INTERNATIONAL JOURNAL OF PHARMACEUTICAL SCIENCES
[ISSN: 0975-4725; CODEN(USA): IJPS00]
Journal Homepage: https://www.ijpsjournal.com

OPEN ACCESS

Review Article

Role of Computational Chemistry in Modern Drug Discovery: Recent Advances, Applications, Challenges, and Future Perspectives

Prakash Sahu*, Sanjay Kumar Gautam, Smita Patel, Govind Yadav

Department of Pharmaceutical Chemistry, GSRM Memorial College of Pharmacy, Lucknow, Uttar Pradesh, India.

ARTICLE INFO

Published: 4 Nov 2026
Keywords:
Computational chemistry, computer-aided drug design (CADD), molecular docking, molecular dynamics, QSAR, pharmacophore modeling, virtual screening, ADMET, artificial intelligence, machine learning, medicinal chemistry
DOI:
10.5281/zenodo.23939619

ABSTRACT

The development of a novel therapeutic agent is a time-consuming and resource-intensive endeavor, traditionally requiring extensive chemical synthesis and biological evaluation. A major bottleneck in this pipeline is the efficient selection of viable lead molecules from vast chemical libraries. Computational chemistry has emerged as a cornerstone of modern drug discovery, enabling the in silico evaluation of molecular properties and target-ligand interactions prior to wet-lab experimentation. This review provides a comprehensive overview of key computational methodologies employed across distinct stages of drug discovery. Structure-based approaches, such as molecular docking and molecular dynamics (MD) simulations, offer structural insights into ligand binding modes, binding affinities, and macromolecular stability. Ligand-based strategies, including Quantitative Structure-Activity Relationship (QSAR) modeling and pharmacophore mapping, facilitate the identification of critical structural features essential for biological activity. Additionally, virtual screening (VS) enables high-throughput filtration of candidate compounds, while in silico ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) profiling aids in the early elimination of pharmacokinetically unfavorable candidates. Recently, the integration of Artificial Intelligence (AI) and Machine Learning (ML)—particularly deep learning and generative algorithms—has advanced molecular property prediction, de novo drug design, and lead optimization. However, computational predictions remain dependent on dataset quality, algorithmic parameters, and underlying physical assumptions, necessitating validation through experimental assays. Herein, we discuss the applications, advantages, structural limitations, and future outlook of computational approaches in pharmaceutical research.

***Corresponding Author:** Prakash Sahu
Address: Department of Pharmaceutical Chemistry, GSRM Memorial College of Pharmacy, Lucknow, Uttar Pradesh, India
Email ✉: prakashsahu42@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.

INTRODUCTION

Drug discovery and development is a complex, multi-stage process involving target identification, lead generation, chemical synthesis, structural optimization, preclinical testing, and clinical evaluation. Given the astronomical size of chemical space (estimated at over 10^{60} small molecules), screening compounds purely through empirical wet-lab experiments is cost-prohibitive and inefficient. Consequently, minimizing redundant experimental trials remains a primary objective in pharmaceutical research.

While classical medicinal chemistry relies heavily on synthesis followed by in vitro and in vivo screening, modern pipelines increasingly integrate Computer-Aided Drug Design (CADD). Computational chemistry enables researchers to model molecular structures, calculate quantum and physicochemical properties, and simulate target-ligand interactions in silico.

Over the past few decades, computational tools have evolved from basic molecular mechanics and classical QSAR models to high-performance molecular dynamics, deep QSAR, and generative AI platforms. Far from replacing experimental validation, computational workflows act as a predictive framework that prioritizes high-probability candidates, optimizes chemical series, and rationalizes observed biological activity.

2. Objectives

The primary objectives of this review are:

* To outline the strategic role of computational chemistry in accelerating modern drug discovery.
* To evaluate core computational paradigms, including structure-based and ligand-based drug design.
* To examine the principles and applications of molecular docking and molecular dynamics simulations.
* To elucidate the utility of QSAR modeling and pharmacophore mapping in lead optimization.
* To discuss virtual screening strategies and cheminformatics tools for chemical library design.
* To address the significance of in silico ADMET and toxicity prediction.
* To highlight recent advancements in AI, machine learning, deep learning, and generative molecular design.
* To critically analyze the current limitations, data dependencies, and validation requirements of CADD.
* To provide insights into emerging trends, such as explainable AI, automated workflows, and federated learning.

3. Overview of Computational Drug Discovery Workflows

Computational methods generally fall into two broad categories based on input data availability:

* Structure-Based Drug Design (SBDD): Employed when the 3D atomic structure of the biological target (protein/nucleic acid) is resolved.
* Ligand-Based Drug Design (LBDD): Applied when target structural data is absent, relying instead on structural and biological data from known active ligands.

A standard linear CADD workflow follows a structured sequence:

Target Identification → Structural/Binding Site Analysis → Library Selection → Virtual Screening → Molecular Docking → Property Filter (ADMET) → Lead Identification → Lead Optimization → Experimental Validation

In practice, this pipeline operates iteratively; experimental findings continuously refine computational algorithms and theoretical parameters.

3. Structure-Based Drug Design (SBDD)

SBDD relies on atomic-resolution target structures derived from X-ray crystallography, Nuclear Magnetic Resonance (NMR) spectroscopy, Cryo-Electron Microscopy (cryo-EM), or high-accuracy computational modeling (e.g., AlphaFold).

Key Features Analyzed in Binding Sites

* Hydrogen bond donor and acceptor networks
* Hydrophobic pockets and aromatic interactions (\pi-\pi, cation--\pi)
* Electrostatic surface potentials and charged residues
* Steric constraints and channel geometries

* Explicit structural water molecules participating in ligand bridging

SBDD guides structural modifications to maximize ligand binding affinity and target selectivity. However, standard SBDD workflows often treat proteins as rigid entities, which may overlook conformational adaptations induced by ligand binding.

4. Ligand-Based Drug Design (LBDD)

When target structures are unavailable, LBDD operates on the principle that structurally or chemically similar compounds elicit similar biological activities (the similar property principle).

Core LBDD Methodologies

* QSAR Modeling: Statistical correlation between molecular descriptors and biological activity.
* Pharmacophore Mapping: Extraction of 3D spatial arrangements of essential interactive features.
* Molecular Fingerprints & Similarity Searching: Quantitative comparison of molecular representations (e.g., Morgan/ECFP fingerprints, Tanimoto coefficients).
* Chemical Clustering: Grouping active compounds into distinct structural families to identify novel scaffolds.

LBDD performance depends heavily on the chemical diversity and range of the training set; models trained on narrow structural classes struggle to generalize to novel chemical spaces.

5. Molecular Docking

Molecular docking predicts the preferred orientation and binding conformation (pose) of a small molecule within a target binding pocket. Algorithms evaluate conformational space and calculate non-covalent interactions using empirical, force-field, or knowledge-based scoring functions.

Evaluated Interactions

* Hydrogen bonding and salt bridges
* Hydrophobic and van der Waals contacts
* \pi--\pi stacking and cation--\pi interactions
* Metal-coordination interactions

Key Applications

* High-throughput virtual screening of compound libraries
* Binding mode identification and structural rationalization of SAR
* Scaffold hopping and lead optimization

Note: Docking scoring functions approximate true binding free energies (\Delta G). A superior docking score does not guarantee in vitro activity due to simplified solvation, entropic, and receptor flexibility parameters.

6. Molecular Dynamics (MD) Simulations

MD simulations incorporate time-dependent atomic motions, allowing researchers to study the conformational stability, flexibility, and thermodynamics of target-ligand complexes in explicit solvent environments.

Standard Analytical Parameters

* Root-Mean-Square Deviation (RMSD): Measures overall structural stability and displacement over time.
* Root-Mean-Square Fluctuation (RMSF): Quantifies localized residue flexibility.
* Radius of Gyration (R_g): Assesses global protein compactness.
* Solvent-Accessible Surface Area (SASA): Evaluates surface exposure to water molecules.
* Hydrogen Bond Occupancy Analysis: Tracks the durability of key binding interactions throughout the trajectory.

Integrating molecular docking with MD trajectory analysis significantly reduces false positives in virtual screening campaigns.

7. Quantitative Structure-Activity Relationship (QSAR)

QSAR converts chemical structures into quantitative mathematical features (descriptors) and correlates them with measured biological endpoints using regression or classification algorithms.

Commonly Used Descriptors

* Physicochemical/Constitutional: Molecular weight, LogP, Polar Surface Area (PSA)
* Topological & Electrostatic: Hydrogen bond donors/acceptors, partial charges, connectivity indices
* 2D/3D Structural: Molecular fingerprints, shape indices, quantum mechanical descriptors

Model validity requires rigorous adherence to OECD guidelines, including internal validation (Q^2), external validation (R^2_{\text {ext}}), and precise definition of the applicability domain.

8. Pharmacophore Modeling

A pharmacophore defines the ensemble of steric and electronic features necessary for optimal supramolecular interaction with a specific biological target.

Core Features

* Hydrogen bond donors (HBD) and acceptors (HBA)
* Hydrophobic centroids (H) and aromatic rings (RA)
* Positive (PI) and negative (NI) ionizable groups

Pharmacophore screening excels at scaffold hopping—discovering structurally distinct active compounds that preserve necessary spatial feature alignments.

9. Virtual Screening (VS) and Cheminformatics

Virtual Screening Paradigms

* Structure-Based Virtual Screening (SBVS): Involves docking millions of library compounds against target binding pockets to rank binding energy.
* Ligand-Based Virtual Screening (LBVS): Filters libraries using pharmacophore queries, fingerprint similarity, or machine learning models.

Cheminformatics Applications

Cheminformatics provides the computational infrastructure to process, analyze, and manage molecular structures and biological datasets.

* Chemical library diversity design and filtering (e.g., Lipinski's Rule of 5, Veber's rules)
* Chemical space exploration via dimensionality reduction (PCA, t-SNE, UMAP)
* Substructure, scaffold, and similarity searching

11. Quantum Chemical Calculations

Quantum mechanics (QM) methods, particularly Density Functional Theory (DFT), calculate electronic structures, highest occupied/lowest unoccupied molecular orbitals (HOMO/LUMO), and charge distributions.

Applications in Drug Design

* Analyzing localized reactivity and polarization upon ligand binding
* Parameterizing non-standard ligands for molecular dynamics force fields
* Modeling enzymatic reaction mechanisms and covalent inhibitor kinetics
* Hybrid QM/MM (Quantum Mechanics/Molecular Mechanics) calculations for catalytic sitesh

12. In Silico ADMET Profiling

Early evaluation of Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) reduces late-stage clinical attrition rates caused by unfavorable pharmacokinetics or toxicity profiles.

| ADMET Property | Computational Predictions & Metrics |

| Absorption | Caco-2 permeability, human intestinal absorption (HIA), aqueous solubility (\log S) |

| Distribution | Plasma protein binding (PPB), Blood-Brain Barrier (BBB) penetration, Volume of Distribution (V_d) |

| Metabolism | Cytochrome P450 (CYP) inhibition/substrate profiling, metabolic soft-spot prediction |

| Excretion | Renal clearance, renal organic cation transporter (OCT) interaction prediction |

| Toxicity | hERG potassium channel inhibition, Ames mutagenicity, hepatotoxicity, rodent carcinogenicity |

13. Artificial Intelligence and Machine Learning in Drug Discovery

AI and ML have transformed drug design by automatically detecting complex, high-dimensional patterns within chemical and biological datasets.

Machine Learning Methods

* Classical ML: Random Forest (RF), Support Vector Machines (SVM), Extreme Gradient Boosting (XGBoost).
* Deep Learning: Deep Neural Networks (DNN), Convolutional Neural Networks (CNN for 3D grids), Graph Neural Networks (GNN for molecular graphs).

Primary Applications

* Biological target prediction and drug-target interaction (DTI) modeling
* Automated SAR profiling and multi-parameter optimization
* In silico drug repurposing to identify new indications for existing drugs

14. Deep Learning, Deep QSAR, and Generative AI

Deep QSAR

Traditional QSAR requires manual selection of molecular descriptors. Deep QSAR automatically learns hierarchical representations directly from raw molecular formats (SMILES strings, molecular graphs, or 3D atomic coordinates), capturing complex non-linear relationships.

Generative Molecular Design

Rather than filtering pre-existing chemical libraries, generative AI creates novel molecules de novo tailored to target specifications.

* Architectures: Variational Autoencoders (VAEs), Generative Adversarial Networks (GANs), Transformer-based language models, and Diffusion models.
* Reinforcement Learning (RL): Guides generation toward desired chemical property profiles (e.g., target affinity, synthetic accessibility, lower toxicity).

15. Target Structure Prediction and Computational Repurposing

Protein Structure Prediction

Tools like AlphaFold, ESMFold, and RosettaFold generate accurate structural models for proteins whose experimental structures remain unresolved. These models provide structural targets for binding pocket detection and docking studies.

Computational Drug Repurposing

Drug repurposing identifies novel therapeutic uses for existing, clinically approved pharmaceuticals. Computational workflows cross-examine drug-target interactomes, disease transcriptomics, and pathway network models to suggest viable off-target repurposing hypotheses.

16. Integration of Computational and Experimental Workflows

Computational chemistry functions best as part of an iterative design-make-test-analyze (DMTA) cycle:

In Silico Prediction → Prioritized Synthesis → In Vitro / In Vivo Assay → Data Feedback & Model Refinement

This feedback loop ensures continuous optimization: experimental discrepancies update computational models, which subsequently produce refined molecular candidates.

17. Practical Applications in Pharmaceutical Chemistry

* Identification of Novel Scaffolds: Discovering distinct structural families for target modulation via virtual screening.
* SAR Rationalization: Visualizing atomic interactions to explain activity differences among related analogs.
* Focused Library Design: Filtering structural libraries based on target compatibility and physicochemical properties.
* Physicochemical Optimization: Balancing molecular weight, lipophilicity, polar surface area, and aqueous solubility.
* Computer-Aided Retrosynthesis: Utilizing AI synthesis planning tools (e.g., SYNTHIA, ASKCOS) to design efficient synthetic routes.

18. Advantages of Computational Chemistry

* Decreased Experimental Screening Costs: Filters out inactive or toxic compounds prior to synthesis.
* Time Efficiency: Accelerates lead identification and optimization phases.

* Rational Molecular Design: Replaces empirical trial-and-error with mechanistic interaction hypotheses.
* Scalable Data Processing: Analyzes multi-million compound databases efficiently.
* Early ADMET Risk Mitigation: Flags pharmacokinetic and safety liabilities early in development.

19. Limitations and Current Challenges

* Data Quality & Bias: Models trained on noisy, inconsistent, or structurally narrow experimental data produce unreliable predictions.
* Simplified Biological Systems: In silico models struggle to simulate solvent entropy, receptor dynamic ensembles, cellular transport, and biological feedback networks fully.
* Docking & Scoring Limitations: Scoring functions often fail to accurately predict binding affinity across diverse ligand structural classes.
* Model Overfitting & Applicability Domain: ML models can overfit training sets and perform poorly when applied to novel chemical space.
* Black-Box AI Models: Deep learning systems often lack mechanistic interpretability, making it difficult to understand the chemical rationale behind predictions.
* Requirement for Experimental Confirmation: Computational hypotheses remain predictive until validated by biophysical, biochemical, and pharmacological assays.

20. Emerging Trends and Future Perspectives

* Explainable AI (XAI): Interpretability tools (e.g., SHAP, LIME, attention maps) that explain the chemical basis of AI predictions.
* Multimodal Data Integration: Models integrating genomic, structural, transcriptomic, and assay data to predict drug activity.
* Autonomous & Closed-Loop Platforms: AI molecular design coupled directly with automated robotic synthesis and high-throughput screening assays.
* Federated Learning: Multi-institutional AI model training across proprietary pharmaceutical datasets without sharing underlying molecular structures.
* Quantum Computing: Potential long-term advancement in accurate electronic structure calculations and complex biomolecular simulations.

CONCLUSION

Computational chemistry has established itself as an indispensable pillar of modern drug discovery. Methodologies such as molecular docking, dynamics simulations, QSAR, pharmacophore mapping, virtual screening, ADMET profiling, and AI algorithms offer complementary tools to guide experimental drug design.

While computational predictions cannot eliminate the need for wet-lab research, integrating in silico modeling with organic synthesis and biological evaluation creates an efficient, rational workflow. Continued advancements in algorithm design, structural biology, and computational power will further refine predictive accuracy, accelerating the discovery of novel therapeutics.

Declarations

Author Contribution

Prakash Sahu: Conceptualization, literature curation, formal analysis, original draft preparation, review, and editing.

Conflict of Interest

The author declares no conflict of interest.

Funding

This research received no external funding.

Acknowledgements

The author acknowledges the academic and institutional support provided by the Department of Pharmaceutical Chemistry, GSRM Memorial College of Pharmacy, Lucknow, Uttar Pradesh, India.

REFERENCES

1. Tropsha A, Isayev O, Varnek A, Schneider G, Cherkasov A. Integrating QSAR modelling and

deep learning in drug discovery: the emergence of deep QSAR. Nat Rev Drug Discov. 2024;23:141–155.

2. Wei H, McCammon JA. Structure and dynamics in drug discovery. Npj Drug Discov. 2024;1:1.
3. Gangwal A, Ansari A, Ahmad I, Azad AK, Sulaiman WMAM. Current strategies to address data scarcity in artificial intelligence-based drug discovery: a comprehensive review. Comput Biol Med. 2024;179:108734.
4. Kawai K, Karuo Y, Tarui A, Sato K, Kataoka M, Omote M. Development of drug discovery platforms using artificial intelligence and cheminformatics. Chem Pharm Bull. 2024;72 (9):794–799.
5. Abbas MKG, Rassam A, Karamshahi F, Abunora R, Abouseada M. The role of AI in drug discovery. ChemBioChem. 2024;25 (14):e202300816.
6. Kokudeva M, Vichev M, Naseva E, Miteva DG, Velikova T. Artificial intelligence as a tool in drug discovery and development. World J Exp Med. 2024;14 (3):96042.
7. Agu PC, Obulose CN. Piquing artificial intelligence towards drug discovery: tools, techniques, and applications. Drug Dev Res. 2024;85 (2):e22159.
8. Catacutan DB, Alexander J, Arnold A, et al. Machine learning in preclinical drug discovery. Nat Chem Biol. 2024;20:960–973.
9. Schneider G. Automating drug discovery. Nat Rev Drug Discov. 2018;17:97–113.
10. Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. Nat Rev Drug Discov. 2004;3:935–949.
11. Sliwoski G, Kothiwale S, Meiler J, Lowe EW Jr. Computational methods in drug discovery. Pharmacol Rev. 2014;66:334–395.
12. Durrant JD, McCammon JA. Molecular dynamics simulations and drug discovery. BMC Biol. 2011;9:71.
13. Leach AR. Molecular Modelling: Principles and Applications. 2nd ed. Pearson Education; 2001.
14. Patrick GL. An Introduction to Medicinal Chemistry. 6th ed. Oxford University Press; 2017.
15. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev. 2001;46:3–26.
16. Vamathevan J, Clark D, Czodrowski P, et al. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov. 2019;18:463–477.
17. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596:583–589.
18. Koirala M, Yan L, Mohamed Z, DiPaola M. AI-integrated QSAR modeling for enhanced drug discovery: from classical approaches to deep learning and structural insight. Int J Mol Sci. 2025;26:9384.
19. Chakraborty C, Bhattacharya M, Lee SS, Wen ZH, Lo YH. The changing scenario of drug discovery using AI to deep learning: recent advancement, success stories, collaborations, and challenges. OMICS J Integr Biol. 2024.
20. Generative artificial intelligence for small molecule drug design. Curr Opin Chem Biol. 2024;81:103175.

HOW TO CITE: Prakash Sahu, Sanjay Kumar Gautam, Smita Patel, Govind Yadav Role of Computational Chemistry in Modern Drug Discovery: Recent Advances, Applications, Challenges, and Future Perspectives, Int. J. of Pharm. Sci., 2026, Vol 4, Issue 11, 703-709. https://doi.org/10.5281/zenodo.23939619