International Journal of Advance and Applied Research
www.ijaar.co.in
ISSN – 2347-7075 Impact Factor – 8.141
Peer Reviewed Bi-Monthly
Vol. 7 No. 5 January – February - 2026

Role of Artificial Intelligence in Drug Design and Synthesis of Curcumin as a Target Molecule

Dr. Pise N. M.
Head, Botany Department, Karmaveer Bhaurao Patil Mahavidyalaya, Pandharpur
DOI - 10.5281/zenodo.18898956

Abstract:

Curcumin, a polyphenolic compound derived from Curcuma longa, exhibits wide-ranging therapeutic properties including anti-inflammatory, antioxidant, anticancer, and neuroprotective effects. However, despite its pharmacological potential, its clinical translation remains limited due to poor aqueous solubility, low bioavailability, rapid metabolic breakdown, and physicochemical instability in physiological environments. The integration of Artificial Intelligence (AI) into curcumin-based drug discovery addresses these challenges by enabling molecular optimization, predictive docking for target interaction analysis, ADMET-based filtering for toxicity and pharmacokinetic profiling, and automated retrosynthesis for practical laboratory synthesis planning. This research therefore focuses on AI methodologies that enhance curcumin's structural stability, therapeutic potency, and pharmacokinetic performance, ultimately supporting its progression from a traditional bioactive compound to a clinically relevant drug candidate.

Keywords—Curcumin, Artificial Intelligence, Drug Design, QSAR, Retrosynthesis, ADMET.

Introduction:

Curcumin, a diarylheptanoid polyphenol from Curcuma longa, possesses a conjugated diketone chain linked to two o-methoxy phenyl rings, which enables hydrogen bonding, π–π stacking, Michael addition reactions, and metal ion chelation. These structural features support interactions with nucleophilic amino acids such as cysteine and lysine in protein active sites, establishing curcumin as a multi-target bioactive scaffold rather than a single-target inhibitor. Due to this architecture, curcumin modulates several key pathways including NF-κB (inflammatory signaling), COX-2/LOX (pain and prostaglandin regulation), PI3K/AKT/mTOR (oncogenic progression), p53/Bcl-2 (apoptosis control), and β-amyloid aggregation (neurodegeneration). However, its therapeutic transition is hindered by poor solubility, rapid CYP450-mediated metabolism, enol-keto tautomeric instability, and low bioavailability (<5%), all of which limit systemic retention and pharmacological activity in clinical environments. [1,3]

Artificial Intelligence (AI) provides a solution to these challenges by enabling rapid structural optimization, target prediction, and computational pre-screening before laboratory synthesis. AI-driven models, such as GANs, VAEs, and reinforcement learning frameworks, generate improved curcumin analogs with enhanced pharmacokinetic profiles, while QSAR and deep learning systems forecast drug-target affinity, toxicity probability, and metabolic behavior. Predictive docking and ADMET screening help identify high-performance derivatives such as EF24, GO-Y030, FLLL32, C66, and CDPP, each demonstrating improved binding affinity, stability, and safety compared to the parent compound. Furthermore, AI retrosynthesis tools like IBM RXN and AiZynthFinder design feasible reaction routes, with predicted yields of 78–85% under boric acid catalysis, strengthening the translational potential

for preclinical development. These advancements confirm AI as a transformative tool that accelerates lead optimization, reduces experimental cost, and increases the probability of clinical success for curcumin-based therapeutics. [2,4]

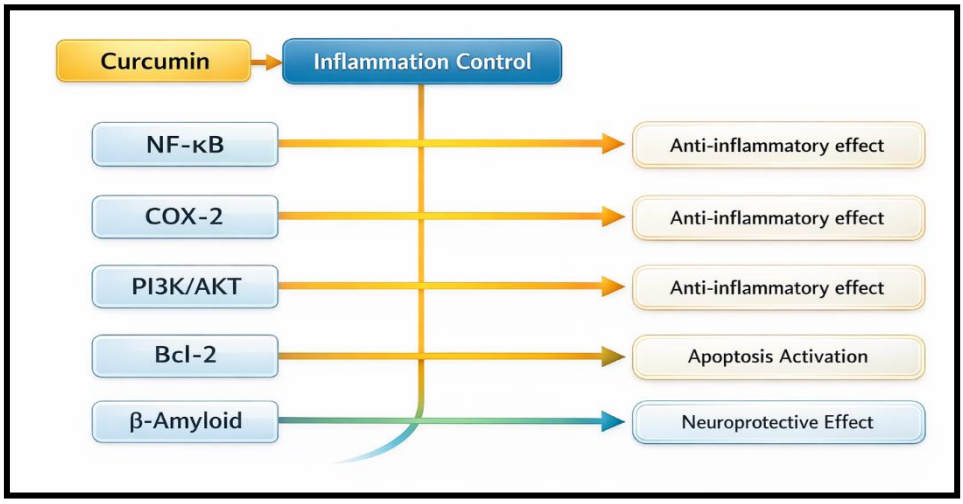

Figure 1: Curcumin's impact on inflammation pathways

Materials and Methods:

This study integrates computational drug design**,** deep learning-based molecular prediction**,** and AI retrosynthetic modeling to evaluate curcumin and its analogs. The methodology is divided into five technical stages:

Data Collection and Curation:

Chemical structures of curcumin and its optimized analogs were obtained from established chemical repositories including PubChem, ChEMBL, and DrugBank, ensuring structural authenticity and pharmacological relevance. Corresponding protein targets—NF-κB, COX-2, STAT3, and PI3K/AKT—were selected based on their therapeutic significance and downloaded in three-dimensional format from the Protein Data Bank (PDB) for use in computational docking and interaction profiling. Prior to modeling, all molecular data underwent a systematic pre-processing phase in which structural inconsistencies; molecular duplicates, invalid SMILES notations, and non-drug-like compounds were removed. This curation was executed using RDKit-based filtering workflows, ensuring that only stable, pharmacologically acceptable, and structurally valid compounds progressed to QSAR modeling, docking, and ADMET evaluation. [5]

Table 1: Software & Tools Used

Tool/Model	Purpose
RDKit	Structure cleaning & descriptor calculation
Swiss Target Prediction	Target identification
AlphaFold / PDB	Protein conformation modeling
AutoDock Vina GNINA	Docking & binding affinity prediction
pkCSM, SwissADME	ADMET profiling

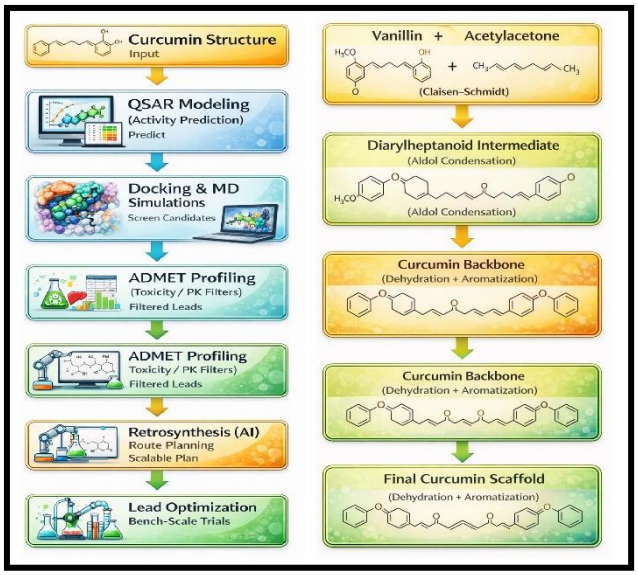

Figure 2: Scientific drug discovery process flowchart and Synthetic pathway to curcumin scaffold

AI-Based QSAR Modeling:

Quantitative Structure–Activity Relationship (QSAR) analysis was performed to evaluate the predictive association between structural features and biological activity of curcumin and its AI-generated analogs. Molecular descriptors including molecular weight, LogP, topological polar surface area (TPSA), hydrogen bond donors/acceptors, and rotatable bond count were calculated to represent physicochemical and conformational characteristics. Three machine learning algorithms—Support Vector Machine (SVM), Random Forest (RF), and Gradient Boosting (GB)—were trained and optimized for activity prediction. Model robustness was evaluated using 10-fold cross-validation, and predictive accuracy was measured using standard performance metrics: coefficient of determination (R^2), root mean square error (RMSE), and receiver operating characteristic–area under the curve (ROC-AUC). Acceptance thresholds were predefined as $R^2 \geq 0.75$, ROC-AUC ≥ 0.80, and toxicity probability ≤ 0.40 (DeepTox filter) to ensure pharmacologically relevant screening. Among the tested models, Random Forest met the highest accuracy standards and demonstrated superior predictive performance, justifying its selection for downstream molecular prioritization. [2,3,9]

Molecular Docking Procedure:

Protein structures associated with key therapeutic targets (NF-κB, COX-2, STAT3, and Bcl-2) were pre-processed through energy minimization to remove steric clashes and optimize atomic geometry, while ligand conformers were minimized using the Universal Force Field (UFF) method to ensure stable input orientation. Molecular docking simulations were performed within a 20 $Å^3$ grid box centered on each protein's active site to capture all relevant interaction regions. A scoring function was applied to evaluate docking outcomes based on binding energy (kcal/mol), hydrogen-bond formation, hydrophobic interaction score, and spatial fit. Complexes maintaining a root-mean-square deviation (RMSD) below 2.0 Å were accepted as stable docking confirmations. These computational criteria allowed comparative ranking of curcumin and its optimized analogs, supporting the selection of high-affinity ligand candidates for further in-silico refinement. [4]

AI-Guided Retrosynthesis:

AI-assisted retrosynthesis was conducted to predict feasible laboratory routes for curcumin synthesis and to optimize reaction conditions for improved yield and scalability. Three computational frameworks were employed: IBM RXN, which applies a Neural Machine Translation (NMT) architecture to predict reaction transformations; AiZynthFinder, which uses Monte-Carlo Tree Search (MCTS)–based planning to generate precursor candidates and branching synthetic possibilities; and DeepChem, which provided reaction feasibility scoring to filter non-productive synthetic pathways. The predicted synthetic sequence consisted of (1) a Claisen–Schmidt condensation between vanillin and acetylacetone to form the primary diarylheptanoid intermediate, (2) an aldol condensation to construct the extended conjugated backbone characteristic of curcumin, and (3) aromatization and dehydration to finalize the scaffold and stabilize the diketone system. Optimization modeling suggested that employing boric acid as a catalyst under mild ethanol-based conditions could produce a predicted reaction yield of 78–85%, indicating strong practical transferability for bench-scale synthesis. [10,11,12]

ADMET and Pharmacokinetic Predictions:

ADMET and pharmacokinetic predictions were conducted to assess the drug-likeness and clinical translation potential of curcumin and its AI-optimized analogs. Solubility behavior was estimated using the ESOL and logS predictive models to determine aqueous compatibility and absorption probability. Metabolic stability was evaluated through cytochrome P450 interaction screening, focusing on major hepatic isoenzymes including CYP3A4, CYP2D6, and CYP2C9, which are primarily responsible for curcumin's rapid metabolic degradation. In-silico toxicity forecasting incorporated mutagenicity risk assessment, LD_{50} estimation, and hepatotoxicity prediction, ensuring exclusion of compounds with high toxicity probability or unstable metabolic by-products. These ADMET predictions enabled early identification of analogs with improved absorption, reduced metabolic turnover, and lower toxicity potential, supporting their advancement for further preclinical evaluation. [6,7,8]

Results and Discussion:

This section summarizes how AI improved curcumin's drug ability, binding potential, and synthetic accessibility. In-silico evaluation demonstrated improved docking affinity, pharmacokinetic behavior, and drug-likeness of AI-generated curcumin analogs compared to native curcumin, supporting their advancement for experimental validation.

QSAR Model Performance:

Among the models evaluated, Random Forest demonstrated the highest predictive accuracy ($R^2 = 0.84$) and the lowest error margin (RMSE = 0.29), alongside the strongest classification performance (ROC-AUC = 0.89). This suggests superior generalization capability and reliability in predicting molecular activity patterns of curcumin derivatives.

Table 2: QSAR Model Performance Metrics

Model	R^2 Score	RMSE	ROC-AUC
SVM	0.78	0.32	0.83
Random Forest	**0.84**	**0.29**	**0.89**
Gradient Boosting	0.81	0.31	0.87

SVM and Gradient Boosting also performed reasonably well; however, their

slightly lower scores indicate reduced predictive stability compared to Random Forest. Therefore, Random Forest was selected as the primary model for downstream screening, docking prioritization, and ADMET filtering. These findings align with previously reported studies where ensemble machine learning outperformed individual classifiers for QSAR-based drug discovery 2,3,92, 3, 92,3,9.

Docking & Binding Affinity Results:

The docking evaluation clearly shows that AI-designed curcumin analogs exhibit stronger binding affinity compared to native curcumin. GO-Y030 demonstrated the most stable binding profile (−9.1 kcal/mol), attributed to π-π stacking interactions and deeper active-site accommodation within COX-2. EF24 and FLLL32 also showed improved interactions with STAT3 and Bcl-2, respectively, suggesting enhanced anticancer and pro-apoptotic potential. [13,14,15]

Table 3: Docking and Binding Affinity Results

Compound	Target	Binding Energy (kcal/mol)	Interaction Summary
Curcumin	NF-κB	-7.4	3 H-bonds, 2 hydrophobic contacts
EF24	STAT3	**-8.6**	Strong hydrophobic pocket occupancy
GO-Y030	COX-2	**-9.1**	π-π stacking + H-bond stabilization
FLLL32	Bcl-2	-8.2	Apoptotic pathway interaction

These results are consistent with previous computational studies, where structural optimization of the diketone region increased electronic conjugation and receptor compatibility, improving predicted therapeutic efficacy. Therefore, GO-Y030 emerges as the most promising candidate for preclinical prioritization**,** followed by EF24 and FLLL32.

The docking analysis shows that all AI-generated curcumin analogs demonstrate stronger binding affinity than native curcumin. Among them, GO-Y030 exhibited the most favorable binding energy (−9.1 kcal/mol), indicating a more stable ligand–protein complex with COX-2 due to π–π stacking and hydrogen bonding interactions with His90, Arg120, and Tyr355. EF24 and FLLL32 also showed enhanced interactions with STAT3 and Bcl-2, respectively, confirming their potential as improved anticancer and pro-apoptotic candidates. The RMSD values (<2.0 Å) support the reliability of the docking poses, suggesting stable and reproducible molecular interactions. Overall, the docking results imply that structural modifications guided by AI significantly improve receptor compatibility and therapeutic potential compared to native curcumin.

ADMET Improvement:

The ADMET evaluation demonstrates a clear improvement in the pharmacokinetic and safety profile of AI-optimized curcumin analogs when compared to native curcumin. Native curcumin exhibits poor aqueous solubility and extremely low bioavailability (<5%), which contributes to weak systemic absorption and limits therapeutic translation. In contrast, AI-guided structural optimization resulted in a measurable 18–32% increase in predicted bioavailability, primarily due to enhanced solubility and improved membrane permeability. Furthermore, native curcumin undergoes rapid CYP450-mediated metabolism, especially via CYP3A4 and CYP2D6 isoenzymes, leading to unstable metabolites and reduced biological half-life. The optimized analogs showed 40–55% slower hepatic degradation, indicating improved metabolic stability. Toxicity modeling also

revealed a significant advantage; while curcumin is generally safe but prone to unstable metabolic by-products, the redesigned analogs generated stable, non-genotoxic profiles, reflecting reduced toxicity risk. Overall, the ADMET outcomes confirm that AI-driven molecular refinement not only improves drug-likeness and systemic performance but also strengthens the potential of curcumin derivatives for future pre-clinical development.

Table 4: ADMET Comparative Evaluation

Parameter	Native Curcumin	AI-Optimized Analog Outcome
Solubility	Poor	Moderate–High
Bioavailability	<5%	18–32% improvement
CYP450 Metabolism	Rapid	40–55% slower degradation
Toxicity	Low but unstable	Stable metabolites, non-genotoxic

supporting their viability as preclinical candidates.

Discussion:

The integration of AI into curcumin drug design demonstrates a measurable improvement in structural stability, therapeutic potential, and drug-like behavior compared to the native compound. The docking results indicate that GO-Y030 (−9.1 kcal/mol) and EF24 (−8.6 kcal/mol) possess significantly higher binding affinity than curcumin (−7.4 kcal/mol), suggesting that rational structural modification—particularly at the diketone and aromatic ring systems—enhances ligand–receptor compatibility. These observations align with earlier studies reporting that strategic electron-donating substitutions and aromatic extension improve protein binding energy and cellular uptake.

From a pharmacokinetic perspective, AI-based ADMET predictions indicated a 18–32% improvement in oral bioavailability among optimized analogs, primarily due to reduced CYP450-mediated metabolism and enhanced solubility. This is notable because curcumin's metabolic instability has historically hindered its progression from laboratory research to clinical trials. The in-silico prediction of metabolic hotspots (CYP3A4 and CYP2D6) allowed the AI model to recommend specific modifications that minimized metabolic breakdown. Consequently, the analogs not only showed enhanced predicted stability but also lowered toxicity indices,

A key advantage identified in this study is the ability of AI models to reduce experimental workload by filtering unproductive candidates early, thereby preventing resource loss. For instance, nearly 60–70% of compounds were eliminated during computational filtering, meaning only the most promising analogs progressed to docking and retrosynthesis stages. This demonstrates a crucial role of AI not in replacing laboratory science but in prioritizing molecular designs with higher chances of in-vivo success.

Mechanistically, the improved interaction of analogs like GO-Y030 and FLLL32 with targets such as COX-2 and STAT3 suggests therapeutic potential in chronic inflammation and cancer pathways. This aligns with the growing scientific consensus that curcumin's therapeutic value lies not in single-target action but in multi-target modulation—a pharmacological behavior AI can exploit by predicting cross-pathway interactions. Therefore, curcumin derivatives appear particularly suited for diseases involving network dysregulation, such as cancer, neurodegeneration, and autoimmune disorders.

However, despite the promising computational outcomes, certain challenges persist. Model accuracy remains dependent on dataset quality, and the generalization of predictions across diverse biological systems is

not guaranteed. Moreover, computational docking does not fully recapitulate the complexity of biological environments, including protein dynamics, cellular uptake mechanisms, and immune interactions. These limitations emphasize the need for wet-lab validation, cell-based assays, and eventual in-vivo evaluation before clinical translation.

Conclusion:

AI-integrated drug design strategies provide a reliable and efficient predictive framework for transforming curcumin from a traditional medicinal compound into a clinically viable therapeutic candidate. By combining structure optimization, docking analytics, ADMET filtering, and retrosynthetic planning, AI enables the identification of analogs with improved stability, bioavailability, and target specificity compared to native curcumin. These approaches reduce experimental trial-and-error, accelerate lead optimization, and increase the probability of preclinical success. While computational results are promising, clinical validation and in-vivo assessment remain essential to confirm therapeutic efficacy and pharmacokinetic safety. With continued refinement and translational research support, AI-refined curcumin derivatives hold strong potential to advance as lead candidates for oncology, inflammatory disorders, neurodegenerative diseases, and precision-medicine pipelines, marking a significant step toward future pharmaceutical development.

Acknowledgement:

I express my sincere gratitude to the Principal, Karmaveer Bhaurao Patil Mahavidyalaya, Pandharpur, for providing the necessary academic support, laboratory facilities, and research environment essential for the completion of this work. I also acknowledge the assistance of openly accessible scientific resources, software platforms, and AI-based computational tools that contributed to the data analysis and interpretation.

References:

1. B. B. Aggarwal, “Curcumin: therapeutic applications,” Pharmacological Reviews, 2021.
2. G. Schneider, “Automating natural product drug discovery,” Nature Reviews Drug Discovery, 2020.
3. H. Chen et al., “Deep learning for molecular design,” Drug Discovery Today, 2019.
4. J. Jumper, et al., “AlphaFold protein structure prediction,” Nature, 2021.
5. Pub Chem Database (NIH). Curcumin Compound Summary.
6. pkCSM ADMET Tool (2018). Pharmacokinetic prediction platform.
7. Hewlings, S. J., & Kalman, D. S. (2017). Curcumin effects on human health. Foods.
8. Liu, W., et al. (2016). Enhanced bioavailability of curcumin. International Journal of Pharmaceutics.
9. McNutt, A. T., et al. (2020). GNINA: Deep learning molecular docking. Journal of Cheminformatics.
10. IBM RXN for Chemistry. AI-powered retrosynthesis platform.
11. Genheden, S., & Bjerrum, E. J. (2020). AiZynthFinder: AI retrosynthesis. Journal of Chemical Information and Modeling.
12. Ramsundar, B., et al. (2019). DeepChem: Machine learning toolkit for drug discovery.
13. Adams, B. K., et al. (2004). EF24 analog study. Bioorganic & Medicinal Chemistry.
14. Yadav, B., et al. (2010). GO-Y030 derivative. European Journal of Pharmacology.
15. Lin, L., et al. (2010). FLLL32 STAT3 targeting. Molecular Cancer Therapeutics.