

RESEARCH

Open Access



# Syn-MolOpt: a synthesis planning-driven molecular optimization method using data-derived functional reaction templates

Xiaodan Yin<sup>1,2</sup>, Xiaorui Wang<sup>1</sup>, Zhenxing Wu<sup>1</sup>, Qin Li<sup>3</sup>, Yu Kang<sup>1</sup>, Yafeng Deng<sup>4</sup>, Pei Luo<sup>5</sup>, Huanxiang Liu<sup>3</sup>, Guqin Shi<sup>6</sup>, Zheng Wang<sup>6</sup>, Xiaojun Yao<sup>3\*</sup>, Chang-Yu Hsieh<sup>1\*</sup> and Tingjun Hou<sup>1\*</sup>

## Abstract

Molecular optimization is a crucial step in drug development, involving structural modifications to improve the desired properties of drug candidates. Although many deep-learning-based molecular optimization algorithms have been proposed and may perform well on benchmarks, they usually do not pay sufficient attention to the synthesizability of molecules, resulting in optimized compounds difficult to be synthesized. To address this issue, we first developed a general pipeline capable of constructing functional reaction template library specific to any property where a predictive model can be built. Based on these functional templates, we introduced Syn-MolOpt, a synthesis planning-oriented molecular optimization method. During optimization, functional reaction templates steer the process towards specific properties by effectively transforming relevant structural fragments. In four diverse tasks, including two toxicity-related (GSK3 $\beta$ -Mutagenicity and GSK3 $\beta$ -hERG) and two metabolism-related (GSK3 $\beta$ -CYP3A4 and GSK3 $\beta$ -CYP2C19) multi-property molecular optimizations, Syn-MolOpt outperformed three benchmark models (Modof, HierG2G, and SynNet), highlighting its efficacy and adaptability. Additionally, visualization of the synthetic routes for molecules optimized by Syn-MolOpt confirms the effectiveness of functional reaction templates in molecular optimization. Notably, Syn-MolOpt's robust performance in scenarios with limited scoring accuracy demonstrates its potential for real-world molecular optimization applications. By considering both optimization and synthesizability, Syn-MolOpt promises to be a valuable tool in molecular optimization.

**Scientific contribution** Syn-MolOpt takes into account both molecular optimization and synthesis, allowing for the design of property-specific functional reaction template libraries for the properties to be optimized, and providing reference synthesis routes for the optimized compounds while optimizing the targeted properties. Syn-MolOpt's universal workflow makes it suitable for various types of molecular optimization tasks.

**Keywords** Molecular optimization, Synthetic planning, Reaction template, Toxicity optimization, Metabolic property optimization

\*Correspondence:

Xiaojun Yao  
xjyao@mpu.edu.mo  
Chang-Yu Hsieh  
kimhsieh@zju.edu.cn  
Tingjun Hou  
tingjunhou@zju.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

## Introduction

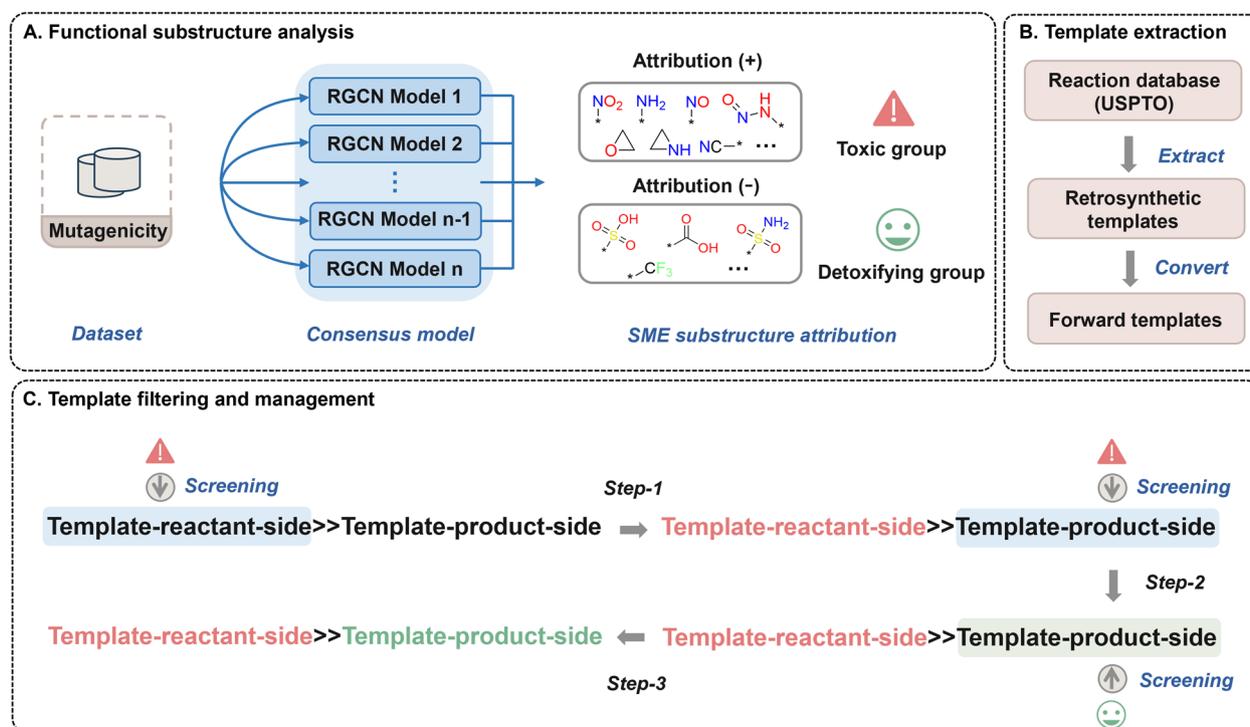
Drug development is a lengthy and complex process, primarily focused on the efficient identification of novel molecules with desirable properties [1–3]. Traditionally, this task heavily relied on human expertise in designing, synthesizing, and evaluating new molecular entities [4]. Although this approach has yielded many scientific breakthroughs, the high costs and time-consuming nature of traditional methods constrain the exploration of the chemical space. However, the recent advent of deep learning (DL) technology has propelled the use of generative models in molecular design, significantly enhancing the efficiency of searching the chemical space for optimal compound structures [5, 6]. Despite notable advancements, examples of these models successfully discovering molecules that satisfy a range of predefined properties remain scarce [7, 8]. Hence, structural modifications of candidate molecules are often essential to further refine their desired properties, emphasizing the crucial role of molecular optimization in drug development.

In recent years, many DL-based molecular optimization algorithms have been proposed, often delivering impressive results on *in silico* benchmarks. However, most of these methods face a severe drawback: insufficient consideration of molecular synthesizability [9, 10], which poses a formidable challenge for physically testing these novel designs. Especially during lead optimization, researchers usually rely on algorithmic predictions to modify the structures of candidate molecules. However, even slight structural changes can complicate synthesis, thereby stalling drug development. Therefore, it is important to prioritize molecular synthesizability in optimization algorithm design, enhancing drug development efficiency and ensuring the practicality of these algorithms. To tackle this issue, some approaches have incorporated synthetic accessibility (SA) scores as a synthesizability metric into molecular optimization algorithms [11, 12]. While this strategy facilitates rapid screening and assessment of compounds, its reliance on predefined chemical and heuristic rules [13] limits prediction accuracy and fails to provide practical synthesis pathways. The challenge of devising feasible synthesis routes for new molecules is a major obstacle in validating molecular optimization algorithms [14]. Fortunately, computer-assisted synthesis planning (CASP) [15–25] offers a solution by inputting the target molecule's structure and outputting feasible reaction pathways. However, due to the exhaustive search within vast chemical spaces, CASP algorithms usually require seconds to minutes to plan the synthesis routes for a target molecule. This renders post-filtering strategies, which separate optimization from synthesis planning, less ideal for molecular optimization workflows [9]. Consequently, an integrated

molecular optimization method incorporating synthesis planning has emerged as a promising alternative, offering a streamlined pathway from calculation to synthesis.

Recently, numerous synthesis planning-driven molecular design strategies have emerged, categorized broadly into template-free and template-based approaches. Template-free methods [26–28] directly model chemical reactions using data-driven algorithms to predict reaction outcomes, offering enhanced flexibility and a wider exploration space. However, compared to template-based methods, they may experience reduced prediction accuracy, particularly when faced with out-of-domain data [10]. Template-based methods [14, 29–32] use reaction templates from scientific literature to quickly identify potential reaction pathways, providing more chemically feasible routes. However, they face challenges due to limited template coverage. For instance, SynNet, a molecular design algorithm developed by Gao et al. [14], utilizes 91 publicly available reaction templates, these templates primarily focused on molecular skeleton construction and peripheral modifications. However, there is a potential concern that these templates may not include functional reaction templates specifically tailored for optimizing specific properties. In drug design, medicinal chemists prefer to analyze molecular properties and reactivity through chemically significant fragments, rather than isolated atoms or bonds. Optimizing lead compounds often involves strategic substitution of bioisosteres and functional groups. This fragment-oriented approach aligns with chemical logic and provides clear guidance for molecular optimization. Therefore, crafting functional reaction templates that can precisely transform or modify specific fragments stands out as a promising avenue for precise optimization of molecular properties. Moreover, on multi-property molecular optimization, amalgamating various optimization goals into a composite objective function is common. This involves determining the relative importance of each goal, yet the ability to explore trade-offs among objectives is limited [33]. In synthesis planning-driven molecular optimization, the strategic utilization of functional reaction templates to transform pertinent fragments integrates structural optimization into the construction of synthesis tree. This approach extends beyond sole reliance on the guidance and constraints of the target composite function, providing a practical solution to address this limitation.

In this study, we back up our claim by introducing Syn-MolOpt, a synthesis planning-driven molecular optimization framework based on data-derived functional reaction templates. Initially, utilizing the substructure mask explanation method (SME) [34], we introduce a general pipeline that can design property-specific functional reaction templates for any property



**Fig. 1** Overview of the functional reaction template library design process. **A** Functional substructure analysis: using SME to analyze the impacts of substructures on specific property; **B** Reaction template extraction: extracting reaction templates from open-source reaction datasets using RDChiral; and **C** Reaction template screening and management: screening reaction templates by substructure matching

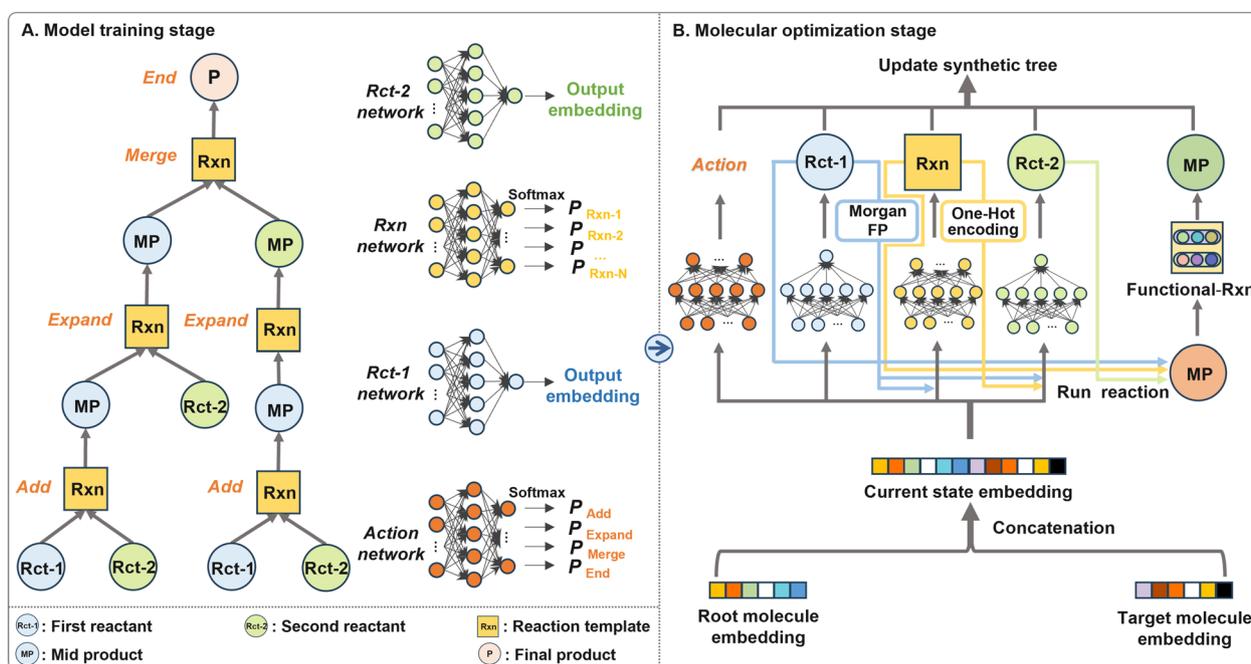
for which a predictive model can be built. During molecule optimization, these templates are used to guide optimization towards specified properties by effectively transforming relevant structural fragments. Currently, most molecular optimization algorithms predominantly focus on enhancing bioactivity and certain well-modeled properties of candidate molecules, such as logP and QED [35–37]. However, the toxicity [38, 39] and metabolic properties [40, 41] of molecules are critical in assessing their safety and clinical potential, and therefore require careful consideration. Therefore, this study identifies two toxicity and two metabolic characteristics as properties requiring optimization. The results demonstrate that our approach exhibits promising capabilities on multi-property optimization compared to existing methods. Moreover, our model demonstrates remarkable robustness in scenarios where the scoring feedback for properties undergoing optimization is not sufficiently accurate, highlighting its potential in practical molecular optimization applications. Syn-MolOpt takes into account both molecular optimization and synthesis, with its universal workflow suitable for optimizing various molecular properties, promising to emerge as an effective tool in molecular optimization.

## Results and discussion

### The overall workflow of Syn-MolOpt

In this study, we introduce Syn-MolOpt, a synthesis planning-driven molecular optimization method based on the data-derived property-specific functional reaction template library (Sect. "The Construction of Domain-Relevant Functional Substructure Datasets and Functional Reaction Template Libraries") to precisely modify or transform specific structural fragments. Syn-MolOpt can provide reference synthetic pathways for proposed molecules with optimized properties. Moreover, the Syn-MolOpt framework can be employed to build optimization models for any property that is amenable to quantitative structure–activity relationship (QSAR) modeling, making it suitable for molecular optimization tasks across various domains. A detailed explanation of the Syn-MolOpt construction process is outlined below.

*Design of functional reaction template library* Figure 1 illustrates the process of creating a library of functional reaction templates specific to certain properties using Syn-MolOpt. With sufficient training data for a property predictor, we can design customized functional reaction templates for that particular property. This process initially relies on a molecular dataset, taking mutagenicity as an example, to construct a consensus model



**Fig. 2** The overall workflow of Syn-MolOpt. **A** Model training stage; and **B** molecular optimization stage. During the model training stage, four neural networks were trained to predict the reaction action, the first reactant, the reaction template, and the second reactant. In the molecular optimization stage, these neural networks were used to predict the aforementioned four reaction elements respectively

for property prediction through the Relational Graph Convolutional Network (RGCN) [42] algorithm. It is important to note that the molecular dataset used for building a consensus model needs to have both a sufficient quantity and high quality of data to ensure effective QSAR modeling. Utilizing our previously introduced perturbation-based SME explanation method [34], the molecules in the dataset are dissected into various substructures, including BRICS substructures [43], Murcko scaffolds [44], and functional groups. Each substructure is then assigned a contribution value, denoting its influence on the specified property, as depicted in Fig. 1A. A description of SME method can be found in the Supplementary Information. This process ultimately results in a mutagenicity functional substructure dataset with attributed values, where a positive attribution signifies a promoting role in mutagenicity, whereas a negative attribution indicates a detoxifying effect. As illustrated in Fig. 1B, general SMARTS retrosynthetic reaction templates are extracted from an open-source reaction dataset using the RDKit wrapper RDChiral [45], and subsequently transformed into forward reaction templates. This study used the USPTO reaction dataset, which has undergone preliminary atom mapping with rxnmapper [46]. Finally, as shown in Fig. 1C, the reaction templates extracted in the previous step are filtered and managed in three steps. Firstly, positively attributed substructures

(toxic groups) are employed to screen the reactant-side of the above reaction templates via substructure matching, selecting those with toxic groups. Secondly, the filtered templates are further screened on the product-side using the same positively attributed substructures, but excluding templates with toxic groups on the product-side. This yields templates that successfully transform mutagenic substructures on the reactant-side. Thirdly, negatively attributed substructures (detoxifying groups) are used to filter the product-side of the reaction templates obtained in the second stage, selecting those with detoxifying groups. The resulting templates contain toxic groups on the reactant-side and detoxifying groups on the product-side, enabling the conversion of toxic groups. It is worth noting that functional templates often emerge from the first two steps, and the need of the third step depends on the availability of suitable templates. Since there might be overlap or encompassing of templates, manual intervention is required to ensure the independence and practicality of each template. Detailed information about the manual intervention can be found in the Supplementary Information.

**Implementation of Syn-MolOpt** As illustrated in Fig. 2, the implementation of Syn-MolOpt consists of two stages: model training and molecular optimization. Syn-MolOpt models the synthesis pathway of a compound as a bottom-up synthesis tree (Fig. 2A), which each step

modeled as a Markov decision process. This is achieved through training four neural networks (Fig. 2A) for the reaction action ( $R_{act}$ ), the first reactant ( $R_{rct-1}$ ), the reaction template ( $R_{rxn}$ ), and the second reactant ( $R_{rct-2}$ ). The performance of these four neural networks are detailed in Table S5 in the Supplementary Information. Subsequently, in the molecular optimization phase (Fig. 2B), these networks are used to predict  $R_{act}$ ,  $R_{rct-1}$ ,  $R_{rxn}$ , and  $R_{rct-2}$ . The predicted reactants,  $R_{rct-1}$  and  $R_{rct-2}$  (or  $R_{rct-1}$  in the case of a unimolecular reaction template), undergo a reaction according to the predicted  $R_{rxn}$ . If the resulting mid-product matches a functional reaction template, functional processing is carried out. The synthesis tree is updated after each reaction step. Beyond utilizing functional reaction templates for structural modifications, a genetic algorithm (GA) is also used to numerically optimize the embeddings of the root molecules of the synthesis tree. The synthesis tree generator then decodes these optimized vectors to produce synthesizable molecules. This iterative process continues until molecules with the desired properties are generated. A detailed description of the model's framework is provided in Sect. "Model architecture".

#### The construction of domain-relevant functional substructure datasets and functional reaction template libraries

To create customized functional reaction template libraries for specific properties, we initially utilized the SME method to build domain-specific functional substructure datasets focused on optimizing desired properties. In this study, the properties targeted for optimization include two types of toxicity and two types of metabolic properties: mutagenicity (Mutag), hERG-related cardiotoxicity (hERG), CYP3A4, and CYP2C19. We constructed four consensus models, each exhibiting outstanding performance on the test set, with the ROC-AUC scores of 0.901, 0.862, 0.913, and 0.909 for Mutag, hERG, CYP3A4, and CYP2C19, respectively. Utilizing these consensus models, SME calculates the contributions of substructures to the predicted outcomes, enabling the construction of datasets comprising substructures that significantly impact the molecular properties under optimization. Functional groups, predefined as molecular substructures, are evaluated for their property contributions based on their attribution values across the entire dataset. We analyzed the attributions of functional groups that occurred at least 10 times in the Mutag, hERG, CYP3A4, and CYP2C19 datasets. Table S1 shows the detailed information of the representative functional groups with positive and negative attributions in each dataset. In the Mutag and hERG datasets, a positive attribution means that the functional group enhances toxicity, whereas a negative attribution

**Table 1** Details of the molecules to be optimized on four multi-property optimization experiments

Experiment	Initial mean score <sup>a</sup>	
GSK3 $\beta$ -Mutag	GSK3 $\beta$ : 0.787 $\pm$ 0.109	Mutag: 0.887 $\pm$ 0.055
GSK3 $\beta$ -hERG	GSK3 $\beta$ : 0.782 $\pm$ 0.105	hERG: 0.807 $\pm$ 0.094
GSK3 $\beta$ -CYP3A4	GSK3 $\beta$ : 0.772 $\pm$ 0.090	CYP3A4: 0.949 $\pm$ 0.021
GSK3 $\beta$ -CYP2C19	GSK3 $\beta$ : 0.767 $\pm$ 0.086	CYP2C19: 0.846 $\pm$ 0.192

<sup>a</sup> The score is the probability value

denotes its detoxifying effect. In the CYP3A4 and CYP2C19 datasets, a positive value suggests the functional group is favorable to metabolic inhibition, while a negative attribution indicates that it helps reduce metabolic inhibition. Our analysis of the functional substructure datasets for Mutag and hERG revealed considerable overlap in both positively and negatively attributed substructures. A similar situation was observed in the functional substructure datasets for CYP3A4 and CYP2C19. Therefore, we integrated the functional substructure datasets for Mutag and hERG, and, following the process of building the functional reaction template library, extracted a set of universal reaction templates that can effectively mitigate both toxicity types. Similarly, we extracted general reaction templates to reduce the inhibitory effects of compounds on CYP3A4 and CYP2C19.

#### The application of Syn-MolOpt on toxicity optimization

Given that many candidate molecules exhibit promising binding affinities but fall short in drug development due to excessive toxicity [47], we conducted multi-property optimization experiments to sustain or enhance the molecular binding affinity to glycogen synthase kinase 3 beta (GSK3 $\beta$ ) while reducing Mutag and hERG risks. Specifically, for molecules that bind well to GSK3 $\beta$  but possess toxicity, the goal is to optimize them into non-toxic molecules that exhibit even greater binding capability to GSK3 $\beta$  (evidenced by a high GSK3 $\beta$  score and low Mutag/hERG scores). To validate the effectiveness of Syn-MolOpt on multi-property optimization, we conducted a comparative analysis with three established models: HierG2G [48], Modof [35], and SynNet [14]. Notably, Modof and HierG2G have been reported to perform well on multi-property optimization tasks (DRD2 and QED), while SynNet is a molecular design strategy driven by synthesis planning. The Mutag and hERG scores predicted by the corresponding consensus models represent the likelihood of the toxicity of a compound. Meanwhile, the GSK3 $\beta$  score predicted through the Therapeutic Data Common (TDC) interface [49] indicates the probability that a compound is an inhibitor of GSK3 $\beta$ . Table 1 provides the details of the molecules to be optimized on four multi-property optimization

**Table 2** Overall comparison on multi-property optimization of GSK3 $\beta$ -Mutag

Methods	Property	Top-1 <sup>a</sup>	Top-10	Top-100	Top-128	Syn-ratio <sup>b</sup> $\uparrow$
Modof	GSK3 $\beta$ $\uparrow$	0.870	0.852 $\pm$ 0.047	0.762 $\pm$ 0.134	0.736 $\pm$ 0.138	0.500
	Mutag $\downarrow$	0.057	<b>0.080 <math>\pm</math> 0.069</b>	0.445 $\pm$ 0.376	0.527 $\pm$ 0.387	
HierG2G	GSK3 $\beta$ $\uparrow$	0.830	0.872 $\pm$ 0.105	0.498 $\pm$ 0.277	0.474 $\pm$ 0.286	0.651
	Mutag $\downarrow$	0.092	0.396 $\pm$ 0.268	0.417 $\pm$ 0.321	0.412 $\pm$ 0.319	
SynNet	GSK3 $\beta$ $\uparrow$	0.860	0.838 $\pm$ 0.022	0.787 $\pm$ 0.036	0.780 $\pm$ 0.038	1.000 (0.695) <sup>c</sup>
	Mutag $\downarrow$	0.080	0.093 $\pm$ 0.045	0.134 $\pm$ 0.068	0.142 $\pm$ 0.072	
Syn-MolOpt	GSK3 $\beta$ $\uparrow$	<b>0.900</b>	<b>0.899 <math>\pm</math> 0.017</b>	<b>0.871 <math>\pm</math> 0.021</b>	<b>0.867 <math>\pm</math> 0.024</b>	<b>1.000 (0.703)</b>
	Mutag $\downarrow$	<b>0.020</b>	0.094 $\pm$ 0.009	<b>0.121 <math>\pm</math> 0.046</b>	<b>0.128 <math>\pm</math> 0.051</b>	

Bold values indicate the highest performance metrics

<sup>a</sup> Top-n is the average score for the top n output molecules

<sup>b</sup> Syn-ratio is synthesizable ratio

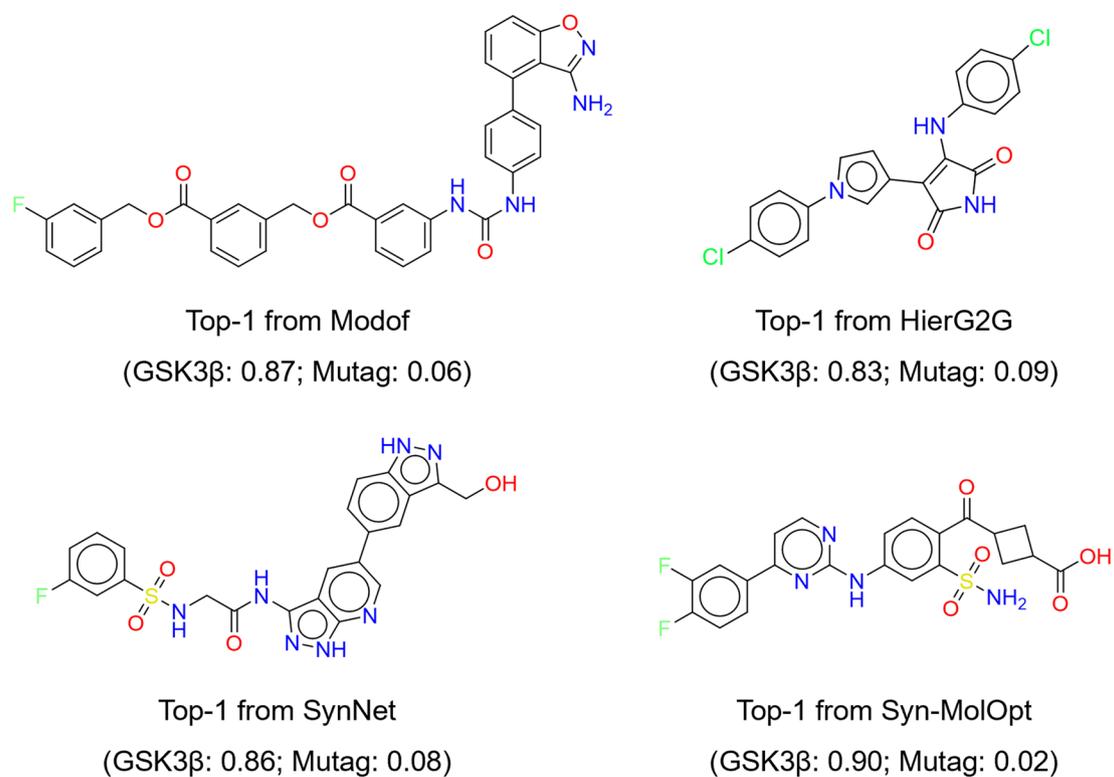
<sup>c</sup> The value in parentheses represents the synthesizable ratio evaluated by AiZynthFinder

experiments, where each experiment involves refining 128 molecules. We examined the performance of each method by assessing the properties of the top-N (N=1, 10, 100, 128) molecules generated by each approach. Additionally, we analyzed the synthesizability of the optimized molecules. For molecules optimized by Modof and HierG2G, we employed the multi-step retrosynthesis tool AiZynthFinder, developed by Genheden et al. [50], to plan their synthesis routes. We set specific search parameters for the retrosynthesis, including a maximum depth of 6, a search time of 120 s, and an iteration of 300 rounds. Molecules with successfully planned synthesis routes by AiZynthFinder were deemed synthesizable, allowing us to determine their synthesizable ratio. As Syn-MolOpt and SynNet can directly provide synthesis routes for their optimized molecules, their synthesizable ratio was inherently rated as 100%. However, to ensure a fair comparison, we also assessed the synthesizability of the molecules optimized by these two methods using AiZynthFinder. Since AiZynthFinder relies on predefined rules and known reactions, it may not always plan synthesis routes for compounds for which Syn-MolOpt (or SynNet) has already identified a route. In addition, to further assess the performance of Syn-MolOpt, we conducted similar experiments on c-Jun N-terminal Kinases-3 (JNK3) as we did on GSK3 $\beta$ , and the results can be found in the Supplementary Information.

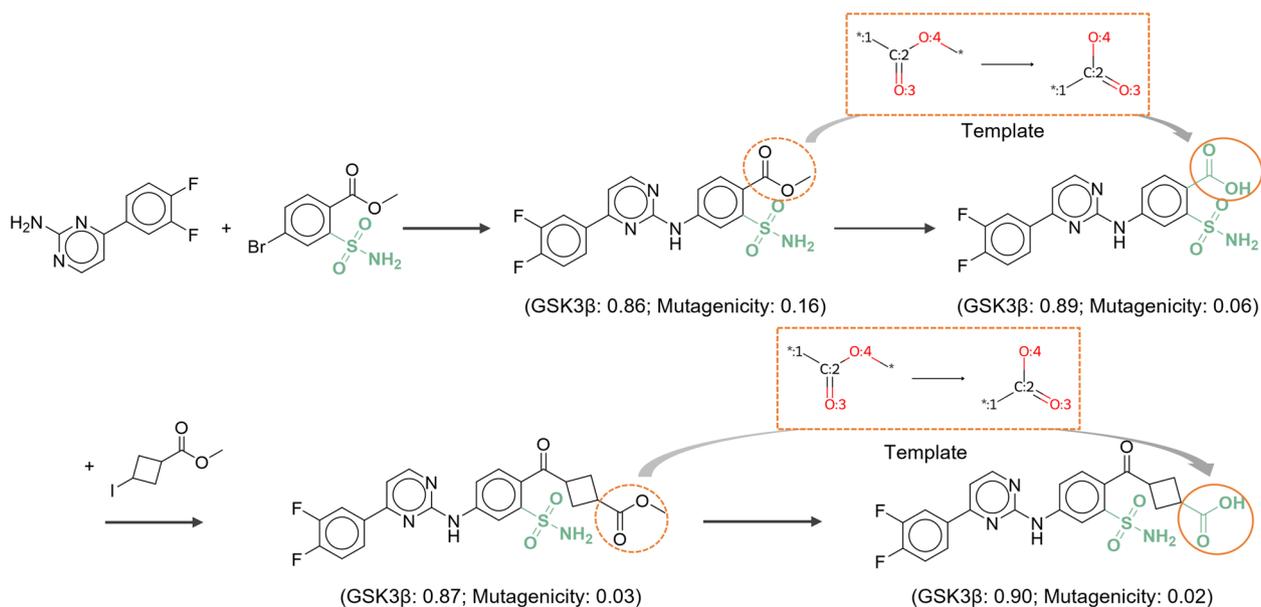
#### Experimental results on multi-property optimization of GSK3 $\beta$ -Mutag

As shown in Table 1, the molecules designated for optimization in the GSK3 $\beta$ -Mutag multi-property optimization experiment already exhibit a favorable average GSK3 $\beta$  score (0.787  $\pm$  0.109). However, they also suffer from a higher average Mutag score (0.887  $\pm$  0.055), indicating a high mutagenic potential. Thus, the goal

of optimization is to mitigate the mutagenic potential (reduce Mutag scores) while enhancing the GSK3 $\beta$  scores simultaneously. Table 2 outlines the GSK3 $\beta$  and Mutag scores of molecules optimized by Syn-MolOpt, Modof, HierG2G, and SynNet. The data show that Syn-MolOpt outperforms other methods in almost all metrics except for the top-10 Mutag score. Moreover, compared to the initial molecules, the top-128 molecules optimized by Syn-MolOpt significantly reduced the average Mutag score (from 0.887  $\pm$  0.055 to 0.128  $\pm$  0.051) and improved the average GSK3 $\beta$  score (from 0.787  $\pm$  0.109 to 0.867  $\pm$  0.024). However, Modof, HierG2G, and SynNet face challenges in effectively reducing the mutagenicity of compounds while maintaining their GSK3 $\beta$  binding affinity. This is evident as the average GSK3 $\beta$  scores of the top-128 molecules optimized by these three methods all showed varying degrees of decline compared to the average GSK3 $\beta$  score of the initial molecules. Furthermore, based on the synthesis route planning by AiZynthFinder, we calculated the synthesizable ratio of the top-128 molecules optimized by Modof and HierG2G, yielding scores of 0.500 and 0.651, respectively. This suggests that a considerable portion of these molecules are challenging to synthesize, posing difficulties for experimental application. In contrast, Syn-MolOpt and SynNet provide synthesizable route references for the optimized compounds, thereby enhancing their practical usability. Additionally, we visualized the structures of the top-1 molecule optimized by each of the four methods. As depicted in Fig. 3, the top-1 molecules optimized by Modof and SynNet respectively contain substructures with detoxifying effects:  $-\text{NC}(=\text{O})\text{N}-$  and  $-\text{NS}(=\text{O})(=\text{O})-$ . The top-1 molecule optimized by Syn-MolOpt, featuring two detoxifying groups (sulfonamide and carboxyl) achieves a GSK3 $\beta$  score of 0.900 and a Mutag score of 0.020. Analysis of its synthetic route in Fig. 4 revealed



**Fig. 3** The top-1 molecule structures optimized by Modof, HierG2G, SynNet, and Syn-MolOpt on the GSK3β-Mutag multi-property optimization experiment



**Fig. 4** The synthetic route of the top-1 molecule optimized by Syn-MolOpt on the GSK3β-Mutag multi-property optimization experiment. The dotted box displays the functional reaction template used

**Table 3** Overall comparison on multi-property optimization of GSK3 $\beta$ -hERG

Methods	Property	Top-1 <sup>a</sup>	Top-10	Top-100	Top-128	Syn-ratio <sup>b</sup> $\uparrow$
Modof	GSK3 $\beta$ $\uparrow$	0.870	0.840 $\pm$ 0.065	0.751 $\pm$ 0.125	0.720 $\pm$ 0.138	0.422
	hERG $\downarrow$	<b>0.084</b>	0.191 $\pm$ 0.125	0.432 $\pm$ 0.326	0.499 $\pm$ 0.339	
HierG2G	GSK3 $\beta$ $\uparrow$	0.920	<b>0.882 <math>\pm</math> 0.094</b>	0.436 $\pm$ 0.266	0.403 $\pm$ 0.275	0.682
	hERG $\downarrow$	0.282	0.385 $\pm$ 0.159	0.312 $\pm$ 0.242	0.310 $\pm$ 0.242	
SynNet	GSK3 $\beta$ $\uparrow$	0.900	0.854 $\pm$ 0.032	0.774 $\pm$ 0.050	0.765 $\pm$ 0.053	1.000 (0.648) <sup>c</sup>
	hERG $\downarrow$	0.111	0.147 $\pm$ 0.073	0.154 $\pm$ 0.087	0.161 $\pm$ 0.097	
Syn-MolOpt	GSK3 $\beta$ $\uparrow$	<b>0.940</b>	0.866 $\pm$ 0.039	<b>0.804 <math>\pm</math> 0.042</b>	<b>0.793 <math>\pm</math> 0.044</b>	<b>1.000 (0.883)</b>
	hERG $\downarrow$	0.105	<b>0.107 <math>\pm</math> 0.064</b>	<b>0.132 <math>\pm</math> 0.073</b>	<b>0.131 <math>\pm</math> 0.071</b>	

Bold values indicate the highest performance metrics

<sup>a</sup> Top-n is the average score for the top n output molecules

<sup>b</sup> Syn-ratio is synthesizable ratio

<sup>c</sup> The value in parentheses represents the synthesizable ratio evaluated by AiZynthFinder

that the sulfonamide derives from commercially available building blocks, while the carboxyl group is introduced through ester bond hydrolysis via a functional reaction template. And with the introduction of carboxyl groups, the Mutag score of the compound gradually decreases, highlighting the effectiveness of functional reaction templates. In Figure S4, we present more examples of functional reaction templates that aid in molecular optimization. In addition, we have visualized the synthetic routes of the top-10 molecules generated by Syn-MolOpt in the Supplementary Information and discussed the feasibility of these routes.

#### Experimental results on multi-property optimization of GSK3 $\beta$ -hERG

In the GSK3 $\beta$ -hERG optimization task, the initial molecules had average GSK3 $\beta$  and hERG scores of 0.782  $\pm$  0.105 and 0.807  $\pm$  0.094 (Table 1), respectively. These scores imply a significant likelihood of hERG-related cardiac toxicity. Table 3 shows the optimization results for four different methods. The data reveals that Syn-MolOpt, barring slightly inferior hERG scores for the top-1 molecule and average GSK3 $\beta$  scores for the top-10 molecules compared to Modof and HierG2G, performed the best on all other performance metrics. Notably, among the top-128 molecules, only Syn-MolOpt succeeded in effectively reducing hERG toxicity while enhancing GSK3 $\beta$  binding affinity. Furthermore, the synthesizability rates for the compounds optimized by Modof and HierG2G were 0.422 and 0.682, respectively.

#### The application of Syn-MolOpt on metabolic property optimization

The experimental results clearly demonstrated that the generic functional reaction template library constructed for mitigating Mutag and hERG toxicity enables

Syn-MolOpt to effectively reduce the Mutag or hERG toxicity of molecules. Cytochrome P450 inhibitors interfere drug metabolism, potentially leading to unsafe drug accumulation in humans. Consequently, optimizing the metabolic properties of candidate molecules is beneficial for enhancing their safety and efficacy. To test the adaptability of our property-specific functional reaction template library, we applied the Syn-MolOpt workflow to a multi-property optimization focusing on binding affinities and metabolic properties. Specifically, given a molecule that binds well to GSK3 $\beta$  but exhibits CYP inhibitory properties, the objective is to optimize it into another molecule that has lower metabolic inhibitory properties and enhanced GSK3 $\beta$  binding (with a higher GSK3 $\beta$  score and a lower CYP3A4/CYP2C19 score). CYP3A4 and CYP2C19 scores are predicted by the corresponding consensus models, representing the likelihood of metabolic inhibition. We also benchmarked Syn-MolOpt's performance against Modof, HierG2G, and SynNet. Initially, in the GSK3 $\beta$ -CYP3A4 multi-property optimization task, the test molecules had GSK3 $\beta$  and CYP3A4 scores of 0.772  $\pm$  0.090 and 0.949  $\pm$  0.021, respectively, as shown in Table 1. The molecular optimization results are summarized in Table 4, revealing that the Syn-MolOpt's average GSK3 $\beta$  score for the top-10 molecules (0.892  $\pm$  0.019) was slightly lower than that of HierG2G (0.899  $\pm$  0.100). However, the molecules optimized by Syn-MolOpt demonstrated the best performance on the other metrics, with the top-1 molecule achieving a GSK3 $\beta$  score of 0.920 and a CYP3A4 score of 0.013. Moreover, in the GSK3 $\beta$ -CYP2C19 optimization task, Syn-MolOpt consistently outperformed the other three methods, except for the CYP2C19 score of the top-ranked molecule. An analysis of the top-128 molecules in the GSK3 $\beta$ -CYP3A4 and GSK3 $\beta$ -CYP2C19 tasks revealed that only Syn-MolOpt achieved the goal of

**Table 4** Overall comparison on multi-property optimization of GSK3 $\beta$ -CYP3A4 and GSK3 $\beta$ -CYP2C19

Task	GSK3 $\beta$ -CYP3A4					
Methods	Property	Top-1 <sup>a</sup>	Top-10	Top-100	Top-128	Syn-ratio <sup>b</sup> $\uparrow$
Modof	GSK3 $\beta$ $\uparrow$	0.880	0.771 $\pm$ 0.094	0.770 $\pm$ 0.112	0.740 $\pm$ 0.117	0.531
	CYP3A4 $\downarrow$	0.041	0.162 $\pm$ 0.121	0.735 $\pm$ 0.335	0.778 $\pm$ 0.310	
HierG2G	GSK3 $\beta$ $\uparrow$	0.880	<b>0.899 <math>\pm</math> 0.100</b>	0.472 $\pm$ 0.296	0.447 $\pm$ 0.303	0.682
	CYP3A4 $\downarrow$	0.147	0.428 $\pm$ 0.243	0.466 $\pm$ 0.340	0.468 $\pm$ 0.335	
SynNet	GSK3 $\beta$ $\uparrow$	0.900	0.810 $\pm$ 0.043	0.734 $\pm$ 0.053	0.721 $\pm$ 0.056	<b>1.000 (0.789)<sup>c</sup></b>
	CYP3A4 $\downarrow$	0.015	0.055 $\pm$ 0.086	0.078 $\pm$ 0.084	0.081 $\pm$ 0.085	
Syn-MolOpt	GSK3 $\beta$ $\uparrow$	<b>0.920</b>	0.892 $\pm$ 0.019	<b>0.842 <math>\pm</math> 0.033</b>	<b>0.836 <math>\pm</math> 0.036</b>	1.000 (0.727)
	CYP3A4 $\downarrow$	<b>0.013</b>	<b>0.033 <math>\pm</math> 0.023</b>	<b>0.046 <math>\pm</math> 0.055</b>	<b>0.058 <math>\pm</math> 0.066</b>	
Task	GSK3 $\beta$ -CYP2C19					
Modof	GSK3 $\beta$ $\uparrow$	0.850	0.828 $\pm$ 0.033	0.705 $\pm$ 0.129	0.682 $\pm$ 0.139	0.406
	CYP2C19 $\downarrow$	<b>0.025</b>	0.103 $\pm$ 0.078	0.346 $\pm$ 0.359	0.423 $\pm$ 0.389	
HierG2G	GSK3 $\beta$ $\uparrow$	0.910	0.859 $\pm$ 0.103	0.551 $\pm$ 0.295	0.539 $\pm$ 0.297	0.587
	CYP2C19 $\downarrow$	0.072	0.246 $\pm$ 0.297	0.474 $\pm$ 0.372	0.487 $\pm$ 0.373	
SynNet	GSK3 $\beta$ $\uparrow$	0.850	0.811 $\pm$ 0.032	0.746 $\pm$ 0.059	0.730 $\pm$ 0.069	<b>1.000 (0.820)<sup>c</sup></b>
	CYP2C19 $\downarrow$	0.065	0.095 $\pm$ 0.063	0.167 $\pm$ 0.124	0.167 $\pm$ 0.134	
Syn-MolOpt	GSK3 $\beta$ $\uparrow$	<b>0.920</b>	<b>0.869 <math>\pm</math> 0.032</b>	<b>0.806 <math>\pm</math> 0.038</b>	<b>0.797 <math>\pm</math> 0.039</b>	1.000 (0.766)
	CYP2C19 $\downarrow$	0.076	<b>0.094 <math>\pm</math> 0.035</b>	<b>0.108 <math>\pm</math> 0.066</b>	<b>0.108 <math>\pm</math> 0.065</b>	

Bold values indicate the highest performance metrics

<sup>a</sup> Top-n is the average score for the top n output molecules

<sup>b</sup> Syn-ratio is synthesizable ratio

<sup>c</sup> The value in parentheses represents the synthesizable ratio evaluated by AiZynthFinder

simultaneously optimizing both target properties compared to the initial molecules. Additionally, a synthesizability analysis indicated that many molecules optimized by Modof and HierG2G were challenging to synthesize in both tasks.

#### The exploration of Syn-MolOpt's potential in real-world molecular optimization scenarios

As mentioned above, similar to the tasks of optimizing for toxicity, the Syn-MolOpt workflow has also demonstrated remarkable performance in optimizing compounds' binding affinity and metabolic properties, confirming the versatility of this molecular optimization approach. In addition to its practicality, Syn-MolOpt integrates molecular optimization with synthesis planning, offering synthesis pathways for the optimized output molecules. Similarly, SynNet can also generate synthesis routes for the output molecules. In the optimization process, SynNet navigates the chemical space and selects compounds via a fitness function (a QSAR model for a target property) in a GA. The accuracy of this fitness function significantly influences the outcomes of molecular optimization. In contrast to SynNet, Syn-MolOpt strategically uses functional reaction templates to guide the structural optimization of compounds during the synthesis tree generation, and subsequently uses the

**Table 5** Performance comparison of the consensus models

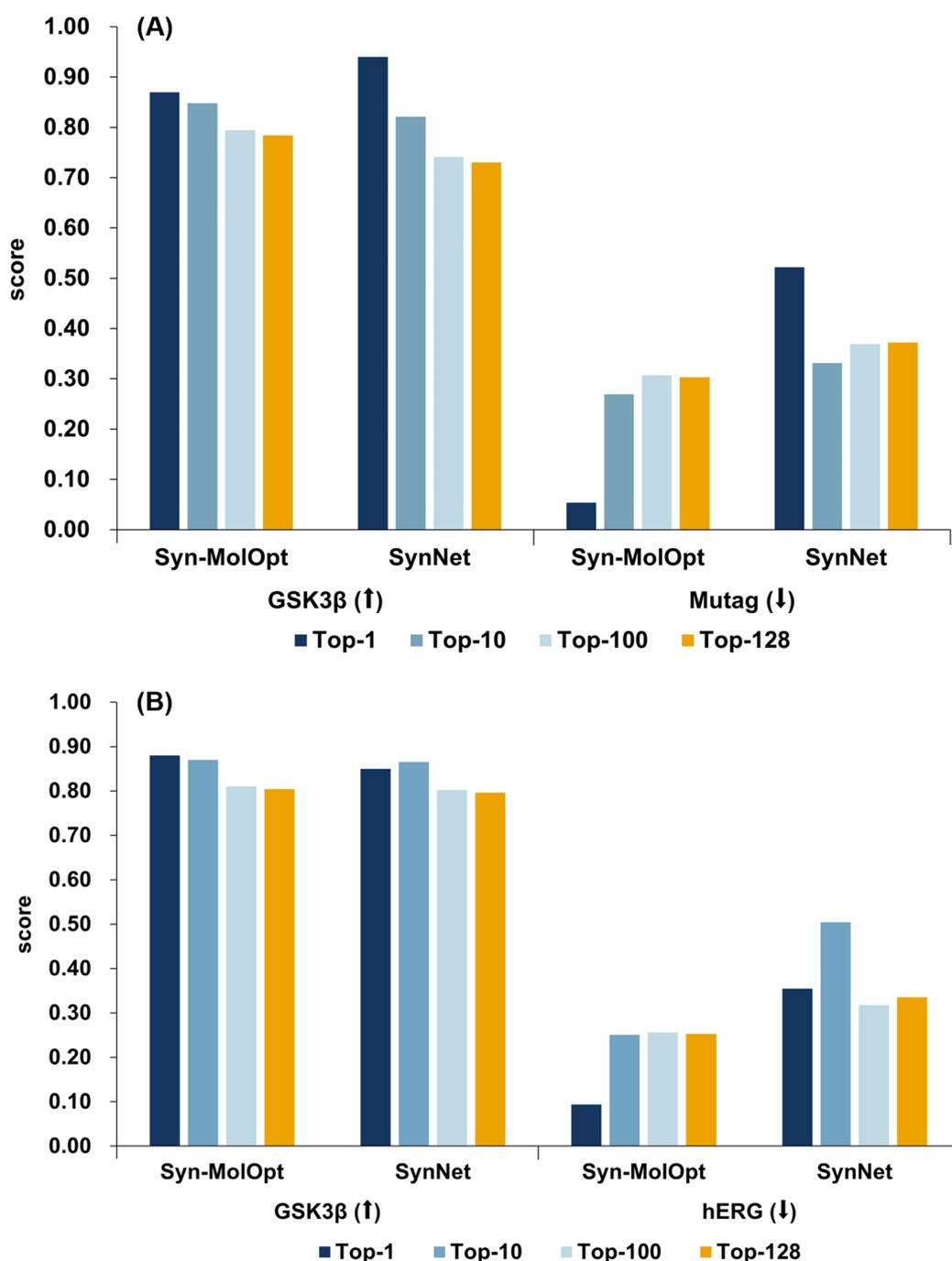
Model	ROC-AUC
Mutag <sub>100%-trn</sub>	0.901
hERG <sub>100%-trn</sub>	0.862
Mutag <sub>20%-trn</sub>	0.837
hERG <sub>20%-trn</sub>	0.773

fitness function to select compounds with desired properties, rather than relying solely on the fitness function. However, in real-world scenarios, the data available for the properties requiring optimization is often insufficient to train a highly accurate QSAR model.

To assess the practical potential of Syn-MolOpt, we simulated a real-world molecular optimization scenario and compared the performances of Syn-MolOpt and SynNet. We respectively extracted 20% of the data from the training sets of the Mutag and hERG toxicity datasets based on scaffold classification to retraining two consensus models (Mutag<sub>20%-trn</sub> and hERG<sub>20%-trn</sub>). As shown in Table 5, compared to the consensus models trained on the entire datasets (Mutag<sub>100%-trn</sub> and hERG<sub>100%-trn</sub>), the performance of the retrained models declined with reduced training data, achieving ROC-AUCs of 0.837 and 0.773, respectively. Using

these retrained consensus models, we reconstructed the functional reaction template library for optimizing Mutag and hERG toxicity. In simulated real-world optimizations of GSK3 $\beta$ -Mutag and GSK3 $\beta$ -hERG, we used Mutag<sub>20%-trn</sub> and hERG<sub>20%-trn</sub> as the fitness functions and implemented the reconstructed functional reaction template library in Syn-MolOpt. The molecules

designated for testing in the GSK3 $\beta$ -Mutag and GSK3 $\beta$ -hERG optimization tasks (Table 1) were then used to evaluate the performance of Syn-MolOpt and SynNet in this specific scenario. To ensure an objective comparison, we assessed the Mutag and hERG scores of the final optimized molecules using the highly accurate Mutag<sub>100%-trn</sub> and hERG<sub>100%-trn</sub> models (Table 5). The



**Fig. 5** Performance comparison of Syn-MolOpt and SynNet in simulated real-world optimizations of **A** GSK3 $\beta$ -Mutag and **B** GSK3 $\beta$ -hERG

optimization results for the GSK3 $\beta$ -Mutag and GSK3 $\beta$ -hERG tasks are presented in Fig. 5A and B. These figures illustrate that in molecular optimization tasks with low-accuracy scoring functions, Syn-MolOpt outperforms SynNet, particularly in toxicity optimization. The top-1 and top-10 molecules generated by Syn-MolOpt exhibit significantly lower Mutag and hERG scores compared to those optimized by SynNet. Specifically, the top-1 and top-10 molecules from Syn-MolOpt had the Mutag scores of 0.054 and 0.269, respectively, while those from SynNet were 0.522 and 0.331. Furthermore, the hERG scores for the top-1 and top-10 molecules from Syn-MolOpt were 0.094 and 0.251, respectively, while those from SynNet were 0.355 and 0.504. This suggests that by strategically transforming fragments using functional reaction templates during synthesis tree generation, Syn-MolOpt achieves more robust performance, even with inaccurate scoring feedback. This highlights Syn-MolOpt's applicability in real-world molecular optimization scenarios.

## Conclusion

In this study, to address the issue that most current molecular optimization algorithms inadequately consider the synthesizability of their optimized molecules, we proposed a synthesis planning-driven molecular optimization scheme, Syn-MolOpt, based on functional reaction templates. This method allows for the design of property-specific functional reaction template libraries for the properties to be optimized, providing reference synthetic routes for the optimized compounds while optimizing the targeted properties. We focus on the optimization of toxicity and metabolic properties, across two toxicity-related (GSK3 $\beta$ -Mutag and GSK3 $\beta$ -hERG) and two metabolism-related (GSK3 $\beta$ -CYP3A4 and GSK3 $\beta$ -CYP2C19) multi-property molecular optimization tasks. Syn-MolOpt demonstrated superior performance compared to three benchmark models, Modof, HierG2G, and SynNet, confirming its effectiveness and versatility. The visualization of synthetic routes for molecules optimized by Syn-MolOpt demonstrates that functional reaction templates can help facilitate molecular optimization. Furthermore, simulated real-world experiments show that Syn-MolOpt, by using functional reaction templates to effectively transform relevant fragments during the synthesis tree generation process rather than relying solely on scoring functions to guide the process, can achieve robust performance even when the scoring feedback on properties to be optimized is not sufficiently accurate. Despite these advancements, there is still room for improvement in Syn-MolOpt. In our current multi-property

molecular optimization experiments, we have constructed functional reaction template libraries for only one property. For instance, in the GSK3 $\beta$ -Mutag optimization task, we have only designed reaction templates specifically for Mutag. Designing more universal functional reaction template libraries for multiple properties could potentially further enhance molecular optimization performance.

## Methods

### Dataset

The Mutag and hERG datasets were collected from our previous work [34], and the CYP3A4 and CYP2C19 datasets were collected from the study of Huang et al. [49]. The details of the above datasets are shown in Table S2. The reaction templates used include property-specific functional reaction templates and 91 publicly available reaction templates from Hartenfeller et al. [51] and Button et al. [52]. The purchasable building blocks were sourced from Enamine Building Blocks (US stock; accessed on June 12, 2023) and then filtered by substructure matching using toxicity alerts (or CYP alerts) to filter out potential blocks that may increase toxicity (or CYP inhibition).

### Model architecture

As shown in Fig. 2, the implementation of Syn-MolOpt consists of two stages: model training and molecular optimization. In Syn-MolOpt, the synthesis pathway of a compound is modeled as a tree structure, referred to as a synthesis tree. A valid synthesis tree, by a series of discrete reaction templates, connects the root node (the final product molecule) to purchasable building blocks through feasible reactions, representing all steps and branches of the target compound's synthesis pathway. We construct the synthesis tree in a bottom-up manner, one reaction step at a time, modeling the construction of the synthesis tree as a Markov decision process. The state transitions of the synthesis tree must satisfy the Markov property:  $P(S_{t+1}|S_t, \dots, S_0) = P(S_{t+1}|S_t)$ . After obtaining a specific intermediate compound, subsequent reaction steps of the synthesis tree can be inferred based on the intermediate compound. We define the intermediate synthesis tree at reaction step  $t$  as  $T_t$ . The implementation of each step in constructing the synthesis tree can be divided into four steps: First, reaction action ( $R_{act}$ ) sampling, with possible action types including "add", "extend", "merge", and "end"; second, sampling of the first reactant ( $R_{rct-1}$ ); then, reaction template sampling ( $R_{rxn}$ ); finally, sampling of the second reactant ( $R_{rct-2}$ ). When  $R_{act} = \text{"Add"}$ , one or two new reactant nodes will be added to  $T_t$ , and a new product node will be generated under the given reaction template. When  $R_{act} = \text{"Expand"}$ , the most

recently added node is treated as the first reactant, and a new product node will be added to  $T_t$  if the given reaction template is for a unimolecular reaction; for a bimolecular reaction template, the second reactant is selected first, then the reaction is carried out according to the reaction template, adding a new reactant node and a product node to  $T_t$ . In the Syn-MolOpt frame only unimolecular and bimolecular reactions are allowed. When  $R_{act} = \text{"Merge"}$ , two root nodes act as reactants in a bimolecular reaction to produce a product molecule according to the given reaction template, adding a new product node to  $T_t$  and merging two synthesis subtrees. When  $R_{act} = \text{"End"}$ , it indicates that the construction of the synthesis tree is complete. Based on the reaction templates, and the pre-processed set of purchasable building blocks, we first generated valid synthesis trees according to the construction method of the synthesis tree described above, then characterized the generated synthesis trees, and trained four neural networks as described before ( $N_{act}$ ,  $N_{rct-1}$ ,  $N_{rxn}$ , and  $N_{rct-2}$ ).

As shown in Fig. 2B, molecular optimization is a conditional synthesis tree generation process, comprising five modules, using the concatenated embeddings of the target molecule to be optimized and the most recently added molecule in the synthesis tree as the current state embedding  $E_{State-T}$  of the synthesis tree. Initially,  $E_{State-T}$  serves as the input for both  $N_{act}$  and  $N_{rct-1}$ , predicting the action type for the reaction step and the embedding of the first reactant  $E_{rct-1}$ , respectively. The embedding of the first reactant serves as a query to select the appropriate first reactant from the pre-processed set of building blocks using k-NN search. Next, the concatenated embedding of  $E_{State-T}$  and  $E_{rct-1}$  is input into  $N_{rxn}$ , which outputs a probability distribution of available reaction templates, and masks inapplicable reaction templates based on the first reactant, thus selecting a suitable reaction template. If the chosen reaction template is for a bimolecular reaction, the concatenated embedding of  $E_{State-T}$ ,  $E_{rct-1}$ , and the embedding of the predicted reaction template  $E_{rxn}$  serve as the input for  $N_{rct-2}$ , outputting the embedding of the second reactant  $E_{rct-2}$ , which is then selected through the k-NN algorithm. Finally, the predicted first and second reactants (or the first reactant in the case of a unimolecular reaction template) react according to the predicted reaction template, and the resulting intermediate is matched with functional reaction templates, if a match is successful, functional processing is carried out. The synthesis tree is updated after each reaction step. Apart from molecular optimization through functional reaction templates during the construction of the synthesis tree, on the other hand, GA is used to perform numerical optimization on the embeddings of the root molecules of the synthesis tree. The GA

operates on Morgan fingerprints with a configuration of 4096 bits and a radius of 2. The mutation is defined as flipping 24 bits in the fingerprint, occurring with a probability of 0.5. The population size is initially set to 128, and the offspring size generated in each iteration is 512. The algorithm runs for a maximum of 50 generations, with an early stop criterion activated if the increase in the population's mean value is less than 0.01 across 10 consecutive generations, signaling convergence. The synthesis tree generator is then used as a decoder to obtain synthesizable molecules corresponding to the optimized vectors, and this process is repeated until the conditions are satisfied.

### Model construction and evaluation

In the construction of drug-likeness models for many real-world scenarios, a common practice is to integrate the prediction results of multiple models to obtain a consensus model. In our study, Mutag score, hERG score, CYP3A4 score and CYP2C19 score were the prediction results of their respective consensus models. When training the consensus model, each dataset is randomly divided into training, validation and test sets in a ratio of 8:1:1. First, 10 RGCN models based on different random seeds are constructed. Subsequently, a consensus model was constructed by integrating these 10 RGCN sub-models. The average of the predictions from the 10 submodels will be used as the output of the final consensus model. The four consensus models are all used for classification tasks, evaluated by the area under the receiver operating characteristic curve (ROC-AUC). When training the  $R_{act}$  network, the  $R_{rct-1}$  network, the  $R_{rxn}$  network, and the  $R_{rct-2}$  network in Syn-MolOpt, Morgan molecular fingerprints and one-hot encoding are used to represent molecules and reaction templates, respectively, as inputs for the MLPs. The length of Morgan molecular fingerprint is 4096 and the radius is 2. Among them, the prediction of  $R_{act}$  and  $R_{rxn}$  are classification tasks,  $R_{rct-1}$  and  $R_{rct-2}$  prediction are regression tasks.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13321-025-00975-9>.

Supplementary Material 1. Table S1. Attributions of functional groups in four datasets; Table S2. Details of the molecules to be optimized on four multi-property optimization experiments related to JNK3; Table S3. Overall comparison of four JNK3-related multi-property optimization tasks; Table S4. Details of the datasets for the four properties to be optimized; Table S5. The performance of the models predicting reaction action, the first reactant, reaction template, and the second reactant.

### Acknowledgements

This work was funded by the National Key R&D Program of China (2024YFA1307500), the Macao Science and Technology Development Fund (Project no: 0056/2020/AMJ), the National Natural Science Foundation of China (22220102001, 92370130), and Shanghai Qilu Pharmaceutical R&D Center.

### Author contributions

X.Y., X.Y., C.H. and T.H. designed the research study. X.Y. and X.W. developed the method and wrote the code. X.Y., X.W., Z.W., Q.L., Y.K., Y.D., P.L., H.L., G.S., and Z.W. performed the analysis. X.Y., X.Y., C.H. and T.H. wrote the paper. All authors read and approved the manuscript.

### Funding

National Key R&D Program of China, 2024YFA1307500; Macao Science and Technology Development Fund, 0056/2020/AMJ; National Natural Science Foundation of China, 22220102001 and 92370130.

### Availability of data and materials

The Mutagenicity, hERG, CYP3A4 and CYP2C19 datasets, and molecular optimization datasets used in this study are available at <https://github.com/xiaodanyin/Syn-MolOpt/tree/main/data>. The molecular optimization code of Syn-MolOpt is developed based on SynNet, and all the codes of Syn-MolOpt are available at <https://github.com/xiaodanyin/Syn-MolOpt>.

### Declarations

#### Competing interests

The authors declare no competing interests.

#### Author details

<sup>1</sup>College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310058, China. <sup>2</sup>Liangzhu Laboratory, Zhejiang University, Hangzhou 311121, China. <sup>3</sup>Faculty of Applied Sciences, Macao Polytechnic University, Macao 999078, China. <sup>4</sup>CarbonSilicon AI Technology Co., Ltd, Hangzhou 310018, Zhejiang, China. <sup>5</sup>Dr. Neher's Biophysics Laboratory for Innovative Drug Discovery, State Key Laboratory of Quality Research in Chinese Medicine, Macau Institute for Applied Research in Medicine and Health, Macau University of Science and Technology, Macao 999078, China. <sup>6</sup>Shanghai Qilu Pharmaceutical R&D Center, 576 Libing Road, Pudong New Area District, Shanghai 310115, China.

Received: 28 September 2024 Accepted: 18 February 2025

Published online: 02 March 2025

### References

- Gao W, Fu T, Sun J et al (2022) Sample efficiency matters: a benchmark for practical molecular optimization. *NeurIPS* 35:21342–21357
- Wang J, Wang X, Sun H et al (2022) ChemistGA: a chemical synthesizable accessible molecular generation algorithm for real-world drug discovery. *J Med Chem* 65:12482–12496
- Wang M, Hsieh C-Y, Wang J et al (2022) Relation: a deep generative model for structure-based de novo drug design. *J Med Chem* 65:9478–9492
- Sanchez-Lengeling B, Aspuru-Guzik A (2018) Inverse molecular design using machine learning: generative models for matter engineering. *Science* 361:360–365
- Kumar A, Loharch S, Kumar S et al (2021) Exploiting cheminformatic and machine learning to navigate the available chemical space of potential small molecule inhibitors of SARS-CoV-2. *Comput Struct Biotechnol J* 19:424–438
- Vatanever S, Schlessinger A, Wacker D et al (2021) Artificial intelligence and machine learning-aided drug discovery in central nervous system diseases: state-of-the-arts and future directions. *Med Res Rev* 41:1427–1473
- Bilodeau C, Jin W, Jaakkola T et al (2022) Generative models for molecular discovery: recent advances and challenges. *Wires Comput Mol Sci* 12:e1608
- Zhavoronkov A, Ivanenkov YA, Aliper A et al (2019) Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nat Biotechnol* 37:1038–1040
- Gao W, Coley CW (2020) The synthesizability of molecules proposed by generative models. *J Chem Inf Model* 60:5714–5723
- Reidenbach D, Coley C W, Yang K (2023) Generating multi-step chemical reaction pathways with black-box optimization. ICLR 2023- MLDD workshop.
- Parrot M, Tajmouati H, da Silva VBR et al (2023) Integrating synthetic accessibility with AI-based generative drug design. *J Cheminform* 15:83
- Yu J, Wang J, Zhao H et al (2022) Organic compound synthetic accessibility prediction based on the graph attention mechanism. *J Chem Inf Model* 62:2973–2986
- Skoraczynski G, Kitlas M, Miasojedow B et al (2023) Critical assessment of synthetic accessibility scores in computer-assisted synthesis planning. *J Cheminform* 15:6
- Gao W, Mercado R, Coley C W (2021) Amortized tree generation for bottom-up synthesis planning and synthesizable molecular design. *arXiv preprint arXiv:2110.06389*.
- Coley CW, Rogers L, Green WH et al (2017) Computer-assisted retrosynthesis based on molecular similarity. *Acs Central Sci* 3:1237–1245
- Coley CW, Thomas DA III, Lummiss JA et al (2019) A robotic platform for flow synthesis of organic compounds informed by AI planning. *Science* 365:1566
- Cook A, Johnson AP, Law J et al (2012) Computer-aided synthesis design: 40 years on. *Wires Comput Mol Sci* 2:79–107
- Corey EJ, Wipke WT (1969) Computer-assisted design of complex organic syntheses: pathways for molecular synthesis can be devised with a computer and equipment for graphical communication. *Science* 166:178–192
- Dai H, Li C, Coley C et al (2019) Retrosynthesis prediction with conditional graph logic network. *NeurIPS* 32.
- Mikulak-Klucznik B, Gołębiewska P, Bayly AA et al (2020) Computational planning of the synthesis of complex natural products. *Nature* 588:83–88
- Schreck JS, Coley CW, Bishop KJ (2019) Learning retrosynthetic planning through simulated experience. *Acs Central Sci* 5:970–981
- Schwaller P, Petraglia R, Zullo V et al (2020) Predicting retrosynthetic pathways using transformer-based models and a hyper-graph exploration strategy. *Chem Sci* 11:3316–3325
- Tetko IV, Karpov P, Van DR et al (2020) State-of-the-art augmented NLP transformer models for direct and single-step retrosynthesis. *Nat Commun* 11:5575
- Wang X, Hsieh C-Y, Yin X et al (2023) Generic interpretable reaction condition predictions with open reaction condition datasets and unsupervised learning of reaction center. *Research* 6:0231
- Yin X, Hsieh C-Y, Wang X et al (2024) Enhancing generic reaction yield prediction through reaction condition-based contrastive learning. *Research* 7:0292
- Bradshaw J, Paige B, Kusner M J et al (2019) A model to search for synthesizable molecules. *NeurIPS* 32
- Bradshaw J, Paige B, Kusner MJ et al (2020) Barking up the right tree: an approach to search over molecule synthesis dags. *NeurIPS* 33:6852–6866
- Korovina K, Xu S, Kandasamy K et al (2020) Chemo: Bayesian optimization of small organic molecules with synthesizable recommendations. *AISTATS* pp 3393–3403
- Gottipati S K, Sattarov B, Niu S et al (2020) Learning to navigate the synthetically accessible chemical space using reinforcement learning. *PMLR* pp 3668–3679
- Horwood J, Noutahi E JAO (2020) Molecular design in synthetically accessible chemical space via deep reinforcement learning. *ACS Omega* 5:32984–32994
- Levin I, Fortunato ME, Tan KL et al (2023) Computer-aided evaluation and exploration of chemical spaces constrained by reaction pathways. *Aiche J* 69:e18234
- Nguyen DH, Tsuda KJ (2021) A generative model for molecule generation based on chemical reaction trees. *arXiv preprint arXiv:2106.03394*
- Fromer JC, Coley CW (2023) Computer-aided multi-objective optimization in small molecule discovery. *Patterns* 4(2):100678. <https://doi.org/10.1016/j.patter.2023.100678>
- Wu Z, Wang J, Du H et al (2023) Chemistry-intuitive explanation of graph neural networks for molecular property prediction with substructure masking. *Nat Commun* 14:2585

35. Chen Z, Min MR, Parthasarathy S et al (2021) A deep generative model for molecule optimization via one fragment modification. *Nat Mach Intell* 3:1040–1049
36. Griffiths R-R, Hernández-Lobato JMJC, s, (2020) Constrained Bayesian optimization for automatic chemical design using variational autoencoders. *Chem Sci* 11:577–586
37. Yang Y, Hsieh C-Y, Kang Y et al (2023) Deep generation model guided by the docking score for active molecular design. *J Chem Inf Model* 63:2983–2991
38. Garrido A, Lepailleur A, Mignani SM et al (2020) hERG toxicity assessment: useful guidelines for drug design. *Eur J Med Chem* 195:112290
39. Segall MD, Barber CJD, d t, (2014) Addressing toxicity risk when designing and selecting compounds in early drug discovery. *Drug Discov Today* 19:688–693
40. De Groot MJ (2006) Designing better drugs: predicting cytochrome P450 metabolism. *Drug Discov Today* 11:601–606
41. Zhang Z, Tang W (2018) Drug metabolism in drug discovery and development. *Acta Pharm Sin B* 8:721–732
42. Schlichtkrull M, Kipf T N, Bloem P et al (2018) Modeling relational data with graph convolutional networks. In *The semantic web: 15th international conference, ESWC 2018, Heraklion, Crete, Greece, June 3–7*, pp 593–607
43. Degen J, Wegscheid-Gerlach C, Zaliani A et al (2008) On the art of compiling and using 'drug-like' chemical fragment spaces. *ChemMedChem* 3:1503
44. Bemis GW, Murcko MAJJ, o m c, (1996) The properties of known drugs. 1 Molecular frameworks. *J Med Chem* 39:2887–2893
45. Coley CW, Green WH, Jensen KF et al (2019) RDChiral: an RDKit wrapper for handling stereochemistry in retrosynthetic template extraction and application. *J Chem Inf Model* 59:2529–2537
46. Schwaller P, Hoover B, Reymond J-L et al (2021) Extraction of organic chemistry grammar from unsupervised learning of chemical reactions. *Sci Adv* 7:eabe4166
47. Tran TTV, Surya WA, Tayara H et al (2023) Artificial intelligence in drug toxicity prediction: recent advances, challenges, and future perspectives. *J Chem Inf Model* 63:2628–2643
48. Jin W, Barzilay R, T (2020) Jaakkola Hierarchical generation of molecular graphs using structural motifs. In *International conference on machine learning*. PMLR, pp 4839–4848
49. Huang K, Fu T, Gao W et al (2021) Therapeutics data commons: machine learning datasets and tasks for drug discovery and development. *arXiv preprint arXiv:2102.09548*
50. Genheden S, Thakkar A, Vá C et al (2020) AiZynthFinder: a fast, robust and flexible open-source software for retrosynthetic planning. *J Cheminform* 12:70
51. Hartenfeller M, Eberle M, Meier P et al (2011) A collection of robust organic synthesis reactions for in silico molecule design. *J Chem Inf Model* 51:3093–3098
52. Button A, Merk D, Hiss JA et al (2019) Automated de novo molecular design by hybrid machine intelligence and rule-driven chemical synthesis. *Nat Mach Intell* 1:307–315

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.