EUROPEAN JOURNAL OF BIOMEDICAL AND PHARMACEUTICAL SCIENCES
http://www.ejbps.com

ISSN 2349-8870
Coden USA: EJBPBO
Volume: 13
Issue: 3
Year: 2026

ARTIFICIAL INTELLIGENCE IN PREFORMULATION AND PHYSICOCHEMICAL PROPERTY PREDICTION: A COMPREHENSIVE REVIEW

Pasam Jyothirmayi[1*], Abbineni Anusha[2], Dr. N. Srinivasa Rao[3], M. Leela satyavathi[4], D. Neha Sri[4], L. Vasavi Bhavani[4], S. K. Asha Begam[4], G. Nandhini[4]

[1,2]Associate Professor, Department of Pharmaceutics, Vikas College of Pharmacy, Affiliated To JNTUK, Vissannapeta, NTR Dist.
[3]Professor and Principal, Vikas College of Pharmacy, Affiliated to JNTUK, Vissannapeta, NTR Dist.
[4]Department of Pharmaceutics, Vikas College of Pharmacy, Affiliated to JNTUK, Vissannapeta, NTR Dist.

***Corresponding Author: Pasam Jyothirmayi**

Associate Professor, Department of Pharmaceutics, Vikas College of Pharmacy, Affiliated to JNTUK, Vissannapeta, NTR Dist.

DOI: https://doi.org/10.5281/zenodo.18795985

How to cite this Article: Pasam Jyothirmayi[1*], Abbineni Anusha [2], Dr. N. Srinivasa Rao[3], M. Leela satyavathi[4], D. Neha Sri[4], L. Vasavi Bhavani[4], S. K. Asha Begam[4], G. Nandhini[4] (2026). Artificial Intelligence In Preformulation And Physicochemical Property Prediction: A Comprehensive Review. European Journal of Biomedical and Pharmaceutical Sciences, 13(3), 36–41.

Article Received on 26/01/2026 **Article Revised on 16/02/2026** **Article Published on 01/03/2026**

ABSTRACT

Artificial intelligence (AI) has gained significant attention in pharmaceutical sciences due to its ability to analyze complex datasets and generate reliable predictions. In preformulation studies, understanding the physicochemical properties of active pharmaceutical ingredients (APIs) is crucial for the development of safe, effective, and stable dosage forms. Conventional experimental approaches, although well established, are often time-consuming and resource intensive. AI-based techniques such as machine learning, deep learning, and quantitative structure–property relationship (QSPR) modeling provide efficient alternatives by enabling early prediction of solubility, stability, permeability, polymorphism, pKa, and lipophilicity. This review elaborates on the principles of AI, its applications in preformulation, commonly used algorithms, data sources, advantages, limitations, regulatory considerations, and future prospects, with emphasis on its relevance in modern drug development.

KEYWORDS: Artificial intelligence; Preformulation; Physicochemical property prediction; Machine learning; Drug development.

1. INTRODUCTION

Preformulation studies form the backbone of pharmaceutical product development. These studies involve the systematic investigation of physical and chemical characteristics of drug substances, including solubility, stability, polymorphism, hygroscopicity, particle size, and drug–excipient compatibility. The information generated during preformulation is essential for selecting suitable dosage forms and optimizing formulation strategies. AI revolutionized the formulations and development of modern pharmaceuticals. With the help of AI, researches can now optimize drug design, develop formulation, and streamline clinical drugs in a much accurate and efficient way.[1]

2. Artificial Intelligence in Pharmaceutical Sciences

Artificial intelligence (AI) refers to a collection of advanced computational techniques that enable computer systems to perform tasks traditionally requiring human intelligence, such as learning from data, pattern recognition, reasoning, and predictive decision-making. In pharmaceutical sciences, AI has gained considerable importance due to its ability to process large and complex datasets and to generate reliable predictions that support various stages of drug discovery and development.[2]

In pharmaceutical research, AI is primarily implemented through machine learning (ML) and deep learning (DL) approaches. Machine learning algorithms learn relationships between molecular descriptors and experimentally measured properties by analyzing historical datasets. These descriptors may include

physicochemical parameters, structural fingerprints, and formulation-related attributes.[3] Deep learning, a more advanced subset of machine learning, employs multilayered neural networks capable of capturing highly non-linear relationships, making it particularly suitable for modeling complex drug behaviors.

The application of AI in pharmaceutical sciences extends across multiple domains, including drug discovery, preformulation studies, formulation optimization, quality control, and regulatory decision support. In the context of preformulation, AI models analyze molecular and experimental data to predict drug behavior such as solubility, stability, permeability, and solid-state properties.[4] By enabling early-stage prediction and informed decision-making, AI reduces reliance on extensive laboratory experimentation and accelerates the overall drug development process.

Figure 1: AI in preformulation.

3. Role AI in preformulation studies

AI tools have significantly accelerated pharmaceutical research. For example, **Adaptive Neuro- Fuzzy Inference Systems (ANFIS)** facilitate excipient selection, simplifying research and shortening drug development timelines. Additionally, in silico models are for predicting aqueous solubility by analyzing factors such as molecular size, shape, and hydrogen-bonding potential.[5]

Preformulation studies for nano based therapeutics

Nano-based therapeutics offer an effective formulation strategy for delivering active drug ingredients dur to their well-controlled morphological design and unique physicochemical properties.[6] Novel drug delivery systems aim to enhance bioavailability, stability, and therapeutic efficacy.

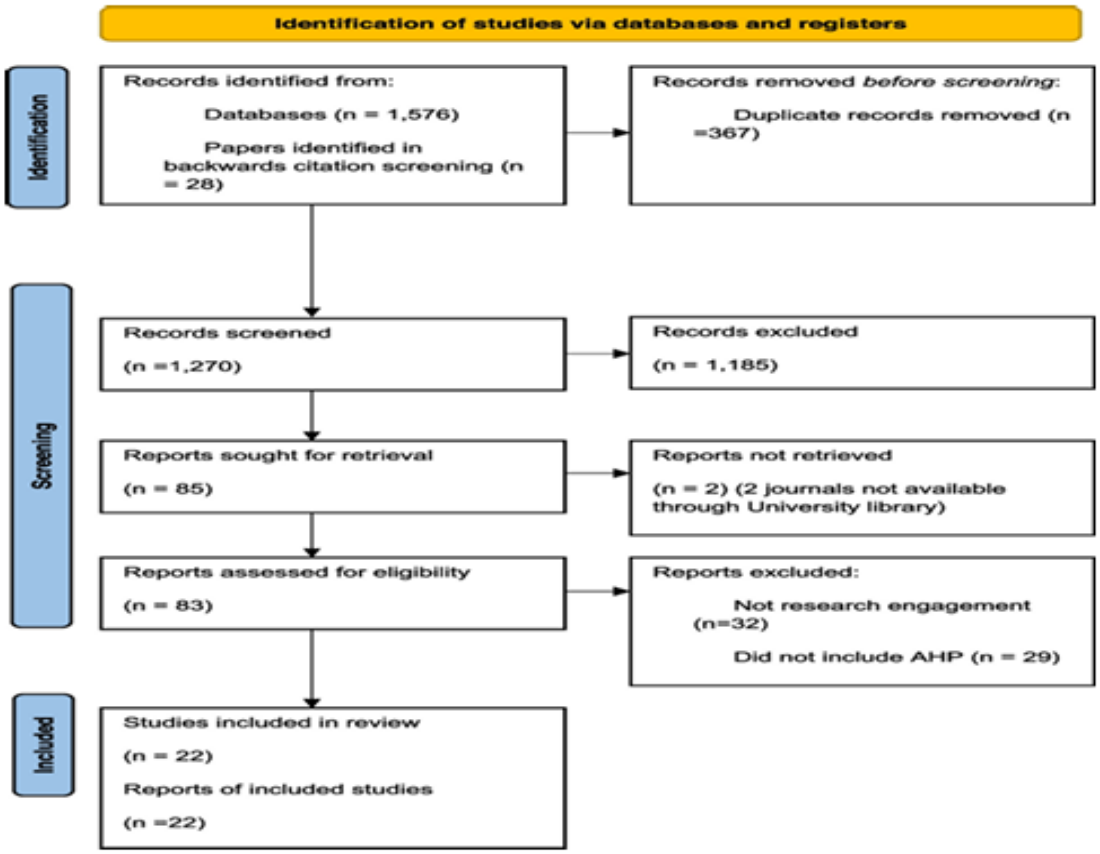

Figure 2: Preformulation studies for nano based therapeutics.

4. Importance of AI in Preformulation Studies

AI supports early-stage decision-making by predicting potential formulation challenges and reducing dependence on extensive experimental screening. This leads to improved efficiency, reduced development time, and cost savings.

Preformulation studies play a crucial role in pharmaceutical development by providing essential information on the physicochemical characteristics of active pharmaceutical ingredients (APIs). These studies guide the selection of suitable dosage forms, excipients, and manufacturing processes. However, conventional preformulation approaches often involve extensive experimental screening, which is time-consuming, costly, and resource-intensive. The integration of artificial intelligence (AI) into preformulation research has significantly improved the efficiency and effectiveness of this critical stage.[7]

AI enables early-stage prediction of key physicochemical properties such as solubility, stability, permeability, polymorphism, and ionization behavior. By analyzing molecular descriptors and experimental datasets, AI-based models can identify potential formulation challenges before laboratory testing is initiated.[8] This predictive capability allows formulation scientists to prioritize promising drug candidates and eliminate unsuitable compounds at an early stage, thereby reducing the risk of late-stage formulation failure.

Another important contribution of AI in preformulation is its ability to support rational decision-making in excipient selection and solid-state optimization. AI models can evaluate drug–excipient compatibility, predict polymorphic transitions, and assess stability under various environmental conditions. This helps in designing robust formulations with improved stability and bioavailability.[9] Furthermore, AI facilitates optimization of formulation strategies such as salt selection, particle size modification, and solubility enhancement techniques.

Overall, the application of AI in preformulation studies enhances development efficiency by reducing experimental workload, development time, and cost. By combining AI-based predictions with experimental validation, pharmaceutical scientists can achieve a more systematic and data-driven approach to formulation development, leading to improved product quality and accelerated drug development timelines.[10]

5. AI Techniques Used in Preformulation

Machine learning algorithms such as random forest, support vector machines, k-nearest neighbors, and artificial neural networks are widely used. Deep learning approaches, including graph neural networks, enable

improved prediction accuracy for complex molecular systems.

Artificial intelligence techniques employed in preformulation studies primarily include machine learning and deep learning approaches, which enable efficient analysis of complex pharmaceutical datasets.[11] These techniques are used to establish relationships between molecular structure, physicochemical properties, and formulation performance, thereby supporting rational formulation design.

Machine learning algorithms such as **random forest (RF)**, **support vector machines (SVM)**, **k-nearest neighbors (k-NN)**, and **artificial neural networks (ANNs)** are widely applied in preformulation research. Random forest models are particularly effective in handling large datasets and identifying non-linear relationships between molecular descriptors and formulation properties. Support vector machines are commonly used for classification and regression tasks, such as predicting solubility class or permeability behavior.[12] The k-nearest neighbors approach predicts drug properties based on structural similarity, while artificial neural networks simulate biological neural systems to model complex and non-linear relationships.

In recent years, **deep learning techniques** have gained increasing importance due to their superior predictive performance. Advanced architectures such as **graph neural networks (GNNs)** represent drug molecules as graphs, where atoms are treated as nodes and bonds as edges. This allows direct learning from molecular structure without the need for extensive manual feature extraction. Deep learning models are particularly useful for predicting complex properties such as polymorphism, solid-state behavior, and multi-parameter physicochemical interactions.[13]

Overall, the application of these AI techniques in preformulation enhances predictive accuracy, reduces experimental dependency, and enables a more systematic and data-driven approach to formulation development. When combined with experimental validation, these models significantly improve the efficiency and reliability of preformulation studies.[14]

6. Prediction of Physicochemical Properties

AI is extensively applied to predict solubility, stability, permeability, polymorphism, pKa, and partition coefficient, which are critical determinants of drug performance and formulation success.

Prediction of physicochemical properties is a central objective of preformulation studies, as these properties directly influence drug stability, bioavailability, manufacturability, and therapeutic performance. Artificial intelligence (AI) has been widely adopted for the accurate and early prediction of critical physicochemical parameters, thereby supporting rational formulation design and reducing reliance on extensive experimental testing.[15]

AI-based models are extensively used to predict **aqueous solubility**, which is a key determinant of oral bioavailability. By analyzing molecular descriptors such as lipophilicity, molecular weight, hydrogen bonding capacity, and crystal lattice energy, AI algorithms can identify solubility limitations at an early stage.[16] This information assists in selecting appropriate solubility enhancement strategies, including salt formation, solid dispersions, and particle size reduction.

Stability prediction using AI focuses on both chemical and physical stability of drug substances. Machine learning and deep learning models evaluate degradation pathways and assess the influence of environmental factors such as temperature, humidity, light, and pH. These predictions help in optimizing formulation composition and selecting suitable storage and packaging conditions to ensure product quality throughout shelf life.[17]

AI is also applied to predict **permeability and absorption characteristics**, which are essential for understanding drug transport across biological membranes. By correlating molecular size, polarity, and ionization behavior with permeability data, AI models support biopharmaceutical classification and dosage form selection. In addition, AI-based prediction of **polymorphism and solid-state properties** assists in identifying potential crystal forms and polymorphic transitions, which can significantly affect solubility, stability, and manufacturability.

Furthermore, AI enables accurate prediction of **pKa and partition coefficient (log P)**, providing insight into the ionization and lipophilicity of drug molecules. These parameters are crucial for predicting drug distribution, membrane permeability, and formulation behavior. Overall, the application of AI in physicochemical property prediction enhances preformulation efficiency, improves formulation success rates, and accelerates the drug development process.

Preformulation Objectives

The pre-formulation tool can be viewed as a decisive tool for making decisions during the drug discovering and development phases. Through knowledge of physicochemical characteristics and how they Effect biological performances enable the choice of possible lead molecules and related drug molecules. It is important to thoroughly understand the physicochemical properties of new drug.[18]

Advantages of AI-Based Preformulation

- Provides rapid and cost-effective prediction of drug physicochemical and biopharmaceutical properties
- Reduces experimental workload by minimizing extensive trial-and-error studies

- Enables early identification of formulation-related risks such as instability and poor solubility
- Supports informed decision-making at early stages of drug development
- Enhances overall formulation success rates and development efficiency

Limitations and Challenges of AI-Based Preformulation

- Strong dependence on the availability of high-quality and reliable datasets
- Limited interpretability of complex AI models, making decision-making less transparent
- Risk of bias due to incomplete, imbalanced, or non-representative data
- Requirement for experimental validation to confirm AI-predicted results
- Need for skilled personnel and computational resources

Application of AI

- The AI system uses physicochemical properties such as solubility, pKa, log P, melting point, and hygroscopicity, which helps in understanding drug behavior before formulation.
- AI predicts physicochemical properties such as solubility, pKa, log P, melting point, and hygroscopicity, which helps in understanding drug behavior before formulation.
- AI helps in predicting stability and degradation pathways of drugs and estimates shelf life under different conditions like temperature, humidity, light, and pH.ai assists in identifying and selecting stable polymorphic and crystal forms of drugs to ensure consistent solubility and bioavailability.[19]

Table 1: Applications of AI in Preformulation.

AI Method	Application Area	Predicted Property
Random Forest	Solubility modeling	Aqueous solubility
SVM	Permeability prediction	Intestinal absorption
ANN	Stability analysis	Degradation tendency
QSPR	Structure–property correlation	pKa, log P

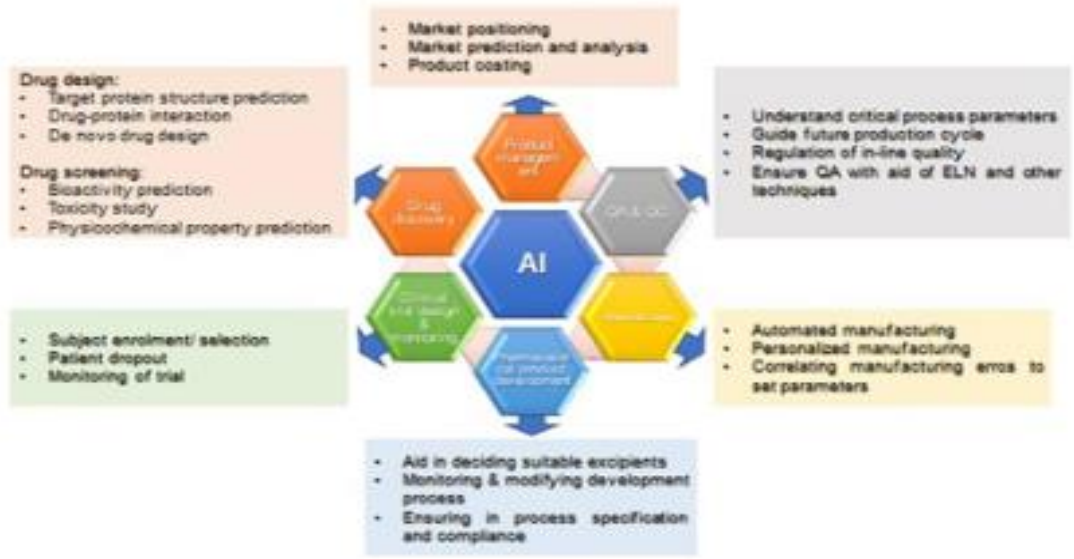

Figure 3: Applications of AI in preformulation.

8. Regulatory Considerations

Regulatory acceptance of AI-based approaches in pharmaceutical development requires a high level of transparency in model design and decision-making processes. Ensuring data integrity, reliability, and traceability is essential to meet regulatory expectations.[20] AI-generated predictions must be properly validated through experimental studies before regulatory submission. In recent years, regulatory agencies have increasingly acknowledged the potential of artificial intelligence to improve efficiency and decision-making in drug development, provided that these systems comply with established quality and validation standards.

10. CONCLUSION

Artificial intelligence has emerged as a powerful tool in preformulation and physicochemical property prediction. Its integration with experimental methods enhances efficiency and innovation in drug development.

REFERENCES

1. 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.
2. Muratov EN, Bajorath J, Sheridan RP, et al. QSAR modeling: achievements and challenges. J Med Chem., 2020; 63: 4047–4069.
3. Tetko IV, Ostermann C, et al. Prediction of aqueous solubility of drug-like compounds. J Chem Inf Model, 2021; 61: 337–346.
4. Paul D, Sanap G, Shenoy S, et al. Artificial intelligence in drug discovery and development. Drug Discov Today, 2021; 26: 80–93.
5. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. N Engl J Med., 2022; 386: 1351–1361.
6. Chen H, Engkvist O, Wang Y, et al. Deep learning for drug property prediction. Brief Bioinform, 2023; 24: bbad012.
7. Gupta R, Badhe Y, et al. Role of artificial intelligence in pharmaceutical formulation development. Pharmaceutics, 2024; 16: 112.
8. Kailash Ahirwar and Rahul Shukla; Pre-formulations studies: A versatile tool in formulation design submitted: October 20 2022, Reviewed: 02 February 2023, published: 23 Feb 2023.
9. Surojit Banerjee, Debadri Banerjee, Deepti Rajput, Vikas Anand; AI and its applications in drug discovery, formulations, development; 31 May 2022.
10. Jun-hung Jiang: Division of molecular pharmaceutics and drug delivery, college of pharmacy, the university of Texas at Austin.
11. Defang Ouyang: State key laboratory of quality research in Chinese medicines, institute of Chinese medical sciences, 2022 Oct 22; 14(11): 2257.
12. Amrendra; Received: 17 July 2025 Accepted: 24 July 2025; Published: 31 July 2025.
13. Anupama Singh, Vikas A. Saharan, Rania; AI in drug formulations and delivery: Benefits, trends, and future Perspectives. (2025)
14. Prediction of physicochemical properties based on neutral network modelling https: //www.sciencedirect.com/science/article/abs/pii/S0169409X03001170
15. AI in optimizing formulations and Excipients. https: //sciensage.info/index.php/JASR/article/download/2477/1717
16. Emerging Artificial intelligence Technologies used in the development of solid dosage forms. https: //pmc.ncbi.nlm.nih.gov/articles/PMC9694557/
17. Artificial intelligence in drug formulations and development https: //pmc.ncbi.nlm.nih.gov/articles/PMC9694557/
18. Agatonovic-Kustrin et al. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical researchJ. Pharm. Biomed. Anal. (2000)
19. Wiercioch M, et al. Pharmaceutical applications of AI. Pharmaceutics, 2021; 13(10): 1571.

Dara S., Dhamercherla S., Jadav S.S., Babu C.M., Ahsan M.J. Machine Learning in Drug Discovery.