



Unlocking New Frontiers: The Role of Generative AI in Drug Discovery from Natural Products – Applications, Challenges, and Future Directions

Sweta Koka

Acropolis Institute of Pharmaceutical Education and Research, Indore, Madhya Pradesh, India

Article info

Received: 11/10/2025

Revised: 10/11/2025

Accepted: 21/12/2025

© IJPLS

www.ijplsjournal.com

Abstract

Natural products derived from plants, fungi, microorganisms, and other biological sources represent a valuable reservoir of structurally diverse and biologically active compounds for pharmaceutical discovery. However, conventional natural-product drug discovery is often time-consuming, resource-intensive, and challenged by complex chemical structures, limited data availability, and difficulties in compound isolation, optimization, and biological characterization. Generative artificial intelligence (GenAI) is emerging as a transformative approach that can complement conventional drug-discovery strategies by learning patterns from chemical and biological datasets and generating novel molecular structures inspired by natural products. This review explores the emerging applications, limitations, and future perspectives of GenAI in natural-product-based drug discovery.

Major applications include AI-driven compound generation using approaches such as variational autoencoders, generative adversarial networks, recurrent neural networks, transformers, diffusion models, and reinforcement learning; predictive modeling of biological activity; integration of genomics, proteomics, and metabolomics data; virtual screening; target identification; and lead optimization. Overall, the convergence of natural-product chemical diversity and generative AI offers a promising pathway to accelerate the discovery and development of innovative therapeutic candidates while complementing established experimental approaches.

Keywords: Generative artificial intelligence; Natural products; Drug discovery; Molecular generation; Virtual screening; Machine learning; Bioactivity prediction; Lead optimization; Multi-omics; Natural-product-inspired therapeutics

Introduction

As the pharmaceutical landscape continues to evolve, the integration of advanced technologies has become crucial in addressing the pressing challenges of drug discovery. Among these innovations, generative artificial intelligence (AI) has emerged as a transformative force, revolutionizing how researchers identify and develop new therapeutic compounds. Generative AI refers to a class of algorithms that can learn from existing data to generate novel outputs, such as chemical structures or biological sequences. This capability is particularly relevant to drug

discovery, where the need for innovative approaches is paramount to overcoming the limitations of traditional methodologies. [1, 2,3,13]

Natural products have long been a cornerstone of pharmaceutical development, serving as the basis for many essential medications.

*Corresponding Author

E.mail: sweta.koka@gmail.com

These biologically active compounds, derived from plants, fungi, and microorganisms, offer a rich source of chemical diversity that can be harnessed for therapeutic purposes. However, the process of isolating and optimizing these compounds into viable drugs is often lengthy and complex. In this context, generative AI presents a unique opportunity to streamline and enhance the discovery process, tapping into the vast potential of natural products while accelerating the identification of promising drug candidates.^[1,2]

Innovation in drug discovery is critical for addressing unmet medical needs and tackling emerging health challenges. With the global burden of diseases continually evolving, researchers are under pressure to develop new therapies more efficiently and effectively. Applying generative AI in this domain not only promises to expedite the discovery process but also aims to reduce the costs of drug development. As we delve deeper into the emerging applications, limitations, and future perspectives of generative AI in drug discovery, it becomes evident that this technology holds the potential to redefine the pharmaceutical landscape, paving the way for novel therapeutic solutions that could benefit patients worldwide.^[2,3,13]

Emerging Applications of Generative AI

The integration of generative AI into drug discovery is revolutionizing the way researchers approach the identification and development of new pharmaceuticals, particularly those derived from natural products. This section will explore the various emerging applications of generative AI in this domain, focusing on AI-driven compound generation, predictive modeling for biological activity, and enhancements in virtual screening processes.^[2,3,4]

AI-driven Compound Generation from Natural Products

Techniques Used in Compound Generation

Generative AI employs various techniques to design novel compounds inspired by natural products. These techniques include deep learning models, such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs), which are capable of learning complex patterns from chemical data. Additionally, generative models like variational autoencoders (VAEs) and

generative adversarial networks (GANs) are increasingly utilized to create new molecular structures that mimic the chemical properties of existing natural compounds. These models can explore vast chemical spaces and suggest unique molecular architectures that may not be readily apparent through traditional design approaches.^[4,5,6,7,8,9]

Case Studies Showcasing Successful Applications

Numerous case studies highlight the successful application of generative AI in developing drug candidates from natural products. For instance, researchers have utilized GANs to generate potential anti-cancer compounds derived from the chemical structures of known natural products, significantly accelerating the discovery process. Another notable case involved the application of VAEs to identify novel antibiotic candidates by mining natural product libraries, leading to the discovery of compounds with enhanced efficacy against resistant bacterial strains. These examples illustrate the potential of generative AI to not only replicate existing natural products but also innovate upon them.^[2,3,4,10]

Impact on Lead Optimization Processes

The incorporation of generative AI into lead optimization processes has profound implications for drug development. By predicting the pharmacokinetic properties and toxicity profiles of newly generated compounds, AI models can streamline the selection of candidates for preclinical testing. This capability reduces the time and costs associated with drug development, allowing researchers to focus their efforts on the most promising candidates. Moreover, generative AI can facilitate a more iterative approach to optimization, enabling continuous refinement of compound structures based on real-time feedback from biological assays.^[7,8,13]

Predictive Modeling for Biological Activity Machine Learning Algorithms in Predicting Efficacy

Predictive modeling plays a crucial role in assessing the biological activity of compounds generated through AI. Machine learning algorithms, including support vector machines (SVMs) and random forests, are employed to

analyze chemical and biological data, allowing researchers to predict the efficacy of new compounds with higher accuracy. By training these models on extensive datasets that include both chemical properties and biological outcomes, researchers can better understand the relationships between molecular structure and biological activity.^[13]

Integrating Omics Data for Better Predictions

The integration of omics data—such as genomics, proteomics, and metabolomics—enhances the predictive capabilities of generative AI in drug discovery. By leveraging multi-omics approaches, researchers can gain deeper insights into the mechanisms of action of compounds and their interactions within biological systems. This holistic view not only improves predictive accuracy but also aids in identifying potential off-target effects and optimizing therapeutic windows.^[2,3,18]

Examples of Successful Predictions

Several successful predictions demonstrate the effectiveness of generative AI in anticipating the biological activity of new compounds. For instance, researchers have successfully predicted the anti-inflammatory properties of novel compounds derived from natural products by utilizing machine learning models trained on historical data. These models have accurately identified lead candidates that exhibit desired biological activity, thereby expediting the drug discovery process.^[2,13]

Enhancing Virtual Screening Processes

Role of Generative Adversarial Networks

Generative adversarial networks (GANs) are transforming virtual screening processes by generating novel compound libraries that can be screened for biological activity. By simulating the chemical space of natural products, GANs can create diverse and innovative compound libraries that enhance the chances of discovering new therapeutics. This approach allows for a more efficient exploration of chemical diversity compared to traditional high-throughput screening methods.^[4,8,9]

Comparison to Traditional Screening Methods

Traditional methods of virtual screening often rely on rigid docking simulations and predefined compound libraries, which can limit the discovery of novel candidates. In contrast, generative AI techniques enable the creation of compounds that may not exist in known databases, thus broadening the scope of potential drug candidates. This flexibility significantly increases the likelihood of identifying unique structures that exhibit desirable biological properties.^[3,4,13]

Future Trends in Virtual Screening Enhancements

As generative AI continues to evolve, future trends in virtual screening will likely include the integration of real-time data feedback loops that refine compound generation based on screening outcomes. Furthermore, advancements in AI algorithms will enhance the predictive power of virtual screening, enabling more precise assessments of compound efficacy and safety profiles. This evolution will not only speed up the drug discovery process but also lead to more successful therapeutic discoveries from natural products.^[8,9,13]

The emerging applications of generative AI in drug discovery from natural products are paving the way for innovative approaches to compound generation, predictive modeling, and virtual screening. These advancements hold great promise for transforming the landscape of pharmaceutical research and development, offering new opportunities to address unmet medical needs.^[2,3,4]

Limitations of Generative AI in Drug Discovery

While the promise of generative AI in drug discovery from natural products is substantial, several limitations hinder its full realization and implementation in the pharmaceutical industry. Understanding these constraints is essential for navigating the challenges ahead and optimizing the use of AI technologies in this vital domain.^[2,13,14]

Data Limitations and Quality Concerns

One of the primary limitations of generative AI in drug discovery is the reliance on high-quality data

for training models. The efficacy of machine learning algorithms, particularly those used in generative AI, is heavily contingent upon the availability of robust datasets. However, in the context of natural products, data can be sparse, inconsistent, or poorly annotated. Moreover, natural product libraries often contain compounds with complex structures and varied biological activities, further complicating data acquisition and quality. Inadequate or biased datasets can lead to suboptimal model training, resulting in flawed predictions and compound generations that may not translate effectively into viable drug candidates. ^[2,11,12,13]

Challenges in Interpretability of AI Models

Another significant hurdle is the interpretability of AI models. Many generative AI techniques, especially those utilizing deep learning frameworks, function as "black boxes," making it challenging for researchers to understand how specific predictions are made. This lack of transparency can be problematic in drug discovery, where understanding the rationale behind a candidate's biological activity or potential side effects is crucial. The inability to interpret AI-generated results can lead to skepticism among researchers and regulatory bodies, hindering the adoption of these advanced techniques in clinical settings. ^[13,14]

Regulatory Hurdles for AI-Driven Drug Discovery

The regulatory landscape presents additional challenges for the integration of generative AI in drug discovery. The regulatory landscape for AI in drug development is evolving, with agencies such as the FDA and EMA issuing discussion papers, reflection papers, and related principles addressing the use of AI in medicinal-product development. This uncertainty can create barriers for pharmaceutical companies looking to implement AI-driven methodologies, as they may face difficulties in meeting compliance standards and securing necessary approvals. Moreover, as AI technologies continue to evolve rapidly, regulators must balance the need for innovation with the imperative to ensure safety and efficacy in drug products. This dynamic can create an

environment of hesitation, stalling the progress of AI applications in drug discovery. ^[15,16]

In summary, while generative AI offers transformative potential for drug discovery from natural products, significant limitations related to data quality, model interpretability, and regulatory frameworks must be addressed. Overcoming these challenges will be crucial for leveraging the full capabilities of AI and ensuring that it can effectively contribute to the advancement of pharmaceutical innovation. ^[2,13,15,16]

Future Perspectives of Generative AI in Drug Discovery

As the field of drug discovery continues to evolve, the integration of generative AI holds significant promise for reshaping methodologies and outcomes. The coming years are poised to witness a confluence of advancements in AI technology and a deeper understanding of biological systems, presenting numerous opportunities for innovation in drug discovery. ^[2,4,9,13]

Potential Advancements in AI Technology

The future of generative AI in drug discovery is likely to be characterized by rapid advancements in algorithms and computational power. Enhanced machine learning techniques, such as deep learning and reinforcement learning, are expected to improve the efficiency and accuracy of compound generation and predictive modeling. Moreover, the integration of diverse data types—including genomic, proteomic, and metabolomic data—can facilitate more holistic approaches to understanding drug interactions and efficacy. As AI systems become increasingly sophisticated, they will be able to simulate complex biological processes, providing researchers with invaluable insights that were previously unattainable. ^[4,7,9,18]

Collaborations Between AI and Traditional Drug Discovery Methods

The future landscape of drug discovery will likely see a synergistic relationship between generative AI and traditional methodologies. By harnessing the strengths of both approaches, pharmaceutical companies can leverage AI's capacity for rapid data analysis and predictive modeling alongside established experimental techniques. This collaborative framework will enhance the drug

discovery pipeline, enabling researchers to validate AI-generated compounds more efficiently and reduce the time and costs associated with bringing new therapeutics to market. Furthermore, interdisciplinary teams comprising AI experts, pharmacologists, and biologists will be essential in fostering innovation and ensuring that AI applications are grounded in biological relevance. [2,13,15,16]

Vision for the Next Decade in Drug Discovery

In envisioning the next decade of drug discovery, one can anticipate a landscape where generative AI is integral to every stage of the process—from initial compound identification to clinical trials. AI-driven platforms may facilitate personalized medicine by predicting patient responses based on genetic and phenotypic data, leading to tailored treatment options that improve patient outcomes. Additionally, the continual refinement of AI models will likely enhance their interpretability, allowing researchers to better understand the underlying mechanisms driving drug interactions. [2,4,18]

Moreover, as ethical considerations surrounding AI usage in healthcare grow, there will be a concerted effort to establish guidelines and best practices that ensure transparency and accountability in AI-driven drug discovery. This evolution will not only bolster public trust in AI applications but also pave the way for regulatory bodies to adapt to the changing technological landscape. [14,15,16]

By merging innovative computational techniques with biological insights and fostering collaborations between disciplines, the pharmaceutical industry can address long-standing challenges and usher in a new era of drug development that is more efficient, effective, and responsive to patient needs. As we look ahead, the role of generative AI in shaping the future of medicine cannot be overstated, paving the way for breakthroughs that could significantly enhance global health outcomes. [2,13,18]

Conclusion

In summary, the integration of generative AI into the realm of drug discovery, particularly when harnessed with the rich diversity of natural products, presents an unprecedented opportunity

for innovation in pharmaceuticals. As we have discussed, generative AI techniques provide novel pathways for compound generation, predictive modeling, and enhancements to virtual screening processes, significantly improving the efficiency and efficacy of drug development. [1,2,3]

The case studies highlighted demonstrate the transformative potential of these technologies, showcasing successful applications that not only streamline lead optimization but also pave the way for new therapeutic avenues. However, the journey is not without its challenges. Data limitations, interpretability issues, and regulatory hurdles present significant barriers that must be addressed to ensure the safe and effective implementation of AI-driven methodologies in drug discovery. [4,7,8,10]

Looking forward, the future of generative AI in this field appears promising. Advancements in AI technology, coupled with collaborative efforts between traditional drug discovery practices and emerging computational methods, could redefine the landscape of pharmaceutical development over the next decade. As we anticipate these changes, it is crucial for stakeholders to engage in dialogue about ethical considerations, data governance, and regulatory frameworks that will support the responsible use of generative AI. [2,9,13,18]

As we continue to explore its applications and address its limitations, there is a palpable sense of excitement about the future possibilities that lie ahead. The intersection of artificial intelligence and drug discovery holds the promise of not only accelerating the development of new therapies but also enhancing our understanding of complex biological systems, ultimately leading to improved health outcomes worldwide. [14,15,16]

References

1. Newman DJ, Cragg GM. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J Nat Prod.* 2020;83(3):770-803.
2. Mullowney MW, Duncan KR, Elsayed SS, et al. Artificial intelligence for natural product drug discovery. *Nat Rev Drug Discov.* 2023;22(11):895-916.
3. Saldívar-González FI, Aldas-Bulos VD, Medina-Franco JL, Plisson F. Natural product drug discovery in the artificial

- intelligence era. Chem Sci. 2022;13(6):1526-1546.
4. Tong X, Liu X, Tan X, et al. Generative models for de novo drug design. J Med Chem. 2021;64(19):14011-14027.
 5. Gómez-Bombarelli R, Wei JN, Duvenaud D, et al. Automatic chemical design using a data-driven continuous representation of molecules. ACS Cent Sci. 2018;4(2):268-276.
 6. Segler MHS, Kogej T, Tyrchan C, Waller MP. Generating focused molecule libraries for drug discovery with recurrent neural networks. ACS Cent Sci. 2018;4(1):120-131.
 7. Olivecrona M, Blaschke T, Engkvist O, Chen H. Molecular de-novo design through deep reinforcement learning. J Cheminform. 2017;9:48.
 8. Blaschke T, Arús-Pous J, Chen H, et al. REINVENT 2.0: an AI tool for de novo drug design. J Chem Inf Model. 2020;60(12):5918-5922.
 9. Alakhdar A, Poczos B, Washburn N. Diffusion models in de novo drug design. J Chem Inf Model. 2024;64(19):7238-7256.
 10. Stokes JM, Yang K, Swanson K, et al. A deep learning approach to antibiotic discovery. Cell. 2020;180(4):688-702.e13.
 11. Zeng X, Zhang P, He W, et al. NPASS: natural product activity and species source database for natural product research, discovery and tool development. Nucleic Acids Res. 2018;46(D1):D1217-D1222.
 12. Zhao H, Yang Y, Wang S, et al. NPASS database update 2023: quantitative natural product activity and species source database for biomedical research. Nucleic Acids Res. 2023;51(D1):D621-D628.
 13. 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.
 14. Jiménez-Luna J, Grisoni F, Schneider G. Drug discovery with explainable artificial intelligence. Nat Mach Intell. 2020;2:573-584.
 15. European Medicines Agency. Reflection paper on the use of artificial intelligence (AI) in the medicinal product lifecycle. EMA/CHMP/CVMP/83833/2023. Amsterdam: EMA; 2024.
 16. U.S. Food and Drug Administration. Using Artificial Intelligence & Machine Learning in the Development of Drug & Biological Products: Discussion Paper and Request for Feedback. Silver Spring (MD): FDA; 2023. Revised February 2025.
 17. Schwaller P, Laino T, Gaudin T, et al. Molecular Transformer: a model for uncertainty-calibrated chemical reaction prediction. ACS Cent Sci. 2019;5(9):1572-1583.
 18. Pun FW, Podolskiy D, Izumchenko E, et al. Target identification and assessment in the era of AI. Nat Rev Drug Discov. 2026;25:534-552.

Cite this article as:

Koka S. (2025). Unlocking New Frontiers: The Role of Generative AI in Drug Discovery from Natural Products – Applications, Challenges, and Future Directions. *Int. J. of Pharm. & Life Sci.*, 16(12):36-41.

Source of Support: Ni

Conflict of Interest: Not declared

For reprints contact: ijplsjournal@gmail.com