lowest-energy isomer retained. Target protein structures were retrieved from the RCSB Protein Data Bank: SRC (PDB 1Y57, 1.91 Å), AKT1 (PDB 4EKL, 2.00 Å), EGFR (PDB 7U9A, 2.60 Å), BCL2 (PDB 6GL8, 1.40 Å) and the HIF-1α PAS-B domain in heterodimer with ARNT PAS-B (PDB 4H6J, 1.52 Å)(Cardoso et al. 2012). Structures were prepared using the Protein Preparation Wizard (Sastry et al. 2013): bond orders were assigned from the Chemical Component Dictionary, hydrogens were added, disulfide bonds created, selenomethionines converted to methionines, missing side chains and loops modelled with Prime, het states generated with Epik at pH 7.4 ± 2.0, hydrogen-bond assignments optimized using PROPKA at pH 7.4, and the structures subjected to restrained minimization to a maximum heavy-atom RMSD of 0.30 Å with the OPLS4 force field. Water molecules more than 5 Å from any het group were removed during preprocessing, and all remaining crystallographic waters were deleted prior to grid generation. For SRC, AKT1, EGFR and BCL2, receptor grids were generated using the co-crystallized ligand to define the active site, with an inner box of 10 Å and an enclosing box accommodating ligands of up to 25 Å in length; no positional or hydrogen-bond constraints were applied. PDB 4H6J contains no co-crystallized ligand, and SiteMap did not identify a high-scoring pocket in the vicinity of the ligand-accessible Cys255 (Dscore 0.442, below the 0.83 threshold conventionally taken to indicate a druggable site). The HIF-1α grid was therefore defined geometrically: Cys255 of the HIF-1α PAS-B chain was selected and the selection expanded to all

complete residues having any atom within 5 Å, yielding a fifteen-residue shell comprising Phe242, Leu243, Ser244, Glu245, Phe252, Ser253, Tyr254, Cys255, Asp256, Glu257, Ile259, Pro267, Leu270 and Leu271 of HIF-1α, together with His378 of the ARNT PAS-B chain; the grid was centered on the centroid of these residues using the same box dimensions. Docking at this site was performed non-covalently, as none of the phytochemicals examined possesses an electrophilic group capable of the covalent modification of Cys255 reported for arylsulfonamide probes. Docking was performed in Glide using sequential Standard Precision and Extra Precision scoring (Friesner et al. 2004; Friesner et al. 2006; Halgren et al. 2004), against the five prioritized hub gene targets (Section 2.4). Where a ligand failed to return an Extra Precision pose, its Standard Precision score is reported for comparison.

The docking protocol was validated by redocking each co-crystallized ligand into its corresponding receptor grid, across all Epik-generated protonation states, with heavy-atom RMSD computed relative to the crystallographic pose. In addition to whole-ligand heavy-atom RMSD analysis, regional RMSD assessment was performed for the SRC kinase co-crystallized ligand (MPZ; PDB ID: 1Y57). The ligand was divided into central-scaffold and peripheral regions to investigate the structural contributions to the observed deviation. Corresponding heavy atoms were compared without independent ligand fitting, using the crystallographic ligand as the reference. Extra Precision RMSD values were 0.52–0.69 Å for AKT1 (4EKL, ligand 0RF), 0.21–0.37 Å for EGFR (7U9A, ligand M19), 0.80–3.70 Å for BCL2 (6GL8, ligand S55746) and 2.89–3.46 Å for SRC (1Y57, ligand MPZ) (Online Resource 1). Reproduction accuracy tracked ligand flexibility rather than target identity: the compact ATP-competitive ligands 0RF and M19 were reproduced to sub-Ångström accuracy, whereas deviation for the larger and more flexible MPZ and S55746 was localized to solvent-exposed substituents. Redocking validation was not possible for HIF-1α, as PDB 4H6J is an apo structure.

3. Results and Discussion

3.1 Comparative Drug-Likeness Screening

Following LC-MS-based phytochemical identification, compounds from the three fractions were evaluated for drug-likeness using SwissADME (Fig. 1; full results in Online Resource 2). In total, 27 compounds met the drug-likeness criterion (no more than one violation across the five rules). The proportion of drug-like compounds was highest in the ethyl acetate (9/24, 37.5%) and residue (12/33, 36.4%) fractions and lowest in the n-butanol fraction (6/23, 26.1%), consistent with the greater polarity and glycoside content of n-butanol-soluble constituents. From these, nine compounds with reported anticancer or related bioactivity were prioritized, together with Silychristin and Marchantin J, which were retained for their documented anticancer-relevant activity, giving eleven phytochemicals for network pharmacology and docking: Pelargonidin, (+)-Jasmonic acid, 3,9-dihydroxy-1-methoxy-2,8-diprenylpterocarpan, and Ethyl-beta-carboline (ethyl acetate); Silychristin, Chalcone, Norharman, and Marchantin J (n-butanol); Secoisolariciresinol, Vindoline, and 4-Acetoxy-3,3',4',7-tetramethoxyflavan (residue). Rotenone, although it had no violations of any of the five rules, was not prioritized because of its well-documented mitochondrial toxicity and neurotoxicity (Betarbet et al. 2000).

3.2 Comparative Target Overlap and Cancer-Type Selection

Predicted protein targets for the eleven prioritized compounds, generated via Swiss Target Prediction, were intersected with disease-associated gene sets for five major cancer types (lung: 406/8,374; liver: 403/8,204; breast: 385/7,395; prostate: 375/6,854; colorectal: 348/5,961). Because disease-associated gene set size varied considerably across cancer types, raw overlap counts were corrected for this using a hypergeometric enrichment test (background N = 20,000 protein-coding genes). All five cancer types showed highly significant target enrichment ($p < 1x10-73$ in every case), confirming that the compound-predicted target pool is genuinely disease-relevant across the five cancers screened rather than for any one in particular. Fold-enrichment relative to chance ranked colorectal (2.34-fold) and prostate (2.19-fold) marginally above lung cancer (1.94-fold), indicating that lung cancer's higher raw overlap count partly reflects its larger disease gene set rather than disproportionately stronger enrichment. Lung cancer was nonetheless selected as the disease context for all subsequent analyses on the basis of its substantially higher global mortality burden relative to the other four cancers screened, and its comparably strong, highly significant enrichment. This finding is consistent with, and extends, prior work establishing *S. jamaicensis's* antioxidant, anti-inflammatory, and antibacterial properties (Ololade et al. 2017), by identifying for the first time a disease-specific mechanistic rationale for its anticancer relevance.

Comparing the three fractions directly, ethyl acetate showed the highest percentage overlap between its predicted target set and lung-cancer-associated genes (299 of 360 targets, 83.06%), followed closely by n-butanol (260 of 315, 82.54%) and residue (236 of 287, 82.23%) (Fig. 2). The narrow spread across fractions (under 1 percentage point) indicates that target relevance to lung cancer was not concentrated in a single fraction but distributed consistently across all three, supporting a whole-extract rather than single-fraction rationale for anticancer activity.

3.3 Hub Gene Convergence

Protein-protein interaction networks (STRING, confidence > = 0.7) were constructed for each fraction's overlapping target set, and hub genes were identified via Maximal Clique Centrality in CytoHubba. Nine genes ranked among the MCC top 10 in at least two fractions (HIF1A, AKT1, BCL2, SRC, EGFR, NFKB1, PPARG, ESR1, and MMP9; Online Resource 3), but only HIF1A ranked in the top 10 of all three. HIF1A therefore emerged as the single hub shared across the chemically distinct compound pools of the three fractions. AKT1 recurred in the ethyl acetate and residue fractions, BCL2 in the ethyl acetate and n-butanol fractions, and SRC and EGFR in the n-butanol and residue fractions; AKT1 was absent from the n-butanol target set altogether. HIF1A, AKT1, SRC, EGFR, and BCL2 were prioritized for docking according to the criteria described in Section 2.4. All five prioritized hub genes have independently documented relevance to lung cancer: HIF1A overexpression is an established prognostic biomarker and therapeutic target in both non-small cell and small cell lung cancer (Yang et al. 2016), while AKT1 and EGFR are among the most frequently reported network-pharmacology hub genes in lung cancer studies, with AKT1 the single most frequently identified target (Merecz-Sadowska et al. 2026). SRC, which

promotes migration, adhesion, and invasion in lung cancer (Ramos et al. 2025), and BCL2, whose overexpression in NSCLC protects tumour cells from apoptosis (Zapata Dongo et al. 2025), complete a hub panel spanning proliferative signaling, hypoxic adaptation, and cell survival, affirming that this convergence reflects genuine disease-relevant biology rather than a computational artefact.

3.4 Gene Ontology Enrichment: Convergence and Divergence

Gene Ontology enrichment (ShinyGO v0.85.1) revealed a consistent structural signature across all three fractions: "membrane raft" was the top enriched Cellular Component term in every fraction (fold enrichment 6.13x ethyl acetate, 7.77x n-butanol, 7.39x residue), and "cellular response to chemical stimulus" dominated Biological Process enrichment in all three (5.24x, 5.33x, 5.48x respectively). Molecular Function enrichment showed a shared kinase-centric signature: catalytic activity on protein and protein kinase activity dominated ethyl acetate and n-butanol, while residue's top Molecular Function term was protein kinase activity specifically (8.87x), consistent with the kinase component (AKT1, SRC, EGFR) of the prioritized hub gene panel.

3.5 KEGG Pathway Enrichment

KEGG enrichment of each fraction's lung-cancer-associated targets identified Pathways in Cancer (hsa05200) as the most significantly enriched pathway in all three fractions (FDR 4.4×10^{-30}, 1.8×10^{-34}, and 3.4×10^{-31}; 7.88-, 9.23-, and 9.25-fold; 54, 55, and 50 genes for the ethyl acetate, n-butanol, and residue fractions, respectively) (Fig. 3). Ten of the twelve cancer-relevant pathways compared in Fig. 3 were enriched in all three fractions. EGFR tyrosine kinase inhibitor resistance (hsa01521) was among the most strongly enriched pathways (14.6-, 17.9-, and 18.5-fold), a finding of direct relevance to non-small cell lung cancer, in which acquired resistance to targeted therapy limits long-term benefit; EGFR, SRC, and BCL2 mapped to this pathway in every fraction. Apoptosis was enriched in all three fractions, most strongly in the n-butanol (17.0-fold) and ethyl acetate (14.8-fold) fractions and less so in the residue fraction (10.1-fold), alongside PI3K-Akt signaling, central carbon metabolism in cancer, and PD-L1/PD-1 checkpoint signaling. Fraction-specific differences were also evident: small cell lung cancer (hsa05222) was enriched only in the ethyl acetate fraction (14.2-fold), and HIF-1 signaling was enriched in the ethyl acetate and residue fractions but was not among the top 50 pathways of the n-butanol fraction. The positions of four of the prioritized hub genes within the Pathways in Cancer reference map are shown in Fig. 4. Together, this convergence on core oncogenic, apoptotic, and drug-resistance pathways, combined with fraction-specific enrichment, supports a mechanistically coherent anticancer relevance distributed across the extract rather than confined to a single fraction's chemistry.

3.6 Comparative Molecular Docking

Docking was performed against the five prioritized hub gene targets using Standard Precision followed by Extra Precision scoring, with each receptor grid validated by redocking of its co-crystallized ligand (Online Resource 1). Reference inhibitor behaviour differed systematically between target classes. At the

three kinase ATP sites, the reference inhibitors returned strong and internally consistent scores, dasatinib at SRC (XP − 9.51 kcal/mol), erlotinib at EGFR (− 9.25 kcal/mol) and ipatasertib at AKT1 (− 11.26 kcal/mol), and these were also the grids at which redocking reproduced the crystallographic pose most accurately. At the two protein–protein interaction surfaces, the reference inhibitors failed to separate from the compound set: venetoclax scored − 5.44 kcal/mol at the BCL2 BH3 groove, below three of the phytochemicals tested, and acriflavine scored − 2.15 kcal/mol at the HIF-1α PAS-B site. An empirical scoring function that cannot rank a potent clinical inhibitor above a plant metabolite cannot be relied upon to rank those metabolites among themselves; HIF1A and BCL2 scores are therefore reported in full (Online Resource 5) but were not used to compare fractions.

Among the validated kinase targets, Silychristin (n-butanol) returned the strongest and most consistent results, with XP docking scores of − 9.43 kcal/mol at SRC and − 9.30 kcal/mol at EGFR, comparable to dasatinib (− 9.51 kcal/mol) and erlotinib (− 9.25 kcal/mol) respectively (Fig. 5). Near-reference binding reproduced at two independently validated targets is more informative than a single favourable score, and Silychristin's multitarget relevance is consistent with an independent network-pharmacology study of silibinin, from the same flavonolignan family, which identified SRC and the PI3K-Akt pathway as key targets (Li et al. 2024). Secoisolariciresinol (residue) was next strongest, scoring − 8.91 kcal/mol at SRC, − 7.74 kcal/mol at AKT1 and − 7.02 kcal/mol at EGFR; this compound has been independently identified as a stable binder of 6-phosphogluconate dehydrogenase in a separate lung cancer (A549) drug discovery study (Khan et al. 2023). In the ethyl acetate fraction, Pelargonidin scored − 7.37 kcal/mol at AKT1 and − 5.90 kcal/mol at EGFR, and remains of interest because independent experimental evidence shows that it directly induces oxidative-stress-mediated apoptosis in A549 lung cancer cells (Yue and Li 2024). No phytochemical exceeded its matched reference inhibitor at any target, which is the expected outcome when the comparator is an optimized clinical agent.

Comparing fractions by their strongest result at a validated target, n-butanol produced the most favourable single score (Silychristin–SRC, − 9.43 kcal/mol), followed by residue (Secoisolariciresinol–SRC, − 8.91 kcal/mol) and ethyl acetate (3,9-dihydroxy-1-methoxy-2,8-diprenylpterocarpan–EGFR, − 6.63 kcal/mol). The n-butanol and residue fractions each therefore yielded a compound binding within approximately 0.6 kcal/mol of a clinical kinase inhibitor, while ethyl acetate did not. As with target overlap and pathway enrichment, however, the differences between fractions remain small relative to the differences between compounds within them, reinforcing the conclusion that anticancer-relevant binding is distributed across the extract rather than concentrated in a single polarity range.

The convergence of the enrichment and docking analyses at EGFR is notable. EGFR tyrosine kinase inhibitor resistance was among the pathways enriched in all three fractions (Section 3.5), and EGFR was also the target at which redocking reproduced the crystallographic pose most accurately (0.21–0.37 Å). That a phytochemical from the n-butanol fraction returns a score at this target comparable to a first-generation EGFR tyrosine kinase inhibitor provides structural support for the pathway-level prediction by an independent method. Because docking scores are computational estimates rather than measured binding affinities, these results identify Silychristin and Secoisolariciresinol as priority candidates for

experimental validation at SRC and EGFR, with Pelargonidin a secondary candidate supported by independent evidence of apoptosis induction in A549 lung cancer cells (Yue and Li 2024).

3.7 Methodological Considerations and Limitations

The divergence in behaviour between target classes observed here reflects a recognized property of empirical docking scoring functions rather than a property of the fractions examined. Such functions are parameterized predominantly on enclosed, buried binding sites and reward the hydrophobic enclosure characteristic of enzyme active sites; protein–protein interaction surfaces present the opposite geometry, being large, shallow and solvent-exposed, with binding driven by extended surface contact rather than deep burial. The use of matched clinical reference inhibitors at every target made this distinction measurable rather than assumed. At the three kinase ATP sites, redocking reproduced the crystallographic pose closely (0.21–0.37 Å at EGFR and 0.52–0.69 Å at AKT1) and the reference inhibitors scored strongly and consistently at both precisions. Neither condition held at the two protein–protein interfaces: venetoclax, which binds BCL2 with sub-nanomolar affinity, returned − 5.44 kcal/mol and ranked below three of the phytochemicals tested, and at HIF-1α the scores of individual ligands shifted by up to 2.1 kcal/mol between two grids differing only in enclosing box size, a magnitude comparable to the differences between compounds. Fraction rankings derived from such data would describe the scoring function rather than the chemistry and were therefore not reported.

The HIF-1α result merits specific comment, since HIF1A was the one hub gene common to all three fractions and would, on topological grounds alone, have been the obvious priority for structure-based follow-up. HIF-1α is a transcription factor rather than an enzyme and possesses no catalytic site; the only ligand-accessible position reported for its PAS-B domain is Cys255, identified through covalent modification (Cardoso et al. 2012). SiteMap assigned this region a Dscore of 0.442, below the 0.83 value conventionally taken to indicate a druggable site, and the fifteen-residue shell surrounding Cys255 contains five polar or charged residues, indicating a partly solvent-exposed rather than enclosed cavity. These computational indicators are consistent with the clinical record: no direct small-molecule inhibitor of HIF-1α has reached advanced development, whereas the paralogous HIF-2α possesses a well-defined buried PAS-B cavity that has been exploited by allosteric inhibitors (Scheuermann et al. 2013). Network centrality and structural tractability are therefore distinct properties of a target, and a hub gene identified by topology alone should not be assumed to be a viable docking target. Reporting this divergence, rather than substituting a more tractable target after the fact, is necessary if network pharmacology prioritization is to be assessable.

Several limitations qualify these findings. Structural analysis of HIF-1α was necessarily restricted to the isolated PAS-B domain in heterodimer with ARNT PAS-B, as no experimental structure of the full-length protein is available; the AlphaFold model of full-length HIF-1α (UniProt Q16665) carries an average pLDDT of 63.66, indicating predominantly low-confidence prediction consistent with extensive intrinsic disorder, and was therefore not used. Redocking of the SRC reference ligand did not meet the 2.0 Å criterion, returning 2.89–3.46 Å across protonation states and both precisions; the pose taken forward

for regional analysis had a whole-ligand RMSD of 3.38 Å. That deviation was not evenly distributed. The central binding scaffold was reproduced to 0.76 Å, whereas the peripheral region and the piperazine-containing portion deviated by 5.00 Å and 6.98 Å respectively, and the Met341 hinge hydrogen bond seen in the crystallographic complex was preserved. Superimposition of the crystallographic and redocked MPZ structures (Online Resource 6) shows the same pattern: close agreement across the core, divergence at the solvent-exposed periphery. The elevated whole-ligand RMSD therefore reflects conformational freedom in a flexible substituent rather than a failure to reproduce the binding mode. SRC results are accordingly interpreted with corresponding caution, although the phytochemicals examined are substantially smaller and more conformationally restricted than that ligand. Finally, docking scores are computational estimates of binding mode and relative affinity rather than measured affinities, and support prioritization rather than conclusions about biological activity; experimental validation in lung cancer cell lines is required before any therapeutic inference can be drawn.

4. Conclusion

This study provides the first comparative network pharmacology evaluation of *Stachytarpheta jamaicensis* solvent fractions against lung cancer. Ethyl acetate, n-butanol, and residue fractions each showed high (> 82%) target overlap with lung-cancer-associated genes and shared HIF1A as a common hub gene, with AKT1, SRC, EGFR, and BCL2 each recurring in two fractions, and were jointly enriched for Pathways in Cancer, apoptosis, and EGFR tyrosine kinase inhibitor resistance, indicating that anticancer-relevant activity is a distributed property of the extract rather than isolated to one fraction. Silychristin (n-butanol) and Secoisolariciresinol (residue) emerged as the strongest lead compounds, Silychristin returning docking scores comparable to dasatinib at SRC and erlotinib at EGFR at redocking-validated binding sites. The docking stage also showed that network centrality does not guarantee structural tractability: HIF1A, the only hub gene common to all three fractions, presented no druggable pocket, and neither HIF1A nor BCL2 permitted reliable comparative scoring, a limitation that applied equally to their clinical reference inhibitors. Several additional drug-like constituents (Online Resource 2) were not prioritized in this study and remain candidates for future evaluation. These findings establish a mechanistic, target-level rationale for the anticancer potential of *S. jamaicensis* and identify concrete priority candidates for in vitro validation in lung cancer cell lines.

Declarations

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

All data supporting the findings of this study are included in the article and its Online Resources. Protein structures are available from the RCSB Protein Data Bank under accession codes 1Y57, 4EKL, 7U9A,

6GL8 and 4H6J. Compound structures are available from PubChem under the CIDs listed in Online Resource 5. Full SwissADME, KEGG enrichment, hub gene, redocking validation and docking score datasets are provided in Online Resources 1-5. Schrodinger Maestro project files are available from the corresponding author on reasonable request.

Statements and Declarations

Competing Interests: The authors declare no competing interests.

Authors' Declaration

The authors hereby declare that the work presented in this article is original and that any liability for claims relating to the content of this article will be borne by them.

Acknowledgement

The authors thank the Department of Biological Sciences (Biochemistry Unit), Al-Hikmah University, Ilorin, for laboratory and computational support.

References

1. Abubakar AR, Haque M (2020) Preparation of medicinal plants: basic extraction and fractionation procedures for experimental purposes. J Pharm Bioallied Sci 12:1–10. https://doi.org/10.4103/jpbs.jpbs_175_19
2. Betarbet R, Sherer TB, MacKenzie G, Garcia-Osuna M, Panov AV, Greenamyre JT (2000) Chronic systemic pesticide exposure reproduces features of Parkinson's disease. Nat Neurosci 3:1301–1306. https://doi.org/10.1038/81834
3. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, Jemal A (2024) Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 74:229–263. https://doi.org/10.3322/caac.21834
4. Cardoso R, Love R, Nilsson CL, Bergqvist S, Nowlin D, Yan J et al (2012) Identification of Cys255 in HIF-1α as a novel site for development of covalent inhibitors of HIF-1α/ARNT PasB domain protein-protein interaction. Protein Sci 21:1885–1896. https://doi.org/10.1002/pro.2172
5. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY (2014) cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol 8(Suppl 4):S11. https://doi.org/10.1186/1752-0509-8-S4-S11
6. Daina A, Michielin O, Zoete V (2017) SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep 7:42717. https://doi.org/10.1038/srep42717

7. Daina A, Michielin O, Zoete V (2019) SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules. Nucleic Acids Res 47(W1):W357–W364. https://doi.org/10.1093/nar/gkz382
8. Egan WJ, Merz KM, Baldwin JJ (2000) Prediction of drug absorption using multivariate statistics. J Med Chem 43:3867–3877. https://doi.org/10.1021/jm000292e
9. Finar IL (1986) Organic Chemistry, vol 1, 5th edn. Longman Scientific and Technical, Harlow
10. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT et al (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem 47:1739–1749. https://doi.org/10.1021/jm0306430
11. Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA et al (2006) Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. J Med Chem 49:6177–6196. https://doi.org/10.1021/jm051256o
12. Garba Y, Bello NU, Bashir M, Mohammed SY (2023) In vitro and in vivo analysis of the antioxidant potentials of methanolic extract of Tamarindus indica leaves. Niger J Biochem Mol Biol 38:104–111. https://doi.org/10.4314/njbmb.v38i3.1
13. Ge SX, Jung D, Yao R (2020) ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics 36:2628–2629. https://doi.org/10.1093/bioinformatics/btz931
14. Ghose AK, Viswanadhan VN, Wendoloski JJ (1999) A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. J Comb Chem 1:55–68. https://doi.org/10.1021/cc9800071
15. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL (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
16. Hopkins AL (2008) Network pharmacology: the next paradigm in drug discovery. Nat Chem Biol 4:682–690. https://doi.org/10.1038/nchembio.118
17. Kanehisa M, Goto S (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 28:27–30. https://doi.org/10.1093/nar/28.1.27
18. Khan GB, Qasim M, Rasul A, Ashfaq UA, Alnuqaydan AM (2023) Identification of lignan compounds as new 6-phosphogluconate dehydrogenase inhibitors for lung cancer. Metabolites 13:34. https://doi.org/10.3390/metabo13010034
19. Kokate CK (1999) Practical Pharmacognosy, 4th edn. Vallabh Prakashan, New Delhi
20. Li P, Wang D, Yang X, Liu C, Li X, Zhang X et al (2024) Anti-tumor activity and mechanism of silibinin based on network pharmacology and experimental verification. Molecules 29:1901. https://doi.org/10.3390/molecules29081901
21. Liew PM, Yong YK (2016) Stachytarpheta jamaicensis (L.) Vahl: from traditional usage to pharmacological evidence. Evid Based Complement Alternat Med 2016:7842340. https://doi.org/10.1155/2016/7842340

22. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 46:3–26. https://doi.org/10.1016/S0169-409X(00)00129-0
23. Merecz-Sadowska A, Sadowski A, Zajdel K, Jecek A, Sitarek P, Zajdel R (2026) A network pharmacology review of plant-derived anticancer compounds in lung, breast, colorectal and prostate cancer. Int J Mol Sci 27:6177. https://doi.org/10.3390/ijms27146177
24. Muegge I, Heald SL, Brittelli D (2001) Simple selection criteria for drug-like chemical matter. J Med Chem 44:1841–1846. https://doi.org/10.1021/jm015507e
25. Ochoa D, Hercules A, Carmona M, Suveges D, Baker J, Malangone C et al (2023) The next-generation Open Targets Platform: reimagined, redesigned, rebuilt. Nucleic Acids Res 51(D1):D1353–D1359. https://doi.org/10.1093/nar/gkac1046
26. Odebiyi OO, Sofowora EA (1978) Phytochemical screening of Nigerian medicinal plants II. Lloydia 41:234–246
27. Ololade ZS, Ogunmola OO, Kuyooro SE, Abiona OO (2017) *Stachytarpheta jamaicensis* leaf extract: chemical composition, antioxidant, anti-arthritic, anti-inflammatory and bactericidal potentials. J Sci Innov Res 6:119–125
28. Ramos R, Sousa C, Vale N (2025) Src and Abl as therapeutic targets in lung cancer: opportunities for drug repurposing. Pharmaceuticals (Basel) 18:1426. https://doi.org/10.3390/ph18101426
29. Sastry GM, Adzhigirey M, Day T, Annabhimoju R, Sherman W (2013) Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. J Comput Aided Mol Des 27:221–234. https://doi.org/10.1007/s10822-013-9644-8
30. Scheuermann TH, Li Q, Ma HW, Key J, Zhang L, Chen R et al (2013) Allosteric inhibition of hypoxia inducible factor-2 with small molecules. Nat Chem Biol 9:271–276. https://doi.org/10.1038/nchembio.1185
31. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D et al (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504. https://doi.org/10.1101/gr.1239303
32. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S et al (2016) The GeneCards suite: from gene data mining to disease genome sequence analyses. Curr Protoc Bioinf 54:1–33. https://doi.org/10.1002/cpbi.5. 1.30.
33. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J et al (2019) STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 47(D1):D607–D613. https://doi.org/10.1093/nar/gky1131
34. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S et al (2021) The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res 49(D1):D605–D612. https://doi.org/10.1093/nar/gkaa1074

35. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD (2002) Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem 45:2615–2623. https://doi.org/10.1021/jm020017n
36. Viktorová J, Dobiasová S, Řehořová K, Biedermann D, Káňová K, Šeborová K et al (2019) Antioxidant, anti-inflammatory, and multidrug resistance modulation activity of silychristin derivatives. Antioxid (Basel) 8:303. https://doi.org/10.3390/antiox8080303
37. Widiyastuti Y, Sholikhah IYM, Haryanti S (2019) Cytotoxic activities of ethanolic and dichloromethane extract of leaves, stems, and flowers of Jarong [*Stachytarpheta jamaicensis* (L.) Vahl.] on HeLa and T47D cancer cell line. AIP Conf Proc 2202:020101. https://doi.org/10.1063/1.5141714
38. Xu A, Hu Z, Qu J et al (2010) Cyclic bisbibenzyls induce growth arrest and apoptosis of human prostate cancer PC3 cells. Acta Pharmacol Sin 31:609–615. https://doi.org/10.1038/aps.2010.37
39. Yang SL, Ren QG, Wen L, Hu JL (2016) Clinicopathological and prognostic significance of hypoxia-inducible factor-1 alpha in lung cancer: a systematic review with meta-analysis. J Huazhong Univ Sci Technolog Med Sci 36:321–327. https://doi.org/10.1007/s11596-016-1586-7
40. Yasuma A, Ichikawa T (1953) Ninhydrin-Schiff and alloxan-Schiff staining: a new histochemical staining method for protein. J Lab Clin Med 41:296–299
41. Yue L, Li Y (2024) Pelargonidin inhibits cell growth and promotes oxidative stress-mediated apoptosis in lung cancer A549 cells. Biotechnol Appl Biochem 71:1195–1203. https://doi.org/10.1002/bab.2621
42. Zapata Dongo RJ, Fontana D, Mologni L, Faya Castillo JE, Infante Varillas SF (2025) Inhibition of the anti-apoptotic protein BCL2 in EML4-ALK cell models as a second proposed therapeutic target for non-small cell lung cancer. PLoS ONE 20:e0308747. https://doi.org/10.1371/journal.pone.0308747

Figures

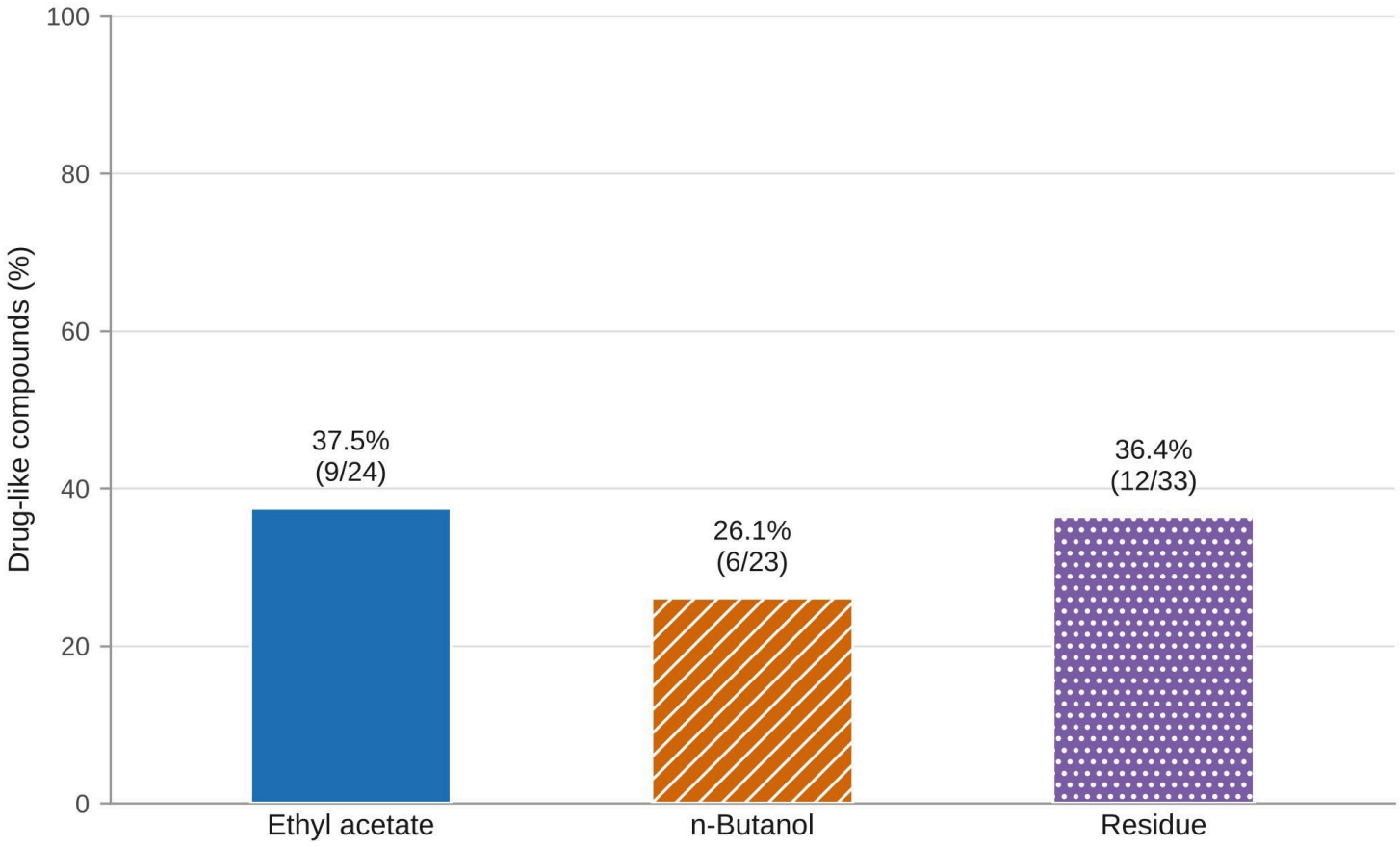


Figure 1

Proportion of drug-like compounds in each solvent fraction (no more than one violation in total across the Lipinski, Ghose, Veber, Egan, and Muegge rules; SwissADME), with drug-like/evaluated counts shown

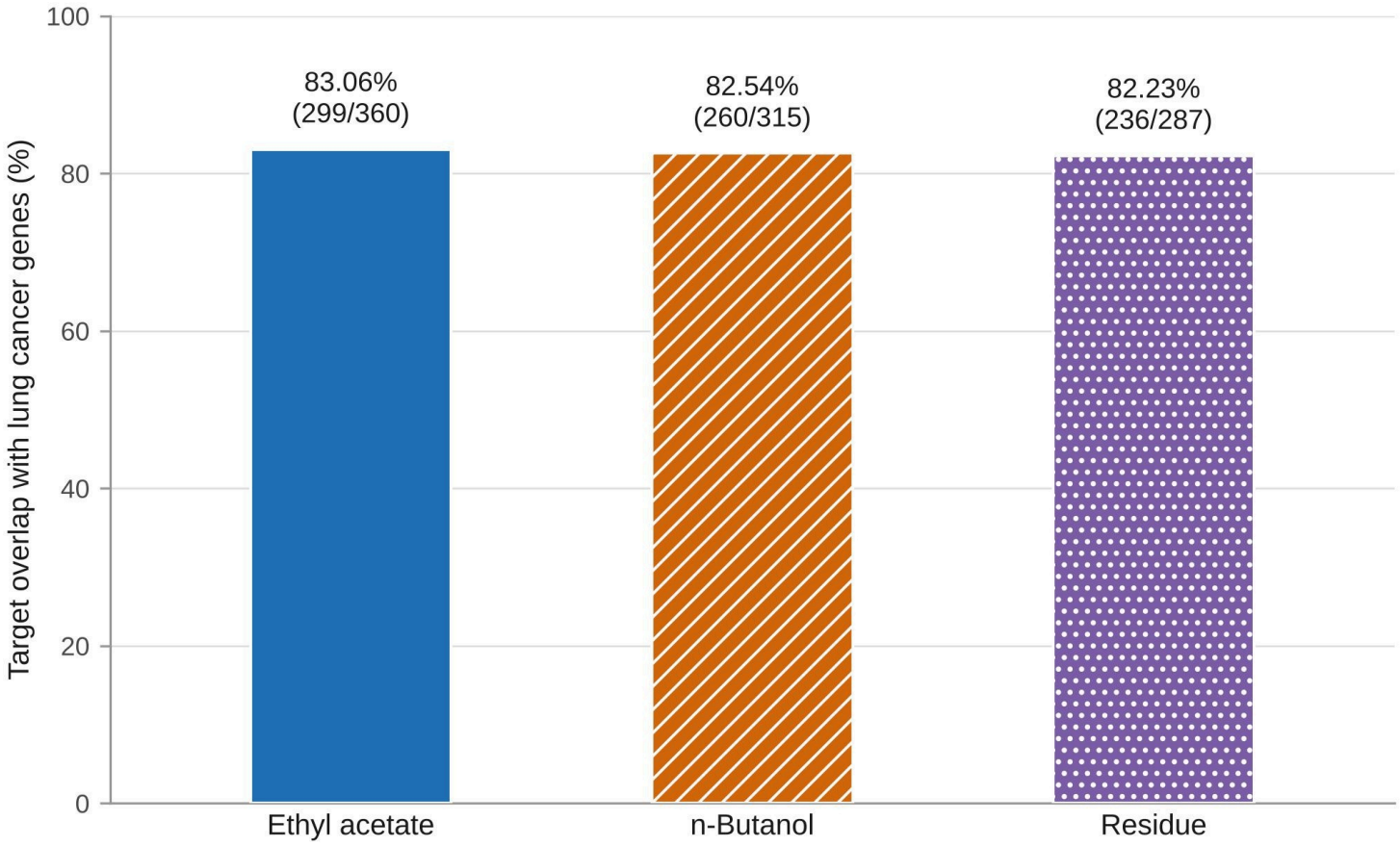


Figure 2

Percentage overlap between each solvent fraction’s predicted target set and lung cancer-associated genes

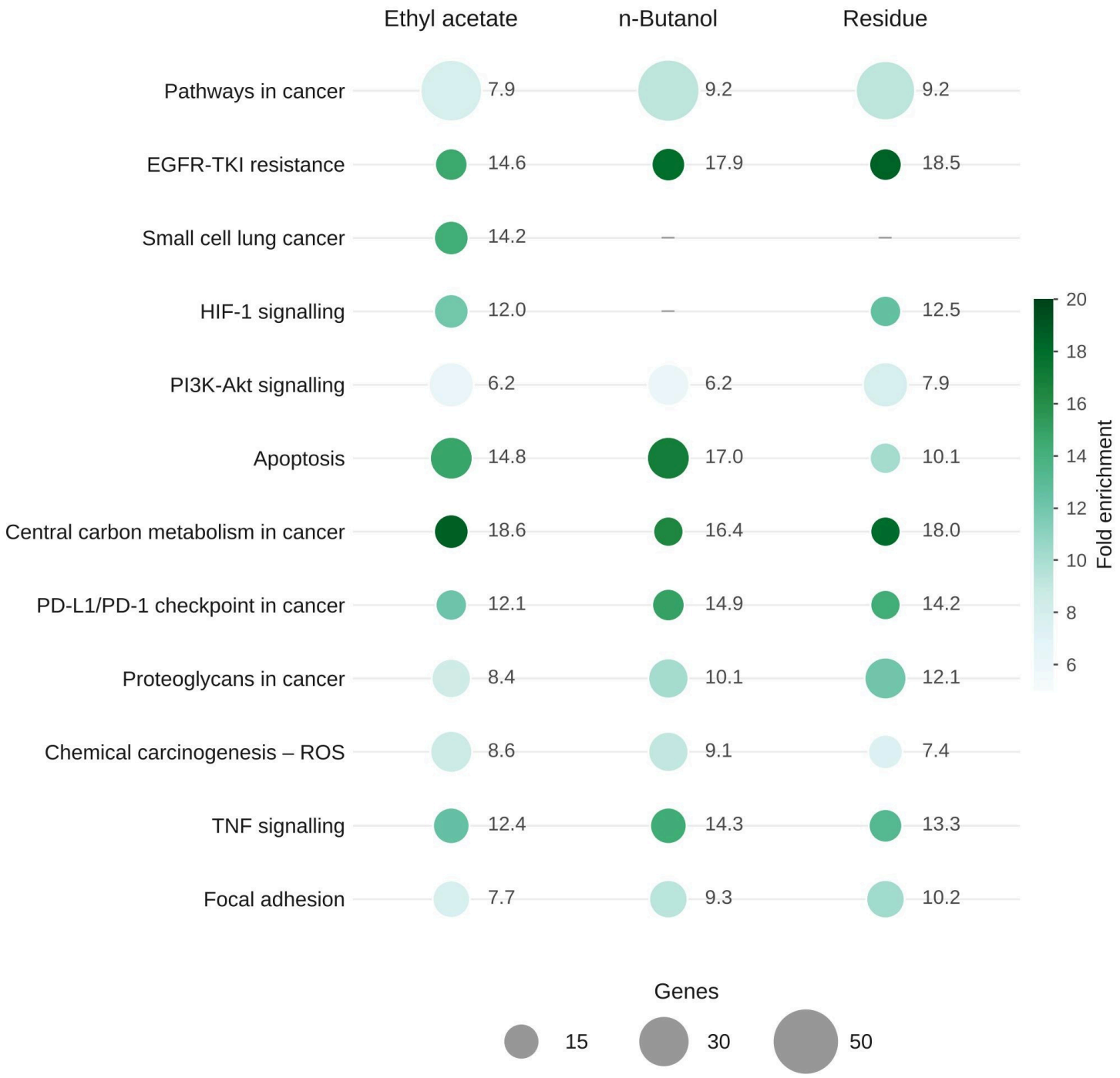


Figure 3

KEGG enrichment of selected cancer-relevant pathways across the ethyl acetate, n-butanol, and residue fractions (ShinyGO v0.85.1; top 50 pathways per fraction). Dot size indicates the number of overlapping genes; colour and labels indicate fold enrichment. All displayed pathways had FDR ≤ 1.2×10^{-9}. – indicates that the pathway was not among the top 50 enriched pathways for that fraction

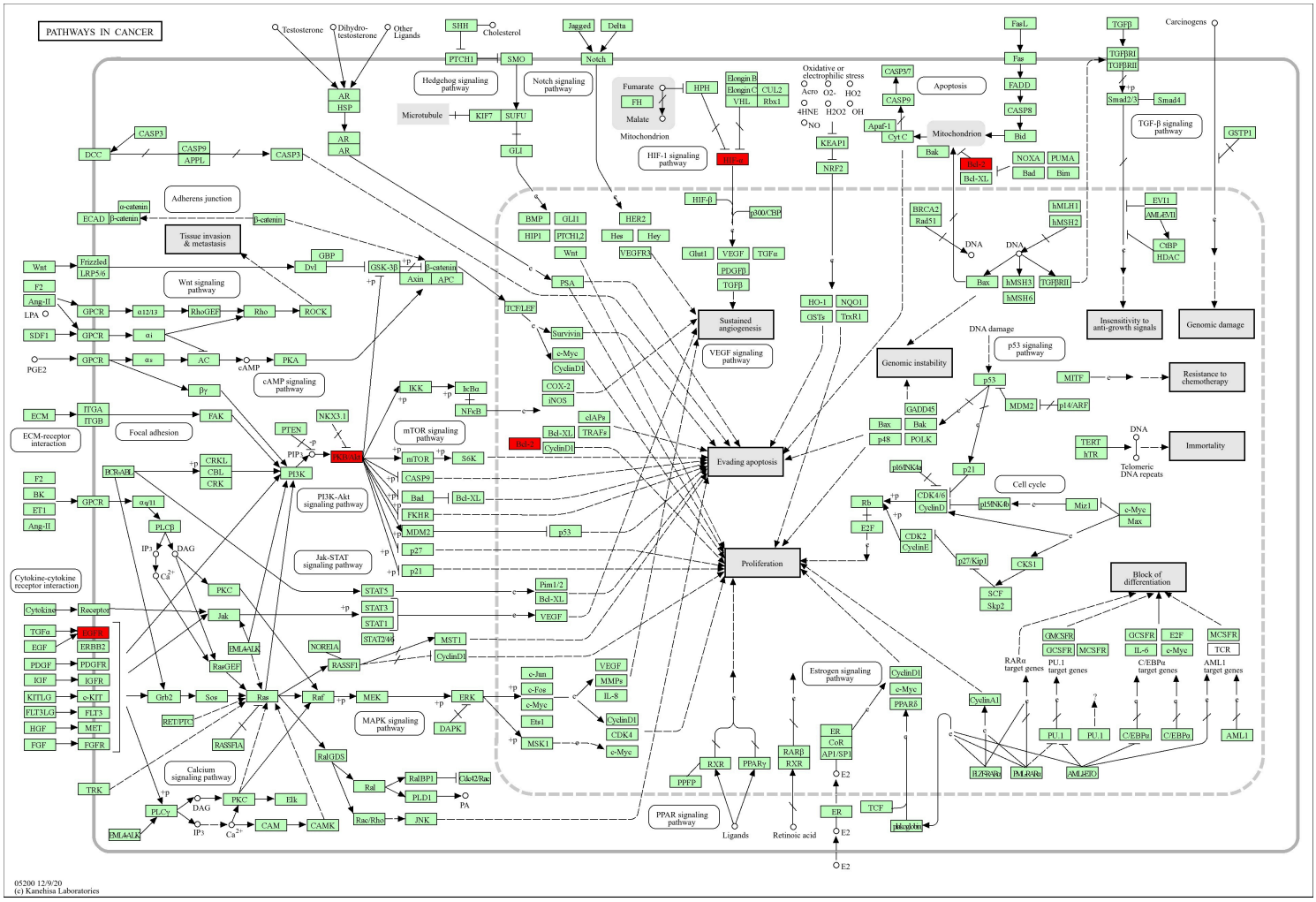


Figure 4

KEGG reference pathway map for “Pathways in Cancer” (hsa05200), with four of the five prioritized hub genes highlighted in red: HIF-α (HIF1A), PKB/Akt (AKT1), Bcl-2 (BCL2), and EGFR. SRC is not a constituent gene of this specific reference map and is therefore not shown. Pathway map reproduced from KEGG (Kanehisa and Goto 2000)

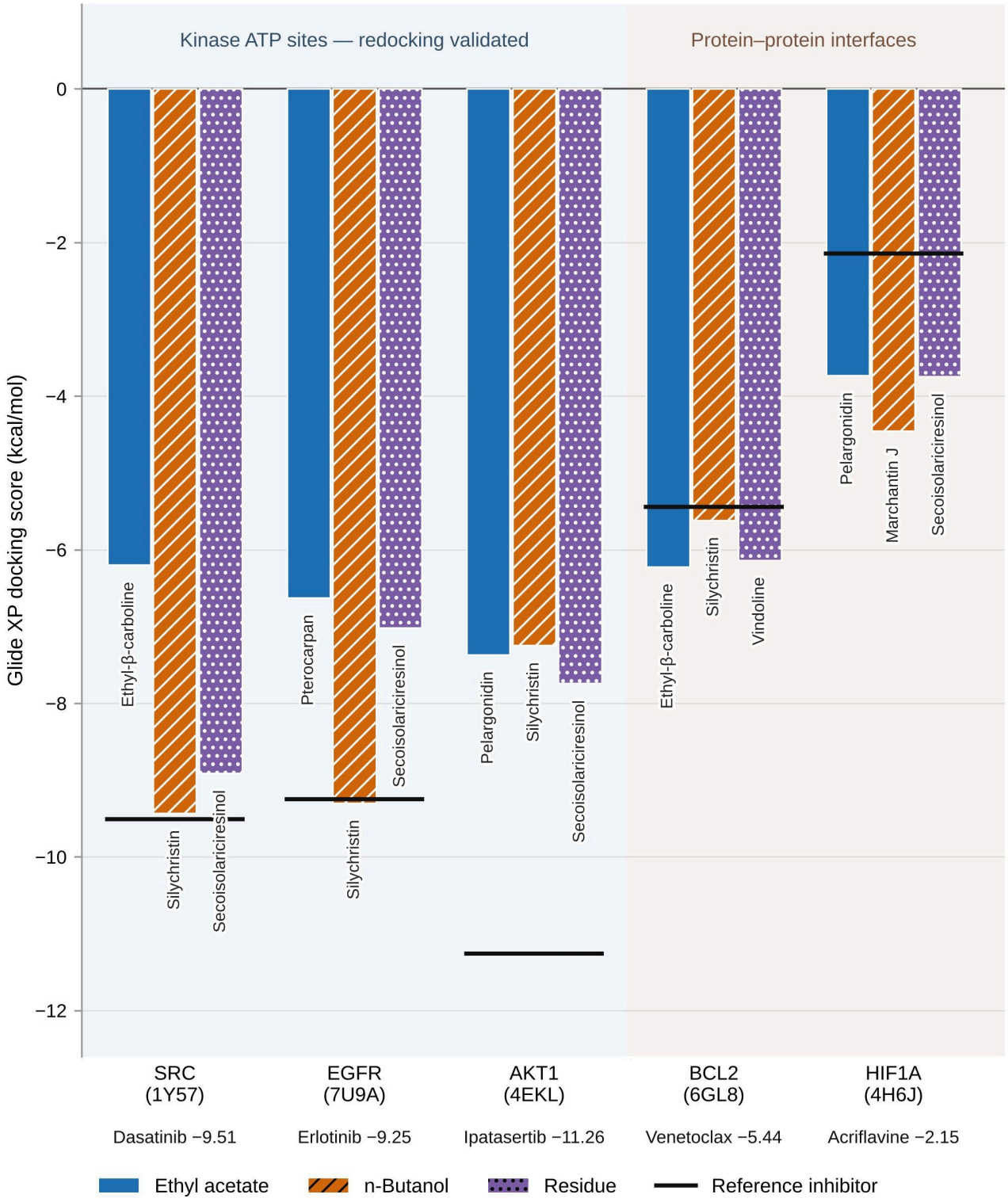


Figure 5

Glide XP docking scores (kcal/mol) for the strongest phytochemical from each solvent fraction at each of the five prioritized hub gene targets, with the matched reference inhibitor shown as a horizontal bar. More negative scores indicate more favourable predicted binding. Targets are grouped by binding-site type: the three kinase ATP sites, at which redocking validation was achieved, and the two protein–protein

interaction surfaces, at which the reference inhibitors did not separate from the compound set and scores were therefore not used for fraction comparison

Supplementary Files

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

- ESM1.xlsx
- ESM2.xlsx
- ESM3.xlsx
- ESM4.xlsx
- ESM5.xlsx
- ESM6.pdf
- SupplementaryInformation.docx