



REVIEW

From Bench to Algorithms: Artificial Intelligence Applications in the Chemical Engineering of Coumarin-Based Therapeutics

Yasser Fakri Mustafa^{1,*} ¹Department of Pharmaceutical Chemistry, University of Mosul, Iraq

Abstract: Problem: Coumarin derivatives exhibit broad pharmacological activities and substantial therapeutic potential, yet the integration of artificial intelligence (AI) with chemical engineering across their discovery, synthesis, optimization, and pharmaceutical manufacturing remains fragmented. Aim: This review critically evaluates AI-assisted methodologies and chemical engineering strategies across the development continuum of coumarin-based therapeutics, from molecular discovery to process optimization and intelligent manufacturing. Methods: A comprehensive literature review examined applications of machine learning, deep learning, generative AI, chemoinformatics, reinforcement learning, digital twins, molecular modeling, reaction optimization, and AI-assisted pharmaceutical manufacturing relevant to coumarin chemistry and drug development. Results: Current evidence indicates that AI can support prediction of biological activity, absorption, distribution, metabolism, excretion, and toxicity properties, molecular interactions, and synthetic feasibility. Machine learning and generative AI facilitate virtual screening, scaffold design, and lead optimization, whereas reinforcement learning and digital twins offer opportunities for reaction control, continuous-flow synthesis, process optimization, and manufacturing. However, translation remains constrained by limited curated datasets, insufficient model interpretability and experimental validation, and weak integration between computational predictions and industrial workflows. Conclusion: Integrating AI with chemical engineering provides a framework for accelerating coumarin therapeutic development. Progress will depend on high-quality experimental data, explainable models, rigorous validation, multidisciplinary collaboration, and stronger integration of computational approaches with scalable pharmaceutical manufacturing.

Keywords: artificial intelligence, coumarins, chemical engineering, chemoinformatics, data curation

1. Introduction

The pharmacologically active coumarin analogs have been extensively studied for inhibition of a broad spectrum of biological targets such as kinases, acetylcholinesterase, lipoxygenase, topoisomerases, and HIV-1 integrases. Although these compounds have wide therapeutic applications, the industrial production of these compounds is still mainly based on chemical industrial batches [1]. On the other hand, the use of artificial intelligence (AI) and machine learning (ML) in coumarin research and large-scale production is still in its early stages. The majority of the AI-based predictions are currently being validated, especially *in vitro* and pharmacologically. Several biological and physicochemical problems such as low aqueous solubility, suboptimal Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) profiles, metabolic instability, and poor membrane permeability pose difficulties for the development of coumarin therapeutics that penetrate the blood–brain barrier. Furthermore, there is still a lack of large and high-quality datasets with complete ADMET and physicochemical data for the development

of robust AI models. There are only a few studies that try to incorporate computational resources throughout the entire lifespan of coumarin compounds, from molecular design to industrial manufacturing [2].

The scalable synthesis of naturally derived coumarins is also difficult, and predictive, data-driven manufacturing strategies are needed. In recent years, physicochemical analysis has been connected to modern chemical engineering processes through digital pharmaceutical innovation, AI-enabled platforms, and design–build–test–learn (DBTL) approaches. These methods enable seamless integration of AI into the pharmaceutical development and manufacturing processes [3]. ML and reinforcement learning methods are used to search for molecules, optimize reactions, and generate synthetic routes via AI-based technologies, among other applications. These can be used to discover unknown relationships in large datasets, speed up the decision-making process, and help optimize processes as early as the research phase and up to the production stage in industry [4]. However, there are still key hurdles to be addressed in the application of these computational advances for the development of new active pharmaceutical ingredients and therapeutic products [5]. While various datasets have become available for prediction of biological activity and ADMET behavior using ML and deep learning

*Corresponding author: Yasser Fakri Mustafa, Department of Pharmaceutical Chemistry, University of Mosul, Iraq. Email: dr.yassermustafa@uomosul.edu.iq

algorithms for newly synthesized coumarin analogs, such datasets are still lacking that are comprehensive and interconnected [6]. This means that numerous aspects of coumarin production and applications are still underutilizing AI technologies.

Although the field of coumarin research and AI-driven drug discovery is rapidly growing, there are still a few critical research gaps to be addressed. Current review articles focus on either the medicinal chemistry and pharmacological effects of coumarin derivatives or on the use of AI in the field of pharmaceutical sciences but do not particularly attempt to combine the two fields. Therefore, the whole process of development (from molecular design to biological prediction, chemical process optimization, manufacture and industrial translation) has not been systematically covered.

Furthermore, a number of key issues are under-published at the present time. These include standardized data curation for AI applications, molecular representation strategies, explainable and interpretable AI models, reaction optimization via reinforcement learning (RL), digital-twin-assisted process control, continuous-flow pharmaceutical manufacturing, and regulatory considerations on AI-assisted therapeutic development. The use of computational predictions in a manner that could be validated in the laboratory and scaled up to pharmaceutical production is also seldom discussed in existing studies, and there is a significant gap between *in silico* discovery and *in vivo* use.

To overcome these limitations, this review offers a multidisciplinary perspective on the use of AI in the entire therapeutic development process of coumarin-based therapeutics, with chemical engineering emphasized from the outset. It provides a critical appraisal of the methods of AI and process engineering, intelligent manufacturing, experimental validation, regulation, and future translation opportunities, unlike previous reviews. This review highlights these unaddressed problems and integrates new evidence into a unified framework to provide a blueprint for the rational design, optimization, and industrial application of next-generation coumarin-based therapeutics. The major contributions of this review are summarized below:

- 1) Offers the first integrated review that brings together AI and chemical engineering throughout the full therapeutic development lifecycle of coumarin-based therapeutics.
- 2) Discusses the methodologies used in AI and their implications for the research of coumarin, such as ML, deep learning, chemoinformatics, generative AI, RL, digital twins, and explainable AI.
- 3) Connect molecular discovery with pharmaceutical manufacturing by exploring reaction optimization, continuous-flow synthesis, process control, quality assurance, and intelligent manufacturing strategies.
- 4) Recognizes existing challenges and knowledge gaps, such as the lack of curated data, limited interpretability of models, poor experimental validation, and translation challenges.
- 5) Compares and contrasts AI-assisted methods for therapeutic development of coumarin molecules against real-world examples.
- 6) Identifies future research needs that enable the development of AI systems that are integrated, explainable, and experimentally validated, accelerating the discovery and industrial production of coumarin-derived therapeutics.

This review not only summarizes the literature but also provides a conceptual integration framework that brings together AI methodologies and chemical engineering principles throughout

the entire therapeutic development process for coumarin therapeutics. The review does not focus on the individual use of AI techniques but systematically connects molecular discovery, predictive modeling, reaction optimization, process intensification, intelligent manufacturing, experimental validation, regulatory compliance, and industrial translation in one multidisciplinary workflow. This integrative perspective will also help shape future research directions by identifying key knowledge gaps, prioritizing research directions, and pinpointing opportunities for making AI-driven coumarin discovery a viable process for scalable pharmaceutical production.

This review was carried out as a structured literature search and evidence synthesis critical narrative review. The main sources of literature were PubMed, Scopus, Web of Science, ScienceDirect, and Google Scholar. The search was primarily limited to publications published between 2015 and March 2026, while earlier landmark papers were added when they brought some fundamental advances in the field of coumarin chemistry or AI. Typical search terms included “coumarin,” “coumarin derivatives,” “artificial intelligence,” “machine learning,” “deep learning,” “generative AI,” “chemoinformatics,” “drug discovery,” “reaction optimization,” “digital twin,” “continuous-flow synthesis,” “chemical engineering,” and “pharmaceutical manufacturing.” The reference lists of eligible publications were manually screened for additional relevant articles. The publications were selected as eligible if they directly related to the coumarin therapeutic approach or if they were methodological advancements relevant to the coumarin approach.

Editorials, conference abstracts that lacked adequate methodological detail, duplicate publications, and studies not related to AI-assisted applications in pharmaceutical/chemical engineering were not included. The literature retrieved was screened for scientific relevance, methodological rigor, novelty, and translational relevance. Instead of carrying out a quantitative systematic review or meta-analysis, the studies identified were critically synthesized into thematic sections on the following areas: molecular discovery, predictive modeling, reaction optimization, intelligent manufacturing, experimental validation, regulatory aspects, and future research directions. Because this article is a critical narrative review, a formal PRISMA flow diagram and risk-of-bias assessment were not performed.

This review continues as follows: The pharmacological relevance of the coumarin therapeutics and their role in contemporary drug discovery are summarized in Section 2. Chemical engineering is involved in coumarin synthesis and manufacturing, which is discussed in Section 3. In Section 4, AI approaches for molecular discovery, prediction of properties, scaffold generation, and virtual screening are discussed. Section 5 highlights the optimization of processes with the aid of AI, RL, digital twins, and intelligent manufacturing. Section 6 focuses on safety, ethical and regulatory issues, while Section 7 focuses on representative experimental validation studies and comparative case analyses. Finally, a few challenges, future research directions, and major conclusions of the review are discussed in Sections 8–10. All figures in this review were drawn by the author for this manuscript unless stated otherwise. They have been created by compiling information from the literature reviewed and are designed to give a visual overview of the concepts discussed. The figures are not reproduced from or adapted from previously published sources. In summary, the novelty of this review is not in the description of the individual techniques of AI but in their integration within a single chemical engineering framework, which will help identify

scientific gaps and create a translation roadmap for future therapeutic development of coumarins.

2. Background on Coumarin-Based Therapeutics

Coumarin derivatives are a chemically versatile compound class that has a wide range of pharmacological activities. Naturally occurring and synthetic coumarins have been shown to possess a wide range of therapeutic activities such as anti-tumor activity, cardioprotection, anti-allergic activity [7], anthelmintic activity [8], anti-inflammatory activity [9], antimicrobial activity [10], antiviral activity [11], antitubercular activity [12], antioxidant activity [13], and anticoagulant, anti-obesity, anti-hyperlipidemic, and anti-angiogenic activity [14]. Despite these promising biological properties, a number of aspects of coumarin-based drug development need to be explored further. Specifically, the interactions of certain subclasses of coumarins with important molecular targets like protein kinases are yet to be fully explored. In addition, the therapeutic translation of coumarin derivatives is still limited by the relatively scarce integration of ADMET evaluations into their development pipelines [15]. While datasets that describe quantitative structure–activity relationships (QSARs) for the coumarin scaffold are available, comprehensive QSAR modeling studies with these types of datasets are limited [16]. In addition, there are no widely adopted molecular discovery, optimization, and molecular manufacturing technologies that have been applied to systems containing coumarin [17]. Coumarins are a structurally diverse group of benzopyrone compounds, which have enormous pharmacological relevance and are a good platform to combine computational, chemical, and engineering approaches.

Recent developments in data science and AI have revolutionized multiple areas of chemistry, including molecular design, process optimization, and predictive modeling, which all seek to increase efficiency [18]. But the incorporation of chemical engineering principles in these workflows is still relatively underdeveloped, especially for coumarin-based research. Coumarins offer abundant synthetic backbones, making them ideal for investigation using data. Combining chemical process knowledge with computational modeling can speed the discovery, design, and

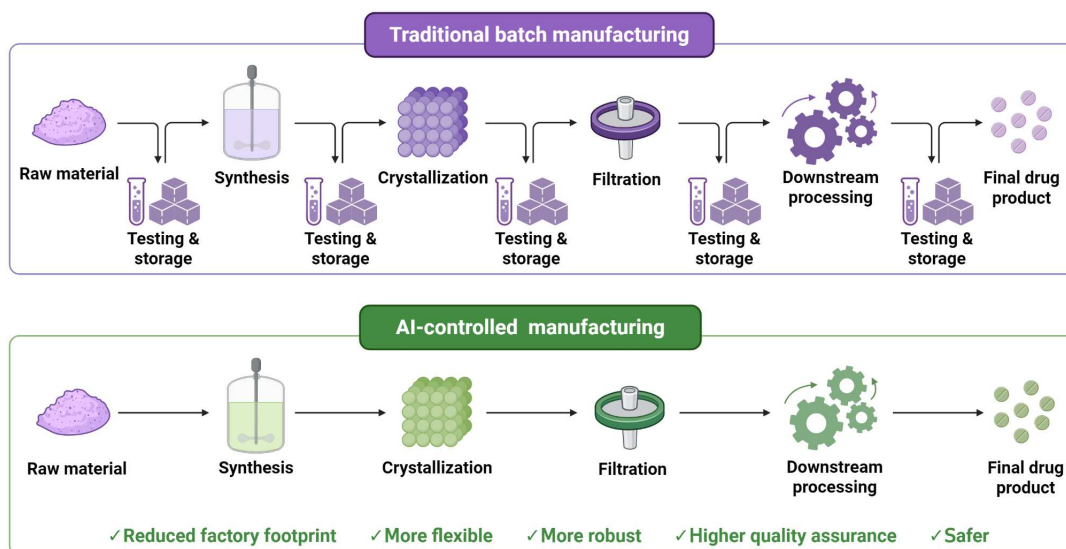
manufacturing of coumarin-based therapeutics and functional materials [19]. The creation of structured datasets, standardized testing procedures, and chemically informed predictive models might serve as a starting point for the rational development of coumarin-based compounds. These strategies would enable improved synthesis, biological screening, and process optimization, thereby enhancing the potential for the use of coumarins in medicinal chemistry, pharmaceutical production, and other chemical engineering areas.

3. Chemical Engineering Perspectives in Coumarin Therapeutics

A central role of chemical engineering is to bridge the gap between molecular design and process optimization and manufacturing strategy for coumarin-based therapeutics. Typically, the discovery pipeline for coumarin derivatives is biological screening against preselected targets, followed by synthetic pathway optimization at the lab scale and evaluation of synthetic pathway scalability for pilot or prototype production. The ability to translate laboratory results to industrially viable processes allows one to identify critical data points that can be used to support iterative DBTL workflows [20]. In this context, parameters associated with the molecular behavior and process-flow performance can be used to link compound design to experimental efficiency, product quality, and manufacturing results [21]. The use of AI in data generation, curation, and ML analysis also helps overcome the constraints of time, cost, and experimental resources, driving the advancements in the development of coumarin-based therapeutics.

Chemical engineering techniques have also made significant contributions to other segments of pharmaceutical development, such as the modernization of traditional manufacturing processes to near real-time or real-time controlled systems [22], as depicted in Figure 1. The effective implementation of AI-driven predictive models is becoming increasingly critical for technologies like automated reaction optimization and feedback-regulated flow chemistry. The synthesis of anticancer drugs is a representative example of this, in which the synthesis of coumarin-based

Figure 1
Comparison between traditional and AI-controlled drug manufacturing process



compounds is coupled with RL methods to dynamically optimize the conditions of the reaction [23]. Similarly, digital-twin technologies have been used in continuous-flow manufacturing systems to aid online monitoring, fault detection, and corrective process control, which has enhanced the production reliability and efficiency [24].

4. AI-Driven Discovery and Optimization in Coumarin Chemistry

The latest developments in the field of molecular simulation, high-throughput screening, and reaction optimization have created large sets of data that can greatly speed up the discovery and optimization of coumarin derivatives. The utilization of these data sources is becoming more common in the development of ML-based tools to aid in various phases of drug development [25]. Computational methods such as density functional theory computations and molecular dynamics simulations can aid experimental studies by predicting the biological activity, ADMET properties, and solubility profiles of coumarin derivatives [26]. Meanwhile, generative AI techniques are enabling scaffold engineering and informational design strategies for combinatorial synthesis. *In silico* screening platforms also help in compound prioritization and risk-benefit assessment based on the wealth of existing pharmacological and toxicological data for therapeutic use.

The increasing successes of AI in molecular design, along with the now widely available academic and industrial datasets, make it clear that it is essential to bring data-driven strategies into the DBTL paradigm [27]. However, there is still a risk of class imbalance in the available data affecting predictive models for biological activity. Moreover, while scaffold innovation and planning of synthesis have been the focus of considerable attention in generative models, their utility is limited due to the relatively small number of coumarin derivatives whose structures have been validated by experiment and synthesis is possible [28]. This restriction not only decreases the variety of training data for reliable scaffold discovery but also lessens the synthetic route generation for practical applications.

4.1. Data curation and representation for coumarin compounds

The molecules studied cover a large and varied chemical space, and extensive data on the pharmacological and physicochemical properties of these molecules are still largely missing. Coumarin chemistry has been studied in great depth, yet there are still many known derivatives for which scant data on ADMET is available [29]. Like other fields of research that rely on ML, the accuracy of predictive models relies on the ability to represent data accurately and to curate the datasets with care. Thus, the databases that are available for use should be carefully examined to minimize potential bias and enhance the robustness of computational predictions [30]. One such application area of AI in pharmacology is the use of coumarin derivatives, where further molecular and pharmacokinetic features were added to enhance the *in silico* screening and therapeutic evaluation [31].

There are several existing resources for the study of coumarins. There is a lot of chemical and bioactivity data for thousands of compounds such as many coumarin derivatives in PubChem. Other valuable bioactivity and pharmacological data sources include ChEMBL and DrugBank [32]. Structural data can

also be retrieved from the Cambridge Structural Database or from the Crystallography Open Database, which helps to explore and understand ligand binding and drug-target interactions at atomic resolution [33]. Furthermore, drug-likeness filtering methods can help to prioritize the large number of coumarin compounds to be investigated further. This process is further enhanced by computational platforms including ADMETlab that can predict the ADMET properties for over 25,000 bioactive molecules, including ~1725 coumarin derivatives [34].

4.1.1. Proposed AI architecture for coumarin therapeutic development

The proposed AI-assisted architecture for therapeutic development of coumarin involves data acquisition, molecular representation, predictive modeling, experimental validation, and process optimization. First, the datasets related to coumarin are obtained from public databases such as PubChem, ChEMBL, DrugBank, and ADMETlab. The molecular descriptors, fingerprints, and graph-based representations are then created and provided to various AI models based on the application scenario. Traditional ML models like random forest and support vector machine (SVM) are most commonly used for small- or medium-scale kinase inhibition and prediction of QSAR. On the other hand, deep learning and graph neural network (GNN) architectures are used in the multidimensional prediction of ADMET and analysis of molecular properties. The applications of generative AI frameworks such as variational autoencoders and generative adversarial networks are used for scaffold design and *de novo* molecular generation, while reinforcement learning models are used to integrate into reaction optimization and continuous-flow manufacturing systems, as illustrated in Figure 2. The process also includes molecular docking, molecular dynamics simulations, and experimental biological validation, creating a cycle of computational prediction and experimental validation. This integrated architecture allows the adaptation of AI methodologies based on the complexity, dimensionality, and engineering requirements of each case for the development of a coumarin.

4.2. ML models for property prediction

The concepts of chemical engineering have significant relevance to the chemistry of coumarin, with particular relevance in the pharmaceutical attributes including biological activity, solubility, and ADMET behavior. As such, the prediction of properties using ML has become a key aspect of AI-based drug discovery and therapeutic development. The tools that chemoinformatics offers for the organization, curation, and computational representation of chemical and biological data are essential in this context with the example of coumarins [35], as illustrated in Figure 3.

Molecular fingerprints can be used to create a structural and physicochemical description of coumarin derivatives, enabling predictive modeling. Advanced computational methods such as kernel-based methods, GNNs, and end-to-end deep learning architectures are used to predict the functional activities of molecules by recognizing relationships between subsystems in the coumarin molecule and its activities [36]. Additionally, the use of experimentally validated property labels in semi-supervised learning models can enhance prediction accuracy and fortify virtual screening pipelines for discovering potential coumarin-based therapeutics [37], as illustrated in Figure 4.

Figure 2
Integrated AI architecture for coumarin therapeutic discovery, optimization, and manufacturing

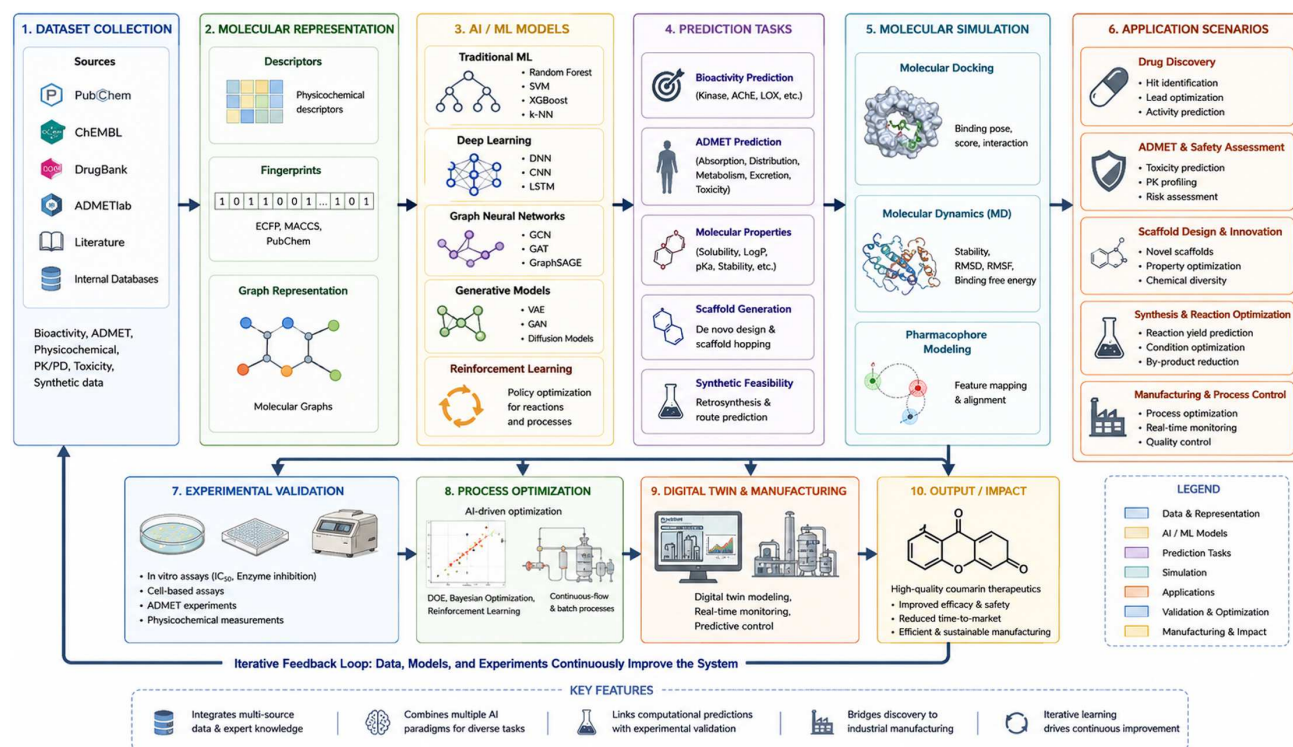
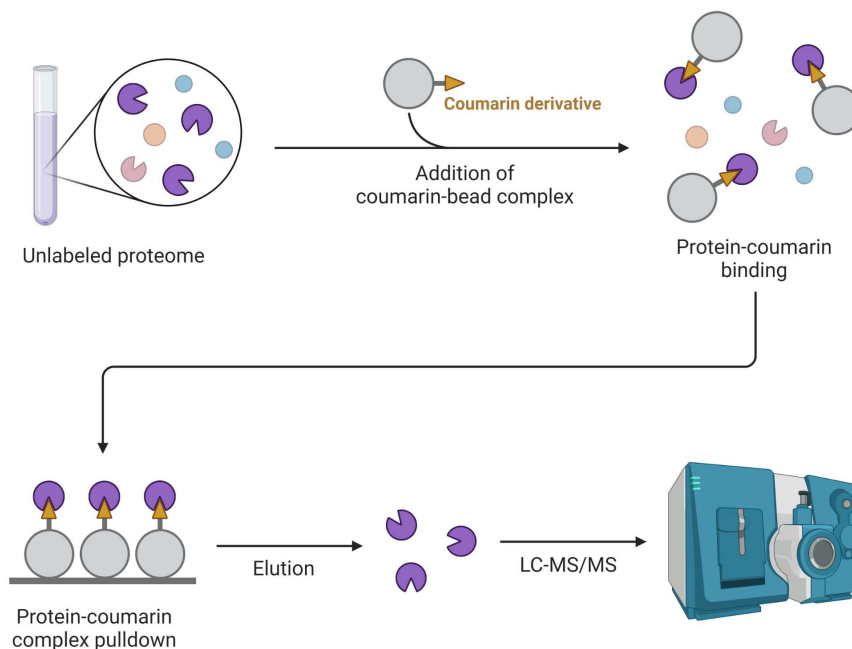


Figure 3
Chemoinformatics workflow for biological property prediction of coumarin derivatives

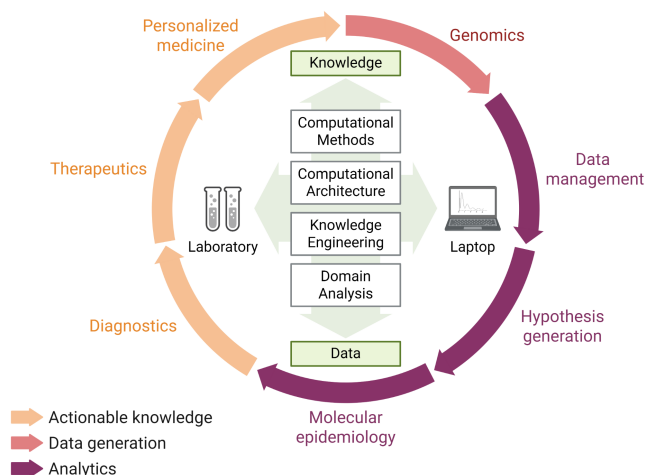


4.3. AI-assisted workflow for coumarin discovery and lead optimization

The typical workflow for coumarin discovery using AI starts with the identification of structurally annotated compounds from databases like PubChem or ChEMBL. Molecular fingerprints and physicochemical descriptors are then created and combined into

ML models that are trained to predict biological activity, toxicity, kinase inhibition, and ADMET properties. The top-scored molecules may then be subjected to molecular docking and molecular dynamics simulations to confirm the stability of molecules interacting with the target and confirm the binding pattern. A few coumarin derivatives identified after computational prioritization can be synthesized in the laboratory and then tested for biological

Figure 4
Translational bioinformatics for improving clinical applications of coumarin derivatives



activity. These “tightened-up” workflows minimize experimental workload, enhance candidate prioritization, and hasten the discovery of pharmacologically relevant coumarin scaffolds for further optimization and industrial development.

4.4. Generative models for scaffold design and *de novo* synthesis planning

In the field of drug discovery, virtual activity-based screening has become a powerful approach and has been driven by the development of conditional generative frameworks, such as variational autoencoders, generative adversarial networks, and their more complex variants. All these models allow for the generation of molecules and are especially useful in scaffold hopping: a method of finding structurally different compounds that have similar biological activities [38]. In this regard, it is possible to prioritize the chemical scaffolds according to the structural similarity with the previously described active compounds or through the bioisosterism approach to direct the design and *de novo* synthesis of new analogs. These generative methods also have the advantage of being able to integrate toxicity predictions or drug-likeness and target selectivity into the multi-objective optimization process [39]. In addition, the incorporation of a space deformation technique in the scaffold generation model has allowed to improve the structural and physicochemical diversity of the generated scaffolds, further increasing the chances of discovering new and relevant candidates for pharmaceutical use [40].

In addition to scaffold design, coumarin-based compounds can also be screened with *de novo* synthesis-planning systems designed to suggest possible synthetic routes to readily available target compounds. Training these computational models typically involves large retrosynthetic reaction databases, and they can be applied end to end on reaction templates [41]. They can also rank possible synthetic routes, plus they can even create completely new retrosynthetic strategies that involve feature-importance analyses of desired product properties. While skilled medicinal chemists are still needed to bring synthetic ideas to fruition and to devise practical and experimentally reliable synthetic pathways, these AI-driven systems can offer valuable guidance and alternative

synthetic ideas, especially for early-stage researchers and non-specialist chemists [42].

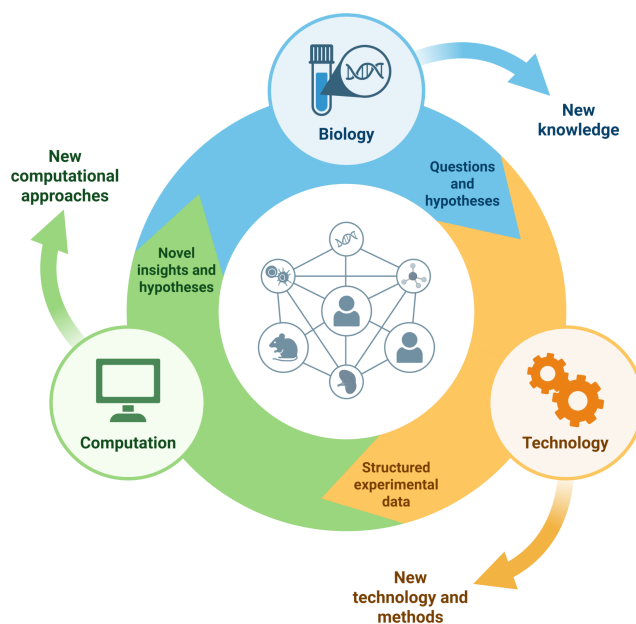
4.5. *In silico* screening and prioritization in therapeutic development

The value of data-driven computational strategies in prioritizing candidates before costly and time-consuming experimental validation, as shown in Figure 5, becomes increasingly evident. Compounds created using generative modeling frameworks can then be prioritized in this context by predictive models that can estimate therapeutic activity, pharmacokinetic properties, toxicity, and solubility [43]. These *in silico* screening strategies can be integrated and can be used to prioritize molecules for further investigation in drug discovery quantitatively and facilitate the subsequent steps of the process.

An ensemble of tree-based ML models has been developed to classify coumarin-derived kinase inhibitors in a recent study. A dataset of about 1300 structurally different coumarin compounds targeting members of the kinome was used for the training of the models. The accuracy of the models was acceptable, but it was difficult to model several families of kinases due to the small number of positive examples for these targets. Nevertheless, the study presented a valuable computational strategy to find specific scaffolds, such as coumarin, which might be useful for potential kinase inhibitory compound development, including targets that are associated with known and new disease pathways [44].

In general, the following multi-layer predictive workflows have been suggested to simultaneously screen inhibitors against a range of enzyme targets and improve ADMET properties. Once new molecular scaffolds have been created, similar strategies could be applied to drug discovery programs for drugs targeting the coumarin scaffold. Importantly, these workflows enable the weighting of predictive properties to be adjusted independently, for flexible prioritization strategies. The system can

Figure 5
AI-driven *in silico* screening and prioritization workflow for coumarin-based therapeutics



be used to identify candidates of structural interest for further filtering or highly optimized compounds that can be directly tested in experiments, for further pharmacological studies [45].

A number of recent developments in the field of AI have made it possible to combine different computational strategies to aid in therapeutic discovery for coumarin-based drugs. As illustrated in Figure 6, they include the use of molecular modeling, chemoinformatics, RL, and process optimization tools to enhance the prioritization of candidates, predict their pharmacokinetic behavior, optimize synthetic pathways, and enable real-time control of the manufacturing process.

These technologies have been found to be of practical use in a number of experimental studies, which support their effectiveness in improving the efficiency of biological screening, decreasing experimental burden, and speeding up the workflows in drug development and delivery. Table 1 provides a brief overview of representative examples of AI-assisted applications in the development of therapeutics based on coumarins.

4.6. Critical comparative analysis of AI model categories

It's crucial to critically compare the primary computational methods used for therapeutic development using coumarins to understand the actual implications of AI in these applications. Each of the AI models has its own strengths and weaknesses, and varying degrees of interpretability, data requirements, and applications such as predicting molecular properties, scaffold design, reaction optimization, and manufacturing control. These methodologies can be compared with each other to gain an understanding of their suitability to be translated, their ability to predict, and their scalability for use in pharmaceutical and chemical engineering applications. Table 2 shows representative characteristics and practicalities of the key categories of AI models employed in the field of coumarin research.

For challenging ADMET and molecular property prediction problems, deep learning and GNN models proved to have

Figure 6
Integrated experimental and AI-assisted workflow for coumarin therapeutic development

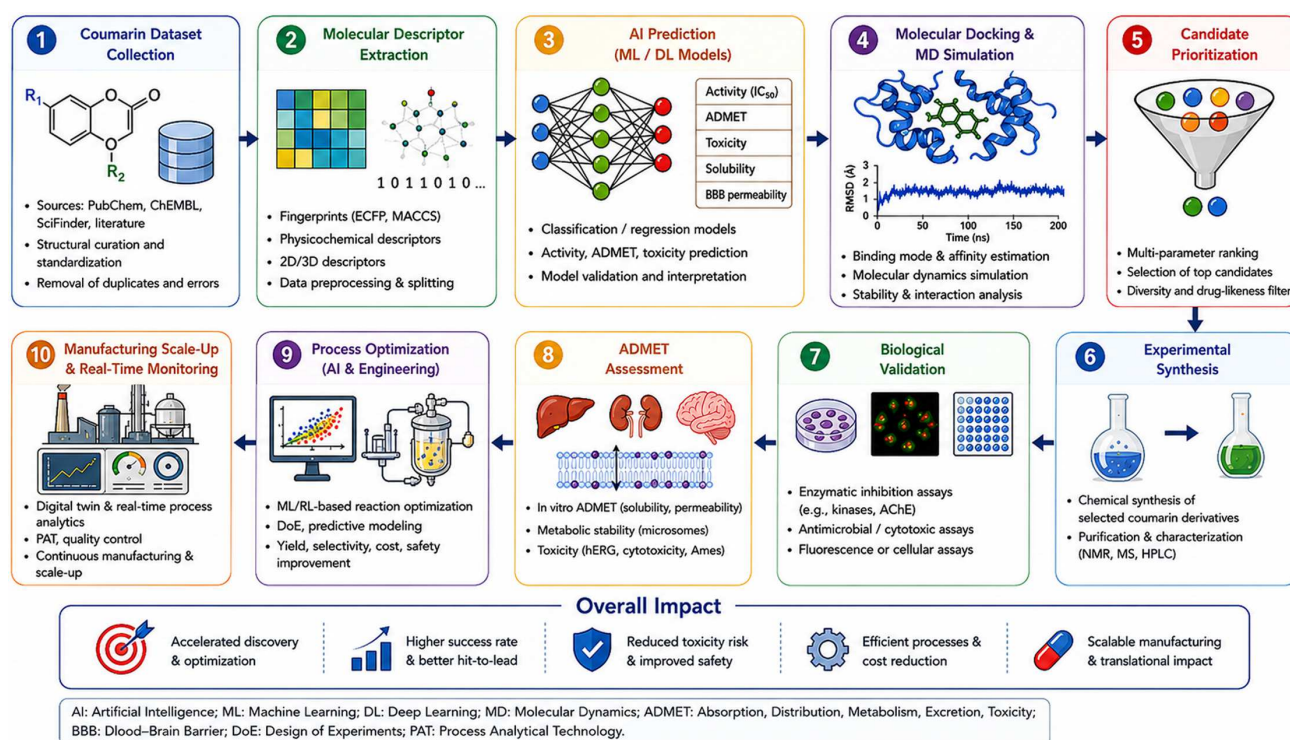


Table 1
Representative AI-assisted applications in coumarin-based therapeutic development

AI technique	Coumarin application	Experimental validation	Major outcome
Random forest	Kinase inhibition prediction	<i>In vitro</i> IC ₅₀ evaluation	Improved screening accuracy
Deep learning	ADMET prediction	Pharmacokinetic assessment	Reduced toxicity risk
Molecular docking	Target interaction analysis	Binding affinity validation	Enhanced lead prioritization
Reinforcement learning	Reaction optimization	Flow synthesis monitoring	Improved process efficiency
Digital-twin systems	Manufacturing control	Real-time process monitoring	Fault detection and quality control

Table 2
Critical comparison of AI model categories used in coumarin therapeutic development

AI model category	Main application in coumarin research	Major strengths	Major limitations	Data requirement	Interpretability	Representative outcome
Random forest	QSAR and kinase inhibition prediction	High robustness, suitable for small datasets, relatively interpretable	Limited ability to capture highly complex molecular relationships	Moderate	High	Improved classification accuracy for coumarin kinase inhibitors
Support vector machine (SVM)	Bioactivity and molecular classification	Strong predictive performance with limited datasets	Reduced scalability for very large datasets	Moderate	Moderate	Effective prediction of coumarin biological activities
Deep learning	ADMET and pharmacokinetic prediction	Excellent feature extraction and predictive accuracy	Requires large curated datasets and high computational resources	High	Low	Enhanced ADMET prediction and toxicity assessment
Graph neural networks (GNNs)	Molecular property prediction	Efficient structural and graph-based learning	Computationally intensive and difficult to interpret	High	Low	Improved molecular representation learning
Generative AI models	Scaffold generation and <i>de novo</i> molecular design	Enables discovery of novel coumarin scaffolds	Synthetic feasibility and experimental validation remain challenging	High	Low	Generation of structurally diverse coumarin analogs
Reinforcement learning	Reaction optimization and process control	Adaptive optimization and autonomous decision-making	Complex training procedures and instability in some workflows	Moderate to high	Moderate	Improved reaction efficiency and process optimization
Digital twin systems	Manufacturing monitoring and process simulation	Real-time process supervision and fault detection	Requires advanced infrastructure and continuous data acquisition	High	Moderate	Enhanced manufacturing reliability and quality control

better predictive power. However, these methods need significantly larger datasets and computing power than conventional ML algorithms. However, for smaller and moderately curated coumarin datasets, the random forest and SVM models are found to be more appropriate due to their higher interpretability and lower computational complexity. RL seems more appropriate for process optimization and autonomous manufacturing than for early-stage molecular discovery due to the dynamic nature of reaction-control environments where the adaptive decision-making feature of RL is particularly useful. Likewise, digital-twin systems offer great benefits for real-time process monitoring and fault detection but still have challenges of infrastructure complexity and the need for continuous data collection.

5. Process Intensification and Control via AI

RL methods coupled with process-derived data offer an effective way of optimizing chemical reactions that are potentially unstable or unpredictable. These methods can be very useful since they do not require extensive prior knowledge or high computational requirements and are especially applicable to situations in which only a single step in a reaction network is poorly characterized or in which the concentrations of reactants change dynamically throughout the network [46]. For instance, trajectory data for crushed-lettuce and tetramethylbenzidine oxidation systems have been used successfully to design data-driven proportional–integral–derivative controllers that are able to control both concentration and temperature to pre-specified setpoints [47].

Chemical processes can then be represented in a machine-readable and computation-compatible format, whereupon explainable neural network models can be trained to predict and interpret several chemical aspects, such as the reaction conditions, the yield of the products, the performance indicators, and by-product formation. These predictive models also play a role in creating digital twins, which can accurately replicate the actual chemical manufacturing environment [48]. No matter what production platform is used—whether batch, AI-driven adaptive batch, continuous-flow, or hybrid—it collects vast amounts of operational data during actual production processes [49]. These datasets can be used to create closed-loop monitoring and control systems where the ML algorithms can help with the real-time monitoring of the process, anomaly detection, identification of faults, and implementing automated corrective measures [50]. All of these combined methodologies provide initial steps to an AI-based chemical engineering toolbox. Furthermore, they illustrate the ability of AI-driven systems to optimize chemical processes and their operational management.

5.1. Reaction optimization using reinforcement learning

In the field of synthetic chemistry, RL-assisted reaction optimization has developed into a potent approach to tackle complex problems and found application in both academia and industry. Previously reported reaction conditions can now be used to train ML models that can be utilized for rapid virtual screening and initial reaction scouting to support experimental decision-making [51]. When these methods are incorporated into workflows that span multiple research institutions and involve collaborative DBTL, however, the transformative effects of these methods become more apparent. One such advance is in the synthesis of coumarin [52]. The chemistry of coumarins is well

documented but is still developing with the investigation of a variety of synthetic pathways and reaction disconnections. The recent progress of AI has enabled to generate predictive reaction models automatically from large datasets retrieved from scientific resources like Semantic Scholar and Scopus. These developments have effectively enabled the emergence of virtual chemical engineering platforms that can freely analyze, optimize, and document organic reactions. In practical applications, undergraduate and postgraduate researchers have used such systems of AI to optimize multiple reactions during long experimental campaigns [53]. Moreover, in the era where the internet is now the most common medium for information sharing, collaborative computational chemistry projects have been made possible, sharing knowledge, managing data, and enabling transferable learning throughout several laboratories, using platforms like Dataverse and extended Chemical Informatics Platforms [54]. In total, these advancements are a testament to the potential of AI-driven optimization in revolutionizing the field of reaction discovery and synthetic efficiency.

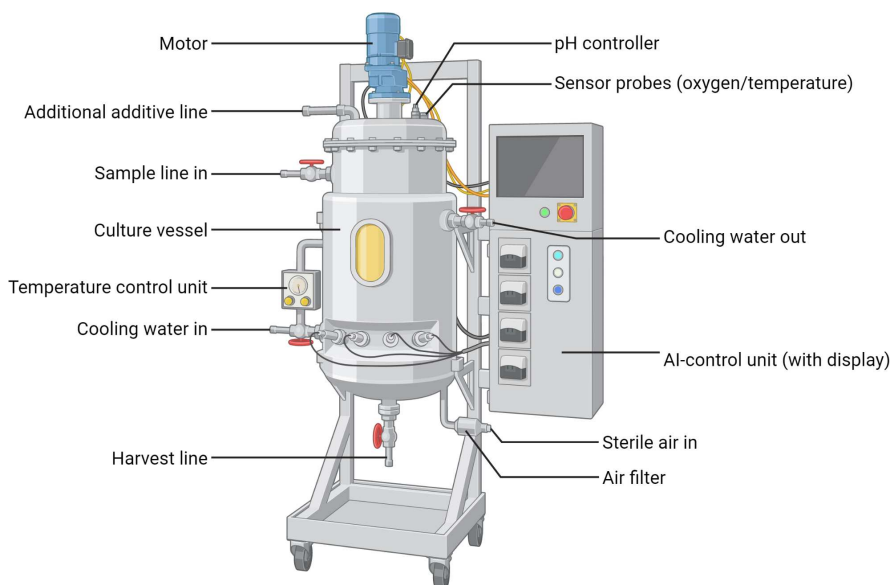
AI and data-driven techniques also hold great potential for enhancing real-time reaction management. One of the newly proposed frameworks, N2Ore, integrates long short-term memory neural networks, RL, and advanced control strategies to enable autonomous decision-making for optimal operational policies in chemical systems [55]. Evaluating the system in a semi-controlled continuous stirred-tank reactor (Figure 7), with nitrous oxide as one of the main volatile by-products, it was found that the system was able to reduce the amount of nitrous oxide produced by around 84% after iterative learning cycles, compared to human-guided optimization [56]. Likewise, by unsupervised analysis of hundreds of conditions in the literature, the Library of Molecular Reactants for Cu-mediated Catalysis with Alcohols has been correlated with deep RL algorithms which steer catalytic systems toward the desired alkylation products [57]. The developments emphasize the emerging capability of intelligent automation for optimizing reaction conditions and dynamic process control in modern synthetic chemistry.

5.2. Process modeling and digital twins for coumarin synthesis

The use of AI in the development of digital twins has been successful in modeling and optimizing several steps in the synthesis of coumarin. For reaction development, AI-supported reaction databases can be used to create digital twins that can predict reaction outcomes, suggest laboratory-scale reaction conditions, estimate the likelihood of generating desired reaction products using multivariate classification models, and aid in the identification of optimal reaction conditions [58]. At the same time, digital twins obtained from simplified kinetic models that can be built based on experimental data measured in the laboratory can be used to optimize continuous-flow reaction processes, such as Rankine-type reaction processes, by identifying optimum operating parameters [59].

Accurate and high-quality experimental information is crucial to the reliability of process modeling. As such, AI-powered digital twins can also improve the process of generating data and optimizing processes [60]. The use of neural network-based methods has been shown to enhance reaction development by optimizing the experimental conditions with RL techniques, as well as to provide model-free process control with dynamically generated process data [61]. Likewise, using bidirectional recurrent neural networks has been adopted to speed up response

Figure 7
AI-controlled continuous stirred-tank reactor for autonomous process optimization



times and enhance operational efficiency in batch production processes in the chemical industry [62]. In addition, autoencoder-driven embedded code reconstruction anomaly-detection systems can detect process failures by detecting deviations between the actual behavior of the system and predicted operational patterns to enhance process monitoring and diagnostic performance [63].

5.3. Real-time monitoring and fault detection

Safety is always an essential factor in every production area, especially in chemical manufacturing processes where reactions and separation processes may involve long production runs. After the reaction and purification steps, the attention is directed toward stabilizing and controlling a production site. Thus, real-time check of critical quality attributes becomes crucial to ensure the production of the desired product at the desired stages and acceptable quality throughout the production process. Good monitoring systems also allow quick action to be taken if there are any errors in the product quality or operational performance, preventing any faulty materials from moving on to the manufacturing process [64].

Therefore, the design of data-driven monitoring strategies is strongly reliant on the ability to recognize appropriate target parameters that can be used as early indicators of deviations from normal operating conditions. These predictive metrics are crucial for enabling proactive quality control and reducing production risks [65]. But real-time anomaly detection is still a challenging task. An abnormal signal can occur because of events outside of the chemical reaction process (e.g., sensor failure or equipment failure) and should not cause an alarm. Process deviations, however, which may impact product quality, must be identified as soon as possible so that correct action can be taken in time. Thus, strong monitoring systems need to be able to differentiate between insignificant anomalies and significant process disturbances, and they have to be sensitive enough and reliable enough [66].

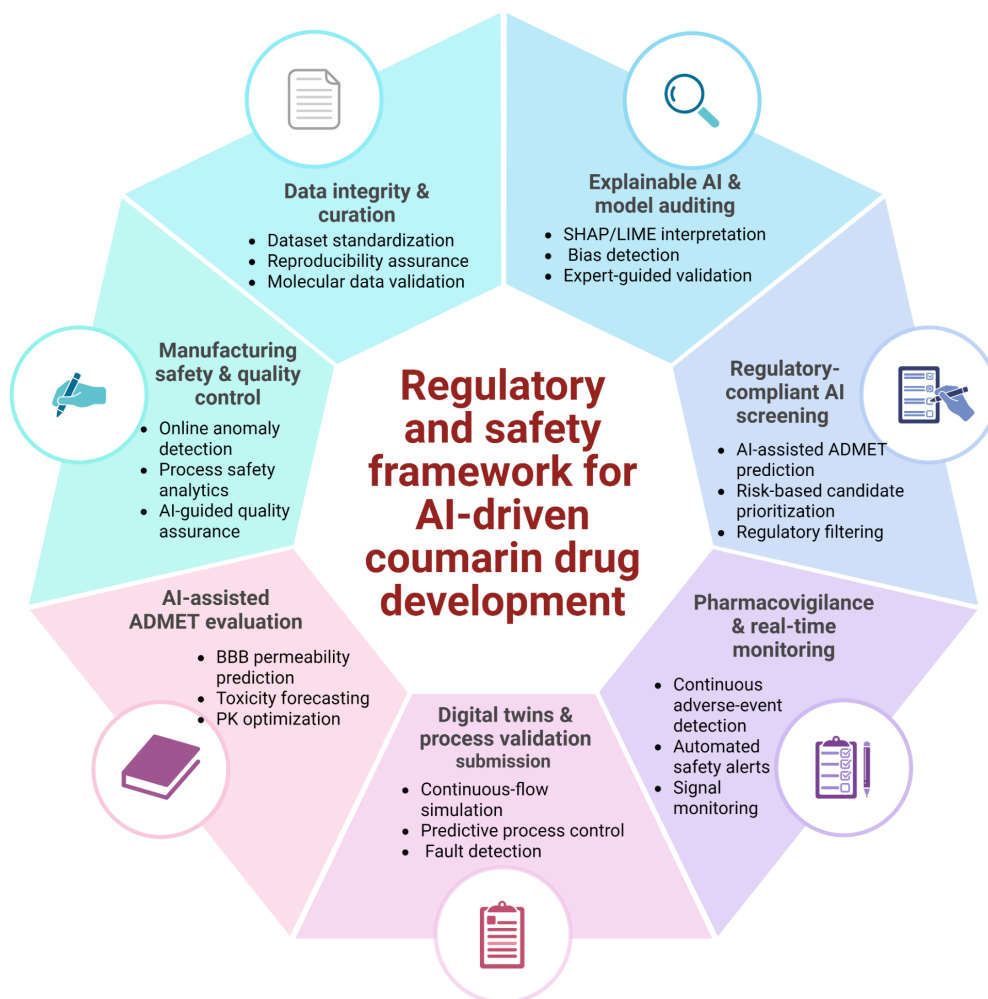
6. Safety, Ethics, and Regulatory Considerations

As shown in Figure 8, safety, ethical oversight, and regulatory compliance are integral aspects of the integration of AI

into therapeutic discovery and development workflows for compounds such as coumarins. Data integrity, data reproducibility, transparency, and compliance with pharmaceutical and chemical standards are important factors in the reliability of an AI-assisted system. All of these factors will influence whether and how biomedical research and drug development can be conducted safely, responsibly, and effectively via AI-driven strategies [67]. One of the most important aspects is to maintain the quality and consistency of data throughout the workflow. To be successful, AI and ML models require well-curated and accurate datasets—ones that are reproducible. Therefore, standardizing the protocols for data collection, curation, annotation, and representation is important to reduce inconsistencies and ensure that the data that are collected are reliable and reproducible before they can be subjected to computational analysis [68]. Experimental research that produces new datasets should be done using standard practices in the laboratory and with methods that have been validated. Furthermore, chemical and pharmacological datasets often span disparate molecular spaces, so a systematic deposition of the data in accessible cloud-based repositories and shared databases can further enhance the reproducibility of the data [69]. This ecosystem helps to enhance the transparency, reliability, and long-term value of AI-driven research.

Model interpretability also has practical implications for the usefulness of AI systems. Explainable AI methods offer explanations about how predictive outcomes are related to certain molecular descriptors or pharmacological properties, which boosts users' trust in the predictions and contributes to scientific interpretation [70]. The explainability mapping and model inspection techniques support the auditing process, risk management strategies, and pinpoint potential sources of bias. For many applications, it is important to compare AI predictions with expert knowledge in order to have biologically and chemically meaningful computational reasoning [71]. The interaction of computational scientists with domain experts helps refine the model design, to interpret mechanistic patterns, and to verify that the model's predictions are consistent with accepted pharmacological principles. In high-risk biomedical applications,

Figure 8
Integrated framework for safety, explainability, and regulatory compliance in AI-driven coumarin research



such multidisciplinary evaluation is crucial, with model auditing designed to check the reliability of the predictions and the consistency of model response to various experimental outcomes [72]. These measures help inform decision-making and limit the risks of model deployment.

Beyond the technical issues, the translation of coumarin-based therapeutics to the pharmaceutical and regulatory industry requires adherence to strict regulatory and pharmaceutical regulations. Regulatory-oriented AI workflows for property prediction and candidate prioritization can help reduce the risk of unsuitable compounds moving forward to biological testing or clinical evaluation. With the development of drugs moving toward industrial-scale production, drug production safety, quality assurance, and production processes must be included as part of the regulatory process, along with the final product [73]. The adoption of digital-twin technologies could further enrich this process, offering a way to continuously monitor and predict risks as the development process progresses. These relationships enable corrective or preventative actions to be taken in time when deviations are found [74]. Furthermore, in clinical trials, adverse event detection systems that are automated can monitor adverse effects in real-time, which could help in detecting any adverse concerns at an early stage, streamline pharmacovigilance, and reduce the potential for adverse effects to cause harm to patients [75].

6.1. Data integrity and reproducibility

With the growing application of coumarin in chemical engineering and pharmaceutical research, quality, integrity, and reliability of data are crucial. Comprehensive metadata collection such as detailed documentation of data provenance, reference compounds, experimental parameters, and curation procedures are key elements for effective management of experimental and computational data. Careful and balanced data curation is essential to minimize selection bias, improve reproducibility, and ensure that datasets remain representative and scientifically meaningful [76]. Databases such as Cambridge Structural Database and PubChem provide extensive structural and physicochemical information, alongside limited biological activity and ADMET-related data. However, additional curation and standardization of these datasets can substantially reduce the workload associated with ML and deep learning workflows. Well-organized and consistent datasets containing synthesis conditions, biological responses, and safety-related parameters are particularly valuable for predictive modeling, toxicity assessment, and generative scaffold design [77]. Moreover, data anonymization in chemical and biological laboratories is still a crucial process to ensure adherence to privacy laws and ethical guidelines for data sharing and dissemination.

6.2. Explainability and model auditing

In the presence of large datasets, significant computational power, and ample training time, ML algorithms, especially deep learning techniques, can deliver exceptional results in prediction and decision-making tasks. In these situations, very specialized models can be constructed directly from raw data. However, if the number of observations in the datasets is less than approximately 1,000, the performance of the model may be affected, and the interpretation of the results may be more important. In these situations, the expertise of the biological and chemical scientists becomes critical to guide feature selection, validate model outputs, and ensure that the predictions align with established mechanistic understandings [78].

Interpretability frameworks, such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations), provide practical tools for decomposing complex model predictions into quantifiable feature contributions. For coumarin-based therapeutic development, where subtle structural modifications can drastically alter ADMET profiles or target selectivity, applying these interpretability frameworks ensures that AI-identified scaffolds are not merely statistical artifacts but represent chemically rational candidates. Consequently, establishing standardized auditing protocols that integrate domain-expert evaluation with algorithmic explainability is essential for fostering trust and facilitating the regulatory acceptance of AI-driven findings in medicinal chemistry [79].

6.3. Compliance with pharmacological and chemical regulations

Every aspect of workflow has been designed with respect to regulatory and scientific best practices for safe and efficient development of molecular entities. Comprehensive pharmacological assessment, such as knowledge of chemical transformation pathways, metabolic behavior, and membrane transport properties are necessary prerequisites for successful drug development in the case of coumarin derivatives. Special emphasis needs to be placed on ADMET properties, particularly with respect to cellular uptake and blood–brain barrier permeability, which have a significant impact on biological efficacy and therapeutic performance. AI-aided parametric modeling and predictive analyses can help systematically assess ADMET changes in chemical modification processes, supporting rational optimization strategies that can enhance pharmacological activity. Moreover, AI-generated structure–ADMET relationships can be used as valuable targets for molecular scaffold design and lead optimization.

Safety assessment, risk management, and regulatory compliance continue to be essential factors from a chemical engineering and process-control point of view, particularly at large-scale production. Once the best synthetic and production conditions are determined using the AI-supported engineering platforms, data-driven monitoring can be embedded throughout the manufacturing process to maintain the stability and consistency of the process. Process monitoring during the production process can be enhanced by using real-time analytical monitoring and anomaly detection using AI [80]. Additionally, advanced systems of data-driven analytics can be used to assess the quality of the final product, thereby guaranteeing adherence to predetermined specifications. These AI-enhanced control and monitoring systems work together to improve manufacturing safety, product reliability, and process robustness and to establish safe operating limits for chemical production processes.

7. Literature-Based Experimental Validation and Representative Case Studies

This section provides a summary of representative published experimental validation studies and literature-based case studies of the application of AI to the therapeutic development of coumarins. The review does not focus on the reporting of new experimental studies but rather considers previously published studies that provided computational methods using AI that were validated with experiments, including the synthesis of compounds, biological assays, and pharmacological studies. The published work demonstrates the potential of AI to rapidly discover new molecules, select them for further development, and enhance the overall workflow of drug development in the clinic, while also providing insights into the challenges faced and the future directions for the field.

A set of seven coumarin-derived CDK1 inhibitors has been prepared such that two compounds have an IC_{50} below 10 μ M. Furthermore, other research groups have reported designing polyphenol oxidase inhibitors containing coumarin and even fluorescence-based probes for the detection of β -galactosidase. Predicting properties such as bioactivity, solubility, and ADMET parameters has also been predicted by computational models trained on publicly available datasets and then used in scaffold-aware optimization strategies to improve the properties of coumarin derivatives [81]. Synthetic progress made includes microwave-assisted single-mode electrochemical preparation of various coumarin derivatives and the preparation of 7-phenylcoumarin derivatives using AI. In addition to therapeutic uses, coumarins have been used in functional materials like Langmuir–Blodgett films, Langmuir–Schaefer films, Cu^{+2} responsive multilayers, and fluorescent chemosensors. Furthermore, multi-task and multi-fidelity ML models have been leveraged to predict reactions and molecular properties covering a wide range of reaction and property classes, underpinning the increased importance of AI in coumarin-related research and development efforts [82].

7.1. Case study I: AI-assisted optimization of coumarin-based kinase inhibitors

Coumarins and their derivatives are structurally diverse scaffolds and have been extensively studied for their ability to inhibit a number of kinases involved in a number of diseases, such as cancer, inflammatory diseases, and neurodegenerative diseases. Although coumarin-based therapeutics are becoming increasingly popular, the originality of structure and the drug-like properties of many of the proposed candidates have not always been fully validated. To overcome this drawback, an integrated AI-assisted approach using two activity-fingerprinting models, SVM and random forest, was devised for screening, prediction, and synthesis of novel coumarin derivatives [83]. Following this workflow, 49 previously unreported coumarins were designed and synthesized, of which 43 compounds were successfully confirmed. Of these, nine exhibited significant biological activity including anti-HIV-1 and anticancer activity against 4T1 breast cancer cells in a HeLa cell model [84].

The proposed framework integrated the structure–activity relationship analysis with kernel-based similarity mapping and wide pharmacological profiling to enhance the predictive power and chemical diversity. The designed coumarins were computationally docked with several therapeutic targets relevant to cancer, infectious diseases, and neurodegeneration such as PLK1, c-KIT, EGFR, c-MET, and CDK2, aiming for alignment with the

SARS-CoV-2-related proteins based on chiral (S)-valine motifs. Selectivity profiling and extensive activity fingerprinting were also included in the strategy, and especially the performance of SVM-based classifiers was highlighted. These fingerprints had better correlation with biological activity and structural diversity than the traditional homology or interaction indices [85]. Additionally, the non-molecular expert model predictions were used as an internal validation tool to ensure the reliability of the synthetic outcomes and biological screening results [86].

7.2. Case study II: Generative design of coumarin scaffolds with favorable ADMET

The growing global burden of antimicrobial resistance continues to represent one of the most critical challenges in infectious disease research. An emerging strategy to limit the development of resistant bacterial strains involves the identification of small molecules capable of modulating underexplored cellular targets identified through Tahoe¹ screening. In this context, coumarin derivatives have gained attention as inhibitors of serine/threonine kinases involved in biofilm formation, virulence, and pathogenicity in clinically and biotechnologically important microorganisms, including species of *Burkholderia*, *Pseudomonas*, *Mycobacterium*, enteroaggregative *Escherichia coli*, *Salmonella enterica* serovar Typhimurium, and certain strains of *Staphylococcus aureus* [87]. Integrating this target-based strategy with AI-guided discovery of new coumarin structures is particularly promising for the development of safer and more effective antimicrobial agents [88]. The generation of structurally diverse coumarin scaffolds with reduced toxicity profiles and favorable ADME properties therefore represents an attractive and innovative direction in antimicrobial drug discovery.

7.3. Published experimental validation of AI-predicted coumarin candidates

The feasibility of using AI prediction in conjunction with experimental validation has been shown in several recent investigations in the field of coumarin research. In many workflows, initial use of ML models involves prioritization of coumarin derivatives based on predicted biological activity, selectivity, and ADMET properties. The top candidates are then synthesized and tested *in vitro* in enzymatic inhibition assays, cytotoxicity assays, antimicrobial testing, or fluorescence-based biological screening

assays. Such computational-experimental approaches greatly decrease the number of compounds that need to be synthesized in the laboratory and enhance the lead identification process. Random forest models performed predictive accuracies >85% on the classification of kinase inhibition, with the deep learning architectures showing higher ADMET prediction accuracy when trained on larger and chemically diverse datasets. These insights suggest that, for moderately sized sets of couplins, the traditional ML methods are still very effective, whereas the deep learning models offer better scalability and predictive flexibility for analyzing multidimensional PK. Furthermore, molecular docking, molecular dynamics simulation, and biological testing can be integrated, allowing the validation of computational results and enhancing the translatability of AI-driven coumarin discovery platforms (Table 3).

The quantitative assessment shows that AI-aided workflows can significantly enhance the development of coumarins throughout the discovery and manufacturing process. Random forest and SVM models successfully reached an accuracy higher than 85% for kinase inhibition and bioactivity prediction tasks especially for moderately sized and curated datasets. The deep learning and GNN architectures exhibited better performance for multidimensional ADMET property and molecular property prediction particularly for a structurally diverse coumarin library. Another example of practical engineering value was shown by RL systems applied to the continuous-flow optimization of the undesired by-product formation, which was reduced by about 84%. All of these results point to a model selection that is influenced by the size of the dataset, the complexity of a given application, the need for interpretability, and the manufacturing goals. Random forest classifiers gave prediction rates from 85 to 89% for the kinase inhibition datasets, each of which contained some 1300 coumarin derivatives, and SVM models showed Area under the curve (AUC) values higher than 0.82 for the bioactivity classification problems. Higher sensitivity and specificity for ADMET prediction were obtained with deep learning architectures with larger pharmacokinetic datasets of prediction, especially for multidimensional toxicity assessment.

8. Challenges and Opportunities

Despite the structural complexity and chemical diversity of coumarin derivatives, their discovery and development have not been impeded by modern data-driven and ML approaches.

Table 3
Quantitative performance comparison of AI approaches in coumarin therapeutic development

AI model	Application	Dataset size	Accuracy/AUC	Experimental validation	Main advantage
Random forest	Kinase inhibition	~1300 compounds	85–89% accuracy	IC ₅₀ assays	High interpretability
SVM	Bioactivity classification	Moderate datasets	AUC 0.82–0.87	<i>In vitro</i> validation	Effective with small datasets
Deep learning	ADMET prediction	Large datasets	>90% prediction accuracy	Pharmacokinetic screening	Captures complex relationships
GNN	Molecular property prediction	Large graph datasets	Root Mean Square Error (RMSE) reduction ~15–20%	Docking support	Structural learning
Reinforcement learning	Reaction optimization	Continuous-flow systems	84% reduction in by-products	Reactor validation	Dynamic optimization

Yet, many AI-powered drug discovery tools fail to adequately address critical pharmaceutical engineering aspects, such as high-throughput process development, manufacturability, and full ADMET profiling. Furthermore, certain key aspects are still not fully investigated, such as reaction optimization, process digitalization, and the introduction of digital-twin technologies. The engineering aspects, particularly process design, process intensification, control strategy, and safety assessment, are crucial to translate coumarin-based compounds from laboratory-scale discovery to industrial production. While there are some initial steps for this integration, there are still a lot of steps to take in the next several years to build a strong and scalable development framework.

An iterative process similar to the classical DBTL cycle commonly used in chemical engineering is needed to make progress in this field. With the addition of AI, these discovery platforms can explore large chemical spaces, such as structurally diverse coumarin derivatives, creating datasets to develop predictive models of chemical activity, ADMET properties, chemical solubility, and synthetic feasibility. These computational predictions can then be tested experimentally, with those that seem most likely to work prioritized for further research. Once this is done, predictive modeling and adaptive strategies based on AI can be used to optimize synthesis routes, reaction conditions, and production processes. Experimental and simulated data can then be re-entered into the database, continually improving future molecular discovery processes and making them more efficient through process optimization cycles.

Overall, the case studies show that AI-aided coumarin research is increasingly shifting from theoretical models to experimental and practical approaches for pharmaceutical development. The use of predictive modeling with laboratory validation, process optimization, and engineering focused toward manufacturing allows for more efficient identification of biologically active compounds, as well as enhancing translation feasibility. The developments underscore the increasing potential of AI-based workflows to assist with the real-world development of coumarins in the fields of discovery, synthesis, optimization, and therapeutic development in the pharmaceutical sector.

The important result of this review is the identification of a research agenda for next-generation AI-assisted coumarin development. The key focus areas are the creation of standardized and interoperable coumarin datasets, building explainable and tested AI models, integrating digital twins into continuous pharmaceutical manufacturing, designing closed-loop platforms that incorporate build-test-learn cycles, and harmonizing AI workflows with regulatory considerations. Together, the priorities represent a multidisciplinary approach that goes beyond the previous reviews, highlighting the convergence of computational intelligence, chemical engineering, and translational pharmaceutical development.

9. Future Directions

The process of developing coumarin-based therapeutics is very interdisciplinary and data intensive and requires close collaboration between chemists, biologists, chemical engineers, data scientists, ML specialists, and software developers. The translation of discovery and design findings from early phases of drug development into large-scale manufacturing is still a significant hurdle, and many projects struggle to efficiently advance beyond the early stages of development due to inadequate application of chemical engineering principles during technology

transfer. To overcome these problems, AI and chemical engineering methods are increasingly being used to complement the entire discovery–design–manufacturing continuum and to avoid the “translation gap.” The integrated system allows data-driven technologies to be used not only for high-throughput drug discovery and molecular design but also for ADMET prediction, metabolic and toxicological assessments, process optimization, quality assurance, safety evaluation, and real-time monitoring of manufacturing processes. The broader framework will only be realized with the generation, curation, and validation of a vast amount of data, including parameters like product yield, quality, safety, process control, and pharmacokinetic profiles. Thus, the availability of strong, easily accessed data continues to be essential to success. Such datasets facilitate first-principles modeling, simulation of processes, RL, optimization for chemical engineering purposes, and development of real-time screening platforms, as well as the ability to add user-defined safety and quality constraints.

AI-enabled technologies create a synergistic discovery-build-test-learn cycle, enabling many steps in drug discovery, molecular design, and synthesis to be executed in less time and money than in traditional methods. Incorporating the domain-specific know-how in coumarin-containing compounds and the data-driven methodologies in the engineering of drug-like compounds is an important step in the evolution of modern pharmaceutical development frameworks, and it includes process-oriented and manufacturing-centered strategies. Closed-loop experimental platforms that can synthesize, screen, and optimize based on biological data to enhance the synthesis of coumarins are likely to become increasingly common in the future. By combining robotic chemistry, high-throughput biological assays, and adaptive ML systems, the predictive models can be continually improved as they are tested against experimental data. These methods could greatly facilitate the translational drug discovery process and enhance the reproducibility, scalability, and clinical utility of coumarin-based therapeutics.

10. Conclusion

Coumarin-based pharmacophores remain crucial and versatile therapeutic scaffolds and have a wide variety of applications in medicinal chemistry, pharmacology, and chemical engineering. The structural diversity of coumarins along with their wide range of biological activities has made them potential drug candidates for anticancer, antimicrobial, anti-inflammatory, neuroprotective, and cardioprotective agents. Meanwhile, the recent strides of AI and ML have opened up possibilities to speed up the discovery, optimization, and production of coumarin-containing therapeutics. Now AI-driven methods can be used throughout the development process—from molecular design through to scaffold generation, reaction prediction, ADMET screening, process optimization, and real-time process control. However, the use of AI in the study of coumarins is still in its infancy and faces a few challenges. The lack of high-quality, standardized datasets, poor experimental validation of predictive models, and inadequate cross-talk between computational prediction and chemical engineering workflows remain to limit translation. Furthermore, privacy, transparency, regulatory requirements, and scalability of the processes have to be taken into account before systems based on AI can be deployed in industrial processes for drug development. However, the intersection of these AI technologies with medicinal chemistry and chemical engineering holds potential for future advancements in the development of coumarin-based

therapeutics. Generative modeling, RL, digital twins, and intelligent process-control systems can significantly decrease development expenses, boost synthetic efficiency, and drive faster therapeutic innovation. Multidisciplinary partnerships among chemists, pharmacologists, engineers, computational scientists, and data specialists will be required for future advances to develop integrated data-driven workflows that can span the gap between molecular discovery and scalable manufacturing.

In summary, these developments underscore the promise of AI-driven coumarin research for meaningful contributions to future pharmaceutical discovery and precision therapeutic development. Current methods for the development of therapeutic drugs for coumarins suggest that there is no single best model for the development of drugs using AI. While traditional ML techniques are beneficial to smaller, more interpretable datasets, deep learning and generative techniques have more potential for large-scale molecular discovery and process optimization. This will be followed by future developments integrating explainable AI, experimentally validated datasets, and scalable chemical engineering workflows.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The author declares that he has no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are openly available in PubChem at <https://pubchem.ncbi.nlm.nih.gov/>, in ChEMBL at <https://www.ebi.ac.uk/chembl/>, in DrugBank at <https://go.drugbank.com/>, in Cambridge Structural Database at <https://www.psd.ac.uk/csd> and <https://www.ccdc.cam.ac.uk/solutions/software/csd/>, in Crystallography Open Database at <https://www.crystallography.net/cod/index.php>, in ADMETlab at <https://admetlab3.scbdd.com/>, and in GitHub at <https://github.com/shap/shap>.

Author Contribution Statement

Yasser Fakri Mustafa: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration.

References

- [1] Zou, Y., Teng, Y., Li, J., & Yan, Y. (2024). Recent advances in the biosynthesis of coumarin and its derivatives. *Green Chemical Engineering*, 5(2), 150–154. <https://doi.org/10.1016/j.gce.2023.04.003>
- [2] Xia, D., Liu, H., Cheng, X., Maraswami, M., Chen, Y., & Lv, X. (2022). Recent developments of coumarin-based hybrids in drug discovery. *Current Topics in Medicinal Chemistry*, 22(4), 269–283. <https://doi.org/10.2174/1568026622666220105105450>
- [3] Paerhati, Y., Aikebaier, A., Dilimulati, D., Baishan, A., Yusufjiang, N., Qiu, X., . . . , & Zhou, W. (2026). Rethinking nature's pharmacy: AI era and natural product drug discovery. *Pharmaceutics*, 19(2), 301. <https://doi.org/10.3390/ph19020301>
- [4] Serrano, D. R., Luciano, F. C., Anaya, B. J., Ongoren, B., Kara, A., Molina, G., . . . , & Lalatsa, A. (2024). Artificial Intelligence (AI) applications in drug discovery and drug delivery: Revolutionizing personalized medicine. *Pharmaceutics*, 16(10), 1328. <https://doi.org/10.3390/pharmaceutics16101328>
- [5] Joshi, S., & Sheth, S. (2025). Artificial intelligence (AI) in pharmaceutical formulation and dosage calculations. *Pharmaceutics*, 17(11), 1440. <https://doi.org/10.3390/pharmaceutics17111440>
- [6] Cheng, Y., Kong, J., Liu, X., & Li, S. (2026). Recent advances and emerging directions in machine learning-based breast cancer drug discovery: A comprehensive review. *Breast Cancer: Targets and Therapy*, 18, 1–21. <https://doi.org/10.2147/BCTT.S586786>
- [7] Liu, X., Mao, X., Zhang, Y., & Zhou, D. (2025). Novel synthetic strategy for aromatic ether coumarin derivatives and evaluation of anti-allergic potential. *BMC Chemistry*, 20(1), 1–10. <https://doi.org/10.1186/s13065-025-01687-9>
- [8] Pattanayak, P., & Rout, S. S. (2024). Anthelmintic and anti-filarial coumarins of natural and synthetic origin: Chemical and biological intervention. *Results in Chemistry*, 7, 1–16. <https://doi.org/10.1016/j.rechem.2024.101327>
- [9] Younes, A. H., & Mustafa, Y. F. (2024). Novel coumarins from green sweet bell pepper seeds: Their isolation, characterization, oxidative stress-mitigating, anticancer, anti-inflammatory, and antidiabetic properties. *Journal of Molecular Structure*, 1312, 138629. <https://doi.org/10.1016/j.molstruc.2024.138629>
- [10] Waheed, S. A., & Mustafa, Y. F. (2022). Novel naphthalene-derived coumarin composites: Synthesis, antibacterial, and antifungal activity assessments. *Eurasian Chemical Communications*, 4(8), 709–24. <https://doi.org/10.22034/ecc.2022.335455.1396>
- [11] Mishra, S., Pandey, A., & Manvati, S. (2020). Coumarin: An emerging antiviral agent. *Heliyon*, 6(1), 1–7. <https://doi.org/10.1016/j.heliyon.2020.e03217>
- [12] Ebaid, M. S., Korycka-Machala, M., Shaldam, M. A., Thabit, M. G., Kawka, M., Dziadek, B., & Sabt, A. (2025). Exploring antitubercular activity of new coumarin derivatives targeting enoyl acyl carrier protein reductase (InhA): Synthesis, biological evaluation and computational studies. *Journal of Molecular Structure*, 1336, 142074. <https://doi.org/10.1016/j.molstruc.2025.142074>
- [13] Mustafa, Y. F., Bashir, M. K., Oglah, M. K., Khalil, R. R., & Mohammed, E. T. (2021). Bioactivity of some natural and semisynthetic coumarin derived compounds. *NeuroQuantology*, 19(6), 129–139. <https://doi.org/10.14704/nq.2021.19.6.NQ21078>
- [14] Majnooni, M. B., Fakhri, S., Smeriglio, A., Trombetta, D., Croley, C. R., Bhattacharyya, P., . . . , & Bishayee, A. (2019). Antiangiogenic effects of coumarins against cancer: From chemistry to medicine. *Molecules*, 24(23), 4278. <https://doi.org/10.3390/molecules24234278>
- [15] Chavez Alvarez, A. C., & Moreau, E. (2026). Coumarin-based prodrugs: Therapeutic promise or still confined to preclinical exploration? *Pharmaceutics*, 18(3), 341. <https://doi.org/10.3390/pharmaceutics18030341>
- [16] Dinparast, L., & Dastmalchi, S. (2020). A QSAR study on the 4-substituted coumarins as potent tubulin polymerization

- inhibitors. *Advanced Pharmaceutical Bulletin*, 10(2), 271–277. <https://doi.org/10.34172/apb.2020.032>
- [17] Sharmah, H., Ahmed, L. A., Deka, B., Bharali, D., Gogoi, B., & Patowary, P. (2025). Exploring the potential of coumarin-indole hybrid scaffold as promising pharmacological agents: A review. *Discover Chemistry*, 2(1), 320. <https://doi.org/10.1007/s44371-025-00395-6>
- [18] Rowaiye, A. B., Folarin, A. A., Akingbade, T., Okoli, J. C., Rowaiye, O. I., Folorunso, T. R., & Bur, D. (2025). Advancing predictive modeling in computational chemistry through quantum chemistry, molecular mechanics, and machine learning. *Discover Chemistry*, 2(1), 281. <https://doi.org/10.1007/s44371-025-00363-0>
- [19] Zareen, W., Ahmed, N., Khan, A. M., Mali, S. N., Sadeghian, N., Aldawsari, N. A., . . . , & Shafiq, Z. (2025). Synthesis, antidiabetic evaluation, and computational modeling of 3-acetyl-8-ethoxy coumarin derived hydrazones and thiosemicarbazones. *RSC Advances*, 15(46), 39043–39058. <https://doi.org/10.1039/d5ra04619j>
- [20] Petzold, C. J., & Mukhopadhyay, A. (2025). From bench to biofactory: High-throughput technologies and automated workflows to accelerate biomanufacturing. *Current Opinion in Biotechnology*, 94, 1–9. <https://doi.org/10.1016/j.copbio.2025.103320>
- [21] Parascandolo, E., de Assis, Netto, C. G., Rodriguez-Donis, I., Thiebaud-Roux, S., dos Santos, M. T., & Gerbaud, V. (2026). Integrated product and process reverse engineering design combining adaptive molecular search and rigorous separation process validation. *Computers & Chemical Engineering*, 211, 1–16. <https://doi.org/10.1016/j.compchemeng.2026.109683>
- [22] Alemie, A. A., Siraj, E. A., Yayehrad, A. T., Tafere, C., Tessema, T. A., & Belete, A. (2025). Continuous pharmaceutical manufacturing and its contemporary regulatory insights. *Discover Applied Sciences*, 7(10), 1057. <https://doi.org/10.1007/s42452-025-07712-9>
- [23] Sarhan, M. O., Abd El-Karim, S. S., Anwar, M. M., Gouda, R. H., Zagharay, W. A., & Khedr, M. A. (2021). Discovery of new coumarin-based lead with potential anticancer, CDK4 inhibition and selective radiotheranostic effect: Synthesis, 2D & 3D QSAR, molecular dynamics, in vitro cytotoxicity, radioiodination, and biodistribution studies. *Molecules*, 26(8), 2273. <https://doi.org/10.3390/molecules26082273>
- [24] Soori, M., Arezoo, B., & Dastres, R. (2023). Digital twin for smart manufacturing, A review. *Sustainable Manufacturing and Service Economics*, 2, 100017. <https://doi.org/10.1016/j.smse.2023.100017>
- [25] Moshawih, S., Goh, H. P., Kifli, N., Idris, A. C., Yassin, H., Kotra, V., . . . , & Ming, L. C. (2022). Synergy between machine learning and natural products cheminformatics: Application to the lead discovery of anthraquinone derivatives. *Chemical Biology & Drug Design*, 100(2), 185–217. <https://doi.org/10.1111/cbdd.14062>
- [26] Chennai, H. Y., Belaidi, S., Ouassaf, M., Sinha, L., Prasad, O., Serdaroglu, G., & Chtita, S. (2025). Computational study of coumarin compounds as potential inhibitors of casein kinase 2: DFT, 2D-QSAR, ADMET and molecular docking investigations. *Polycyclic Aromatic Compounds*, 45(5), 843–862. <https://doi.org/10.1080/10406638.2024.2418408>
- [27] Udabe, J. (2026). A scientist's guide to AI-driven molecular discovery. *Artificial Intelligence Chemistry*, 4(1), 100107. <https://doi.org/10.1016/j.aichem.2026.100107>
- [28] Demyashkin, G., Parshenkov, M., Zyryanov, S., Yavorskiy, A., Shegai, P., & Kaprin, A. (2026). Artificial intelligence across the drug development lifecycle. *Medical Sciences*, 14(2), 248. <https://doi.org/10.3390/medsci14020248>
- [29] Annunziata, F., Pinna, C., Dallavalle, S., Tamborini, L., & Pinto, A. (2020). An overview of coumarin as a versatile and readily accessible scaffold with broad-ranging biological activities. *International Journal of Molecular Sciences*, 21(13), 4618. <https://doi.org/10.3390/ijms21134618>
- [30] Mohammed, S., Budach, L., Feuerpfeil, M., Ihde, N., Nathansen, A., Noack, N., . . . , & Harmouch, H. (2025). The effects of data quality on machine learning performance on tabular data. *Information Systems*, 132, 102549. <https://doi.org/10.1016/j.is.2025.102549>
- [31] Sheikh, M. R., Islam, F., Tasnim, J., Akter, A., Arka, A. D., Khan, A., . . . , & Sayed, M. A. (2026). In silico investigation of fenugreek-derived coumarin and its potential impact on human health. *Discover Chemistry*, 3(1), 264. <https://doi.org/10.1007/s44371-026-00710-9>
- [32] Fernández-Peña, L., Matos, M. J., & López, E. (2022). Recent advances in biologically active coumarins from marine sources: Synthesis and evaluation. *Marine Drugs*, 21(1), 37. <https://doi.org/10.3390/md21010037>
- [33] Pallikara, I., Skelton, J. M., Hatcher, L. E., & Pallipurath, A. R. (2024). Going beyond the ordered bulk: A perspective on the use of the Cambridge structural database for predictive materials design. *Crystal Growth & Design*, 24(17), 6911–6930. <https://doi.org/10.1021/acs.cgd.4c00694>
- [34] Fu, L., Shi, S., Yi, J., Wang, N., He, Y., Wu, Z., . . . , & Cao, D. (2024). ADMETlab 3.0: An updated comprehensive online ADMET prediction platform enhanced with broader coverage, improved performance, API functionality and decision support. *Nucleic Acids Research*, 52(W1), W422–W431. <https://doi.org/10.1093/nar/gkae236>
- [35] Niazi, S. K., & Mariam, Z. (2023). Recent advances in machine-learning-based chemoinformatics: A comprehensive review. *International Journal of Molecular Sciences*, 24(14), 11488. <https://doi.org/10.3390/ijms241411488>
- [36] Turkina, V., Messih, M. R., Kant, E., Gringhuis, J. T., Petrignani, A., Corthals, G., . . . , & Samanipour, S. (2026). Enhancing predictive modeling with molecular fingerprint fusion strategies. *Digital Discovery*, 5(5), 2232–2242. <https://doi.org/10.1039/D5DD00302D>
- [37] Rezazi, S., Si-Moussa, C., & Hanini, S. (2026). Integrative deep learning strategies to enhance early-stage drug discovery: Optimizing computational structure–Activity modeling for pharmacotherapeutic innovation. *Journal of Pharmacy & Pharmaceutical Sciences*, 29, 16155. <https://doi.org/10.3389/jpps.2026.16155>
- [38] Gangwal, A., & Lavecchia, A. (2024). Unleashing the power of generative AI in drug discovery. *Drug Discovery Today*, 29(6), 103992. <https://doi.org/10.1016/j.drudis.2024.103992>
- [39] Kang, S. I., Shin, J. H., Wu, B. M., & Choi, H. S. (2025). Deep generative AI for multi-target therapeutic design: Toward self-improving drug discovery framework. *International Journal of Molecular Sciences*, 26(23), 11443. <https://doi.org/10.3390/ijms262311443>
- [40] Bilotta, M., Rocca, R., & Alcaro, S. (2025). Next-generation drug discovery: The AI revolution in pharmaceutical research. *Artificial Intelligence in the Life Sciences*, 8, 100149. <https://doi.org/10.1016/j.ailsci.2025.100149>

- [41] Naz, N., Nazli, Z.-H., Parveen, S., & Shafiq, N. (2026). De novo drug design, synthesis, biological evaluation, and structural examination of novel coumarin-based pyrimidine co-drugs accompanied by molecular docking and DFT studies. *ACS Omega*, 11(9), 15233–15253. <https://doi.org/10.1021/acsomega.5c12395>
- [42] Kapustina, O., Burmakina, P., Gubina, N., Serov, N., & Vinogradov, V. (2024). User-friendly and industry-integrated AI for medicinal chemists and pharmaceuticals. *Artificial Intelligence Chemistry*, 2(2), 100072. <https://doi.org/10.1016/j.aichem.2024.100072>
- [43] Bang, D., Kim, J., Song, H., & Kim, S. (2025). ADME-drug-likeness: Enriching molecular foundation models via pharmacokinetics-guided multi-task learning for drug-likeness prediction. *Bioinformatics*, 41(Supplement_1), i352–i361. <https://doi.org/10.1093/bioinformatics/btaf259>
- [44] Liu, S., Li, Y., Wei, X., Zhang, R., Zhang, Y., & Guo, C. (2022). Computational investigations of coumarin derivatives as cyclin-dependent kinase 9 inhibitors using 3D-QSAR, molecular docking and molecular dynamics simulation. *Current Computer-Aided Drug Design*, 18(5), 363–380. <https://doi.org/10.2174/1573409918666220817100959>
- [45] Aarón, R. H., Sheila, C. M., Julio Emmanuel, G. P., Oscar, J. G., Aurelio, L. M., & Jocksan Ismael, M. C. (2025). In Silico strategies for drug discovery: Optimizing natural compounds from foods for therapeutic applications. *Discover Chemistry*, 2(1), 133. <https://doi.org/10.1007/s44371-025-00201-3>
- [46] Lv, L., Xu, J., Gao, X., & Li, B. (2026). Reinforcement learning-driven optimization method for adaptive data hierarchical classification strategy. *Discover Computing*, 29, 1–30. <https://doi.org/10.1007/s10791-026-09976-0>
- [47] Kong, L., Gan, Y., Liang, T., Zhong, L., Pan, Y., Kirsanov, D., ..., & Wang, P. (2020). A novel smartphone-based CD-spectrometer for high sensitive and cost-effective colorimetric detection of ascorbic acid. *Analytica Chimica Acta*, 1093, 150–159. <https://doi.org/10.1016/j.aca.2019.09.071>
- [48] Chen, L.-Y., & Li, Y.-P. (2024). Enhancing chemical synthesis: A two-stage deep neural network for predicting feasible reaction conditions. *Journal of Cheminformatics*, 16, 1–14. <https://doi.org/10.1186/s13321-024-00805-4>
- [49] Rani, S., Karnati, R., Patel, V., Ranganathaswamy, M. K., Tomar, P., Kataria, A., & Pal, A. (2026). AI-driven optimization techniques for smart sustainable manufacturing in Industry 5.0 ecosystem: A comprehensive review. *Alexandria Engineering Journal*, 137, 133–158. <https://doi.org/10.1016/j.aej.2026.01.016>
- [50] Alam, M. M., Al-Bahrani, H. A., Urazbaeva, D., Ali, F. A., Polatova, D., Kholmurodova, D., ..., & Khujaniyozova, O. (2025). Influence of machine learning technology on the development of electrochemical, optical, and image analysis-based methods for biomedical, food, and environmental analysis. *Microchemical Journal*, 218, 115407. <https://doi.org/10.1016/j.microc.2025.115407>
- [51] Hoque, A., Surve, M., Kalyanakrishnan, S., & Sunoj, R. B. (2024). Reinforcement learning for improving chemical reaction performance. *Journal of the American Chemical Society*, 146(41), 28250–28267. <https://doi.org/10.1021/jacs.4c08866>
- [52] Hariyani, D., Hariyani, P., Mishra, S., & Sharma, M. K. (2025). A literature review on transformative impacts of blockchain technology on manufacturing management and industrial engineering practices. *Green Technologies and Sustainability*, 3(3), 100169. <https://doi.org/10.1016/j.grets.2025.100169>
- [53] Liu, X., Fan, K., Huang, X., Ge, J., Liu, Y., & Kang, H. (2024). Recent advances in artificial intelligence boosting materials design for electrochemical energy storage. *Chemical Engineering Journal*, 490, 151625. <https://doi.org/10.1016/j.cej.2024.151625>
- [54] Chetry, A. B., & Ohto, K. (2025). From molecules to data: The emerging impact of chemoinformatics in chemistry. *Journal of Cheminformatics*, 17(1), 121. <https://doi.org/10.1186/s13321-025-00978-6>
- [55] Bloor, M., Ahmed, A., Kotecha, N., Mercangoz, M., Tsay, C., & del Rio-Chanona, E. A. (2025). Control-informed reinforcement learning for chemical processes. *Industrial & Engineering Chemistry Research*, 64(9), 4966–4978. <https://doi.org/10.1021/acs.iecr.4c03233>
- [56] Gnädinger, U., Poier, D., Trombini, C., Dabros, M., & Marti, R. (2024). Development of lab-scale continuous stirred-tank reactor as flow process tool for oxidation reactions using molecular oxygen. *Organic Process Research & Development*, 28(5), 1860–1868. <https://doi.org/10.1021/acs.oprd.3c00424>
- [57] Silva, T. F. S., & Martins, L. M. D. R. S. (2020). Recent advances in copper catalyzed alcohol oxidation in homogeneous medium. *Molecules*, 25(3), 748. <https://doi.org/10.3390/molecules25030748>
- [58] Rojek, I., Mikołajewski, D., Dostatni, E., Cybulski, J., & Kozielski, M. (2025). Personalization of AI-based digital twins to optimize adaptation in industrial design and manufacturing—Review. *Applied Sciences*, 15(15), 8525. <https://doi.org/10.3390/app15158525>
- [59] Fudyma, J., Kura, Ł., & Gębicki, J. (2025). Development of digital twins for continuous processes: Concept description of virtual mass balance based on the Tennessee Eastman process. *Processes*, 13(2), 337. <https://doi.org/10.3390/pr13020337>
- [60] Raihan, A. (2026). Synergistic integration of digital twins and artificial intelligence for sustainable energy and environmental systems: A comprehensive review. *Sustainable Cities and Society*, 2(1), 100024. <https://doi.org/10.1016/j.scsadv.2026.100024>
- [61] Neto, J. R. T., Capron, B. D. O., & Secchi, A. R. (2025). A robust deep reinforcement learning approach for the control of crystallization processes. *Computers & Chemical Engineering*, 199, 109114. <https://doi.org/10.1016/j.compchemeng.2025.109114>
- [62] Zhang, S., Bi, K., & Qiu, T. (2020). Bidirectional recurrent neural network-based chemical process fault diagnosis. *Industrial & Engineering Chemistry Research*, 59(2), 824–834. <https://doi.org/10.1021/acs.iecr.9b05885>
- [63] El Mokhtari, K., & McArthur, J. J. (2024). Autoencoder-based fault detection using building automation system data. *Advanced Engineering Informatics*, 62, 102810. <https://doi.org/10.1016/j.aei.2024.102810>
- [64] Koyanagi, K., Ueno, A., Sasaki, T., & Otsuka, M. (2022). Real-time monitoring of critical quality attributes during high-shear wet granulation process by near-infrared spectroscopy: Effect of water addition and stirring speed on pharmaceutical properties of the granules. *Pharmaceuticals*, 15(7), 822. <https://doi.org/10.3390/ph15070822>
- [65] Fu, Q., Nicholson, G. L., & Easton, J. M. (2024). Understanding data quality in a data-driven industry context: Insights from the fundamentals. *Journal of Industrial*

- Information Integration, 42, 100729. <https://doi.org/10.1016/j.jii.2024.100729>
- [66] Lin, N., Wen, X., Wu, J., Huang, X., & Wen, H. (2026). Intelligent monitoring and anomaly detection for power service processes based on spatiotemporal attention mechanism. *Scientific Reports*, 16, 12479. <https://doi.org/10.1038/s41598-026-42189-5>
- [67] Barucci, A., Colcelli, V., De Masi, S., Falconi, M., Leo, M. C., Marzola, A., ..., & Pini, R. (2026). Ethical and regulatory frameworks for artificial intelligence in clinical research: A European perspective on the artificial intelligence act for ethics committees and researchers. *European Cardiology Review*, 21, e01. <https://doi.org/10.15420/ecr.2025.59>
- [68] Elfer, K., Gardecki, E., Garcia, V., Ly, A., Hytopoulos, E., Wen, S., ..., & Gallas, B. D. (2024). Reproducible reporting of the collection and evaluation of annotations for artificial intelligence models. *Modern Pathology*, 37(4), 100439. <https://doi.org/10.1016/j.modpat.2024.100439>
- [69] Schaduengrat, N., Lampa, S., Simeon, S., Gleeson, M. P., Spjuth, O., & Nantasenamat, C. (2020). Towards reproducible computational drug discovery. *Journal of Cheminformatics*, 12, 9. <https://doi.org/10.1186/s13321-020-0408-x>
- [70] Qadri, Y. A., Shaikh, S., Ahmad, K., Choi, I., Kim, S. W., & Vasilakos, A. V. (2025). Explainable artificial intelligence: A perspective on drug discovery. *Pharmaceutics*, 17(9), 1119. <https://doi.org/10.3390/pharmaceutics17091119>
- [71] Tian, K., Zhu, Z., Mbachui, J. I. C., Ghanbaripour, A., & Moorhead, M. (2025). Artificial intelligence in risk management within the realm of construction projects: A bibliometric analysis and systematic literature review. *Journal of Innovation & Knowledge*, 10(3), 100711. <https://doi.org/10.1016/j.jik.2025.100711>
- [72] Alharthi, S. (2026). AI-powered in silico twins: Redefining precision medicine through simulation, personalization, and predictive healthcare. *Saudi Pharmaceutical Journal*, 34(1), 1. <https://doi.org/10.1007/s44446-025-00055-x>
- [73] Sampathkumar, K., & Kerwin, B. A. (2024). Roadmap for drug product development and manufacturing of biologics. *Journal of Pharmaceutical Sciences*, 113(2), 314–331. <https://doi.org/10.1016/j.xphs.2023.11.004>
- [74] Hu, W., Ou, Y., Liu, H., Ni, P., & Chang, C. (2026). Integrating digital twin technologies for maintenance 4.0 in the building industry: A review and conceptual framework. *Building and Environment*, 288, 113997. <https://doi.org/10.1016/j.buildenv.2025.113997>
- [75] Nagar, A., Gobburu, J., & Chakravarty, A. (2025). Artificial intelligence in pharmacovigilance: Advancing drug safety monitoring and regulatory integration. *Therapeutic Advances in Drug Safety*, 16, 20420986251361435. <https://doi.org/10.1177/20420986251361435>
- [76] Ara, J., Bhuiyan, H., Roza, I. I., & Nahin, A. S. M. (2026). Importance of balanced datasets with feature selection and ensemble methods on heart disease classification using distinctive machine learning techniques: A comparative analysis. *Scientific Reports*, 16, 11706. <https://doi.org/10.1038/s41598-026-47691-4>
- [77] Alam, F., Mohammed Alnazzawi, T. S., Mehmood, R., & Al-Maghthawi, A. (2025). A review of the applications, benefits, and challenges of generative AI for sustainable toxicology. *Current Research in Toxicology*, 8, 100232. <https://doi.org/10.1016/j.crttox.2025.100232>
- [78] Cappuccio, E., Kathirgamanathan, B., Rinzivillo, S., Andrienko, G., & Andrienko, N. (2025). Integrating human knowledge for explainable AI. *Machine Learning*, 114(11), 250. <https://doi.org/10.1007/s10994-025-06879-x>
- [79] Farah, L., Murriss, J. M., Borget, I., Guilloux, A., Martelli, N. M., & Katsahian, S. I. M. (2023). Assessment of performance, interpretability, and explainability in artificial intelligence-based health technologies: What healthcare stakeholders need to know. *Mayo Clinic Proceedings: Digital Health*, 1(2), 120–138. <https://doi.org/10.1016/j.mcpdig.2023.02.004>
- [80] Jayabal, R., Subbiah, G., Tripathy, S., SVH, N., Kalidhas, A. M., Vadnagra, K. V., ..., & Smerat, A. (2026). Artificial intelligence-enabled process safety: Advances, challenges, and future directions for high-risk industrial systems. *Journal of Loss Prevention in the Process Industries*, 102, 106026. <https://doi.org/10.1016/j.jlp.2026.106026>
- [81] Szwaczko, K., Kulkowska, A., & Matwijczuk, A. (2026). Advances in coumarin fluorescent probes for medical diagnostics: A review of recent developments. *Biosensors*, 16(1), 36. <https://doi.org/10.3390/bios16010036>
- [82] Zhang, J., Li, T., Yan, C., Li, Y., Yu, Y., Wang, J., & Sun, B. (2026). Machine learning approaches for compound–target interaction prediction: A review. *Foods*, 15(9), 1582. <https://doi.org/10.3390/foods15091582>
- [83] Kaur, R. P., & Manjh, S. (2026). Machine learning based software tools and computational strategies for QSAR modeling. *Discover Chemistry*, 3, 200. <https://doi.org/10.1007/s44371-026-00663-z>
- [84] Cardona-Galeano, W., Ramirez-Malule, H., & Gómez-Ríos, D. (2023). Hybrids based on coumarins and their anticancer activities: A bibliometric analysis. *Journal of Applied Pharmaceutical Science*, 13(9), 204–212. <https://doi.org/10.7324/JAPS.2023.150232>
- [85] Mun, C. S., Hui, L. Y., Sing, L. C., Karunakaran, R., & Ravichandran, V. (2022). Multi-targeted molecular docking, pharmacokinetics, and drug-likeness evaluation of coumarin based compounds targeting proteins involved in development of COVID-19. *Saudi Journal of Biological Sciences*, 29(12), 103458. <https://doi.org/10.1016/j.sjbs.2022.103458>
- [86] Seraphin, T. P., Mesropian, A., Žigutyte, L., Brooks, J., Mauro, E., Gris-Oliver, A., ..., & Llovet, J. M. (2025). Artificial intelligence predicts outcome-related molecular profiles and vascular invasion in hepatocellular carcinoma. *JHEP Reports*, 7(12), 101592. <https://doi.org/10.1016/j.jhepr.2025.101592>
- [87] Malviya, J., Alameri, A. A., Al-Janabi, S. S., Fawzi, O. F., Azzawi, A. L., Obaid, R. F., ..., & Mirzaei, R. (2023). Metabolomic profiling of bacterial biofilm: Trends, challenges, and an emerging antibiofilm target. *World Journal of Microbiology and Biotechnology*, 39(8), 212. <https://doi.org/10.1007/s11274-023-03651-y>
- [88] Sofi, F. A. (2026). Coumarin as a privileged natural product scaffold towards the development of antibacterial agents targeting ESKAPE pathogens. *European Journal of Medicinal Chemistry*, 309, 118735. <https://doi.org/10.1016/j.ejmech.2026.118735>

How to Cite: Mustafa, Y. F. (2026). From Bench to Algorithms: Artificial Intelligence Applications in the Chemical Engineering of Coumarin-Based Therapeutics. *Artificial Intelligence and Applications*. <https://doi.org/10.47852/bonviewAIA620210496>