Journal of Drug Discovery and Health Sciences

journal home page : https://jddhs.com/index.php/jddhs/index

Research Article

Redefining the CADD Paradigm: Harnessing Artificial Intelligence from Target Discovery to Lead Optimization

Omkar Rai[1], Siddela Johnson[2*], Soma Sekhar Pulamarasetti[1], Kovvada Vandana[2], Bongu Ramya Priya[2]

[1]*Department of Pharmacy, B. R. Nahata College of Pharmacy, Mandsaur University, Mandsaur, Madhya Pradesh, India*
[2]*AU College of Pharmaceutical Sciences (AUCOPS), Andhra University, Visakhapatnam, Andhra Pradesh, India*

ARTICLE INFO

Article history:
Received: 30 April, 2026
Revised: 09 May, 2026
Accepted: 25 June, 2026
Published: 30 June, 2026

Keywords:
Artificial intelligence; computer-aided drug design; virtual screening; generative chemistry; lead optimization

DOI:
10.21590/jddhs.03.02.08

ABSTRACT

Artificial intelligence (AI) is reshaping computer-aided drug design (CADD) by enabling data-driven analysis and generation across the early drug-discovery continuum. Conventional CADD approaches—including molecular docking, pharmacophore modelling, quantitative structure-activity relationship analysis and molecular dynamics—remain valuable, but their predictive performance can be constrained by simplified scoring functions, incomplete structural information and the enormous size of chemical space. Machine learning (ML), deep learning, graph neural networks, protein language models, geometric learning and generative models increasingly complement these methods. AI can support disease and target prioritization, protein-structure prediction, binding-site analysis, drug-target interaction prediction, virtual screening, de novo molecular generation, ADMET prediction, retrosynthetic planning and multi-parameter lead optimization. Recent advances such as AlphaFold3, deep-learning docking, ultra-large-library screening and generative molecular design have expanded the scope of structure-based and ligand-based discovery. Importantly, AI-generated candidates must be assessed for potency, selectivity, physicochemical properties, pharmacokinetics, toxicity, synthetic accessibility and uncertainty before experimental testing. Data leakage, bias, activity cliffs, distribution shift, limited prospective validation and interpretability remain significant barriers. This review presents an integrated AI-CADD framework from target identification to lead optimization and emphasizes the design-make-test-analyze cycle as the practical bridge between computational prediction and experimental medicinal chemistry.

Introduction

Drug discovery is a high-risk, data-intensive process in which decisions made during target selection and hit identification strongly influence downstream attrition. Computational methods have long been used to reduce experimental search space, but the emergence of large chemical and biological databases, high-performance computing and modern ML has changed the scale at which CADD can operate. AI is now applied to disease biology, target prioritization, structural prediction, virtual screening, molecular generation, property prediction and synthesis planning (Paul *et al*., 2021; Jiménez-Luna *et al*., 2021; Vamathevan *et al*., 2019). The most important conceptual change is that AI does not need to replace conventional CADD. Instead, it can augment docking, molecular dynamics, QSAR, pharmacophore modelling and medicinal-chemistry reasoning by learning patterns from experimental data. Deep-learning models can operate on molecular graphs, sequences and three-dimensional structures, while generative systems can

***Corresponding Author:** Siddela Johnson

Address: AU College of Pharmaceutical Sciences (AUCOPS), Andhra University, Visakhapatnam, Andhra Pradesh, India

Email ✉**:** siddelajohnson@gmail.com

Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Table 1: Comparison of conventional and AI-enabled CADD across the drug-discovery workflow.

Drug-discovery stage	*Conventional CADD*	*AI-enabled CADD*	*Expected benefit*
Target identification	Pathway/literature analysis	Omics integration, knowledge graphs, target ranking	Prioritized biological targets
Structure prediction	Homology/experimental structures	Protein foundation models	Expanded structural coverage
Hit identification	Docking/pharmacophores	Deep screening and DTI prediction	Higher-throughput hit prioritization
Molecule generation	Analogue/fragment design	Generative AI and reinforcement learning	Broader chemical exploration
Lead optimization	Manual SAR cycles	Multi-objective ML and active learning	Faster DMTA iteration
Developability	Later-stage ADMET testing	Early ADMET/synthesis prediction	Reduced avoidable attrition

propose new chemical matter rather than merely rank existing compounds (Chen *et al*., 2018; Askr *et al*., 2023). The objective of this review is to critically connect AI methods with the practical stages of CADD, beginning with disease and target identification and extending through structure prediction, virtual screening, de novo design, ADMET assessment, retrosynthesis and lead optimization. Particular attention is given to recent advances published from 2021 to 2026, while selected foundational studies are retained because they established methods still used in current AI-CADD workflows.

Materials and Methods

As a review article, this manuscript does not report original laboratory experiments. The literature was synthesised through a narrative review of peer-reviewed studies and recognised preprints on artificial intelligence-enabled computer-aided drug design, with emphasis on work published between 2021 and 2026 and inclusion of selected foundational studies that established methods still in current use. Sources were selected to cover each stage of the AI-CADD pipeline addressed in the Results and Discussion: data and computational foundations, target identification, protein structure prediction, binding-site analysis, drug-target interaction and affinity prediction, virtual screening and docking, QSAR/property prediction, de novo molecular generation, hit-to-lead progression, ADMET and developability prediction, retrosynthesis, lead optimization, multi-objective optimization, model validation, and translational case studies. Findings are synthesised and critically discussed below.

Results and Discussion

Data And Computational Foundations of AI-CADD

AI-CADD depends on reliable biological and chemical data. Major resources include ChEMBL for bioactivity, DrugBank for drug and target information, BindingDB for experimentally determined binding affinities, the Protein Data Bank for macromolecular structures, PubChem for chemical information and the Therapeutic Target Database for target druggability information (Zdrazil *et al*., 2023; Knox *et al*., 2024; Liu *et al*., 2007; Burley *et al*., 2022; Kim, 2021; Zhou *et al*., 2024). The quality, assay context, endpoint definition and chemical diversity of these datasets directly affect model reliability. Molecules can be encoded as fingerprints, physicochemical descriptors, SMILES strings, molecular graphs or three-dimensional conformations. Graph-based models represent atoms as nodes and bonds as edges, allowing the model to learn structural features rather than relying entirely on handcrafted descriptors (Kearnes *et al*., 2016; Gilmer *et al*., 2017; Yang *et al*., 2019). Protein inputs may include sequences, structures, residue graphs and learned embeddings. Multimodal models can integrate ligand, protein and biological context. A critical issue is data leakage. Random train/test splits may place highly similar analogues in both sets and therefore overestimate prospective performance. Scaffold splits, temporal splits, target-held-out evaluations and external validation provide more realistic estimates. Models should also report calibration and applicability-domain information rather than relying on a single accuracy metric (Wu *et al*., 2018; Hasselgren and Oprea, 2024).

AI For Target Identification and Validation

Target identification is increasingly supported by AI-based integration of genomic, transcriptomic, proteomic, phenotypic and clinical data. Network models can identify disease-associated genes and proteins, while knowledge-graph approaches can connect targets to diseases, pathways, phenotypes and known drugs. The therapeutic value of a target, however, depends not only on disease association but also on causal evidence, tractability, tissue distribution, safety and the availability of a pharmacologically addressable binding site (You *et al*., 2022; Tabana *et al*., 2023; Pun *et al*., 2026). AI can assist target assessment by integrating multiple evidence streams and prioritizing targets with convergent biological support. The updated Therapeutic

Target Database provides druggability information that can be combined with AI-driven ranking. Human genetic evidence is particularly valuable because it can increase confidence that modulation of a target is mechanistically linked to disease (Finan *et al*., 2017; Zhou *et al*., 2024). AI-based cellular imaging and phenotypic analysis can also identify disease-relevant morphological signatures. Such systems are most powerful when image-derived phenotypes are linked to molecular perturbation data and followed by experimental target deconvolution. The central principle is that AI can prioritize hypotheses, but target validation remains an experimental and pharmacological task.

AI-Enabled Protein Structure Prediction

Structure-based CADD has historically depended on experimentally determined protein structures. AlphaFold2 demonstrated that deep learning can predict many protein structures with remarkable accuracy, while RoseTTAFold provided an independent multi-track approach (Jumper *et al*., 2021; Baek *et al*., 2021). AlphaFold3 extended the paradigm by predicting complexes involving proteins, nucleic acids, small molecules and ions, providing a broader structural foundation for AI-CADD (Abramson *et al*., 2024). These predictions can increase the number of proteins accessible to structure-based screening, but a predicted structure is not necessarily equivalent to an experimentally observed ligand-bound state. Flexible loops, cryptic pockets, induced fit and alternative conformations may remain difficult. Consequently, predicted structures should be combined with structural homologues, pocket analysis, molecular dynamics and experimental information whenever available (Burley *et al*., 2021; Özçelik *et al*., 2023).

Binding-Site and Druggability Analysis

AI can identify potential binding pockets from protein coordinates, residue environments and geometric features. Deep-learning and geometric models can learn spatial relationships that are difficult to capture with conventional descriptors. Binding-site prediction is especially useful for allosteric targets and proteins for which ligand-bound structures are unavailable. The practical objective is not merely to identify any cavity, but to prioritize sites with suitable volume, enclosure, physicochemical complementarity, conservation or disease relevance and an opportunity for selectivity. Structural AI should therefore be integrated with experimental ligand information and conventional pocket characterization rather than treated as a stand-alone decision system (Özçelik *et al*., 2023; Pun *et al*., 2026).

Drug-Target Interaction and Affinity Prediction

Drug-target interaction (DTI) prediction estimates whether a molecule interacts with a protein and may also estimate binding affinity. DeepPurpose and related architectures demonstrate how deep neural networks can learn joint representations of chemical structures and protein sequences (Huang *et al*., 2021). Graph convolutional approaches have also been applied to drug response and molecular interactions (Nguyen *et al*., 2022). Modern models increasingly incorporate three-dimensional information. TANKBind uses geometric constraints for protein-ligand structure and affinity prediction, while diffusion and geometric-learning methods have expanded the range of protein-ligand pose prediction (Lu *et al*., 2022; Corso *et al*., 2022). These methods can rapidly prioritize molecules but should be evaluated carefully on scaffold- and target-held-out datasets.

Table 2: Major artificial-intelligence model families used in computer-aided drug design and their typical applications.

AI model family	*Typical CADD task*	*Representative input*	*Main strength*
Random Forest / Gradient Boosting	QSAR, ADMET, activity prediction	Descriptors, fingerprints	Strong baseline and relatively interpretable
Convolutional Neural Network	3D binding-site learning	Voxelized structures / molecular grids	Learns spatial patterns
Graph Neural Network	Molecular property and activity prediction	Molecular graphs	Captures atom-bond relationships
Transformer / LLM	Molecular representation, reaction and literature modelling	SMILES, sequences, text	Long-range contextual representation
VAE / GAN	De novo molecular generation	Latent molecular representations	Explores chemical space
Diffusion model	3D and molecular generation	Graphs, coordinates, conditions	Flexible conditional generation
Reinforcement learning	Goal-directed optimization	Molecular states and rewards	Optimizes multiple objectives
Hybrid / ensemble models	Virtual screening and rescoring	Structure + ligand + assay data	Combines complementary signals

AI-Driven Virtual Screening and Molecular Docking

Virtual screening can be divided into ligand-based and structure-based approaches. AI models can rapidly pre-filter ultra-large libraries before more expensive docking and physics-based calculations. Deep Docking demonstrated an AI-enabled workflow for screening very large chemical libraries and thereby reducing the computational burden associated with conventional docking (Gentile *et al*., 2022). GNINA introduced deep learning into molecular docking and scoring, while AutoDock Vina 1.2.0 expanded conventional docking capabilities and remains useful as a transparent baseline (McNutt *et al*., 2021; Eberhardt *et al*., 2021). More recently, RosettaVS demonstrated an AI-assisted structure-based screening platform designed for high-throughput virtual screening (Zhou *et al*., 2024). A robust workflow therefore combines rapid AI filtering with interpretable docking, rescoring and experimental confirmation.

AI for Qsar and Molecular Property Prediction

QSAR modelling remains one of the most practical AI applications in CADD. Deep neural networks, message-passing networks and graph transformers can learn relationships between chemical structure and potency or physicochemical properties. MoleculeNet established a benchmark framework for molecular machine learning and emphasized the importance of standardized evaluation (Wu *et al*., 2018). Modern molecular property models can support potency, solubility, permeability, metabolic stability, plasma protein binding, CYP inhibition and toxicity prediction. Chemprop provides a widely used message-passing framework for molecular property prediction, while ADMETlab 2.0 integrates multiple ADMET endpoints (Heid *et al*., 2024; Xiong *et al*., 2021). Prediction should be coupled with uncertainty and applicability-domain assessment because model performance may deteriorate sharply for novel chemical scaffolds.

AI for DE Novo Molecular Generation

Generative AI changes the discovery problem from selecting molecules from an existing library to constructing new chemical structures. Deep generative approaches include variational autoencoders, recurrent neural networks, reinforcement learning, graph generative models and, more recently, diffusion and transformer-based systems (Cheng *et al*., 2021; Wang *et al*., 2022; Zeng *et al*., 2022). REINVENT demonstrated practical reinforcement-learning-based de novo design, while subsequent systems increasingly condition generation on target activity and multiple molecular properties (Blaschke *et al*., 2020). Three-dimensional interaction-guided generation is particularly promising because it directly incorporates protein-ligand geometry (Zhung *et al*., 2024). Generative AI should not be judged solely by novelty or validity. A useful molecule must balance predicted potency, selectivity, physicochemical properties, ADMET behaviour, synthetic accessibility and chemical stability. Multi-objective optimization is therefore central to practical generative chemistry.

AI for Hit Identification and Hit-To-Lead

AI can identify chemically diverse hits by combining similarity searching, activity prediction, docking and generative design. The discovery of halicin using deep learning demonstrated that AI can identify antibacterial chemical matter with structures distinct from conventional antibiotic classes (Stokes *et al*., 2020). Deep-learning-driven kinase inhibitor design similarly illustrated the possibility of rapidly exploring chemical space around a selected target (Zhavoronkov *et al*., 2019). During hit-to-lead development, AI can prioritize analogues with improved ligand efficiency, selectivity and physicochemical properties. The greatest benefit is obtained when predictions are embedded in an experimental cycle in which newly measured compounds are incorporated into the next model iteration.

Ai for Admet and Developability Prediction

A lead compound must possess more than target potency. Absorption, distribution, metabolism, excretion and toxicity can eliminate otherwise attractive candidates. AI can model permeability, solubility, metabolic stability, CYP inhibition, transporter interactions, cardiotoxicity, hepatotoxicity and mutagenicity. ADMETlab 2.0 integrates multiple ADMET predictions, while recent practical work emphasizes careful dataset definition, calibration and prospective use of ADME models during lead optimization

Table 3: AI-assisted virtual-screening strategies and practical considerations.

Approach	*Input*	*Typical output*	*Key limitation*
Classical docking	Protein + ligand	Pose and docking score	Scoring-function limitations
Deep-learning docking	3D protein/ligand	Pose and learned score	Generalization
Ligand-based ML	Molecular descriptors/graphs	Activity ranking	Training-chemistry dependence
DTI prediction	Protein + molecule	Interaction/affinity prediction	Target and assay bias
Ultra-large screening	Very large libraries	Prioritized subset	Possible false negatives
Hybrid screening	AI + docking + physics	Consensus ranking	Higher complexity

Table 4: Multi-parameter objectives commonly integrated into AI-guided lead optimization.

Optimization objective	*AI strategy*	*Typical endpoint*	*Medicinal-chemistry interpretation*
Potency	QSAR/GNN/3D affinity	IC50, EC50, Kd	Increase target activity
Selectivity	Multi-task learning	Off-target panel	Reduce unwanted pharmacology
Solubility	Regression/classification	Aqueous solubility	Improve formulation/exposure
Permeability	ML/graph models	PAMPA, Caco-2	Balance polarity and exposure
Metabolic stability	ADME models	Microsomal half-life	Reduce rapid clearance
Toxicity	Multi-endpoint ML	hERG, hepatotoxicity, genotoxicity	Remove liabilities early
Synthetic accessibility	Retrosynthesis/SA scoring	Route feasibility	Prefer practical compounds
Overall developability	Multi-objective optimization	Composite utility	Balance competing properties

(Xiong *et al*., 2021; Rich, 2024). The appropriate use of AI is to reduce avoidable synthesis of compounds with predictable liabilities, not to replace experimental pharmacokinetic and toxicological studies.

AI-Assisted Retrosynthesis and Synthetic Accessibility

Synthetic accessibility is a major constraint on generative drug design. Retrosynthesis algorithms propose disconnections that can be evaluated against reaction databases and chemical rules. AiZynthFinder combines neural-network reaction prediction with Monte Carlo tree search, and its later versions incorporate lessons from industrial deployment (Genheden *et al*., 2020; Saigiridharan *et al*., 2024). The Molecular Transformer demonstrated that reaction prediction can be formulated as a sequence-to-sequence translation problem, with uncertainty estimation providing additional decision support (Schwaller *et al*., 2019). Chemformer extended transformer-based retrosynthesis to multistep planning, illustrating the growing role of large reaction datasets (Mercado *et al*., 2024). In AI-CADD, synthetic feasibility should be incorporated before final compound prioritization.

AI for Lead Optimization

Lead optimization is inherently multi-objective: potency, selectivity, solubility, permeability, metabolic stability, safety and synthetic accessibility must be improved simultaneously. AI can learn structure-activity relationships and propose modifications likely to improve one or more properties. Recent approaches distinguish goal-directed generation from structure-directed optimization, including fragment replacement, linker design, scaffold hopping and side-chain decoration (Xia *et al*., 2024; Zhang *et al*., 2024). Deep Lead Optimization frameworks provide increasingly practical strategies for medicinal chemists (Zhang *et al*., 2024). AI-guided ADME models can further reduce the number of compounds synthesized solely to discover predictable liabilities (Rich, 2024).

Multi-Objective Optimization and Active Learning

Multi-objective optimization is essential because improving one property can worsen another. Bayesian optimization, reinforcement learning and active learning can identify compounds that either maximize predicted utility or provide informative experimental data. Active learning is particularly useful when experimental data are expensive and training datasets are small (Current status of active learning for drug discovery, 2021). The optimal next compound is not necessarily the molecule with the highest predicted potency. A compound with moderate predicted activity but high uncertainty may provide greater information and improve the next generation of the model. This principle links AI directly to the design-make-test-analyze cycle.

Explainable Ai, Uncertainty And Model Validation

Interpretability is essential for medicinal chemistry because computational predictions should generate chemically testable hypotheses. Explainable AI can highlight influential molecular substructures, protein residues or learned features, but explanations should not automatically be interpreted as mechanisms. Uncertainty estimation is equally important. Ensemble methods, calibration, conformal prediction and applicability-domain analysis can identify predictions that should be treated cautiously. Benchmarking should use scaffold and temporal splits where appropriate and should include external prospective datasets. The increasing availability of high-capacity models makes rigorous validation more—not less—important (Hasselgren and Oprea, 2024; Askr *et al*., 2023).

Prospective Evidence and Translational Impact

The field is moving from retrospective benchmark performance toward prospective drug-discovery evidence. A major recent example is the generative-AI-discovered TNIK inhibitor program reported in a randomized phase 2a trial for idiopathic pulmonary fibrosis, providing an important translational milestone for AI-enabled

Table 5: Recommended evaluation measures for AI-enabled CADD models and prospective validation

CADD prediction task	*Useful metric*	*What it measures*	*Recommended validation consideration*
Classification	ROC-AUC, PR-AUC, F1	Class discrimination	Use scaffold/external splits and report class balance
Regression / QSAR	RMSE, MAE, R^2	Prediction error and variance explained	Report uncertainty and external performance
Virtual screening	EF, BEDROC, enrichment	Early retrieval of active compounds	Use realistic active/inactive distributions
Ranking	Spearman / Kendall correlation	Agreement in ordering	Test on unseen chemical series
Molecular generation	Validity, uniqueness, novelty	Chemical quality and diversity	Also assess synthesizability
Docking / pose prediction	RMSD, pose success rate	Pose reproduction	Consider receptor flexibility
ADMET prediction	AUROC, RMSE/MAE	Endpoint prediction quality	Prefer assay-specific external validation
Prospective discovery	Hit rate / experimental success	Real-world utility	Predefine criteria and preserve blind tests

Table 6: Major barriers to reliable AI-CADD and recommended mitigation strategies

Challenge	*Why it matters*	*Recommended mitigation*
Data leakage	Inflated benchmark performance	Scaffold/temporal/external splits
Dataset bias	Poor transfer to underrepresented chemistry	Diverse datasets and reweighting
Distribution shift	Novel targets/scaffolds can break models	Applicability domain and uncertainty
Activity cliffs	Small structural changes can cause large effects	Local SAR and analogue-aware validation
Protein flexibility	Static structures miss relevant conformations	Ensembles, MD and experimental structures
Interpretability	Black-box predictions are hard to audit	Attribution and chemical review
Synthetic feasibility	Generated molecules may be impractical	Retrosynthetic planning and expert review
Prospective validation	Retrospective metrics may not translate	Prospective synthesis and blinded evaluation
Reproducibility/IP	Data and provenance may be unclear	Versioned data, code and provenance

discovery (Xu *et al*., 2025). Such examples should be interpreted carefully: the success of a program reflects integrated medicinal chemistry, pharmacology, clinical development and experimental validation, not AI alone. The emerging lesson is that AI is most credible when it produces experimentally validated compounds and measurable improvements in discovery efficiency rather than only high retrospective benchmark scores (Wilczok and Zhavoronkov, 2025).

Challenges, Ethical Issues and Future Directions

The major challenges include biased training data, inconsistent assay measurements, data leakage, activity cliffs, distribution shift, synthetic infeasibility, protein flexibility, limited interpretability and inadequate prospective validation. Intellectual-property issues, provenance of training data and reproducibility also require attention. Future AI-CADD platforms are likely to combine foundation models for proteins and molecules, three-dimensional generative models, multimodal biological knowledge, active learning and automated experimentation. Large language models may increasingly serve as orchestration layers that connect databases, prediction tools, retrosynthesis systems and laboratory platforms. However, human oversight will remain essential for selecting biologically meaningful objectives and evaluating experimental evidence (Hasselgren and Oprea, 2024; Kanakala *et al*., 2024; Pun *et al*., 2026).

Conclusion

AI is becoming an important enabling layer for CADD, connecting target biology, structural modelling, virtual screening, molecular generation, ADMET prediction, synthesis planning and lead optimization. Its greatest value lies in integrating diverse data and accelerating iterative hypothesis testing. Reliable AI-CADD therefore requires high-quality data, rigorous external validation,

uncertainty estimation and experimental confirmation. The most promising future is a human-guided, AI-accelerated design-make-test-analyze cycle that combines computational scale with medicinal-chemistry judgement and laboratory evidence.

Acknowledgement

The author acknowledges the scientific community and publicly available databases and computational resources that have supported advances in AI, structural biology, cheminformatics and drug discovery.

Conflict of interest

The author declares no conflict of interest.

Funding

No specific funding was received for preparation of this review.

Author Contribution

The author conceived the review, conducted the literature analysis, prepared and revised the manuscript, and approved the final version.

References

1. Paul, D., Sanap, G., Shenoy, S., Kalyane, D., Kalia, K., and Tekade, R.K. (2021). Artificial intelligence in drug discovery and development. Drug Discovery Today, 26(1):80-93. https://doi.org/10.1016/j.drudis.2020.10.010
2. Jiménez-Luna, J., Grisoni, F., Weskamp, N., and Schneider, G. (2021). Artificial intelligence in drug discovery: recent advances and future perspectives. Expert Opinion on Drug Discovery, 16(9):949-959. https://doi.org/10.1080/17460441.2021.1909567
3. Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., Li, B., Madabhushi, A., Shah, P., Spitzer, M., and Zhao, S. (2019). Applications of machine learning in drug discovery and development. Nature Reviews Drug Discovery, 18:463-477. https://doi.org/10.1038/s41573-019-0024-5
4. Chen, H., Engkvist, O., Wang, Y., Olivecrona, M., and Blaschke, T. (2018). The rise of deep learning in drug discovery. Drug Discovery Today, 23(6):1241-1250. https://doi.org/10.1016/j.drudis.2018.01.039
5. Askr, H., Elgeldawi, E., Aboul Ella, H., Elshaier, Y.A.M.M., Gomaa, M.M., and Hassanien, A.E. (2023). Deep learning in drug discovery: an integrative review and future challenges. Artificial Intelligence Review, 56(7):5975-6037. https://doi.org/10.1007/s10462-022-10306-1
6. Nag, S., Baidya, A.T.K., Mandal, A., Mathew, A.T., Das, B., Devi, B., and Kumar, R. (2022). Deep learning tools for advancing drug discovery and development. 3 Biotech, 12:110. https://doi.org/10.1007/s13205-022-03165-8
7. Pandey, M., Fernandez, M., Gentile, F., Isayev, O., Tropsha, A., Stern, A.C., and Cherkasov, A. (2022). The transformational role of GPU computing and deep learning in drug discovery. Nature Machine Intelligence, 4:211-221. https://doi.org/10.1038/s42256-022-00463-x
8. Hasselgren, C., and Oprea, T.I. (2024). Artificial intelligence for drug discovery: Are we there yet? Annual Review of Pharmacology and Toxicology, 64:527-550. https://doi.org/10.1146/annurev-pharmtox-040323-040828
9. Cerchia, C., and Lavecchia, A. (2023). New avenues in artificial-intelligence-assisted drug discovery. Drug Discovery Today, 28(4):103516. https://doi.org/10.1016/j.drudis.2023.103516
10. Chen, W., Liu, X., Zhang, S., and Chen, S. (2023). Artificial intelligence for drug discovery: resources, methods, and applications. Molecular Therapy - Nucleic Acids, 31:691-702. https://doi.org/10.1016/j.omtn.2023.02.019
11. Wang, K., Huang, Y., Wang, Y., and others. (2024). Recent advances from computer-aided drug design to artificial intelligence drug design. RSC Medicinal Chemistry, 15(12):3978-4000. https://doi.org/10.1039/D4MD00522H
12. Kanakala, G.C., Devata, S., Chatterjee, P., and Priyakumar, U.D. (2024). Generative artificial intelligence for small molecule drug design. Current Opinion in Biotechnology, 89:103175. https://doi.org/10.1016/j.copbio.2024.103175
13. Khan, M.K., Raza, M.R., Shahbaz, M., Hussain, I., Khan, M.F., Xie, Z., Shah, S.S.A., and Tareen, A.K. (2024). The recent advances in the approach of artificial intelligence towards drug discovery. Frontiers in Chemistry, 12:1408740. https://doi.org/10.3389/fchem.2024.1408740
14. Agu, P.C. (2024). Piquing artificial intelligence towards drug discovery: tools, techniques, and applications. Drug Development Research. https://doi.org/10.1002/ddr.22159
15. Sharma, V., Singh, A., Chauhan, S., Sharma, P.K., Chaudhary, S., Sharma, A., Porwal, O., and Fuloria, N.K. (2024). Role of artificial intelligence in drug discovery and target identification in cancer. Current Drug Delivery, 21(6):870-886. https://doi.org/10.2174/1567201821666230905090621
16. You, Y., Lai, X., Pan, Y., Zheng, H., Vera, J., Liu, S., Deng, S., and Zhang, L. (2022). Artificial intelligence in cancer target identification and drug discovery. Signal Transduction and Targeted Therapy, 7:156. https://doi.org/10.1038/s41392-022-00994-0
17. Tabana, Y., Babu, D., Fahlman, R., Siraki, A.G., and Barakat, K. (2023). Target identification of small molecules: an overview of the current applications in drug discovery. BMC Biotechnology, 23:44. https://doi.org/10.1186/s12896-023-00815-4
18. Pun, F.W., Podolskiy, D., Izumchenko, E., Mortlock, A., Oprea, T.I., Scheibye-Knudsen, M., Fortney, K., Morgen, E., Ren, F., and Zhavoronkov, A. (2026). Target identification and assessment in the era of AI. Nature Reviews Drug Discovery, 25:534-552. https://doi.org/10.1038/s41573-026-01412-8
19. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., and Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. Nature, 596:583-589. https://doi.org/10.1038/s41586-021-03819-2
20. Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G.R., Wang, J., Cong, Q., Kinch, L.N., Schaeffer, R.D., Millán, C., Park, H., Adams, C., Glassman, C.R., DeGiovanni, A., Pereira, J.H., Rodrigues, A.V., van Dijk, A.A., Ebrecht, A.C., Opperman, D.J., Kihara, D., Mentzer, A.J., Kawakami, E., Singh, R., and Baker, D. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. Science, 373(6557):871-876. https://doi.org/10.1126/science.abj8754
21. Abramson, J., Adler, J., Dunger, J., Evans, R., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., Bodenstein, S.W., Evans, D.A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., Beattie, C., Bertolli, O., Bridgland, A., Cherepanov, A., Congreve, M., Cowen-Rivers, A.I., Cowie, A., Figurnov, M., Fuchs, F.B., Gladman, H., Jain, R., Khan, Y.A., Low, C.M.R., Perlin, K., Potapenko, A., Savy, P., Singh, S., Stecula, A., Thillaisundaram, A., Tong, C., Yakneen, S., Zhong, E.D., Zielinski, M., Žídek, A., Bapst, V., Kohli, P., Jaderberg, M., Hassabis, D., and Jumper, J.M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature, 630:493-500. https://doi.org/10.1038/s41586-024-07487-w

22. Burley, S.K. (2021). Impact of structural biologists and the Protein Data Bank on small-molecule drug discovery and development. Journal of Biological Chemistry, 296:100559. https://doi.org/10.1016/j.jbc.2021.100559
23. Burley, S.K., Bhikadiya, C., Bi, C., Bittrich, S., Chen, L., Crichlow, G.V., *et al*. (2022). RCSB Protein Data Bank: celebrating 50 years of the PDB with new tools for understanding and visualizing biological macromolecules in 3D. Protein Science, 31:187-208. https://doi.org/10.1002/pro.4213
24. Burley, S.K. (2021). Open access data: a cornerstone for artificial intelligence approaches to protein structure prediction. Structure, 29(6):515-520. https://doi.org/10.1016/j.str.2021.04.010
25. Özçelik, R., van Tilborg, D., Jiménez-Luna, J., and Grisoni, F. (2023). Structure-based drug discovery with deep learning. ChemBioChem, 24(13):e202200776. https://doi.org/10.1002/cbic.202200776
26. Zhou, G., Rusnac, D.-V., Park, H., Canzani, D., Nguyen, H.M., Stewart, L., Bush, M.F., Nguyen, P.T., Wulff, H., Yarov-Yarovoy, V., Zheng, N., and DiMaio, F. (2024). An artificial intelligence accelerated virtual screening platform for drug discovery. Nature Communications, 15:7761. https://doi.org/10.1038/s41467-024-52061-7
27. Zhung, W., Kim, H., and Kim, W.Y. (2024). 3D molecular generative framework for interaction-guided drug design. Nature Communications, 15:2688. https://doi.org/10.1038/s41467-024-47011-2
28. McNutt, A.T., Francoeur, P., Aggarwal, R., Masuda, T., Meli, R., Ragoza, M., Sunseri, J., and Koes, D.R. (2021). GNINA 1.0: molecular docking with deep learning. Journal of Cheminformatics, 13:43. https://doi.org/10.1186/s13321-021-00522-2
29. Eberhardt, J., Santos-Martins, D., Tillack, A.F., and Forli, S. (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. https://doi.org/10.1021/acs.jcim.1c00203
30. Gentile, F., Yaacoub, J.C., Gleave, J., Fernandez, M., Ton, A.-T., Ban, F., Stern, A., and Cherkasov, A. (2022). Artificial intelligence-enabled virtual screening of ultra-large chemical libraries with deep docking. Nature Protocols, 17(3):672-697. https://doi.org/10.1038/s41596-021-00659-2
31. Corso, G., Stärk, H., Jing, B., Barzilay, R., and Jaakkola, T. (2022). DiffDock: diffusion steps, twists, and turns for molecular docking. arXiv. https://doi.org/10.48550/arXiv.2210.01776
32. Lu, W., Wu, Q., Zhang, J., Rao, J., Li, C., and Zheng, S. (2022). TANKBind: trigonometry-aware neural networks for drug-protein binding structure prediction. bioRxiv. https://doi.org/10.1101/2022.06.06.495043
33. Huang, K., Fu, T., Glass, L.M., Zitnik, M., Xiao, C., and Sun, J. (2021). DeepPurpose: a deep learning library for drug-target interaction prediction. Bioinformatics, 36(22-23):5545-5547. https://doi.org/10.1093/bioinformatics/btaa1005
34. Nguyen, T., Nguyen, G.T.T., Nguyen, T., and Le, D.-H. (2022). Graph convolutional networks for drug response prediction. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 19(1):146-154. https://doi.org/10.1109/TCBB.2021.3060430
35. Heid, E., Greenman, K.P., Chung, Y., Li, S.C., Graff, D.E., Vermeire, F.H., Wu, H., Green, W.H., and McGill, C.J. (2024). Chemprop: a machine learning package for chemical property prediction. Journal of Chemical Information and Modeling, 64(1):9-17. https://doi.org/10.1021/acs.jcim.3c01250
36. Xiong, G., Wu, Z., Yi, J., Fu, L., Yang, Z., Hsieh, C., Yin, M., Zeng, X., Wu, C., Lu, A., Chen, X., Hou, T., and Cao, D. (2021). ADMETlab 2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties. Nucleic Acids Research, 49(W1):W5-W14. https://doi.org/10.1093/nar/gkab255
37. Daina, A., Michielin, O., and Zoete, V. (2017). SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Scientific Reports, 7:42717. https://doi.org/10.1038/srep42717
38. Cheng, Y., Gong, Y., Liu, Y., Song, B., and Zou, Q. (2021). Molecular design in drug discovery: a comprehensive review of deep generative models. Briefings in Bioinformatics, 22(6):bbab344. https://doi.org/10.1093/bib/bbab344
39. Wang, M., Wang, Z., Sun, H., Wang, J., Shen, C., Weng, G., Chai, X., Li, H., Cao, D., and Hou, T. (2022). Deep learning approaches for de novo drug design: an overview. Current Opinion in Structural Biology, 72:135-144. https://doi.org/10.1016/j.sbi.2021.10.001
40. Zeng, X., Wang, F., Luo, Y., Kang, S.-G., Tang, J., Lightstone, F.C., Fang, E.F., Cornell, W., Nussinov, R., and Cheng, F. (2022). Deep generative molecular design reshapes drug discovery. Cell Reports Medicine, 3(12):100794. https://doi.org/10.1016/j.xcrm.2022.100794
41. Blaschke, T., Arús-Pous, J., Chen, H., Margreitter, C., Tyrchan, C., Engkvist, O., Papadopoulos, K., and Patronov, A. (2020). REINVENT 2.0: an AI tool for de novo drug design. Journal of Chemical Information and Modeling, 60(12):5918-5922. https://doi.org/10.1021/acs.jcim.0c00915
42. Popova, M., Isayev, O., and Tropsha, A. (2018). Deep reinforcement learning for de novo drug design. Science Advances, 4(7):eaap7885. https://doi.org/10.1126/sciadv.aap7885
43. Stokes, J.M., Yang, K., Swanson, K., Jin, W., Cubillos-Ruiz, A., Donghia, N.M., MacNair, C.R., French, S., Carfrae, L.A., Bloom-Ackermann, Z., Tran, V.M., Chiappino-Pepe, A., Badran, A.H., Andrews, I.W., Chory, E.J., Church, G.M., Brown, E.D., Jaakkola, T.S., Barzilay, R., and Collins, J.J. (2020). A deep learning approach to antibiotic discovery. Cell, 180(4):688-702.e13. https://doi.org/10.1016/j.cell.2020.01.021
44. Zhavoronkov, A., Ivanenkov, Y.A., Aliper, A., Veselov, M.S., Aladinskiy, V.A., Aladinskaya, A.V., Terentiev, V.A., Polykovskiy, D.A., Kuznetsov, M.D., Asadulaev, A., *et al*. (2019). Deep learning enables rapid identification of potent DDR1 kinase inhibitors. Nature Biotechnology, 37:1038-1040. https://doi.org/10.1038/s41587-019-0224-x
45. Schwaller, P., Laino, T., Gaudin, T., Bolgar, P., Hunter, C.A., Bekas, C., and Lee, A.A. (2019). Molecular Transformer: a model for uncertainty-calibrated chemical reaction prediction. ACS Central Science, 5(9):1572-1583. https://doi.org/10.1021/acscentsci.9b00576
46. Genheden, S., Thakkar, A., Chadimová, V., Reymond, J.-L., Engkvist, O., and Bjerrum, E. (2020). AiZynthFinder: a fast, robust and flexible open-source software for retrosynthetic planning. Journal of Cheminformatics, 12:70. https://doi.org/10.1186/s13321-020-00472-1
47. Saigiridharan, L., Hassen, A.K., Lai, H., Torren-Peraire, P., Engkvist, O., and Genheden, S. (2024). AiZynthFinder 4.0: developments based on learnings from 3 years of industrial application. Journal of Cheminformatics, 16:57. https://doi.org/10.1186/s13321-024-00860-x
48. Mercado, R., Genheden, S., *et al*. (2024). Do Chemformers dream of organic matter? Evaluating a transformer model for multistep retrosynthesis. Journal of Chemical Information and Modeling. https://doi.org/10.1021/acs.jcim.3c01685
49. Zhang, O., Lin, H., Zhang, H., Zhao, H., Huang, Y., Hsieh, C.-Y., Pan, P., and Hou, T. (2024). Deep lead optimization: leveraging generative AI for structural modification. Journal of the American Chemical Society, 146(46):31357-31370. https://doi.org/10.1021/jacs.4c11686
50. Xia, Y., Wang, Y., Wang, Z., and Zhang, W. (2024). A comprehensive review of molecular optimization in artificial intelligence-based drug discovery. Quantitative Biology, 12(1):15-29. https://doi.org/10.1002/qub2.30
51. Rich, A.S. (2024). Machine learning ADME models in practice: four guidelines from a successful lead optimization case study. ACS Medicinal Chemistry Letters. https://doi.org/10.1021/acsmedchemlett.4c00290
52. Xu, Z., Ren, F., Wang, P., Cao, J., Tan, C., Ma, D., Zhao, L., Dai, J., Ding, Y., Fang, H., Li, H., and others. (2025). A generative AI-discovered TNIK inhibitor for idiopathic pulmonary fibrosis: a randomized phase 2a trial. Nature Medicine, 31:2602-2610. https://doi.org/10.1038/s41591-025-03743-2
53. Wilczok, D., and Zhavoronkov, A. (2025). Progress, pitfalls, and impact of AI-driven clinical trials. Clinical Pharmacology & Therapeutics. https://doi.org/10.1002/cpt.3542

54. Ma, X., Liu, B., Li, B., Tong, H.H.Y., and Zhang, Q. (2025). Integrated machine learning and structure-based virtual screening identify osimertinib as a TNIK inhibitor for idiopathic pulmonary fibrosis. Journal of Chemical Information and Modeling, 65(19):10673-10687. https://doi.org/10.1021/acs.jcim.5c01521
55. Zdrazil, B., Felix, E., Hunter, F., Manners, E.J., Blackshaw, J., Corbett, S., de Veij, M., Ioannidis, H., Mendez Lopez, D., Mosquera, J.F., Magarinos, M.P., Bosc, N., Arcila, R., Kizilören, T., Gaulton, A., Bento, A.P., Adasme, M.F., Monecke, P., Landrum, G.A., and Leach, A.R. (2024). The ChEMBL Database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods. Nucleic Acids Research, 52(D1):D1180-D1192. https://doi.org/10.1093/nar/gkad1004
56. Knox, C., Wilson, M., Klinger, C.M., Franklin, M., Oler, E., *et al.* (2024). DrugBank 6.0: the DrugBank Knowledgebase for 2024. Nucleic Acids Research, 52(D1):D1265-D1275. https://doi.org/10.1093/nar/gkad976
57. Liu, T., Lin, Y., Wen, X., Jorissen, R.N., and Gilson, M.K. (2007). BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities. Nucleic Acids Research, 35(Database issue):D198-D201. https://doi.org/10.1093/nar/gkl999
58. Kim, S. (2021). Exploring chemical information in PubChem. Current Protocols, 1(8):e217. https://doi.org/10.1002/cpz1.217
59. Zhou, Y., Zhu, F., and colleagues. (2024). TTD: Therapeutic Target Database describing target druggability information. Nucleic Acids Research, 52(D1):D1465-D1477. https://doi.org/10.1093/nar/gkad751
60. Finan, C., Gaulton, A., Krishnan, E., Wang, Y., Whirl-Carrillo, M., *et al.* (2017). The druggable genome and support for target identification and validation in drug development. Science Translational Medicine, 9(383):eaag1166. https://doi.org/10.1126/scitranslmed.aag1166
61. Zitnik, M., Agrawal, M., and Leskovec, J. (2018). Modeling polypharmacy side effects with graph convolutional networks. Bioinformatics, 34:i457-i466. https://doi.org/10.1093/bioinformatics/bty294
62. Kearnes, S., McCloskey, K., Berndl, M., Pande, V., and Riley, P. (2016). Molecular graph convolutions: moving beyond fingerprints. Journal of Computer-Aided Molecular Design, 30:595-608. https://doi.org/10.1007/s10822-016-9938-8
63. Gilmer, J., Schoenholz, S.S., Riley, P.F., Vinyals, O., and Dahl, G.E. (2017). Neural message passing for quantum chemistry. Proceedings of the 34th International Conference on Machine Learning, 70:1263-1272. https://doi.org/10.48550/arXiv.1704.01212
64. Yang, K., Swanson, K., Jin, W., Coley, C., Eiden, P., Gao, H., Guzman-Perez, A., Hopper, T., Kelley, B., Mathea, M., Palmer, A., Settels, V., Jaakkola, T., Jensen, K., and Barzilay, R. (2019). Analyzing learned molecular representations for property prediction. Journal of Chemical Information and Modeling, 59(8):3370-3388. https://doi.org/10.1021/acs.jcim.9b00237
65. Wu, Z., Ramsundar, B., Feinberg, E.N., Gomes, J., Geniesse, C., Pappu, A.S., Leswing, K., and Pande, V. (2018). MoleculeNet: a benchmark for molecular machine learning. Chemical Science, 9:513-530. https://doi.org/10.1039/C7SC02664A
66. Sliwoski, G., Kothiwale, S., Meiler, J., and Lowe, E.W. Jr. (2014). Computational methods in drug discovery. Pharmacological Reviews, 66:334-395. https://doi.org/10.1124/pr.112.007336
67. Durrant, J.D., and McCammon, J.A. (2011). Molecular dynamics simulations and drug discovery. BMC Biology, 9:71. https://doi.org/10.1186/1741-7007-9-71
68. Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., and Olson, A.J. (2009). AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. Journal of Computational Chemistry, 30:2785-2791. https://doi.org/10.1002/jcc.21256
69. Trott, O., and Olson, A.J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of Computational Chemistry, 31:455-461. https://doi.org/10.1002/jcc.21334
70. Ertl, P., and Schuffenhauer, A. (2009). Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. Journal of Cheminformatics, 1:8. https://doi.org/10.1186/1758-2946-1-8
71. Lipinski, C.A., Lombardo, F., Dominy, B.W., and Feeney, P.J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced Drug Delivery Reviews, 46:3-26. 00129-0). https://doi.org/10.1016/S0169-409X(00)00129-0
72. Veber, D.F., Johnson, S.R., Cheng, H.-Y., Smith, B.R., Ward, K.W., and Kopple, K.D. (2002). Molecular properties that influence the oral bioavailability of drug candidates. Journal of Medicinal Chemistry, 45:2615-2623. https://doi.org/10.1021/jm020017n
73. Tetko, I.V., Karpov, P., and Godin, G. (2020). State-of-the-art augmented intelligence for drug discovery. Nature Reviews Drug Discovery. https://doi.org/10.1038/s41573-020-00075-9
74. Mamoshina, P., Vieira, A., Putin, E., and Zhavoronkov, A. (2016). Applications of deep learning in biomedicine. Molecular Pharmaceutics, 13:1445-1454. https://doi.org/10.1021/acs.molpharmaceut.5b00982
75. Schneider, G. (2018). Automating drug discovery. Nature Reviews Drug Discovery, 17:97-113. https://doi.org/10.1038/nrd.2017.232
76. Segler, M.H.S., Preuss, M., and Waller, M.P. (2018). Planning chemical syntheses with deep neural networks and symbolic AI. Nature, 555:604-610. https://doi.org/10.1038/nature25978
77. Lim, J., Ryu, S., Kim, J.W., and Kim, W.Y. (2018). Molecular generative model based on conditional variational autoencoder for de novo molecular design. Journal of Cheminformatics, 10:31. https://doi.org/10.1186/s13321-018-0286-7
78. Olivecrona, M., Blaschke, T., Engkvist, O., and Chen, H. (2017). Molecular de novo design through deep reinforcement learning. Journal of Cheminformatics, 9:48. https://doi.org/10.1186/s13321-017-0235-x
79. Jin, W., Barzilay, R., and Jaakkola, T. (2018). Junction tree variational autoencoder for molecular graph generation. Proceedings of the 35th International Conference on Machine Learning, 2323-2332. https://doi.org/10.48550/arXiv.1802.04364
80. Brown, N., Fiscato, M., Segler, M.H.S., and Vaucher, A.C. (2019). GuacaMol: benchmarking models for de novo molecular design. Journal of Chemical Information and Modeling, 59:1096-1108. https://doi.org/10.1021/acs.jcim.8b00839
81. Segler, M.H.S., Kogej, T., Tyrchan, C., and Waller, M.P. (2018). Generating focused molecule libraries for drug discovery with recurrent neural networks. ACS Central Science, 4:120-131. https://doi.org/10.1021/acscentsci.7b00512
82. Stärk, H., Ganea, O., Pattanaik, L., Barzilay, R., and Jaakkola, T. (2022). EquiBind: geometric deep learning for drug binding structure prediction. Proceedings of the 39th International Conference on Machine Learning. https://doi.org/10.48550/arXiv.2202.05146
83. Mullowney, M.W., Duncan, K.R., Elsayed, S.S., Garg, N., van der Hooft, J.J.J., Martin, N.I., Meijer, D., *et al.* (2023). Artificial intelligence for natural product drug discovery. Nature Reviews Drug Discovery, 22:895-916. https://doi.org/10.1038/s41573-023-00774-7
84. He, B., Guo, J., Tong, H.H.Y., and To, W.M. (2024). Artificial intelligence in drug discovery: a bibliometric analysis and literature review. Mini-Reviews in Medicinal Chemistry, 24(14):1353-1367. https://doi.org/10.2174/0113895575271267231123160503
85. Abubakar, M.L., Kapoor, N., Sharma, A., Gambhir, L., Jasuja, N.D., and Sharma, G. (2024). Artificial intelligence in drug identification and validation: a scoping review. Drug Research, 74(5):208-219. https://doi.org/10.1055/a-2306-8311
86. Dhudum, R., Ganeshpurkar, A., and Pawar, A. (2024). Revolutionizing drug discovery: a comprehensive review of AI applications. Drugs and Drug Candidates, 3(1):148-171. https://doi.org/10.3390/ddc3010009
87. Karade, D., *et al.* (2023). AIDrugApp: artificial intelligence-based web application for virtual screening of inhibitors against SARS-CoV-2. Journal of Experimental & Theoretical Artificial Intelligence, 35(3):395-443. https://doi.org/10.1080/0952813X.2022.2058619

88. Zhou, G., Gao, Z., Ding, Q., Zheng, H., Xu, H., Wei, Z., Zhang, L., and Ke, G. (2023). Uni-Mol: a universal 3D molecular representation learning framework. International Conference on Learning Representations. https://doi.org/10.48550/arXiv.2303.13873
89. Lu, S., Gao, Z., He, D., Zhang, L., and Ke, G. (2024). Data-driven quantum chemical property prediction leveraging 3D conformations with Uni-Mol+. Nature Communications, 15:7104. https://doi.org/10.1038/s41467-024-51171-7
90. Watson, J.L., Juergens, D., Bennett, N.R., Trippe, B.L., Yim, J., Eisenach, H.E., Ahern, W., Borst, A.J., *et al*. (2023). De novo design of protein structure and function with RFdiffusion. Nature, 620:1089-1100. https://doi.org/10.1038/s41586-023-06415-8
91. Roy, R., and Al-Hashimi, H.M. (2024). AlphaFold3 takes a step toward decoding molecular behavior and biological computation. Nature Structural & Molecular Biology, 31:997-1000. https://doi.org/10.1038/s41594-024-01350-2
92. Zhang, O., Lin, H., and colleagues. (2026). AutoLead: an LLM-guided Bayesian optimization framework for multi-objective lead optimization. Journal of Chemical Information and Modeling. https://doi.org/10.1021/acs.jcim.6c00148
93. Lee JW, Maria-Solano MA, Vu TNL, *et al*. Big data and artificial intelligence (AI) methodologies for computer-aided drug design (CADD). Biochem Soc Trans. 2022;50(1):241-252. https://doi.org/10.1042/BST20211240
94. Bajorath J. Deep machine learning for computer-aided drug design. Front Drug Discov. 2022;2:829043. https://doi.org/10.3389/fddsv.2022.829043
95. Gangwal A, Ansari A, Ahmad I, *et al*. Generative artificial intelligence in drug discovery: basic framework, recent advances, challenges, and opportunities. Front Pharmacol. 2024;15:1331062. https://doi.org/10.3389/fphar.2024.1331062
96. Gangwal A, Lavecchia A. Unleashing the power of generative AI in drug discovery. Drug Discov Today. 2024;29(7):103992. https://doi.org/10.1016/j.drudis.2024.103992

HOW TO CITE THIS ARTICLE: Rai O, Johnson S, Pulamarasetti SS, Vandana K, Priya BR. Redefining the CADD Paradigm: Harnessing Artificial Intelligence from Target Discovery to Lead Optimization. J. of Drug Disc. and Health Sci. 2026;3(2):46-55. **DOI:** 10.21590/jddhs.03.02.08