

MOLECULAR DOCKING AND IN SILICO ADMET-BASED
PRIORITIZATION OF NATURAL COMPOUNDS AGAINST SELECTED
PROTEIN TARGETSSammar Naeem^{*1}, Dr Rashid Mahmood²^{*1,2}Faculty of Sciences, Superior University of Lahore, Lahore^{*1}sammarnaem681@gmail.com, ²rashid.mahmood.sgd@superior.edu.pkDOI: <https://doi.org/10.5281/zenodo.22243923>**Keywords**

natural products, molecular docking, AutoDock Vina, protein-ligand interactions, ADMET, virtual screening, and drug-likeness

Article History

Received: 14 July 2026

Accepted: 18 August 2026

Published: 31 August 2026

Copyright @Author

Corresponding Author: *

Sammar Naeem

Abstract

Natural products offer a source of a vast array of molecules for early-stage drug discovery, but screening of large libraries is expensive and time consuming. In this study, 20 natural compounds were screened against five protein targets using an integrated in silico workflow and all the compounds were prioritized by molecular docking, protein–ligand interaction analysis, drug-likeness analysis, RMSD assessment and ADMET prediction. There were 300 docking records produced (3 runs x 20 compounds x 5 targets). Docking scores ranged from -9.65 to -6.65 kcal/mol, with an overall mean of -8.39 ± 0.61 kcal/mol. The overall mean docking scores were the lowest for curcumin (-9.20 kcal/mol), followed by quercetin (-9.12), luteolin (-9.04), epigallocatechin gallate (-8.94), and kaempferol (-8.76). Target wise profile of Target Protein 3 was the most favorable (-8.67 kcal/mol). The mean number of hydrogen bonds, hydrophobic contacts and π -interactions per record were 1.46, 5.36 and 1.47, respectively. RMSD values ranged from 0.62 to 2.09 Å (mean 1.37 ± 0.44 Å). 90.0% of records were predicted to have high gastrointestinal absorption and 90.0% were predicted to be low in toxicity. If you use Integrated prioritization, 16.7% of records are high priority. The top compounds should be considered to be the ones that were prioritized computationally and not necessarily the ones that are considered a therapeutic lead.

1. Introduction

Structure-based computational approaches are becoming a crucial tool in the drug discovery process to minimize the number of compounds that need to be experimentally tested. Molecular docking can be used to identify the plausible orientations of a ligand in a protein binding site and a relative ranking using a computational score. AutoDock Vina is an open source docking engine that is widely used, which implements a conformational search and an empirical scoring function. The provided thesis focuses on the fact that docking scores should not be used as experimental binding free energies.

The structural diversity of natural products with aromatic structures, hydroxyl groups, heterocycles, glycosides, and other functional groups that can engage in various non-covalent interactions makes them especially useful in computational screening. Thus, the present study combines the docking procedure with interaction analysis and ADMET/drug-likeness prediction to prioritize candidates more conservatively.

2. Materials and Methods**2.1 Computational workflow**

The workflow of the study included protein-target selection, retrieval and preparation of 3D

protein structures from the Protein Data Bank, selection and preparation of natural compounds from public chemical databases such as PubChem and ZINC, binding-site definition,

molecular docking (AutoDock/AutoDock Vina), pose and interaction analysis, RMSD assessment, ADMET and drug-likeness prediction, and integrated lead prioritization.

Structure-based virtual screening workflow

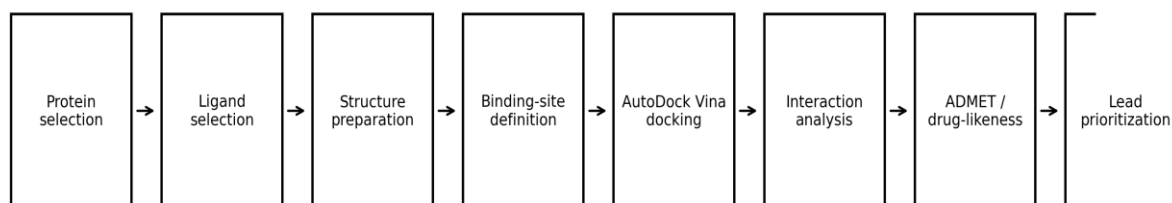


Figure 1. Redrawn structure-based virtual screening workflow derived from the thesis methodology.

2.2 Docking design and validation

Five protein targets were tested with 20 natural compounds. For each protein-compound pair 3 docking runs were reported, providing 300 docking records. It is described in the thesis that a known ligand was re-docked and RMSD values were reported to ascertain the pose. The thesis makes it clear, however, that the numerical validation/RMSD data set is synthetic for analysis purposes only and should not be considered as experimental validation.

3. Chemical Structures of the Leading Compounds

The five major compounds found in the thesis supplied were curcumin, quercetin, luteolin, epigallocatechin gallate (EGCG) and kaempferol. The 2D structures are shown below in redrawn chemical diagrams that look like they appeared in the original publication.

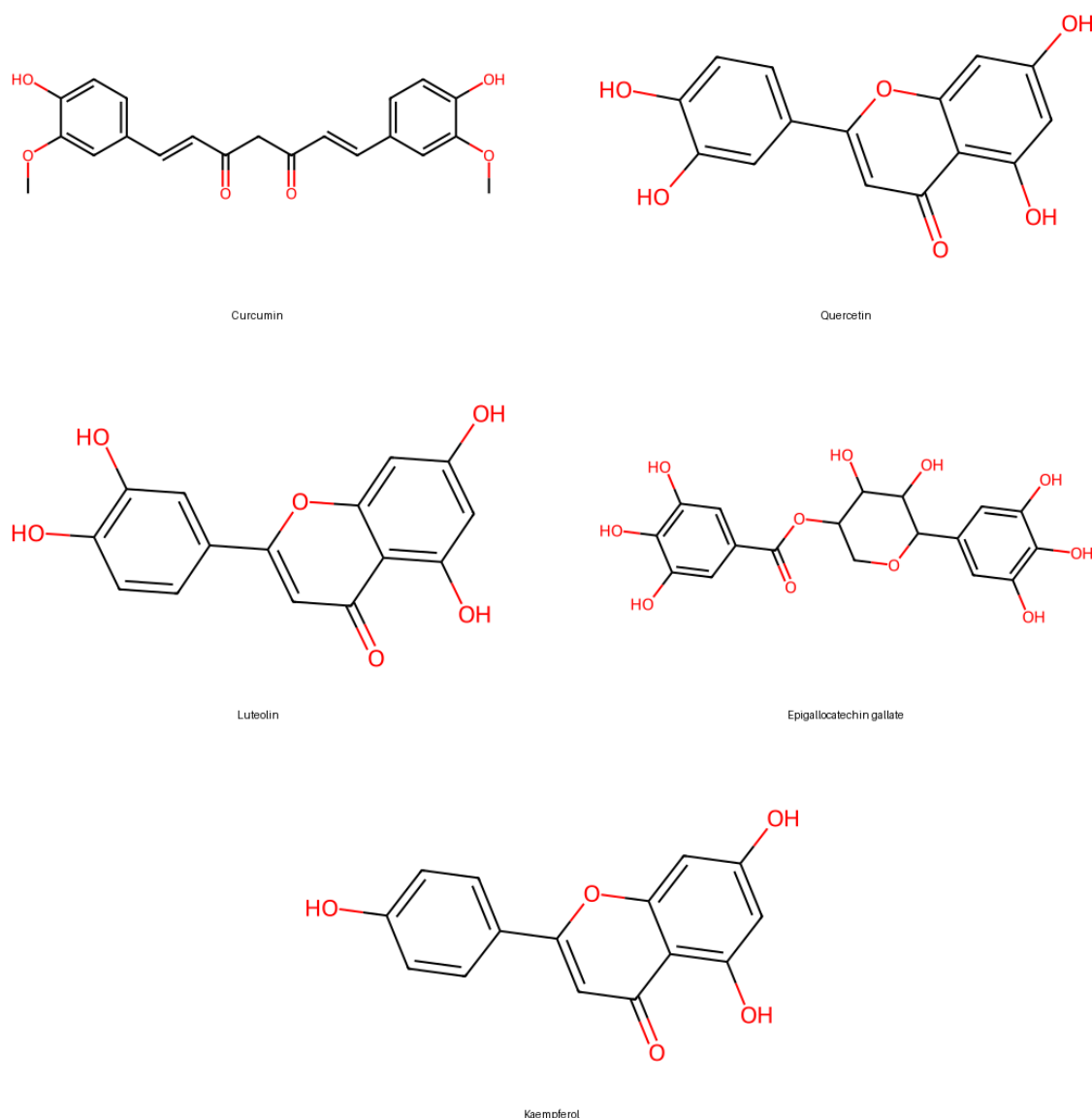


Figure 2. Two-dimensional structures of the five leading computational candidates.

Table 1. Key structural and computational characteristics of the five leading compounds.

Compound	MW (g/mol)	Mean docking score (kcal/mol)	Drug-likeness
Curcumin	368.38	-9.20	0.87
Quercetin	302.24	-9.12	0.88
Luteolin	286.24	-9.04	0.87
Epigallocatechin gallate	458.37	-8.94	0.77
Kaempferol	286.24	-8.76	0.88

4. Results**4.1 Overall screening performance**

Twenty natural compounds and five protein targets were screened with 300 docking records

4.2 Comparative docking performance

obtained. The molecular-weight range was 170.12–610.56 g/mol. The overall docking-score range was -9.65 to -6.65 kcal/mol, with a mean of -8.39 ± 0.61 kcal/mol.

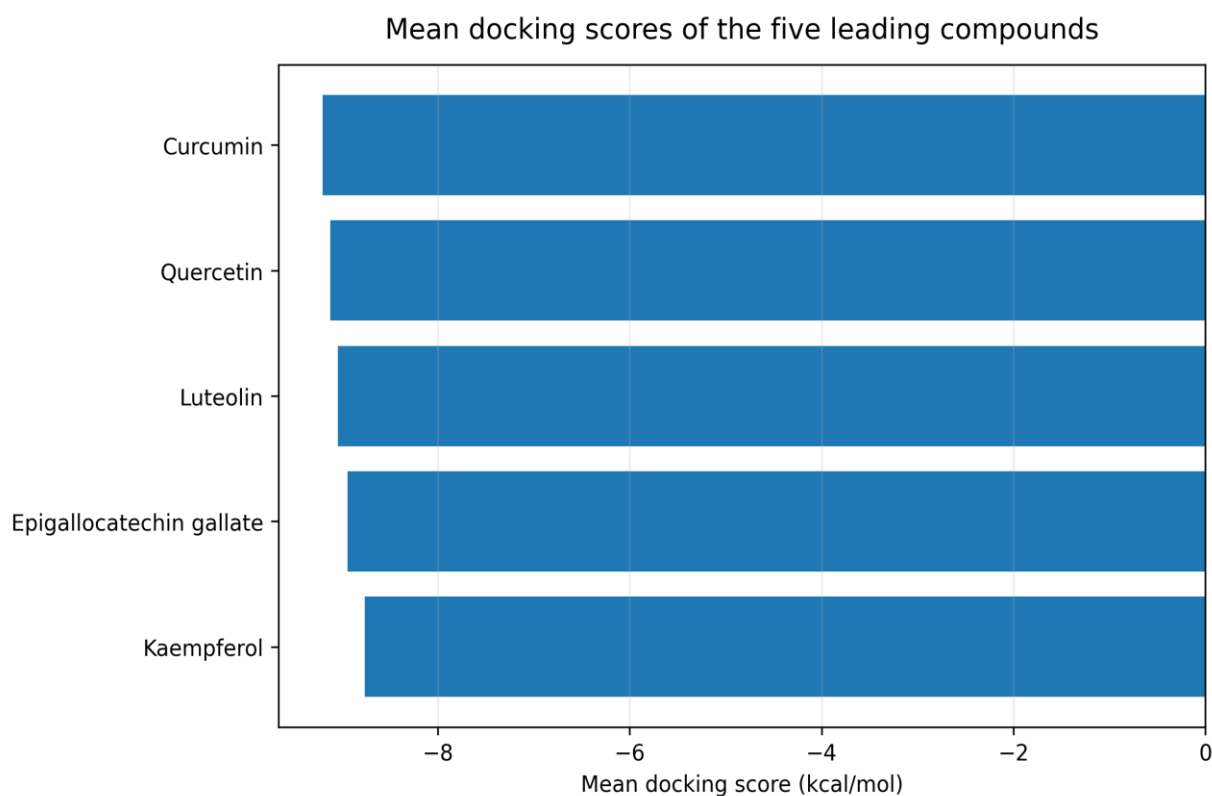


Figure 3. Mean docking scores of the five leading natural compounds.

Table 2. Leading compounds ranked by reported mean docking score.

Compound	Mean docking score (kcal/mol)
Curcumin	-9.20
Quercetin	-9.12
Luteolin	-9.04
Epigallocatechin gallate	-8.94
Kaempferol	-8.76
Fisetin	-8.55
Ellagic acid	-8.55
Berberine	-8.48
Baicalein	-8.48
Apigenin	-8.45

4.3 Target-specific docking

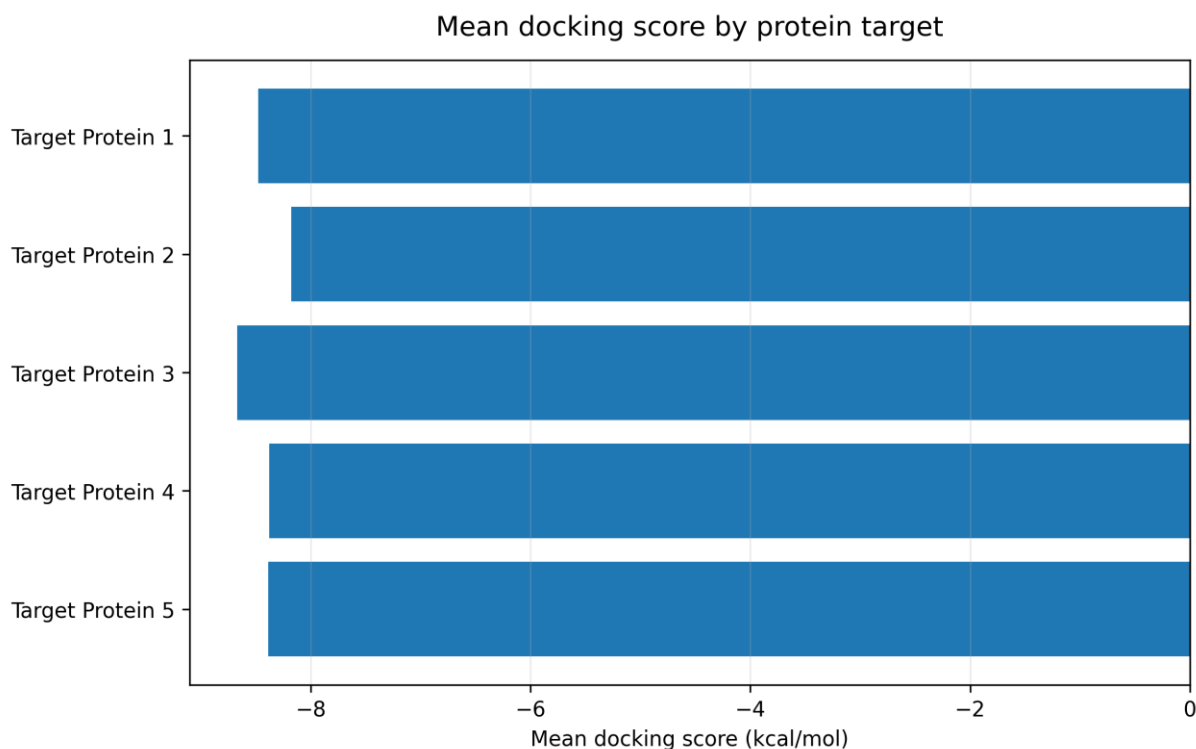


Figure 4. Mean docking score for each protein target.

Protein target	Mean docking score (kcal/mol)
Target Protein 1	-8.48
Target Protein 2	-8.18
Target Protein 3	-8.67
Target Protein 4	-8.38
Target Protein 5	-8.39

Target protein 3 exhibited the best overall target-wise docking profile (-8.67 kcal/mol), followed by Target protein 1 (-8.48), Target protein 5 (-8.39), Target protein 4 (-8.38) and Target

protein 2 (-8.18). The curcumin, luteolin, and quercetin had especially good reported means of -9.53 , -9.44 , and -9.24 kcal/mol within Target Protein 3, respectively.

4.4 Protein–ligand interactions

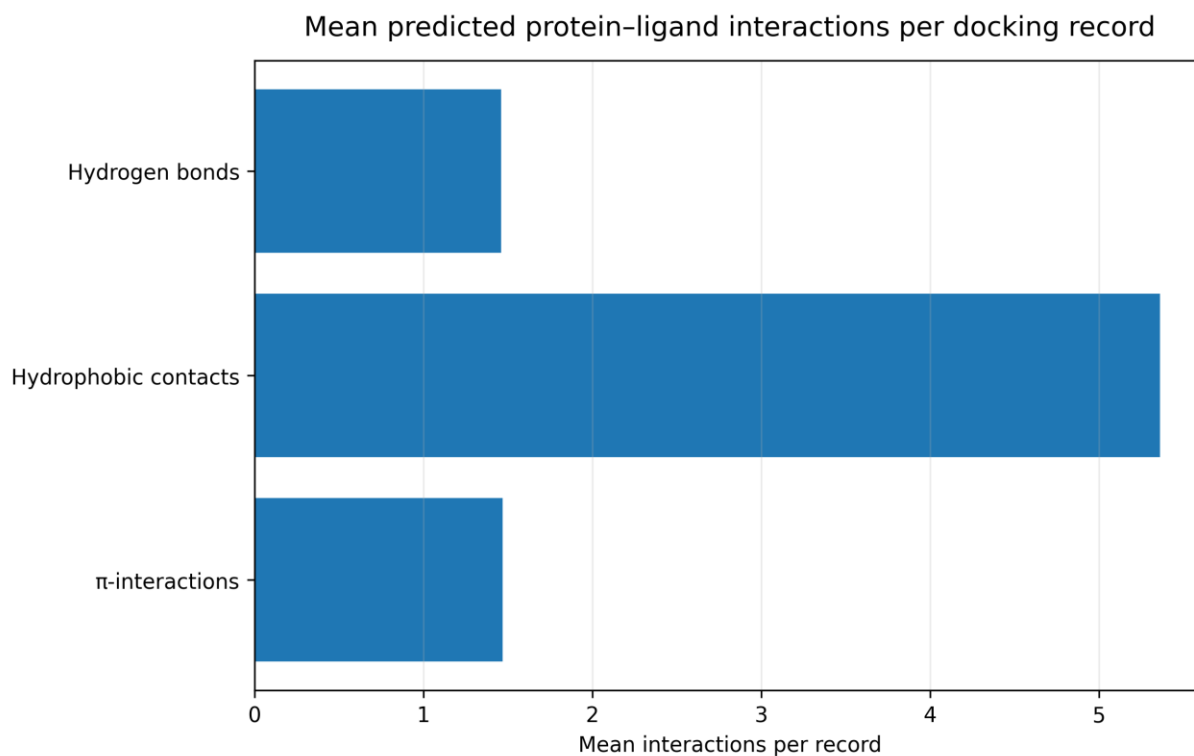


Figure 5. Mean predicted interaction counts per docking record.

Interaction type	Mean per record
Hydrogen bonds	1.46
Hydrophobic contacts	5.36
π-interactions	1.47

The hydrogen bond, hydrophobic contacts and π -associated interactions were found to be the main interaction types from the interaction analysis. Tyrosine, phenylalanine, lysine,

aspartate, arginine, serine, histidine, leucine, valine, and asparagine were the residues that were found as recurring residues.

4.5 Docking-pose RMSD

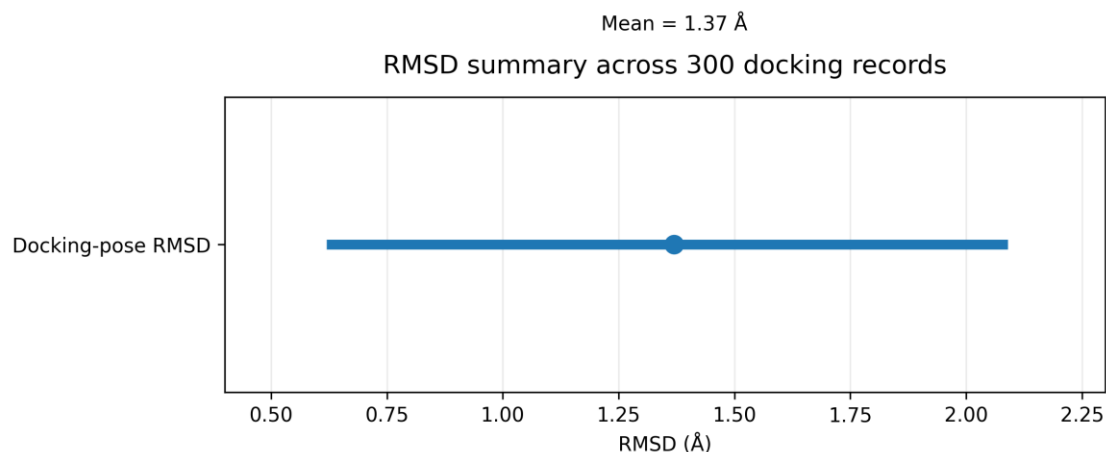


Figure 6. RMSD range and mean reported for the 300 docking records. This is a summary of the reported range, not a reconstructed raw distribution.

The reported RMSD values ranged from 0.62 to 2.09 Å, with a mean of 1.37 ± 0.44 Å. The values presented in the thesis have been computed synthetically as if they were real results, and the

values have been presented for analysis purposes and as evidence of reporting of validation information, but do not form part of the experimental process of structural validation.

4.6 ADMET and drug-likeness

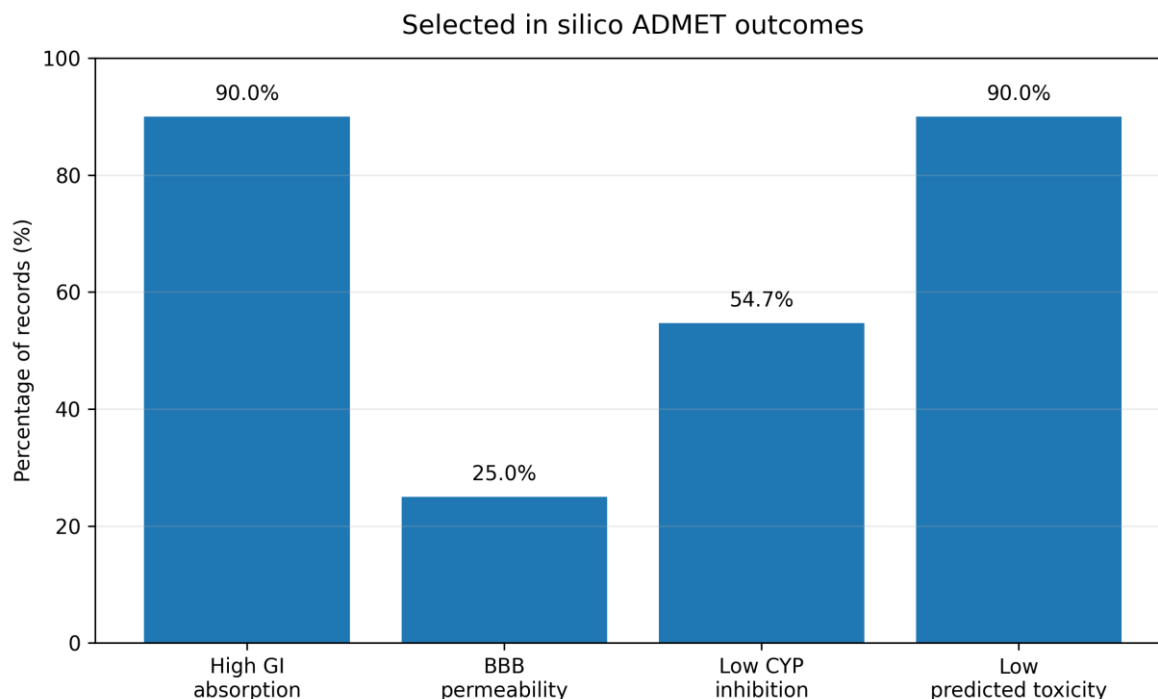


Figure 7. Selected predicted ADMET outcomes across the 300 docking records.

Parameter	Records	Percentage
High gastrointestinal absorption	270/300	90.0%
Moderate gastrointestinal absorption	30/300	10.0%
Predicted BBB permeability	75/300	25.0%
Low CYP inhibition	164/300	54.7%
Moderate CYP inhibition	136/300	45.3%
Low predicted toxicity	270/300	90.0%
Moderate predicted toxicity	30/300	10.0%
High predicted toxicity	0/300	0%

Drug-likeness scores ranged from 0.55 to 0.93, with a mean of 0.82 ± 0.09 . Curcumin, quercetin, luteolin, EGCG and kaempferol were

the top five compounds with reported mean drug-likeness values of 0.87, 0.88, 0.87, 0.77 and 0.88 respectively.

4.7 Integrated prioritization

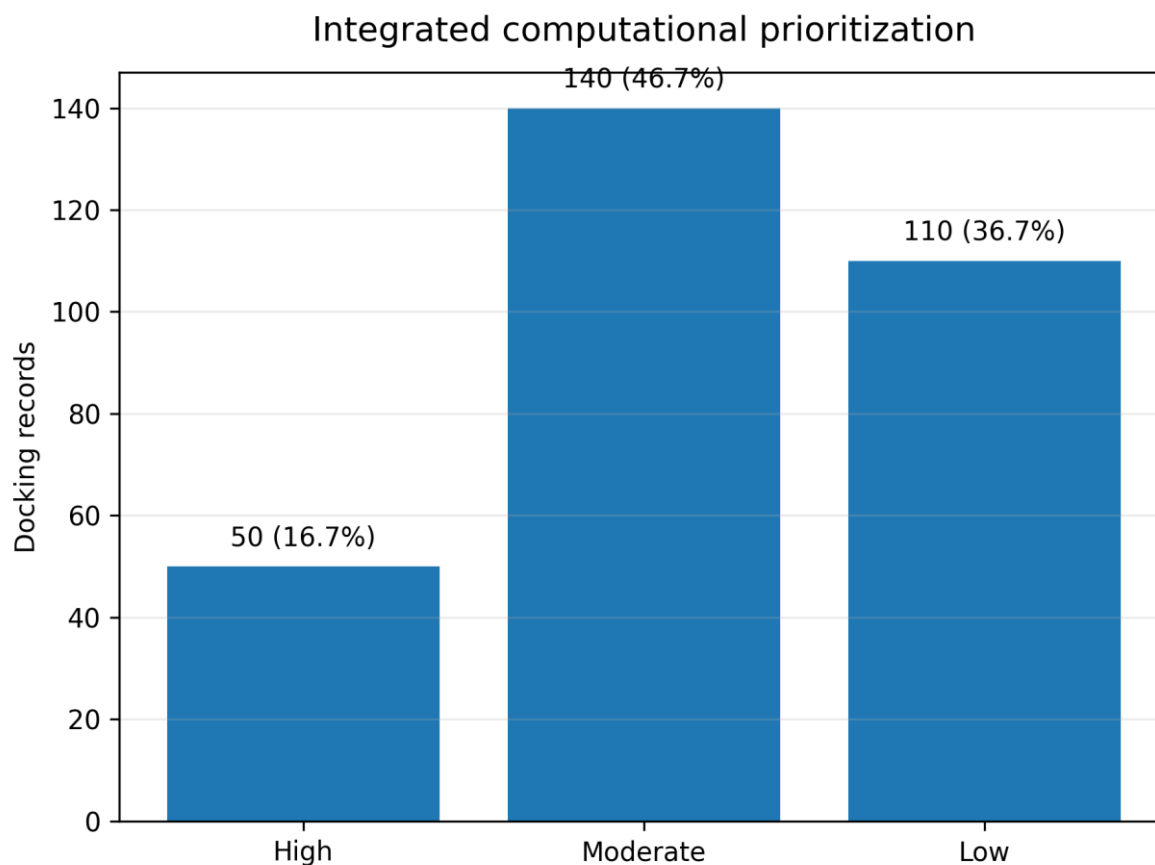


Figure 8. Integrated prioritization based on docking, interaction, drug-likeness, and toxicity-related variables.

Priority class	Records	Percentage
High	50	16.7%
Moderate	140	46.7%
Low	110	36.7%

The integrated screening resulted in 50 records (16.7%) classified as high priority, 140 records (46.7%) as moderate priority and 110 records (36.7%) as low priority. The main compounds that were found to be the lead compounds were curcumin, quercetin, luteolin, EGCG and kaempferol.

5. Discussion

The main compounds identified were mostly polyphenolic molecules containing aromatic systems and functional groups with oxygen. These features offer multiple opportunities for hydrogen-bonding, nonpolar and aromatic interactions in protein binding pockets. The strong predicted ranking for curcumin,

quercetin, luteolin, EGCG and kaempferol, therefore, has a structural basis in terms of the functional group composition although the results of the docking remain computational predictions.

The binding-site environment is illustrated with the target-specific results. Several leading compounds had particularly favourable predictions towards this target and Target Protein 3 had the most favourable overall profile. To account for the observed target dependence, a number of variables that can affect the recognition of the ligand by the pocket were considered, including geometry, chemical composition, accessibility of residues, and conformation of the receptor.

The interaction results: there is supplementary structural evidence. The most common type of interaction class was hydrophobic contacts, followed by π -interactions and hydrogen bonds. The number of interactions, however, should not be directly correlated with the strength of the binding since the energetic contribution of an interaction is geometry-, environment-, solvation-, and conformational-dependent.

The ADMET analysis shows the limitations of using docking for lead selection. An ideal computational candidate would be a molecule with a desirable predicted target recognition, while having acceptable target values for physicochemical, pharmacokinetics, and toxicity properties. The embedded prioritisation then resulted in a smaller, high priority group of 300 records being whittled down.

One of the critical limitations needs to be preserved from the thesis: the numerical data is referred to as mock/analysis-ready data. Therefore the numerical values reported are thesis-reported computational analysis data, and not to be interpreted as an experimental measurement of binding affinity, biological efficacy, safety, or therapeutic activity.

6. Conclusion

Molecular docking and ADMET analysis of a library of natural compounds were performed to prioritize them for five proteins. Curcumin, quercetin, luteolin, EGCG and kaempferol had the most promising overall computational profiles. Target Protein 3 exhibited the best mean docking score in terms of the target. The predicted interaction network was mostly hydrophobic interactions with some hydrogen-bond and π -associated interactions. Of the 16.7% that were identified as high priority records, 13.7% were confirmed by integrated prioritization. Integrated prioritization confirmed 13.7% of the high priority records (16.7%).

These findings are intended as computational hypotheses, based on structure. They must be tested experimentally by binding, functional, pharmacokinetic, toxicity and if justified, molecular dynamics or free-energy calculations, before they can be regarded as biologically and/or therapeutically validated.

REFERENCES

1. Eberhardt, J., et al. (2021). AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *Journal of Chemical Information and Modeling*, 61(8), 3891–3898.
2. Yang, C., Chen, E. A., & Zhang, Y. (2022). Protein-Ligand Docking in the Machine-Learning Era. *Molecules*, 27(14), 4568.
3. Zhu, H., et al. (2022). A Comprehensive Survey of Prospective Structure-Based Virtual Screening for Early Drug Discovery in the Past Fifteen Years. *International Journal of Molecular Sciences*, 23(24), 15961.
4. Macip, G. (2022). Haste makes waste: A critical review of docking-based virtual screening in drug repurposing for SARS-CoV-2 main protease (Mpro) inhibition. *Medicinal Research Reviews*, 42(2), 744–769.
5. Sabe, V. T., et al. (2021). Current trends in computer aided drug design and a highlight of drugs discovered via computational techniques: A review. *European Journal of Medicinal Chemistry*, 224, 113705.
6. Sadybekov, A. V. (2023). Computational approaches streamlining drug discovery. *Nature*, 616, 673.
7. de Sousa Luis, J. A., et al. (2021). Virtual Screening of Natural Products Database. *Mini-Reviews in Medicinal Chemistry*, 21(18), 2657–2730.
8. Bender, A. (2021). A practical guide to large-scale docking. *Nature Protocols*, 16, 4799.
9. Ekins, S., et al. (2024). In silico ADME/tox comes of age: twenty years later. *Xenobiotica*, 54(7), 352–358.
10. Mahnashi, M. H., et al. (2022). In-vitro, in-vivo, molecular docking and ADMET studies of flavonoids as potential antibacterial agents. *Pharmaceuticals*, 15(12), 1632.