RSC Medicinal Chemistry

ROYAL SOCIETY OF CHEMISTRY

EDITORIAL

View Article Online
View Journal | View Issue

Cite this: *RSC Med. Chem.*, 2024, **15**, 3284

Introduction to the themed collection on 'AI in Medicinal Chemistry'

Jian Zhang, *[a] Ola Engkvist[bc] and Gerhard Hessler[d]

DOI: 10.1039/d4md90035a

rsc.li/medchem

The exploration and development of novel pharmaceutical agents require substantial time and financial investment, especially during the lead optimization phase of drug discovery. In recent years, advancements in artificial intelligence (AI) have emerged as pivotal tools in expediting various aspects of medicinal chemistry research. For example, AI algorithms are instrumental from the initial stages of hit confirmation to sophisticated late-stage route optimization, significantly enhancing the efficiency of retrosynthesis and the selection of reaction conditions for new coupling partners. Furthermore, the generation of novel compounds with an optimized property profile presents a critical challenge in medicinal chemistry. Employing AI-based methods offers promising avenues for designing potent compounds with therapeutic potential. Meanwhile, deep learning methods also contribute to critical property identification associated with cells.

We are pleased to present a compilation of reviews and original research articles that span various important facets of AI, along with experimental evidence. The collection highlights advanced approaches for prediction of molecular properties. While Fraslish *et al.* (**https://doi.org/10.1039/D4MD00325J**) illustrate the potential of "DeltaClassifier", Chen *et al.* (**https://doi.org/10.1039/D4MD00423J**) use a transformer architecture to select R-groups for advancing analogue series. The contribution from McCorkindale *et al.* (**https://doi.org/10.1039/D3MD00719G**) employs a consensus approach between a random forest and a Gaussian process model to deconvolute low yield compounds from low potency in screening results of crude reaction mixtures. Furthermore, the concept of counterfactuals (CFs) in machine learning is utilized for multitasking predictions across different classes of protein kinase inhibitors by Lamens and Bajorath (**https://doi.org/10.1039/D4MD00128A**).

Despite considerable progress in integrating AI within medicinal chemistry, numerous algorithms still need further exploration or development. The following case studies, supported by experimental evidence, highlight the potential for these AI methods to capture future medicinal chemistry advancements:

1. Retrosynthesis: the AiZynth tool has exerted considerable influence on drug discovery processes. Shields *et al.* (**https://doi.org/10.1039/D3MD00651D**) illustrate very well the progress that has been made in synthesis prediction during the last few years as well as highlighting the improvement opportunities for the future. A very important field in the future will be to combine AI and chemistry automation. Atz *et al.* (**https://doi.org/10.1039/D4MD00196F**) have applied geometric deep learning to predict optimal reaction conditions for Suzuki coupling, one of the most common reactions in medicinal chemistry. Together, these applications show the power of artificial intelligence to enhance the efficiency of synthesis prediction.

2. Compound generation and prediction: utilizing chemical language models (CLMs) for encoding and tokenization schemes enhances compound representation and generation. AI has mainly been applied to small molecules. Orsi and Reymond (**https://doi.org/10.1039/D4MD00159A**) show in an instructive example of how large language models (LLMs) can be employed to encode antimicrobial peptides (AMPs) and predict their biological properties, including antimicrobial efficacy and hemolysis potential.

3. Compound design: approaches in compound design include the use of molecular pairing techniques to process extensive inhibition data. In a noteworthy application, Hazemann *et al.* (**https://doi.org/10.1039/**

[a] *Medicinal Chemistry and Bioinformatics Center, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China. E-mail: jian.zhang@sjtu.edu.cn*
[b] *Molecular AI, Discovery Sciences, R&D AstraZeneca Gothenburg, Sweden*
[c] *Department of Computer Science and Engineering, Chalmers University of Technology Gothenburg, Sweden*
[d] *Research Platform, Integrated Drug Discovery, Sanofi, Industriepark Höchst, 65926 Frankfurt am Main, Germany*

D4MD00106K) combined generative molecular design, deep reinforcement learning and pharmacophore matching to identify the main protease (Mpro) inhibitors for SARS-CoV-2.

4. Cell painting assay: an important future research field is to combine AI with complex assay readouts to better understand biology. Tandon *et al.* (**https://doi.org/10.1039/D4MD00107A**) used explainable machine learning (XML) to delineate critical properties associated with lysosomotropism in compounds, enhancing the interpretability of high-content imaging assays.

5. Biological evaluation: machine learning-based virtual screening (VS) techniques have been instrumental in discovering inhibitors for the Son of Sevenless 1 (SOS1) protein. Additionally, a machine learning enhanced yield assay deconfounder has been developed by Duo *et al.* (**https://doi.org/10.1039/D4MD00063C**) to distinguish between low yield and low potency in screening assays, thereby identifying potential false negatives more effectively.

This curated collection of algorithms, substantiated by experimental evidence, covers the recent strides made in the integration of AI within medicinal chemistry. The discussions encompass both the challenges and the research opportunities within this rapidly evolving field. These works illuminate the varied impacts of AI on medicinal chemistry, including aspects such as compound optimization, compound representation, compound generation, compound toxicity prediction, and bioinformation. As these selected articles demonstrate the progression of present innovations, it is evident that the application of AI algorithms in medicinal chemistry will continue to expand. Future advancements are anticipated to further enhance the capabilities in this domain, thereby creating new insights for research and application in medicinal chemistry.