Inventum Biologicum: *An International Journal of Biological Research*
2026, Vol. 6, Issue 3, pp. 9-13

Inventum Biologicum

Journal homepage: www.journals.worldbiologica.com/ib

Review paper

Artificial Intelligence as a Catalyst for Structure Based Drug Design

Farooq Ahmad Mir [a] *

[a] *Associate Professor, Department of Chemistry, Higher Education, UT of J&K, India*

ARTICLE INFO

Keywords

artificial intelligence
machine learning
generative models
de novo molecular design
virtual screening

DOI
10.5281/ib-2520626

***Corresponding author**
Farooq Ahmad Mir
Email
mirfarooq219@gmail.com

ABSTRACT

Drug discovery remains one of the most expensive and time-consuming endeavors in modern science, with the development of a single approved therapeutic typically requiring more than a decade of research and several billion dollars in investment. Over the past five years, artificial intelligence (AI) and machine learning (ML) have moved from peripheral computational aids to central drivers of the medicinal chemistry pipeline, reshaping how targets are identified, how hits are generated, and how lead compounds are optimized. This review synthesizes recent advances in AI-enabled drug discovery, with particular emphasis on structure-based and generative molecular design. We examine the principal categories of AI methodology in current use, including deep neural networks for bioactivity prediction, graph-based generative models for de novo molecule design, and transformer architectures for retrosynthesis planning. We further discuss landmark case studies demonstrating translational impact, most notably the discovery of the antibiotic halicin through deep learning-guided screening, and describe how structure-based generative frameworks are being benchmarked against traditional docking approaches. Finally, we critically evaluate persistent challenges relating to data quality, model interpretability, and regulatory acceptance, and outline directions likely to define the next decade of AI-augmented medicinal chemistry.

1. Introduction

The translation of a biological insight into an approved medicine is a long and attrition-heavy process. Industry estimates place the average cost of bringing a new drug to market between two and five billion dollars, with development timelines commonly exceeding twelve years (Mozaffari et al., 2025). The majority of this cost is incurred not in the discovery phase itself but in late-stage clinical failure, much of which stems from inadequate efficacy, unforeseen toxicity, or poor pharmacokinetic behavior that could, in principle, have been anticipated earlier in the pipeline. It is precisely this inefficiency that artificial intelligence has been recruited to address.

Artificial intelligence refers broadly to computational systems capable of performing tasks that would otherwise require human cognition, while machine learning, a subset of AI, refers to algorithms that improve their performance through exposure to data rather than explicit programming (Khan et al., 2024). In the context of medicinal chemistry, AI and ML techniques are now applied across nearly every stage of the discovery pipeline: identifying and validating druggable targets, screening vast chemical libraries in silico, generating novel molecular scaffolds de novo, predicting synthetic routes, and forecasting toxicity and pharmacokinetic liabilities before a single compound is synthesized (Ferreira & Carneiro, 2025).

This review focuses specifically on the structure-based and generative dimensions of this transformation. We first outline the limitations of traditional discovery workflows, then introduce the major classes of AI methodology now in use, present illustrative case studies of translational success, and finally examine the barriers that continue to limit the field's full potential.

2. The Limitations of Traditional Drug Discovery

Classical medicinal chemistry relies heavily on iterative synthesis and testing: chemists design a candidate compound based on known structure-activity relationships, synthesize it, test it against a biological target, and use the result to inform the next design cycle. When protein structural information is available, structure based drug design allows this cycle to be guided by molecular docking and free energy calculations rather than intuition alone (Mozaffari et al., 2025). However, both approaches are fundamentally constrained by the size of chemical space that can be practically explored. The number of synthesizable drug-like molecules has been estimated to exceed 10^60, a figure that dwarfs the capacity of any laboratory to test experimentally, or even to enumerate computationally using traditional physics-based methods alone.

Compounding this combinatorial challenge is the high attrition rate of candidates that do progress to clinical testing. Failures rooted in poor absorption, unexpected off-target toxicity, or lack of translatability from preclinical models to human physiology account for the majority of costs in the drug development pipeline (Mozaffari et al., 2025). These structural inefficiencies have created strong incentive for computational methods capable of narrowing the search space earlier and with greater predictive fidelity than physics-based simulation alone can offer.

Table 1 Principal Categories of Artificial Intelligence Applied Across the Drug Discovery Pipeline

Pipeline Stage	Representative AI/ML Technique	Illustrative Application	Key Source
Target identification & validation	Graph neural networks; knowledge-graph mining	Prioritizing disease-associated proteins from multi-omics data	Mozaffari et al., 2025
Virtual screening	Deep convolutional networks; docking-score prediction	Ranking large compound libraries for likely bioactivity	Khan et al., 2024
De novo molecule design	Generative adversarial networks; variational autoencoders; diffusion models	Generating novel scaffolds conditioned on a target binding pocket	Harris et al., 2023
Retrosynthesis planning	Transformer sequence models	Proposing feasible synthetic routes for generated molecules	Ferreira & Carneiro, 2025
Toxicity & ADMET prediction	Ensemble and deep learning classifiers	Flagging hepatotoxicity or cardiotoxicity liabilities pre-synthesis	Khan et al., 2024
Antimicrobial discovery	Deep learning bioactivity classifiers	Screening molecular libraries for novel antibiotic activity	Stokes et al., 2020

3. Structure Based and Generative Molecular Design

Structure based drug design (SBDD) uses three-dimensional information about a target protein, typically derived from X-ray crystallography, cryo-electron microscopy, or increasingly from AI-based structure prediction, to guide the design of molecules that complement the target's binding site. The advent of highly accurate protein structure prediction has dramatically expanded the number of targets amenable to this approach, including many for which no experimental structure previously existed (Mozaffari et al., 2025).

Building on this structural foundation, generative models have emerged as a distinct and increasingly dominant paradigm. Rather than screening a pre-existing library, generative approaches such as variational autoencoders, generative adversarial networks, and more recently diffusion-based architectures learn the statistical distribution of drug-like molecules and sample novel structures directly, often conditioned on the geometry of a target binding pocket (Harris et al., 2023). This shifts the discovery task from a search problem to a generation problem, in principle allowing chemists to specify desired properties and receive candidate structures optimized against multiple objectives simultaneously, including predicted binding affinity, synthetic accessibility, and favorable pharmacokinetic profiles.

Benchmarking work comparing one, two, and three-dimensional generative approaches has shown that while 3D pocket conditioned generative models can produce chemically plausible poses, there remain substantial trade offs between predicted binding affinity, physical realism, and synthetic tractability, underscoring that generative fluency does not

automatically translate into experimentally validated potency (Harris et al., 2023). This finding has tempered some of the early enthusiasm surrounding purely generative pipelines and reinforced the value of hybrid workflows that combine generative proposal with physics-based or experimental validation.

3.1 Retrosynthesis and Synthetic Accessibility

A molecule generated in silico is only useful if it can be synthesized. Transformer-based sequence models, originally developed for natural language processing, have been adapted to predict feasible retrosynthetic routes by treating chemical reactions as a translation problem between product and precursor representations (Ferreira & Carneiro, 2025). Integrating retrosynthesis prediction directly into generative design loops allows synthetic accessibility to be scored alongside potency and selectivity, reducing the proportion of computationally attractive but practically unmakeable candidates that reach medicinal chemists.

4. Case Studies of Translational Impact

Perhaps the most widely cited demonstration of AI's translational potential in drug discovery is the identification of halicin, a compound originally investigated as a diabetes drug candidate, which was found through a deep learning classifier trained to predict antibacterial activity from molecular structure. The model was used to screen a library of over one hundred million molecules for structures predicted to inhibit the growth of Escherichia coli through a mechanism distinct from those of existing antibiotic classes, ultimately identifying a compound with activity against a broad range of drug-resistant pathogens in vitro and in mouse infection models (Stokes et al., 2020). This case is frequently invoked as proof of concept that AI-guided screening can surface structurally novel bioactive compounds that conventional screening approaches, biased toward known chemotypes, would be unlikely to prioritize.

Beyond antimicrobial discovery, AI-guided approaches are increasingly embedded in oncology and rare disease programs, where they are used to prioritize targets from complex multi-omic datasets, to model protein-protein interactions relevant to previously undruggable targets, and to support hit-to-lead optimization by predicting the impact of chemical modifications on binding affinity and metabolic stability before synthesis (Khan et al., 2024; Ferreira & Carneiro, 2025).

Table 2 Selected Examples Illustrating the Translational Impact of AI in Drug Discovery

Example	AI Approach	Outcome	Reference
Halicin	Deep learning bioactivity classifier screening ~107 compounds	Identification of a structurally novel antibiotic active against multidrug-resistant pathogens	Stokes et al., 2020
Pocket-conditioned generative benchmarking	Comparative evaluation of 1D/2D/3D generative architectures	Identified trade-offs between predicted affinity, physical realism, and synthesizability	Harris et al., 2023
Structural bioinformatics pipelines	AI-assisted structure prediction integrated with molecular dynamics	Enabled rational inhibitor design against thrombosis-related targets	Mozaffari et al., 2025
Cloud-based AI/ML pipeline review	Survey of deep learning, GNNs, and transformers across the discovery pipeline (2019-2024)	Synthesized comparative strengths/limitations for pharmaceutical implementation	Ferreira & Carneiro, 2025

5. Challenges and Limitations

Despite substantial progress, several structural challenges continue to temper the pace of adoption. First, model performance is fundamentally bounded by the quality, size, and representativeness of training data; many publicly available bioactivity datasets are biased toward well-studied target classes and contain limited negative data, which can inflate apparent predictive performance while masking poor generalization to genuinely novel chemotypes (Khan et al., 2024). Second, interpretability remains a persistent concern, particularly for deep neural architectures whose internal decision processes are difficult to relate back to established structure-activity principles, complicating both scientific validation and regulatory review (Ferreira & Carneiro, 2025). Third, the computational demands of training and deploying large generative and structure-prediction models can be substantial, and access to the requisite infrastructure is not evenly distributed across academic and industry groups, raising concerns about equitable participation in AI-driven discovery (Mozaffari et al., 2025).

Ethical and governance questions have also emerged alongside technical ones. Concerns regarding data privacy, the provenance of training datasets, and appropriate human oversight of AI-generated recommendations in a clinical development context all require continued attention as these tools move closer to regulatory and patient-facing applications (Khan et al., 2024).

6. Regulatory Considerations

Regulatory agencies have begun to issue horizon-scanning assessments of AI/ML applications across the medicines lifecycle, acknowledging both the potential efficiency gains and the need for robust validation standards before AI-derived evidence can be incorporated into formal regulatory submissions. Best-practice recommendations for model validation and reporting are emerging but remain inconsistently adopted across the field, suggesting that harmonization of standards will be an important determinant of how quickly AI-enabled discovery translates into approved medicines.

7. Future Directions

- Integration of multimodal data (genomic, structural, clinical, and real-world evidence) into unified predictive models spanning target identification through clinical trial design.
- Development of hybrid physics-informed generative architectures that combine the sampling efficiency of deep generative models with the mechanistic grounding of molecular dynamics and free-energy methods.
- Improved benchmarking standards, including prospective (rather than purely retrospective) validation of generative design pipelines against experimentally synthesized compounds.
- Greater emphasis on explainable AI methods that allow medicinal chemists to audit and trust model-generated design rationales.
- Expansion of AI-guided approaches into historically underserved therapeutic areas, including antimicrobial resistance and neglected tropical diseases, where commercial incentives for traditional discovery are limited.

8. Integrating AI into the Medicinal Chemist's Workflow

A recurring theme across recent reviews is that AI tools deliver the greatest value not as standalone replacements for medicinal chemistry judgment, but as components embedded within an iterative, human-supervised design cycle (Ferreira & Carneiro, 2025). In practice, this typically takes the form of a closed-loop process: a generative or virtual screening model proposes a prioritized set of candidate structures; a medicinal chemist reviews these proposals for synthetic tractability, intellectual property considerations, and consistency with established structure-activity relationships; a subset is synthesized and tested; and the resulting experimental data is fed back into the model to refine subsequent rounds of generation. This active-learning framework allows relatively small numbers of experimental data points to meaningfully improve model predictions over successive cycles, partially mitigating the data scarcity concerns that limit purely retrospective model training (Khan et al., 2024).

Organizationally, this has required pharmaceutical and biotechnology companies to build cross-functional teams that combine computational scientists, structural biologists, and medicinal chemists in far closer working proximity than was typical of earlier discovery paradigms, in which computational support was often consulted only after a chemical series had already been established. Academic groups have mirrored this trend, with several widely used benchmarking datasets and open-source modeling frameworks emerging specifically to lower the barrier for medicinal chemistry laboratories without dedicated machine learning infrastructure to adopt these tools (Harris et al., 2023).

Table 3 Comparative Considerations for Traditional versus AI-Augmented Discovery Workflows

Dimension	Traditional Workflow	AI-Augmented Workflow
Chemical space explored per cycle	Limited by synthetic and screening capacity (typically hundreds to low thousands of compounds)	Millions to billions of virtual candidates evaluated computationally before synthesis
Reliance on existing structural data	High; design typically anchored to known chemotypes and prior SAR	Can generate genuinely novel scaffolds conditioned on target structure alone
Speed of design-test-analyze cycle	Weeks to months per iteration	Computational proposal generated in hours to days; experimental testing remains rate-limiting
Principal bottleneck	Chemist bandwidth and screening throughput	Data quality, model interpretability, and experimental validation capacity

9. Conclusion

Artificial intelligence has transitioned from an auxiliary computational tool to a central pillar of contemporary medicinal chemistry, reshaping how targets are prioritized, how molecules are designed, and how synthetic and safety liabilities are anticipated. Landmark successes such as the discovery of halicin demonstrate that these approaches can yield genuinely novel, clinically relevant chemical matter, while ongoing benchmarking work highlights that substantial methodological maturation is still required before generative design can be considered a reliable substitute for experimental validation. The coming

decade is likely to be defined not by AI replacing medicinal chemists, but by increasingly tight integration of computational generation, physics-based validation, and experimental synthesis into a single iterative discovery loop, one in which the distance between molecule and medicine continues to shrink.

10. A Note on Reproducibility

A related and often underappreciated concern is reproducibility. Because many AI-driven discovery claims rely on proprietary datasets, internal computational infrastructure, or model weights that are not publicly released, independent verification of reported performance is frequently difficult. Community efforts to establish shared benchmarking datasets and standardized evaluation protocols, of the kind used to compare one, two, and three-dimensional generative architectures for structure based design, represent an important corrective, allowing claims of predictive or generative superiority to be assessed on a level playing field rather than taken at face value (Harris et al., 2023). Continued investment in open benchmarking infrastructure is likely to be as consequential to the field's long-term credibility as any single modeling advance.

References

1. Ferreira, F. J. N., & Carneiro, A. S. (2025). AI-driven drug discovery: A comprehensive review. ACS Omega. https://doi.org/10.1021/acsomega.5c00549
2. Harris, C., Didi, K., Jamasb, A. R., Joshi, C. K., Mathis, S. V., Lio, P., & Blundell, T. (2023). Beyond affinity: A benchmark of 1D, 2D, and 3D methods reveals critical trade-offs in structure-based drug design. arXiv. https://arxiv.org/abs/2308.07413
3. Khan, M. K., Raza, M., Shahbaz, M., Hussain, I., Khan, M. F., Xie, Z., Shah, S. S. A., Tareen, A. K., Bashir, Z., & Khan, K. (2024). The recent advances in the approach of artificial intelligence (AI) towards drug discovery. Frontiers in Chemistry, 12, Article 1408740. https://doi.org/10.3389/fchem.2024.1408740
4. Mozaffari, S., Moen, A., Ng, C. Y., Nicolaes, G. A. F., & Wichapong, K. (2025). Structural bioinformatics for rational drug design. Research and Practice in Thrombosis and Haemostasis. https://doi.org/10.1016/j.rpth.2025.102691
5. 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., & Collins, J. J. (2020). A deep learning approach to antibiotic discovery. Cell, 180(4), 688-702. https://doi.org/10.1016/j.cell.2020.01.021