



ARTICLE ID: BTS0626005

COMPUTATIONAL DRUG DESIGN-IN SILICO REVOLUTION REVIEWD: THE EVOLUTION, INNOVATION, AND INFLUENCE ON MODERN DRUG DISCOVERY

Talluri Sarada Mrinalini.

Assistant professor, KVSRR Siddhartha College of Pharmaceutical Sciences, Vijayawada, Andhra Pradesh, India.

Corresponding Mail ID: mrinalini.talluri@gmail.com

RECEIVED: 15 July 2025 | **REVISED & ACCEPTED:** 09 November 2025.

KEYWORDS: In-silico drug design, Computer-aided drug discovery (CADD), Virtual screening, Molecular docking, ADMET prediction, AlphaFold, AI in pharmacology, Computational pharmacology, QSAR modelling.

ABSTRACT:

In-silico drug design evolved from a supportive technique into a core component of pharmaceutical research. This review provided a comprehensive overview of the computational drug discovery landscape, emphasizing innovations in methodologies, AI integration, and software capabilities. Breakthroughs such as AlphaFold 3, 3DMolFormer, and DiffDock significantly advanced protein structure prediction, docking accuracy, and generative molecule design. Artificial intelligence and machine learning enabled large-scale virtual screening, de novo compound generation, and predictive ADMET modelling with improved precision and efficiency. Widely adopted tools—both commercial (e.g., Schrödinger, MOE, Optibrium) and open-source (e.g., AutoDock, PyRx, DataWarrior)—integrated molecular docking, pharmacophore modelling, QSAR, molecular dynamics, and explainable AI for simulating drug–target interactions prior to synthesis. This review contrasted computational and experimental methods, highlighting benefits such as scalability, cost reduction, and early toxicity prediction, while also acknowledging limitations like biological oversimplification and the necessity for experimental validation. Case studies illustrated real-world applications in drug repurposing and hit identification. Emerging directions included quantum simulations, personalized virtual trials, and federated AI platforms—further reinforcing the transformative role of in silico tools in drug discovery.

INTRODUCTION

The field of drug discovery has witnessed a paradigm shift over the past decade, with computational approaches—collectively referred to as in silico methods—playing a central role in identifying, optimizing, and validating potential drug candidates. These methods have gained unprecedented momentum between 2020 and 2025 due to rapid advances in artificial intelligence (AI), deep learning, structural

bioinformatics, and molecular simulations. The COVID-19 pandemic further underscored the urgency for faster, cost-effective, and scalable drug discovery tools, catalyzing the widespread adoption of computational platforms for virtual screening, de novo molecule generation, ADMET prediction, and protein–ligand interaction modelling. In contrast to traditional experimental methods, in silico

**ARTICLE ID: BTS0626005**

techniques allow rapid screening of millions to billions of compounds, reduce wet-lab costs, and provide insights into molecular mechanisms even before synthesis. The emergence of tools such as AlphaFold 3 for accurate protein structure prediction, DiffDock for AI-driven docking, and generative models like 3DMolFormer for novel molecule design has dramatically enhanced the precision and utility of computational drug design pipelines. Despite these advances, computational methods are not without limitations. Challenges remain in predicting biological complexity, synthetic accessibility, and ensuring model generalizability. This review provides a comprehensive overview of modern in silico drug design, comparing it with experimental approaches, highlighting popular software platforms, showcasing literature-based examples, and discussing both the strengths and the inherent gaps of these methods. It aims to inform researchers and practitioners about current trends and inspire hybrid strategies that integrate computation with experimental validation to accelerate drug development.

LITERATURE REVIEW:**Establishing the Computational Core: Structure-Based Techniques**

Between 2010 and 2015, in silico drug design heavily relied on traditional molecular modelling techniques such as molecular docking, pharmacophore modelling, and molecular dynamics (MD) simulations. These methods enabled researchers to predict how drug candidates interact with their biological targets at the atomic level. Studies during this period emphasized the use of structure-based drug design (SBDD) as a central tool in rational drug discovery, leveraging known 3D structures of proteins from X-ray

crystallography or NMR data. For instance, the wide adoption of tools like AutoDock, GOLD, and Schrödinger's Glide allowed high-throughput virtual screening (HTVS) campaigns against infectious diseases, cancer targets, and neurological disorders. Researchers began integrating molecular dynamics (via GROMACS or AMBER) to simulate protein-ligand complexes, allowing predictions of binding stability and conformational flexibility over time. The foundational half of the decade was driven by rapid utilization of structural data, with computational tools streamlining early-phase hit discovery and lead optimization processes.

RISE OF LIGAND-BASED MODELLING AND DRUG REPURPOSING

As limitations in structural data and computational resources emerged, ligand-based approaches such as Quantitative Structure–Activity Relationship (QSAR) and pharmacophore modelling gained renewed attention. These methods used known bioactive molecules to model features required for activity and were particularly useful in screening when protein targets were unknown or undruggable. During this phase, drug repurposing became a powerful strategy to reduce time and cost in pharmaceutical development. Literature began highlighting computational repositioning frameworks such as similarity-based, inference-based, and network pharmacology approaches. For example, Li et al. (2016) reviewed repositioning efforts using drug-target-disease networks and public databases like DrugBank, ChEMBL, and CTD. Tools like DRAR-CPI and SEA were used to infer new targets for known drugs. The mid-decade phase emphasized data-driven drug repurposing and target inference,

**ARTICLE ID: BTS0626005**

showcasing computational power in maximizing the therapeutic value of existing chemical space.

INTEGRATION AND BENCHMARKING OF COMPUTATIONAL PIPELINES

This period saw a surge in efforts to integrate structure- and ligand-based models with cheminformatics and predictive toxicology. Studies by Bajorath and others emphasized the need for data curation, standardized benchmarking, and interpretability of QSAR models. ADMET (absorption, distribution, metabolism, excretion, and toxicity) prediction tools such as admetSAR, SwissADME, and pkCSM became essential components of computational workflows, aiding early pharmacokinetic filtering. Researchers increasingly focused on multi-parametric optimization strategies combining potency, selectivity, and safety profiles. Simultaneously, molecular simulations became more advanced, incorporating coarse-grained models and enhanced sampling techniques. Literature reviews (e.g., Muegge et al., 2017) highlighted hybrid pipelines combining docking, MD, and machine learning for better predictive performance and lower false positives. Integration of modelling layers and focus on predictive safety shaped this era, emphasizing robust, reproducible pipelines essential for translational outcomes.

ADVENT OF ARTIFICIAL INTELLIGENCE IN DRUG DESIGN

The final years of the decade heralded the rise of deep learning (DL) and generative models in drug discovery. Academic and industrial research rapidly adopted variational autoencoders (VAEs), recurrent neural networks (RNNs), and graph neural

networks (GNNs) to model chemical space and predict molecular properties. In 2020, Sanchez-Lengeling and Aspuru-Guzik reviewed the potential of deep generative models to produce novel drug-like molecules beyond known libraries.

Meanwhile, the application of natural language processing (NLP) to SMILES strings led to the development of tools like ChemBERTa and Mol2Vec, allowing semantic understanding of molecules. Reviews (e.g., Gawehn et al., 2018) categorized AI applications into target prediction, bioactivity modelling, ADMET profiling, and retrosynthetic planning. The explosion of AI research also revived focus on explainability and the need for transparent, interpretable predictions, leading to the emergence of Explainable AI (XAI) frameworks. The decade culminated in a technological leap, with AI and generative modelling extending the capabilities of in silico techniques into creative and autonomous drug design.

RECENT ADVANCES IN COMPUTATIONAL DRUG DESIGN

Li et al. (2024) introduced a hybrid quantum computing pipeline for drug discovery, outperforming classical simulations in modelling complex molecular interactions and covalent bonding, thereby accelerating therapeutic development. In parallel, 3DMolFormer (Hu et al., 2025) proposed a dual-channel transformer model for unified protein-ligand docking and pocket-aware 3D drug design, addressing challenges in 3D representation and data scarcity through pretraining and reinforcement learning. Defrock (Corso et al., 2022) reframed docking as a diffusion problem, achieving superior pose prediction by navigating translational, rotational, and torsional spaces with greater precision. These AI-



ARTICLE ID: BTS0626005

enhanced tools offer scalable, accurate alternatives to traditional workflows. Earlier, Ou-Yang et al. (2012) emphasized integrating docking, ADMET, MD simulations, and explainable AI into comprehensive pipelines. Collectively, these studies illustrate the transition from

rule-based to learning-based frameworks, with quantum computing and deep generative models redefining the drug discovery paradigm and offering promising solutions to biological complexity and screening inefficiency.

POPULAR SOFTWARE TOOLS:

CATEGORY	TOOL NAME	ORIGIN	KEY FEATURES
Molecular Docking & Virtual Screening	AutoDock / Vina	Scripps Research	Open-source, widely used for ligand docking.
	GOLD	CCDC	Flexible docking with customizable scoring.
	Schrödinger Glide	Schrödinger Inc.	Commercial docking with high precision.
	DOCK	UCSF	Academic tool for structure-based screening.
	DiffDock	MIT	AI-based diffusion model for accurate docking.
ADMET Prediction & Drug-Likeness	SwissADME	Swiss Institute of Bioinformatics	Web-based tool for drug-likeness and ADME.
	pkCSM	University of Cambridge	Graph-based pharmacokinetics predictions.
	admetSAR	Chinese Academy of Sciences	Online tool for ADMET risk profiling.
	PreADMET	PharmaAlgorithm	Predicts ADME and toxicity parameters.
Molecular Dynamics Simulations	GROMACS	University of Groningen	Fast, open-source MD engine for biomolecules.
	AMBER	University of California	Widely used MD suite for force field studies.
	CHARMM	Harvard University	Advanced force field and MD simulations.
	Desmond	Schrödinger Inc.	GPU-accelerated high-speed MD simulations.
AI-Based / Generative Design Tools	DeepChem	Stanford	ML library for deep learning in chemistry.
	AlphaFold / AlphaFold2	DeepMind	AI model for 3D protein structure prediction.



ARTICLE ID: BTS0626005

Integrated Drug Discovery Platforms	MoleculeChef / MolGPT	AI4Science	Deep generative models for molecule creation.
	REINVENT	AstraZeneca	Reinforcement learning for de novo design.
	3DMolFormer	Tsinghua Univ. (Hu et al.)	Dual-channel transformer for docking and 3D design.
	KNIME	Open-source	Modular cheminformatics and ADMET pipelines.
	MOE	Chemical Computing Group	Full-featured SBDD, QSAR, and visualization tool.
	BIOVIA Discovery Studio	Dassault Systèmes	GUI for modeling, docking, and simulation.
	Insilico Medicine Platform	Insilico Medicine	AI-driven drug discovery across development stages.

ADVANTAGES:

FEATURE	IN SILICO METHODS	EXPERIMENTAL METHODS
Scale & Speed	Virtual screens of 10^8 – 10^9 molecules; generative models yield novel scaffolds in hours to days	HTS limited to 10^5 – 10^6 ; synthesis and assays take weeks/months
Novelty & Diversity	De novo molecule generation explores chemical space beyond known libraries	Dependent on existing synthetic routes, limited diversity
Binding Affinity Prediction	FEP/TI via Schrödinger, MD; DiffDock docking with accurate pose prediction	Biophysical assays (SPR, ITC) but low throughput
Structural Insight	AlphaFold 3 enables accurate modelling of protein–ligand and macromolecular complexes	Requires experimental crystallography, cryo-EM, highly resource-intensive
Explainability	XAI frameworks within platforms like Exscientia, Insilco, Schrödinger deliver interpretable ML predictions	Validation via assays lacks predictive model transparency
Cost Efficiency	Significantly reduces expenses by avoiding unnecessary synthesis and biological testing of poor candidates.	Requires investment in reagents, instrumentation, cell lines, animals, and skilled labour.



ARTICLE ID: BTS0626005

Compound Library Size	Can virtually screen billions of molecules by using generative models or combinatorial libraries.	Limited to the size of physically available chemical libraries.
Early Filtering of Drug-Likeness	ADMET prediction tools (e.g., SwissADME, pkCSM) help eliminate non-viable compounds early.	Drug-likeness and safety are often discovered later, post-synthesis and after in vitro/in vivo studies.
Customization and Flexibility	Supports personalized drug design, scaffold hopping, and virtual analogue generation.	Experimental alteration of scaffolds is time-consuming and requires skilled synthetic chemists.
Mechanistic Insight	Molecular dynamics and docking provide atomistic insight into target-ligand interactions and predict binding mechanisms.	Biochemical assays give functional outputs but may lack structural resolution.
Scalability and Automation	Easily scalable across cloud servers and high-performance computing systems; automated with AI/ML.	Limited by equipment throughput and lab personnel; automation is expensive.
Drug Repurposing Potential	Computational tools can mine databases and network pharmacology to predict new indications for approved drugs.	Requires new hypothesis-driven experiments and clinical trials for each indication.
Integration with Omics Data	Seamless integration with genomics, proteomics, and transcriptomics for target identification (e.g., bioinformatics pipelines).	Omics-driven drug discovery is less common and requires specialized equipment and expertise.
Ethical Considerations	Reduces or avoids animal testing in early discovery phases.	Often involves animal models for toxicity and efficacy testing.

LIMITATIONS

- Inaccurate protein structures or scoring functions can mislead predictions especially when docking into homology models. Tools like AutoDock or DiffDock may still mispredict poses in novel targets.
- AI models often overfit to training sets; predictive performance drops across unseen chemical classes.
- Generative designs may propose molecules that are chemically impractical to synthesize or patent-restricted.

- Many virtual hits fail to show biochemical activity; follow-up assays are essential.
- AI-driven predictions, while promising, still face hurdles in regulatory approval, particularly around safety and off-target profiling.

SCOPE AND FUTURE PERSPECTIVES

- Datasets from DOCKSTRING and SAMPL provide robust challenge



- frameworks for comparing docking and binding predictions
- Coupling MD with QM (e.g. OpenMM/PySCF) and nascent quantum computing approaches promise more accurate energy modelling
 - Platforms like Recursion OS and Pharma.AI pipelines enable massive data sharing and computation across institutions
 - XAI frameworks to clarify drug toxicity predictions, off-target effects, and ADMET outcomes will gain regulatory traction.
 - PBPK and ML-driven pharmacometric models tailored on patient-specific data could usher in patient-centric in silico drug prediction.

CONCLUSION

in silico drug design has advanced from supporting early screening to enabling AI-driven de novo design and comprehensive prediction of drug behaviour. Tools like AlphaFold 3, DiffDock, 3DMolFormer, and platforms including Schrödinger, Insilico Medicine, and AutoDock form a mature ecosystem that increasingly rivals traditional experimental workflows in speed and diversity. While limitations remain, especially around biological realism and experimental validation, hybrid pipelines combining computation and lab testing are becoming the new gold standard accelerating drug discovery while reducing cost and time.

REFERENCES

1. Li W, Yin Z, Li X, et al. A hybrid quantum computing pipeline for real-world drug discovery. *Sci Rep.* 2024;14:16942.

ARTICLE ID: BTS0626005

<https://doi.org/10.1038/s41598-024-67897-8>

2. Hu X, Liu G, Chen C, Zhao Y, Zhang H, Liu X. 3DMolFormer: A dual-channel framework for structure-based drug discovery. *arXiv.* 2025 Feb 7;arXiv:2502.05107.

3. Corso G, Stärk H, Jing B, Barzilay R, Jaakkola T. DiffDock: Diffusion steps, twists, and turns for molecular docking. *arXiv.* 2022 Oct 4;arXiv:2210.01776.

4. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature.* 2021;596(7873):583–589.

5. Zhavoronkov A, et al. Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nat Biotechnol.* 2019;37:1038–1040.

6. Muegge I, Mukherjee P. An overview of recent trends in molecular docking and scoring. *J Comput Aided Mol Des.* 2016;30(10):845–865.

7. Bajorath J. Integration of virtual and high-throughput screening. *Nat Rev Drug Discov.* 2015;14:319–320.

8. Schneider G. Automating drug discovery. *Nat Rev Drug Discov.* 2018;17(2):97–113.

9. Walters WP, Murcko MA. Assessing the impact of generative AI in medicinal chemistry. *Nat Biotechnol.* 2020;38(2):143–145.

10. Sanchez-Lengeling B, Aspuru-Guzik A. Inverse molecular design using machine learning: Generative



ARTICLE ID: BTS0626005

models for matter engineering. *Science*. 2018;361(6400):360–365.

11. Brown N, et al. GuacaMol: Benchmarking models for de novo molecular design. *J Chem Inf Model*. 2019;59(3):1096–1108.

12. Zang C, et al. MoleculeChef: A generative model for multi-objective molecular optimization. *Mach Learn Struct Biol*. 2021.

13. Yang K, et al. Analyzing learned molecular representations for property prediction. *J Chem Inf Model*. 2019;59(8):3370–3388.

14. Lipinski CA, et al. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev*. 2001;46(1–3):3–26.

15. Pires DEV, Blundell TL, Ascher DB. pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J Med Chem*. 2015;58(9):4066–4072.

16. Daina A, Michielin O, Zoete V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep*. 2017;7:42717.

17. Keiser MJ, et al. Predicting new molecular targets for known drugs. *Nature*. 2009;462(7270):175–181.

18. Ou-Yang SS, Lu JY, Kong XQ, et al. Computational drug discovery. *Acta Pharmacol Sin*. 2012;33:1131–1140.

19. Walters WP, Stahl MT, Murcko MA. Virtual screening—An overview. *Drug Discov Today*. 1998;3(4):160–178.

20. Grebner C, Matter H, Plowright AT. Automating drug design with predictive algorithms. *Drug Discov Today Technol*. 2020;35:19–28.