# SYNTHESIS-CONSTRAINED MOLECULAR DESIGN WITH DIRECT OPTIMIZATION OF REACTION CONDITIONS

**Víctor Sabanza-Gil, Matéo Tatzber, Jeff Guo, Jeremy Luterbacher & Philippe Schwaller**
Laboratory of Artificial Chemical Intelligence (LIAC)
& Laboratory of Sustainable and Catalytic Processing (LPDC)
& National Centre of Competence in Research (NCCR) Catalysis
École Polytechnique Fédérale de Lausanne (EPFL)
Lausanne, Switzerland
{victor.sabanzagil, philippe.schwaller}@epfl.ch

## ABSTRACT

Generative molecular design promises to accelerate drug and materials discovery through the proposal of optimized candidates for given molecular properties. However, current models still struggle to incorporate key synthetic constraints like reaction conditions in their designs, which hampers experimental validation. In this work, we show it is possible to directly optimize reaction conditions as part of the molecular optimization process. We utilize a reaction condition prediction model to annotate the retrosynthetic pathways of generated molecules. The generative model is then rewarded based on the alignment of these predicted conditions with user-defined constraints. We demonstrate the ability to avoid restricted solvents such as dimethylformamide (DMF) and dichloromethane (DCM) or enforce the presence of desired solvents like water compared to experiments without reaction condition constraints. Our results indicate that the model learns a distribution shift toward reaction classes compatible with these constraints without compromising property optimization. Finally, we show the utility of the framework in an *in silico* experiment where we design potential fungicides following green chemistry principles. Overall, our contribution helps to bring molecular design closer to laboratory reality and accelerate the validation of proposed molecules.

## 1 INTRODUCTION

The discovery of novel molecules is a fundamental scientific challenge, essential for the development of next-generation therapeutics, agrochemicals and functional materials. Despite its importance, molecular discovery remains a long and expensive process depending on resource-intensive experimental validation (DiMasi et al., 2016). In recent years, machine learning (ML) generative molecular design has emerged as a promising option to efficiently propose molecules with optimal properties and accelerate this development pipeline (Du et al., 2024). Despite several validated examples (Zhavoronkov et al., 2019; Xu et al., 2025), synthesizing the generated candidates remains a practical challenge, and many generative models fail to propose molecules with feasible synthetic pathways (Gao & Coley, 2020; Stanley & Segler, 2023).

Several synthesis-constrained generative models have been recently proposed to address molecule synthesizability. Different approaches to solve this issue exist. Some models incorporate synthesis heuristics as an objective in the optimization (Hassen et al., 2025). A more chemically intuitive approach models the bottom-up assembly of molecular building blocks and chemical reactions to build the desired targets while also building their synthetic route (Swanson et al., 2024; Cretu et al., 2024; Gao et al., 2024). Last, some methods use retrosynthesis models as optimization oracles to directly reward the model based on the paths found for the generated molecules (Parrot et al., 2023; Guo & Schwaller, 2025; Dekleva et al., 2025). Although, some of these models incorporate key synthetic constraints like reaction type control (Guo et al., 2025; Dekleva et al., 2025), there is still a lack of consideration of the full synthetic elements that a chemical route needs in order to be executed in the lab.

Reaction conditions are key among the synthetic parameters that describe a laboratory synthesis. They usually define a reaction feasibility and yield, and several ML models have been proposed to predict suitable conditions for given reactions (Ball et al., 2025). Regarding molecular design, some methods have recently explored the consideration of reaction conditions in the reaction types to promote diversity in parallel chemistry setups (Fromer et al., 2025; Le Vaillant et al., 2025). Although they can promote molecular diversity by considering reaction conditions, these approaches are *post-hoc*, where molecules are generated and filtered in two separate stages. In addition, the granularity of reaction conditions is limited by the reaction template definition, and two reactions sharing the same name may need different conditions depending on the substrates. Therefore, direct consideration of reaction conditions during the generative process is still lacking.

In this work, we introduce a framework to directly optimize reaction conditions as part of the generative molecular design task. We couple a sample-efficient language-based molecular generative model with a multi-step retrosynthesis model and a reaction condition prediction model to reward those molecules whose synthesis routes meet user-specified reaction conditions. We show how this setup allows the simultaneous optimization of target properties like molecular docking or Quantitative Estimate of Drug-likeness (QED) (Bickerton et al., 2012), while shifting the proposed synthesis routes towards conditions avoiding undesired solvents or enforcing specific solvents. We also show how condition optimization implicitly shifts the reaction type distribution of the synthesis routes from the generated molecules. Finally, we propose an *in silico* experiment to generate novel agrochemicals while meeting key sustainability criteria like the avoidance of specific solvents or the use of biomass-derived building blocks in the synthesis.

Our contributions are as follows:

1. We design a reward function to steer a molecular generative model towards synthesizable molecules with user-specified reaction conditions.

2. We show how the framework can effectively push generated molecules towards desired synthesis conditions, shifting the distribution of reaction types and reaction conditions with respect to the baseline case.

3. Through a use case based on agrochemical molecular design, we show how the reward can be mixed with other synthetic constraints to achieve sustainable-by-design candidate molecules.

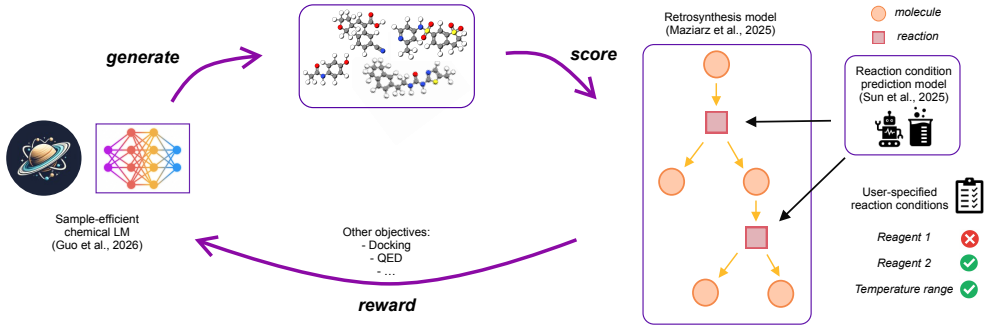


Figure 1: Directly optimizing reaction conditions in synthesis-constrained molecular design. In this framework, we employ Saturn (Guo et al., 2026), a sample-efficient molecular generative language model, to generate a batch of molecules. Syntheseus (Maziarz et al., 2025b) is used to extract the retrosynthetic tree of the generated molecules. We then annotate all the reactions with the top-1 predicted conditions by a reaction condition model (QUARC, (Sun et al., 2025)), and reward the generated molecules based on user-defined conditions constraints (in this case, avoid a specific set of reagents, enforce specific reagents and/or enforce a given temperature range). The resulting molecules satisfy reaction condition constraints and are optimized for other target objectives.

## 2 RELATED WORK

**Synthesis-constrained Generative Molecular Design.**

Molecular design must integrate synthesizability as an intrinsic constraint within the generative process to ensure the synthetic feasibility of the generated candidates (Gao & Coley, 2020; Stanley & Segler, 2023). While traditional virtual screening ensures molecule accessibility by exhaustively scoring commercially available libraries (Lyu et al., 2019; Bender et al., 2021), this paradigm struggles to scale with the emergence of reaction-enumerated spaces. Modern combinatorial libraries, often exceeding $10^{12}$ candidates (Warr, 2021), render brute-force evaluation computationally prohibitive.

Alternative methods introduce synthesizability as part of the generative framework. Early methods employed Metropolis (Vinkers et al., 2003) or genetic algorithms (Spiegel & Durrant, 2020) to steer reaction enumeration from basic building blocks. More recently, deep learning approaches have addressed synthesizability following three main ideas. First, structural heuristics like SAscore (Ertl & Schuffenhauer, 2009) have been used as part of the optimization objective to penalize complex functional group arrangements (Hassen et al., 2025). However, these proxies fail at times to distinguish granular molecular motifs that make some structures unfeasible (Gao & Coley, 2020). Second, *bottom-up* generative models enforce synthesizability via an inductive bias, iteratively assembling molecules from building blocks. This paradigm—implemented through several methods like Monte Carlo Tree Search (Swanson et al., 2024), Reinforcement Learning (Gottipati et al., 2020; Wang et al., 2024), GFlowNets (Koziarski et al., 2024; Cretu et al., 2024; Seo et al., 2024), and Transformers (Luo et al., 2024; Gao et al., 2024; Lee et al., 2025)—simultaneously yields the candidate molecule and its synthetic trajectory. Finally, top-down approaches utilize external retrosynthetic models to evaluate generated structures. While computationally more intensive than heuristics, retrosynthetic workflows provide a robust assessment of feasibility and have been successfully integrated with molecular generative models (Guo & Schwaller, 2025; 2026; Guo et al., 2025; Dekleva et al., 2025; Parrot et al., 2023; Descamps et al., 2025).

**ML-reaction condition prediction**. Reaction condition prediction can be seen as the inverse problem of reaction property prediction (Ball et al., 2025). Given a set of reactants, we want to get the reagents (molecules that do not contribute with heavy atoms to the final product), catalyst, solvent, temperature, and quantities to be used in the reaction setup. As opposed to reaction prediction, condition prediction is a many-to-many problem —because several condition may be suitable for a given reaction— that reflects its inherent complexity. This problem is increased by the sparsity and bias of the datasets at hand. In general, we can distinguish general and local models depending on the dataset size and task (Raghavan et al., 2023). Local models focus on a single reaction type and a well-defined set of substrates and conditions (Marcou et al., 2015; Genheden et al., 2022). General models use datasets derived from literature (e.g. USPTO (Lowe, 2012), Reaxys (Elsevier, 2026)) grouping hundreds of different reaction types (Gao et al., 2018; Dobbelaere et al., 2024; Wang et al., 2025; Sun et al., 2025). In the context of this work, general condition prediction models are used to map the conditions of the reactions proposed by multi-step retrosynthesis models.

## 3 METHODS

In this section, we describe each component of the generative framework: Saturn (Guo et al., 2026) (generative model), Syntheseus (Maziarz et al., 2025b) (interface for retrosynthesis models), QUARC (Sun et al., 2025) (reaction condition prediction model) and the reward formulation (how we make reaction condition constraints learnable in the synthesizable molecular generative design tasks). We also describe the experimental setup used in the development experiments.

**Saturn: molecular generative model.** Saturn (Guo et al., 2026) is a molecular generative chemical language model based on SMILES (Weininger, 1988). It is built on the Mamba architecture (Gu & Dao, 2023) and the Augmented Memory mechanism (Guo & Schwaller, 2024). Through a reinforcement learning (RL) loop using an augmented likelihood loss function based on REINVENT (Olivecrona et al., 2017), Saturn achieves state-of-the-art performance in several molecular optimization benchmarks, and it is able to directly optimize expensive oracles. In this approach, we *incentivize* a general-purpose model without previous inductive biases through direct rewarding of

the molecules to satisfy challenging design objectives. See appendix A for more information about the RL algorithm, and appendix B for more details about model pretraining and hyperparameters.

**Syntheseus: retrosynthesis framework.** Given a target molecule, retrosynthesis models propose a directed graph (*synthesis route*) with nodes and edges representing molecules and reactions, respectively. In this work, we use Syntheseus (Maziarz et al., 2025b) which is a wrapper around single-step retrosynthesis models and search algorithms, allowing mix-and-match formulations to tackle multi-step retrosynthesis. Unless otherwise stated, we use the Enamine EU September 2025 representing commercially available molecules (184,907 in size). Through Syntheseus, we use the MEGAN graph-edits (Sacha et al., 2021) model re-trained by the Syntheseus authors on USPTO-50k (Schneider et al., 2016b; Dai et al., 2019). The search algorithm is Retro* (Chen et al., 2020) with a search time limit of 180 seconds or when the first route is found, whichever occurs first. MEGAN was chosen for development purposes due to its fast inference, but more high-performing models such as Retro-Chimera (Maziarz et al., 2025a) could be readily integrated in the framework.

**QUARC: reaction condition prediction model.** We utilize QUARC to annotate all chemical transformations extracted from the Syntheseus synthesis routes with top-1 predicted reaction conditions. QUARC formalizes reaction condition prediction as a four-stage autoregressive task (Sun et al., 2025). The model sequentially predicts agents, reaction temperature range, reactant stoichiometry, and reagent amounts, conditioned on the reactants and preceding outputs. All non-reactant species −including catalysts, solvents and reagents− are unified under the "agents" category, comprising a vocabulary of 1,376 distinct elements. Agent prediction is modeled as a multi-label classification task, whereas temperature range, reactant amount, and agent amount are modeled as a binned classification task. The model is trained on the Pistachio database version 2023.02 (Software, 2023), comprising around 5M unique reactions extracted from patents. Notably, the task remains highly challenging: top-1 end-to-end prediction accuracy is approximately 4%, while top-1 agent prediction reaches 25%. These performance metrics are often close to popularity-based baselines, underscoring the inherent complexity of reaction condition prediction.

**Reward formulation**. Building on prior work that integrates synthesis constraints in the generated molecules via retrosynthetic mapping (Guo & Schwaller, 2025; 2026; Guo et al., 2025), we extend this framework to optimize reaction conditions explicitly. We employ QUARC to annotate each reaction step within the predicted synthesis pathway of a generated molecule *m* where a retrosynthetic route is found. For every reaction *r* in the synthesis pathway $\mathcal{S}_m$, we assign it the top-1 set of reagents $\mathcal{A}_r$ (comprising solvents, catalyst, and reagents) and the temperature range $T_r$ predicted by QUARC. To align the generation with practical laboratory constraints, we introduce a composite condition reward $R_{condition}$ defined by three user-specified objectives:

- **Avoid reagent** ($\mathbb{1}_{\text{avoid}}$): returns 1 if no reaction step contains a reagent from a forbidden set of reagents $\mathcal{A}_{\text{unwanted}}$ (i.e., $\mathcal{A}_r \cap \mathcal{A}_{\text{unwanted}} = \emptyset$ for all $r \in \mathcal{S}_m$), and 0 otherwise.
- **Enforce reagent** ($\mathbb{1}_{\text{enforce}}$): returns 1 if all reaction steps contain a required reagent or set of reagents $\mathcal{A}_{\text{required}}$ (i.e., $\mathcal{A}_r \cap \mathcal{A}_{\text{required}} \neq \emptyset$ for all $r \in \mathcal{S}_m$), and 0 otherwise.
- **Enforce temperature** ($\mathbb{1}_{\text{temp}}$): returns 1 if the temperatures of all reaction steps fall within the specified temperature bin $T_{\text{required}}$ (i.e., $T_{\text{r}} = T_{required}$ for all $r \in \mathcal{S}_m$), and 0 otherwise.

The final condition reward is computed as the product of these three cases (if one of them is not specified by the user, it is set to 1 by default):

$$R_{condition}(m) = \mathbb{1}_{\text{avoid}} \cdot \mathbb{1}_{\text{enforce}} \cdot \mathbb{1}_{\text{temp}} \tag{1}$$

The final scalar reward acts as a multiplier to the other optimization objectives, effectively steering the generative model towards regions of chemical space accessible via the desired reaction conditions within the synthetic routes. Notice that this reward intrinsically contains the requirement that a molecule is synthesizable, and if no route is found for one molecule by the retrosynthesis model, the molecule directly receives a reward of 0. In addition, this reward can be expanded with other synthetic constraints like building block enforcing (using the TANGO reward (Guo & Schwaller, 2026)) or specific reaction types constraints (Guo et al., 2025; Dekleva et al., 2025) by simply multiplying these rewards.

**Experimental setup.** We run development experiments to compare the optimization with and without reaction condition constraints. We use a drug design-related case study previously proposed in a synthesis-constrained generative model, Reaction-GFlowNet (Koziarski et al., 2024). The objective comprises QuickVina2-GPU-2.1 (Trott & Olson, 2010; Alhossary et al., 2015; Tang et al., 2024) docking scores against an ATP-dependent Clp protease (Mabanglo et al., 2023). We also introduce QED as an additional constraint to avoid exploitation of the docking oracle due to highly lipophilic molecules (Bickerton et al., 2012; Arnott & Planey, 2012). Finally, we introduce synthesizability as a constraint following previous work (Guo & Schwaller, 2025). Our multi-parameter optimization (MPO) baseline is described by the following reward function:

$$R_{baseline}(m) = [DockingScore(m) \cdot QED(m) \cdot R_{synthesizability}(m)]^{1/3} \in [0, 1] \tag{2}$$

where $DockingScore(m)$ is the docking score computed with QuickVina2-GPU-2.1, $QED(m)$ is the QED of the molecule, and $R_{synthesizability}$ is the raw binary synthesizability. Synthesizability is defined following Guo & Schwaller (2025), and it is set to 1 if the retrosynthesis model finds a synthesis route for the generated molecule, and 0 otherwise. This baseline only considers synthesizability, and ignores synthetic constraints. $DockingScore(m)$ raw values are reshaped to fall under the [0,1] range using a reverse sigmoid transformation with $k$ = 0.15 and low ($l$) and high ($h$) boundaries at -16 and 0, respectively, according to the following expression:

$$f(x) = \frac{1}{1 + 10^{k \cdot \frac{10\left(x - \frac{h+l}{2}\right)}{h-l}}} \tag{3}$$

In order to show how direct reaction condition optimization is possible, we design two scenarios where condition optimization would be desirable. The first scenario avoids the use of dimethylformamide (DMF) and dichloromethane (DCM) as solvents and keeps the temperature range at room temperature (20-30°C). The idea is to avoid two undesired solvents and enforce room temperature according to the 12 Principles of Green Chemistry (Anastas & Eghbali, 2010). The second scenario avoids the aforementioned organic solvents and mandates the use of water across all reaction steps, incentivizing the discovery of synthetic routes compatible with benign aqueous media (in this case, we do not enforce any specific temperature range). The main purpose of these experiments is to show how reaction conditions can be selectively steered through direct rewarding. The MPO in these two cases is:

$$R_{enforcing}(m) = [DockingScore(m) \cdot QED(m) \cdot R_{condition}(m)]^{1/3} \in [0, 1] \tag{4}$$

Where $R_{condition}(m)$ is the reaction condition reward, which enforces synthesizability according to the desired reaction conditions (no DMF/DCM and water, respectively).

For each experiment, we run 10,000 oracle calls across 5 independent seeds to ensure robustness. We run each experiment on a single NVIDIA L40S GPU.

## 4 Results and discussion

### 4.1 Optimization of reaction conditions enables synthesis-constrained molecular generation.

Figure 2a displays the top-3 most frequent reaction conditions observed in the synthetic routes of generated molecules, pooled across the 5 seeds. We compare the baseline experiment against experiments employing condition constraints, where only molecules satisfying the specific constraints are considered. For the baseline, which generates molecules without condition constraints, we annotate all reactions in the resulting routes using QUARC. In this baseline set, the two most prevalent reaction conditions utilize DCM and DMF as solvents, respectively. Upon directly optimizing to avoid these solvents, the resulting generated molecules successfully exclude them from their synthetic routes. In the experiment avoiding DMF and DCM, the most common conditions shift to reactions using different solvents under ambient temperature. When specifically enforcing water, all reactions in the generated routes utilize it as a solvent while simultaneously avoiding the unwanted solvents. A similar trend is observed in the distribution of the top-20 most frequent conditions,

as detailed in appendix C. As a remark, even though some valid reaction conditions may be frequent in the baseline case, the condition optimization can steer the generated molecules towards a different chemical space that does not directly match these previously common conditions —for example, ”dioxane, HCl” is the third most frequent condition in the baseline, but it is not frequent when avoiding DMF and DCM. Figure 2b illustrates the evolution of the docking reward throughout the optimization process. The trajectories demonstrate that, despite the imposition of strict reaction condition constraints, the model effectively optimizes other MPO objectives.

Figure 3 illustrates the joint distribution of docking scores and QED for the generated molecular sets. The baseline experiments successfully access regions of the chemical space where both targets are optimized, yielding candidates such as the example molecule with a docking score of -11.5 and QED of 0.93. Conversely, optimization was less pronounced in experiments excluding DMF and DCM, likely due to the contraction of the accesible chemical space under these reaction constraints. Results for the water-enforced solvent constraint revealed a bimodal distribution: one mode or group of molecules mirrored the restricted optimization observed in the DMF/DCM-avoidance case, while the second showed more docked-optimized molecules at the cost of lower average QED. This behaviour suggests the existence of a Pareto front, where improving docking performance within the water-compatible reaction space necessitates a degradation in drug-likeness.

Table 1 presents quantitative metrics for the generated molecules, reporting the average count of synthesizable candidates, docking scores, QED values and number of unique Bemis-Murcko scaffolds (Bemis & Murcko, 1996). As anticipated, the unconstrained baseline yields a higher number of synthesizable molecules compared to the experiments constraining reaction conditions (avoiding DMF/DCM or enforcing water). In addition, the baseline achieves superior values for QED and docking (where lower scores indicate better binding), and a higher number of unique Bemis-Murcko scaffolds —reflecting higher molecular diversity. This trend —together with the joint distribution of docking scores and QED— highlights the inherent tradeoff in synthesis-constrained MPO tasks: imposing strict synthetic constraints inevitably reduces the yield of candidates that simultaneously satisfy physicochemical objectives. Overall, the results demonstrate that the reaction condition reward effectively shifts the reaction condition distribution, enabling the direct optimization of different molecular properties while successfully enforcing specified conditions.

Figure 2c exemplifies the predicted synthetic pathways for top-scoring molecules from the baseline and the condition-constrained experiments. The baseline pathway relies on prohibited solvents, specifically utilizing DMF and DCM for amide coupling and Boc deprotection, respectively. In contrast, the molecule generated under strict constraints employs an alternative solvent in a urea formation, effectively circumventing the restricted solvents. This explicit control ensures adherence to user-defined specifications and facilitates the translation of generated designs to potential experimental validation.

### 4.2 Conditions constraints implicitly steer reaction type distributions

Figure 4 presents the distribution of the ten most frequent reaction types, classified using NameRxn (Software), for each experimental setting. Imposing constraints on reaction conditions induces a marked shift in the reaction type distribution relative to the baseline. The baseline model is dominated by amide couplings and N-Boc deprotections, transformations that typically utilize DMF and DCM as solvents. By penalizing these solvents, the generative process implicitly steers the chemical space toward alternative reaction classes compatible with the imposed constraints. Specifically, avoiding DMF and DCM promotes sulfonamide formation, reductive amination, and urea formation (exemplified in Figure 2c). Conversely, enforcing water as the solvent favors Suzuki couplings, which are chemically compatible with aqueous media. To determine if this optimization merely exploits statistical biases in the training data, we monitored the prevalence of ”dominant” reaction types —defined as those comprising $>10\%$ of the training instances for a given condition. The training data (around 1.36 M reaction examples) was extracted using the code provided by the QUARC authors (Sun et al., 2025). As shown in figure 4b, the proportion of these dominant reaction types in each generated batch of molecules remains stable or decreases during the optimization. This suggests that the model does not simply converge on frequent reaction type-condition pairs by default, but rather explores diverse synthetic pathways to satisfy the specific condition constraints. Collectively, these results demonstrate that the condition-based reward function not only optimizes

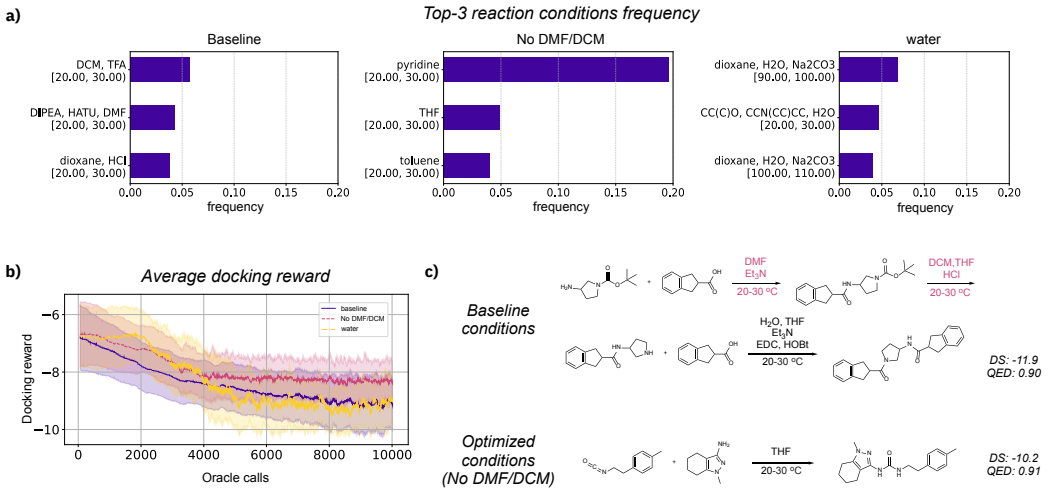


Figure 2: Condition optimization experiments. a) Frequency of the three most common reaction conditions in all reaction steps for the generated molecules pooled across seeds for each experiment (for constrained experiments, only valid routes are included). b) Average docking reward of the generated molecules throughout the optimization (more negative values correspond to better docking scores). c) Example generated molecules with their corresponding synthetic routes and conditions. In the baseline case, the unwanted conditions are highlighted. DS: docking score

Table 1: Results for generative molecular design experiments with condition enforcing. The mean and standard deviation across 5 seeds (0-4 inclusive) are reported. The number of replicates (out of 5) with at least 1 generated molecule that is synthesizable with enforced reaction conditions is reported with **N**. The number of molecules (pooled across all successful replicates) are partitioned into different docking score thresholds and statistics reported. The total number of molecules in each pool across the 5 seeds is denoted by **M**. For the docking score intervals, we report the scores, QED, and number of unique Bemis-Murcko scaffolds values (BMS).
[a] The value here denotes how many molecule are solvable by the retrosynthesis model, without any synthetic constraint.

| Experiment | Synthesizability | | Docking Score Intervals (QED & BMS Annotated) | | |
|---|---|---|---|---|---|
| | Non-solved | Solved (Enforced) | DS < -10 | -10 < DS < -9 | -9 < DS < -8 |
| Baseline (N=5) | 2609 ± 92 | 7373 ± 90[a] | -10.49 ± 0.37 (M=2354)<br>0.87 ± 0.06<br>1785 | -9.47 ± 0.28 (M=8490)<br>0.88 ± 0.06<br>5362 | -8.56 ± 0.28 (M=14261)<br>0.87 ± 0.07<br>8829 |
| No DMF/DCM (N=5) | 3218 ± 128 | 3776 ± 235 | -10.28 ± 0.24 (M=69)<br>0.80 ± 0.12<br>61 | -9.34 ± 0.23 (M=1462)<br>0.86 ± 0.08<br>1004 | -8.48 ± 0.27 (M=7490)<br>0.88 ± 0.08<br>4423 |
| Water (N=5) | 4100 ± 249 | 3171 ± 233 | -10.39 ± 0.28 (M=2082)<br>0.72 ± 0.06<br>977 | -9.61 ± 0.27 (M=4078)<br>0.74 ± 0.05<br>2075 | -8.49 ± 0.29 (M=2733)<br>0.79 ± 0.09<br>1803 |

reaction conditions but also implicitly modulates the underlying reaction type distribution without necessarily converging to a dataset-driven mode collapse.

Optimizing reaction conditions functions as an implicit mechanism for reaction type enforcement, mediated by the underlying condition prediction model. This concept parallels recent strategies where reaction types are explicitly constrained via binary (Guo et al., 2025) or continuous rewards (Dekleva et al., 2025). By restricting the condition space, the generative model is effectively confined to a subspace of compatible reaction types. This coupling may be particularly advantageous for designing reaction plates in High-Throughput Experimentation (HTE). In HTE, the parallel execution of reactions requires shared operating conditions (e.g., solvent, temperature) rather than merely identical reaction classifications. Consequently, our approach may help to maximize the diversity

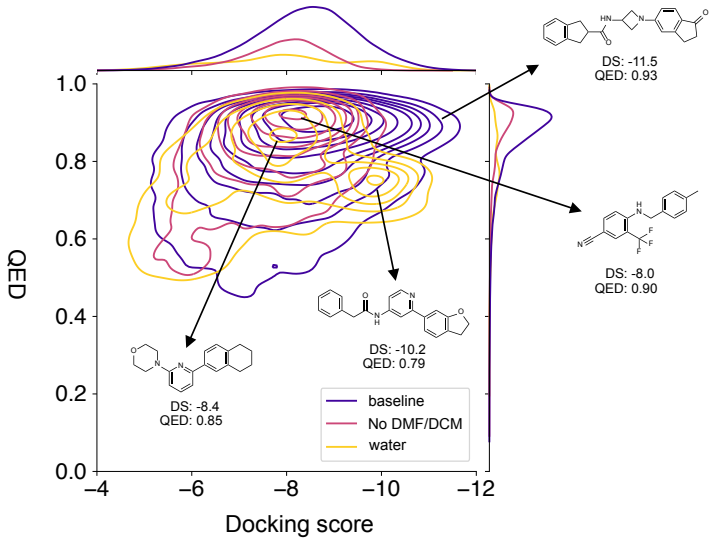


Figure 3: Docking score vs QED distribution for the generated molecules in each experiment. In the experiments enforcing reaction constraints, only molecules satisfying the constraints are considered. Example molecules from the runs are shown, together with their annotated docking score (DS) and QED.

of HTE plates while simultaneously optimizing molecular properties. This integrates plate design directly into the generative loop, offering a substantial improvement over post-hoc filtering methods employed in previous frameworks (Fromer et al., 2025; Le Vaillant et al., 2025).

A primary limitation of this approach arises from the inherent correlation between the condition prediction model and the training datasets. Current models often exhibit a bias toward dominant conditions in the training data (e.g., "popularity dominance"), rather than generalizing based on granular chemical features or reactive groups (Ball et al., 2025). Enhancing these models would facilitate more granular reward signals and a more flexible, feature-specific assignment of conditions. Furthermore, strict binary condition constraints introduce reward sparsity. As the generative process is steered toward restricted chemical spaces, the model encounters frequent negative rewards, potentially hindering efficient exploration. To address this, future work could implement continuous reward shaping to smooth the optimization landscape, analogous to the TANGO approach used for enforcing specific building blocks in the synthesis tree of the generated molecule (Guo & Schwaller, 2026).

### 4.3 Use case: Sustainable-by-design agrochemicals generation with multiple synthetic constraints

We demonstrate the utility of our framework through a practical molecular design use case aimed at discovering sustainable agrochemical candidates. While ML has recently been applied to identify novel pesticide candidates (Yang et al., 2025; Zhang et al., 2025; Kong et al., 2024), the industrial adoption of these molecules is often hindered by synthetic complexity and environmental footprint. Transitioning toward "greener" synthetic routes —exemplified by the process optimization of florylpicoxamid (Babij et al., 2020)— is essential for modern crop protection.

In this *in silico* experiment, we define a MPO objective that integrates fungicidal potency with stringent sustainability constraints (Figure 5). To assess potency, we utilize a docking oracle targeting the Succinate Dehydrogenase (SDH) protein structure, consistent with recent efforts to develop SDH inhibitors (SDHIs) for fungi pests (Zhang et al., 2025). Additionally, we add the following orthogonal synthetic constraints to adhere to Green Chemistry principles (Anastas & Eghbali, 2010):

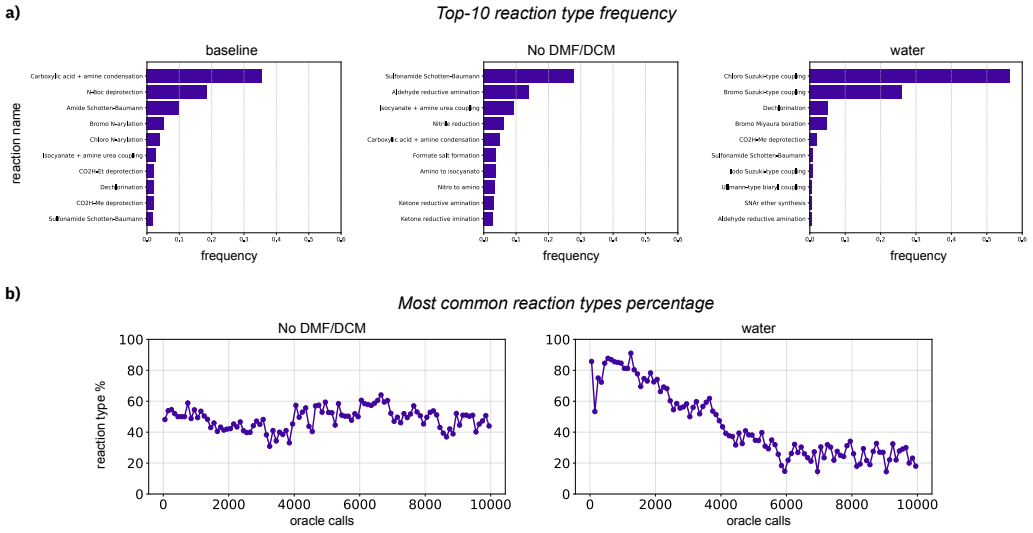


Figure 4: Reaction type shift under condition constraints. a) Top-10 reaction type distribution for each development experiment. b) Evolution of the percentage of dominant reaction types for each enforced condition in each generated batch along the optimization. For each specific reaction condition, dominant reaction types are the ones that appear in more than 10% of the training examples associated with that condition.

1. **Reaction Conditions**: Elimination of hazardous or problematic solvents (DMF, DCM) and a preference for ambient temperature regimes ($\sim$25°C).
2. **Reaction Types**: Avoidance of atom-inefficient protection/deprotection sequences using previous reaction type constraints (Guo et al., 2025).
3. **Feedstock Sustainability**: Enforced integration of biomass-derived building blocks (e.g., lignin-derived aromatics) as primary starting materials using TANGO reward (Guo & Schwaller, 2026).

The MPO objective is described in detailed in Appendix D. We also ran a benchmark to evaluate how Syntheseus default models perform in finding routes for pesticide-like molecules (see Appendix F). Benchmark results suggest that the default model we used in the development experiments was performant enough to be used in this scenario. We therefore did not change the retrosynthetic model in this task.

The optimization was conducted across five independent seeds with a budget of 10,000 oracle evaluations. Success (molecules satisfying all the constraints) was achieved in 60% of the runs, reflecting the increased complexity of navigating heavily constrained synthetic landscapes. Notably, only a subset of the biomass-derived library (2 out of 26 blocks) was incorporated into the final candidates, suggesting a trade-off between structural diversity and the stringent requirement for specific biomass precursors.

Figure 5 illustrates the top-performing candidates found for each biomass building block alongside their predicted synthetic pathways. The model successfully identified routes utilizing *trans*-sinapic acid and ferulic acid —renewable lignin derivatives (Bertella & Luterbacher, 2020)— while adhering to the solvent and temperature constraints. However, the experiment also highlights fundamental challenges in the current state-of-the-art ML predictors: certain steps exhibit potential chemoselectivity issues, and the underlying condition-prediction oracles occasionally fail to output complete condition profiles (see last step of the upper pathway —which misses the solvent— or the Suzuki couplings —which are missing the metal catalyst). These results emphasize that the fidelity of the generative output is intrinsically bound by the accuracy of the predictive synthetic oracles. In addition, the predicted reagents may still not be fully desirable, as in the case of phosphorous oxychloride in the first step of the second route, which is a toxic and corrosive reagent. Stricter sustainability constraints could be enforced in future experiments, depending on the requirements of the specific

scenario. Nevertheless, this use case validates the ability of our framework to synthesize complex, multi-objective constraints into a coherent synthesis-aware molecular discovery pipeline.

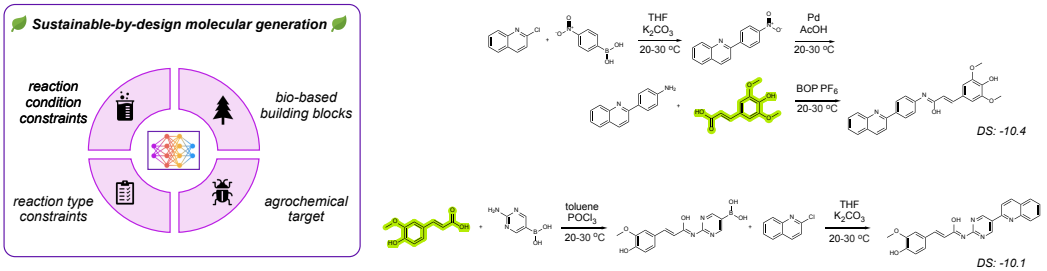


Figure 5: Sustainable-by-design generation of agrochemicals. We combine our reaction condition reward with previously reported features (reaction enforcing (Guo et al., 2025) and building block enforcing (Guo & Schwaller, 2026)) to design potential fungicides with sustainable synthetic pathways. The top generated molecule for each biomass building block and their synthetic routes are shown. Biomass-derived enforced building blocks are highlighted in green. DS: docking score.

## 5 Conclusion

Synthesizability is a remaining challenge in generative molecular design. Despite recent advances in the field, current models struggle to incorporate realistic synthetic constraints in the generative workflows. In this work, we propose a framework to directly optimize reaction conditions while generating property-optimized molecules. We reward molecules after mapping their retrosynthetic routes with a reaction condition prediction model and checking if all the conditions meet the user-specified criteria. Development experiments show how it is possible to shift the distribution of reaction conditions according to specific constraints (avoidance of DMF/DCM, and enforcing of water) while optimizing for other properties. The resulting optimization also generates an implicit shift in the reaction type distribution, similar to what was observed in previous work that directly rewarded the presence of specific reaction types (Guo et al., 2025; Dekleva et al., 2025). Finally, we show the utility of the framework in an *in silico* experiment designing potential fungicides while guaranteeing the sustainability of the generated synthetic routes. Overall, this work opens the possibility to a faster experimental validation of the generated candidates through the incorporation of realistic, laboratory conditions in the generative process.

## Code availability

The generative framework code is available on GitHub: `https://github.com/schwallergroup/saturn`

## Acknowledgments

This work was created as part of NCCR Catalysis (grant number 180544), a National Centre of Competence in Research funded by the Swiss National Science Foundation. V.S.G acknowledges support from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement N° 945363.

## References

Amr Alhossary, Stephanus Daniel Handoko, Yuguang Mu, and Chee-Keong Kwoh. Fast, accurate, and reliable molecular docking with quickvina 2. *Bioinformatics*, 31(13):2214–2216, 2015.

Paul Anastas and Nicolas Eghbali. Green chemistry: Principles and practice. *Chem. Soc. Rev.*, 39 (1):301–312, 2010. ISSN 1460-4744. doi: 10.1039/b918763b. URL http://dx.doi.org/10.1039/B918763B.

John A Arnott and Sonia Lobo Planey. The influence of lipophilicity in drug discovery and design. *Expert opinion on drug discovery*, 7(10):863–875, 2012.

Nicholas R Babij, Nakyen Choy, Megan A Cismesia, David J Couling, Nicole M Hough, Peter L Johnson, Jerzy Klosin, Xiaoyong Li, Yu Lu, Elizabeth O McCusker, Kevin G Meyer, James M Renga, Richard B Rogers, Kenneth E Stockman, Nicola J Webb, Gregory T Whiteker, and Yuanming Zhu. Design and synthesis of florylpicoxamid, a fungicide derived from renewable raw materials. *Green Chem.*, 2020.

Matthew Ball, Dragos Horvath, Thierry Kogej, Mikhail Kabeshov, and Alexandre Varnek. Predicting reaction conditions: a data-driven perspective. *Chemical Science*, 16(38):17523–17541, 2025.

BCPC. Compendium of pesticide common names, 2021. URL http://www.bcpcpesticidecompendium.org/. Database right BCPC (maker), first published 1995. Accessed: 2026-03-24.

Guy W. Bemis and Mark A. Murcko. The properties of known drugs. 1. molecular frameworks. *Journal of Medicinal Chemistry*, 39(15):2887–2893, January 1996. ISSN 1520-4804. doi: 10.1021/jm9602928. URL http://dx.doi.org/10.1021/jm9602928.

Brian J. Bender, Stefan Gahbauer, Andreas Luttens, Jiankun Lyu, Chase M. Webb, Reed M. Stein, Elissa A. Fink, Trent E. Balius, Jens Carlsson, John J. Irwin, and Brian K. Shoichet. A practical guide to large-scale docking. *Nature Protocols*, 16(10):4799–4832, September 2021. ISSN 1750-2799. doi: 10.1038/s41596-021-00597-z. URL http://dx.doi.org/10.1038/s41596-021-00597-z.

Stefania Bertella and Jeremy S. Luterbacher. Lignin functionalization for the production of novel materials. *Trends in Chemistry*, 2(5):440–453, May 2020. ISSN 2589-5974. doi: 10.1016/j.trechm.2020.03.001. URL http://dx.doi.org/10.1016/j.trechm.2020.03.001.

G Richard Bickerton, Gaia V Paolini, Jérémy Besnard, Sorel Muresan, and Andrew L Hopkins. Quantifying the chemical beauty of drugs. *Nature chemistry*, 4(2):90–98, 2012.

Esben Jannik Bjerrum. Smiles enumeration as data augmentation for neural network modeling of molecules, 2017. URL https://arxiv.org/abs/1703.07076.

Darko Butina. Unsupervised data base clustering based on daylight's fingerprint and tanimoto similarity: A fast and automated way to cluster small and large data sets. *Journal of Chemical Information and Computer Sciences*, 39(4):747–750, June 1999. ISSN 0095-2338. doi: 10.1021/ci9803381. URL http://dx.doi.org/10.1021/ci9803381.

Binghong Chen, Chengtao Li, Hanjun Dai, and Le Song. Retro*: learning retrosynthetic planning with neural guided a* search. In *International conference on machine learning*, pp. 1608–1616. PMLR, 2020.

Shuan Chen and Yousung Jung. Deep retrosynthetic reaction prediction using local reactivity and global attention. *JACS Au*, 1(10):1612–1620, 2021. doi: 10.1021/jacsau.1c00246. URL https://doi.org/10.1021/jacsau.1c00246. PMID: 34723264.

Miruna Cretu, Charles Harris, Julien Roy, Emmanuel Bengio, and Pietro Liò. Synflownet: Towards molecule design with guaranteed synthesis pathways. *arXiv preprint arXiv:2405.01155*, 2024.

Hanjun Dai, Chengtao Li, Connor Coley, Bo Dai, and Le Song. Retrosynthesis prediction with conditional graph logic network. *Advances in Neural Information Processing Systems*, 32, 2019.

Dominik Dekleva, Alexey Voronov, Jon Paul Janet, Albin Ekborg, Jure Borišek, Martina H Rambaher, and Hannes H Loeffler. Synthesizability via reward engineering: Expanding generative molecular design into synthetic space. 2025.

Clarisse Descamps, Vincent Bouttier, Juan Sanz García, Quentin Perron, and Hamza Tajmouati. Growing and linking optimizers: Synthesis-driven molecule design. *ChemRxiv*, 2025.

Joseph A DiMasi, Henry G Grabowski, and Ronald W Hansen. Innovation in the pharmaceutical industry: New estimates of R&D costs. *J. Health Econ.*, 47:20–33, May 2016.

Maarten R. Dobbelaere, István Lengyel, Christian V. Stevens, and Kevin M. Van Geem. Rxn-insight: fast chemical reaction analysis using bond-electron matrices. *Journal of Cheminformatics*, 16(1), March 2024. ISSN 1758-2946. doi: 10.1186/s13321-024-00834-z. URL http://dx.doi.org/10.1186/s13321-024-00834-z.

Yuanqi Du, Arian R Jamasb, Jeff Guo, Tianfan Fu, Charles Harris, Yingheng Wang, Chenru Duan, Pietro Liò, Philippe Schwaller, and Tom L Blundell. Machine learning-aided generative molecular design. *Nat. Mach. Intell.*, 6(6):589–604, June 2024.

Elsevier. Reaxys. https://new.reaxys.com/, 2026. Accessed: January 31, 2026.

Peter Ertl and Ansgar Schuffenhauer. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *Journal of Cheminformatics*, 1(1), June 2009. ISSN 1758-2946. doi: 10.1186/1758-2946-1-8. URL http://dx.doi.org/10.1186/1758-2946-1-8.

Jenna C Fromer, Alexandra D Volkova, and Connor W Coley. Optimal compound downselection to promote diversity and parallel chemistry. *J. Chem. Inf. Model.*, 65(12):5989–5997, June 2025.

Hanyu Gao, Thomas J. Struble, Connor W. Coley, Yuran Wang, William H. Green, and Klavs F. Jensen. Using machine learning to predict suitable conditions for organic reactions. *ACS Central Science*, 4(11):1465–1476, November 2018. ISSN 2374-7951. doi: 10.1021/acscentsci.8b00357. URL http://dx.doi.org/10.1021/acscentsci.8b00357.

Wenhao Gao and Connor W. Coley. The synthesizability of molecules proposed by generative models. *Journal of Chemical Information and Modeling*, 60(12):5714–5723, April 2020. ISSN 1549-960X. doi: 10.1021/acs.jcim.0c00174. URL http://dx.doi.org/10.1021/acs.jcim.0c00174.

Wenhao Gao, Shitong Luo, and Connor W Coley. Generative artificial intelligence for navigating synthesizable chemical space. *arXiv preprint arXiv:2410.03494*, 2024.

Samuel Genheden, Agnes Mårdh, Gustav Lahti, Ola Engkvist, Simon Olsson, and Thierry Kogej. Prediction of the chemical context for buchwald-hartwig coupling reactions. *Molecular Informatics*, 41(8), February 2022. ISSN 1868-1751. doi: 10.1002/minf.202100294. URL http://dx.doi.org/10.1002/minf.202100294.

Sai Krishna Gottipati, Boris Sattarov, Sufeng Niu, Yashaswi Pathak, Haoran Wei, Shengchao Liu, Simon Blackburn, Karam Thomas, Connor Coley, Jian Tang, et al. Learning to navigate the synthetically accessible chemical space using reinforcement learning. In *International conference on machine learning*, pp. 3668–3679. PMLR, 2020.

Albert Gu and Tri Dao. Mamba: Linear-time sequence modeling with selective state spaces. *arXiv preprint arXiv:2312.00752*, 2023.

Jeff Guo and Philippe Schwaller. Augmented memory: Sample-efficient generative molecular design with reinforcement learning. *JACS Au*, 2024.

Jeff Guo and Philippe Schwaller. Directly optimizing for synthesizability in generative molecular design using retrosynthesis models. *Chemical Science*, 16(16):6943–6956, 2025.

Jeff Guo and Philippe Schwaller. Tango: direct optimization of constrained synthesizability for generative molecular design. *Nature Computational Science*, pp. 1–11, 2026.

Jeff Guo, Jon Paul Janet, Matthias R. Bauer, Eva Nittinger, Kathryn A. Giblin, Kostas Papadopoulos, Alexey Voronov, Atanas Patronov, Ola Engkvist, and Christian Margreitter. Dockstream: a docking wrapper to enhance de novo molecular design. *Journal of Cheminformatics*, 13 (1), November 2021. ISSN 1758-2946. doi: 10.1186/s13321-021-00563-7. URL http://dx.doi.org/10.1186/s13321-021-00563-7.

Jeff Guo, Víctor Sabanza-Gil, Zlatko Jončev, Jeremy S Luterbacher, and Philippe Schwaller. Generative molecular design with steerable and granular synthesizability control. *arXiv preprint arXiv:2505.08774*, 2025.

Jeff Guo, Junwu Chen, Anthony GX-Chen, and Philippe Schwaller. Sample-efficient generative molecular design using memory manipulation. *Nature Machine Intelligence*, pp. 1–12, 2026.

Alan Kai Hassen, Martin Šícho, Yorick J. van Aalst, Mirjam C. W. Huizenga, Darcy N. R. Reynolds, Sohvi Luukkonen, Andrius Bernatavicius, Djork-Arné Clevert, Antonius P. A. Janssen, Gerard J. P. van Westen, and Mike Preuss. Generate what you can make: achieving in-house synthesizability with readily available resources in de novo drug design. *Journal of Cheminformatics*, 17(1), March 2025. ISSN 1758-2946. doi: 10.1186/s13321-024-00910-4. URL http://dx.doi.org/10.1186/s13321-024-00910-4.

Sunghwan Kim, Jie Chen, Tiejun Cheng, Asta Gindulyte, Jia He, Siqian He, Qingliang Li, Benjamin A Shoemaker, Paul A Thiessen, Bo Yu, Leonid Zaslavsky, Jian Zhang, and Evan E Bolton. Pubchem 2023 update. *Nucleic Acids Research*, 51(D1):D1373–D1380, October 2022. ISSN 1362-4962. doi: 10.1093/nar/gkac956. URL http://dx.doi.org/10.1093/nar/gkac956.

Yijin Kong, Cong Zhou, Du Tan, Xiaoyong Xu, Zhong Li, and Jiagao Cheng. Discovery of potential neonicotinoid insecticides by an artificial intelligence generative model and structure-based virtual screening. *J. Agric. Food Chem.*, 72(10):5145–5152, March 2024.

Michał Koziarski, Andrei Rekesh, Dmytro Shevchuk, Almer van der Sloot, Piotr Gaiński, Yoshua Bengio, Cheng-Hao Liu, Mike Tyers, and Robert A Batey. Rgfn: Synthesizable molecular generation using gflownets. *arXiv preprint arXiv:2406.08506*, 2024.

Franck Le Vaillant, Luidgi Gromat, Clémentine Pescheteau, Nicolas Ducrot, Aurélien Demilly, Jean-Christophe Meillon, Nicolas Do Huu, and Quentin Perron. Thinking outside the library: cluster synthesis of diverse molecules on a single robotic platform. *Chem. Sci.*, November 2025.

Seul Lee, Karsten Kreis, Srimukh Prasad Veccham, Meng Liu, Danny Reidenbach, Saee Paliwal, Weili Nie, and Arash Vahdat. Rethinking molecule synthesizability with chain-of-reaction. *arXiv preprint arXiv:2509.16084*, 2025.

Hannes H. Loeffler, Jiazhen He, Alessandro Tibo, Jon Paul Janet, Alexey Voronov, Lewis H. Mervin, and Ola Engkvist. Reinvent 4: Modern ai–driven generative molecule design. *Journal of Cheminformatics*, 16(1), February 2024. ISSN 1758-2946. doi: 10.1186/s13321-024-00812-5. URL http://dx.doi.org/10.1186/s13321-024-00812-5.

Daniel Mark Lowe. *Extraction of chemical structures and reactions from the literature*. PhD thesis, 2012.

Shitong Luo, Wenhao Gao, Zuofan Wu, Jian Peng, Connor W Coley, and Jianzhu Ma. Projecting molecules into synthesizable chemical spaces. *Proc. 41st International Conference on Machine Learning*, 2024.

Jiankun Lyu, Sheng Wang, Trent E. Balius, Isha Singh, Anat Levit, Yurii S. Moroz, Matthew J. O'Meara, Tao Che, Enkhjargal Algaa, Kateryna Tolmachova, Andrey A. Tolmachev, Brian K. Shoichet, Bryan L. Roth, and John J. Irwin. Ultra-large library docking for discovering new chemotypes. *Nature*, 566(7743):224–229, February 2019. ISSN 1476-4687. doi: 10.1038/s41586-019-0917-9. URL http://dx.doi.org/10.1038/s41586-019-0917-9.

Mark F Mabanglo, Keith S Wong, Marim M Barghash, Elisa Leung, Stephanie HW Chuang, Afshan Ardalan, Emily M Majaesic, Cassandra J Wong, Shen Zhang, Henk Lang, et al. Potent clpp agonists with anticancer properties bind with improved structural complementarity and alter the mitochondrial n-terminome. *Structure*, 31(2):185–200, 2023.

G. Marcou, J. Aires de Sousa, D. A. R. S. Latino, A. de Luca, D. Horvath, V. Rietsch, and A. Varnek. Expert system for predicting reaction conditions: The michael reaction case. *Journal of Chemical Information and Modeling*, 55(2):239–250, February 2015. ISSN 1549-960X. doi: 10.1021/ci500698a. URL http://dx.doi.org/10.1021/ci500698a.

Krzysztof Maziarz, Guoqing Liu, Hubert Misztela, Austin Tripp, Junren Li, Aleksei Kornev, Piotr Gaiński, Holger Hoefling, Mike Fortunato, Rishi Gupta, and Marwin Segler. Chemist-aligned retrosynthesis by ensembling diverse inductive bias models. *arXiv preprint arXiv:2412.05269*, 2025a.

Krzysztof Maziarz, Austin Tripp, Guoqing Liu, Megan Stanley, Shufang Xie, Piotr Gaiński, Philipp Seidl, and Marwin HS Segler. Re-evaluating retrosynthesis algorithms with syntheseus. *Faraday Discussions*, 256:568–586, 2025b.

Marcus Olivecrona, Thomas Blaschke, Ola Engkvist, and Hongming Chen. Molecular de-novo design through deep reinforcement learning. *Journal of cheminformatics*, 9:1–14, 2017.

Maud Parrot, Hamza Tajmouati, Vinicius Barros Ribeiro da Silva, Brian Ross Atwood, Robin Fourcade, Yann Gaston-Mathé, Nicolas Do Huu, and Quentin Perron. Integrating synthetic accessibility with ai-based generative drug design. *Journal of Cheminformatics*, 15(1), September 2023. ISSN 1758-2946. doi: 10.1186/s13321-023-00742-8. URL http://dx.doi.org/10.1186/s13321-023-00742-8.

Priyanka Raghavan, Brittany C. Haas, Madeline E. Ruos, Jules Schleinitz, Abigail G. Doyle, Sarah E. Reisman, Matthew S. Sigman, and Connor W. Coley. Dataset design for building models of chemical reactivity. *ACS Central Science*, 9(12):2196–2204, December 2023. ISSN 2374-7951. doi: 10.1021/acscentsci.3c01163. URL http://dx.doi.org/10.1021/acscentsci.3c01163.

Mikołaj Sacha, Mikołaj Błaz, Piotr Byrski, Paweł Dabrowski-Tumanski, Mikołaj Chrominski, Rafał Loska, Paweł Włodarczyk-Pruszynski, and Stanisław Jastrzebski. Molecule edit graph attention network: modeling chemical reactions as sequences of graph edits. *Journal of Chemical Information and Modeling*, 61(7):3273–3284, 2021.

Nadine Schneider, Nikolaus Stiefl, and Gregory A. Landrum. What's what: The (nearly) definitive guide to reaction role assignment. *Journal of Chemical Information and Modeling*, 56 (12):2336–2346, December 2016a. ISSN 1549-960X. doi: 10.1021/acs.jcim.6b00564. URL http://dx.doi.org/10.1021/acs.jcim.6b00564.

Nadine Schneider, Nikolaus Stiefl, and Gregory A Landrum. What's what: The (nearly) definitive guide to reaction role assignment. *Journal of chemical information and modeling*, 56(12):2336–2346, 2016b.

Seonghwan Seo, Minsu Kim, Tony Shen, Martin Ester, Jinkyoo Park, Sungsoo Ahn, and Woo Youn Kim. Generative flows on synthetic pathway for drug design. *arXiv preprint arXiv:2410.04542*, 2024.

NextMove Software. Namrxn.

NextMove Software. Pistachio reaction database. https://www.nextmovesoftware.com/pistachio.html, 2023.

Jacob O. Spiegel and Jacob D. Durrant. Autogrow4: an open-source genetic algorithm for de novo drug design and lead optimization. *Journal of Cheminformatics*, 12(1), April 2020. ISSN 1758-2946. doi: 10.1186/s13321-020-00429-4. URL http://dx.doi.org/10.1186/s13321-020-00429-4.

Megan Stanley and Marwin Segler. Fake it until you make it? generative de novo design and virtual screening of synthesizable molecules. *Curr. Opin. Struct. Biol.*, 82(102658):102658, October 2023.

Xiaoqi Sun, Jiannan Liu, Babak Mahjour, Klavs F. Jensen, and Connor W. Coley. Data-driven recommendation of agents, temperature, and equivalence ratios for organic synthesis. *Chem. Sci.*, 16:18176–18189, 2025. doi: 10.1039/D5SC04957A. URL http://dx.doi.org/10.1039/D5SC04957A.

Kyle Swanson, Gary Liu, Denise B Catacutan, Autumn Arnold, James Zou, and Jonathan M Stokes. Generative ai for designing and validating easily synthesizable and structurally novel antibiotics. *Nature Machine Intelligence*, 6(3):338–353, 2024.

Shidi Tang, Ji Ding, Xiangyu Zhu, Zheng Wang, Haitao Zhao, and Jiansheng Wu. Vina-gpu 2.1: towards further optimizing docking speed and precision of autodock vina and its derivatives. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2024.

Oleg Trott and Arthur J Olson. Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, 31(2):455–461, 2010.

H. Maarten Vinkers, Marc R. de Jonge, Frederik F. D. Daeyaert, Jan Heeres, Lucien M. H. Koymans, Joop H. van Lenthe, Paul J. Lewi, Henk Timmerman, Koen Van Aken, and Paul A. J. Janssen. Synopsis: Synthesize and optimize system in silico. *Journal of Medicinal Chemistry*, 46(13): 2765–2773, May 2003. ISSN 1520-4804. doi: 10.1021/jm030809x. URL http://dx.doi.org/10.1021/jm030809x.

Mingyang Wang, Shuai Li, Jike Wang, Odin Zhang, Hongyan Du, Dejun Jiang, Zhenxing Wu, Yafeng Deng, Yu Kang, Peichen Pan, et al. Clickgen: Directed exploration of synthesizable chemical space via modular reactions and reinforcement learning. *Nature Communications*, 15 (1):10127, 2024.

Zihan Wang, Kangjie Lin, Jianfeng Pei, and Luhua Lai. Reacon: a template- and cluster-based framework for reaction condition prediction. *Chemical Science*, 16(2):854–866, 2025. ISSN 2041-6539. doi: 10.1039/d4sc05946h. URL http://dx.doi.org/10.1039/D4SC05946H.

Wendy Warr. Report on an nih workshop on ultralarge chemistry databases. May 2021. doi: 10.26434/chemrxiv.14554803.v1. URL http://dx.doi.org/10.26434/chemrxiv.14554803.v1.

David Weininger. Smiles, a chemical language and information system. 1. introduction to methodology and encoding rules. *Journal of chemical information and computer sciences*, 28(1):31–36, 1988.

Shufang Xie, Rui Yan, Junliang Guo, Yingce Xia, Lijun Wu, and Tao Qin. Retrosynthesis prediction with local template retrieval, 2023. URL https://arxiv.org/abs/2306.04123.

Zuojun Xu, Feng Ren, Ping Wang, Jie Cao, Chunting Tan, Dedong Ma, Li Zhao, Jinghong Dai, Yipeng Ding, Haohui Fang, Huiping Li, Hong Liu, Fengming Luo, Ying Meng, Pinhua Pan, Pingchao Xiang, Zuke Xiao, Sujata Rao, Carol Satler, Sang Liu, Yuan Lv, Heng Zhao, Shan Chen, Hui Cui, Mikhail Korzinkin, David Gennert, and Alex Zhavoronkov. A generative AI-discovered TNIK inhibitor for idiopathic pulmonary fibrosis: a randomized phase 2a trial. *Nat. Med.*, 31(8):2602–2610, August 2025.

Ruoqi Yang, Biao Li, Jin Dong, Zhuomei Cai, Hongyan Lin, Fan Wang, and Guangfu Yang. Reinforcement learning-based generative artificial intelligence for novel pesticide design. *J. Adv. Res.*, 78:179–190, December 2025.

Yuan Zhang, Jianqi Chai, Ling Li, Wenqian Zhao, Jinwen Ji, Yanru Li, Ziyu Wang, Yuanyuan Chen, Liangyun Zhang, Zhihui Xu, Chunlong Yang, and Cong Pian. Design and optimization of novel succinate dehydrogenase inhibitors against agricultural fungi based on transformer model. *Mol. Divers.*, August 2025.

Alex Zhavoronkov, Yan A Ivanenkov, Alex Aliper, Mark S Veselov, Vladimir A Aladinskiy, Anastasiya V Aladinskaya, Victor A Terentiev, Daniil A Polykovskiy, Maksim D Kuznetsov, Arip Asadulaev, Yury Volkov, Artem Zholus, Rim R Shayakhmetov, Alexander Zhebrak, Lidiya I Minaeva, Bogdan A Zagribelnyy, Lennart H Lee, Richard Soll, David Madge, Li Xing, Tao Guo, and Alán Aspuru-Guzik. Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nat. Biotechnol.*, 37(9):1038–1040, September 2019.

## A SATURN: REINFORCEMENT LEARNING ALGORITHM

The main framework of Saturn uses a pre-trained language model on a dataset of molecules represented as SMILES (PubChem (Kim et al., 2022) in this case) and applies RL fine-tuning. See appendix B for more details about pretraining. The RL optimization is based on the REINVENT algorithm (Olivecrona et al., 2017; Loeffler et al., 2024) with additional modifications. During the optimization, the pre-trained model is defined as the Prior and has frozen weights. Every generated batch of SMILES is passed through the Prior, returning a log-likelihood. The Agent is initialized also using the weights of the pre-trained model, but these are updated during the RL loop. According to REINVENT, the Augmented Likelihood is defined as:

$$\log\pi_{\text{Augmented}}(x) = \log\pi_{\text{Prior}}(x) + \sigma R(x) \tag{5}$$

where $R$ is any reward function that can assign a scalar reward $\in [0, 1]$ to an input SMILES. $\sigma$ represents a scalar function that can modulate the contribution of the reward signal. RL is done by using the following loss function:

$$L(\theta) = (\log\pi_{\text{Augmented}} - \log\pi_{\theta_{\text{Agent}}})^2 \tag{6}$$

that minimizes the squared distance between the Augmented and the Agent log-likelihoods. The Augmented Likelihood is a linear combination of the frozen Prior and the reward, so the objective is maximizing the reward while avoiding that the Agent deviates too far from the pre-trained distribution. $\theta$ indicates that the Agent is parametrized by the neural network (Mamba (Gu & Dao, 2023) in this case).

Finally, Saturn implements an Augmented Memory mechanism (Guo & Schwaller, 2024) to improve the sample efficiency of the model. A replay buffer stores the top-100 SMILES by reward generated during the optimization. The Agent is updated each epoch both on the newly generated batch of SMILES and the high-reward SMILES stored in the replay buffer. This update is done N times (each one called an Augmentation Round), during which each SMILES in the buffer is randomized (Bjerrum, 2017) (augmented) by generating an equivalent sequence for the same molecule. The Agent is therefore updated with different representations of the same structures, and this method increases sample efficiency. Diversity filters are also applied in the reply buffer following REINVENT (Loeffler et al., 2024) to penalize Bemis-Murcko (Bemis & Murcko, 1996) scaffolds that have been generated more than 10 times. This feature improves exploration and increases the diversity of the generated molecules.

## B SATURN PRETRAINING

All the models in this work used the same pretrained prior model from a previous work on synthesis-constrained molecular design (Guo & Schwaller, 2026). This section contains a copy of the information from the original publication for completeness. The full data pre-processing and pretraining pipeline started from the raw PubChem which was downloaded from `https://ftp.ncbi.nlm.nih.gov/pubchem/Compound/Extras/`. The exact file is "CID-SMILES.gz". **Note that there is now a new version of PubChem. We previously downloaded the version close to Aug. 7, 2024.**

The exact pre-processing steps along with the SMILES remaining after each step are:

1. Raw PubChem - 118,563,810
2. De-duplication - 118,469,904
3. Standardization (charge and isotope handling) based on `https://github.com/MolecularAI/ReinventCommunity/blob/master/notebooks/Data_Preparation.ipynb`. All SMILES that could not be parsed by RDKit were removed - 109,128,315
4. Tokenize all SMILES based on REINVENT's tokenizer: `https://github.com/MolecularAI/reinvent-models/blob/main/reinvent_models/reinvent_core/models/vocabulary.py`

5. Keep SMILES $\leq$ 80 tokens, 150 $\leq$ molecular weight $\leq$ 650, number of heavy atoms $\leq$ 40, number of rings $\leq$ 8, Size of largest ring $\leq$ 8, longest aliphatic carbon chain $\leq$ 4 - 97,667,549
6. Removed SMILES containing the following tokens (due to undesired chemistry, low token frequency, and redundancy): [Br+2], [Br+3], [Br+], [C+], [C-], [CH+], [CH-], [CH2+], [CH2-], [CH2], [CH], [C], [Cl+2], [Cl+3], [Cl+], [ClH+2], [ClH2+2], [ClH3+3], [N-], [N@+], [N@@+], [NH+], [NH-], [NH2+], [NH3+], [NH], [N], [O+], [OH+], [OH2+], [O], [S+], [S-], [S@+], [S@@+], [S@@], [S@], [SH+], [SH-], [SH2], [SH4], [SH], [S], [c+], [c-], [cH+], [cH-], [c], [n+], [n-], [nH+], [nH], [o+], [s+], [sH+], [sH-], [sH2], [sH4], [sH], [s] - **88,618,780**

The final vocabulary contained 35 tokens (2 extra tokens were added, indicating $< START >$ and $< END >$) and carbon stereochemistry tokens were kept. Saturn (Guo et al., 2026) uses the Mamba (Gu & Dao, 2023) architecture and we used the default hyperparameters in the code-base. With the vocabulary size of 35, the model has 5,265,408 parameters. Saturn was pre-trained for 5 steps, with each step consisting of a full pass through the dataset. The model was pre-trained on a workstation with an NVIDIA RTX 3090 GPU and AMD Ryzen 9 5900X 12-Core CPU. The pre-training parameters were:

1. Training steps = 5
2. Seed = 0
3. Batch size = 512
4. Learning rate = 0.0001
5. Randomize (Bjerrum, 2017) every batch of SMILES

Relevant metrics of the pre-trained model (final model checkpoint) are:

1. Average negative log-likelihood (NLL) = 30.914
2. Validity (10k) = 98.74%
3. Uniqueness (10k) = 98.73%
4. Wall time = 106 hours (takes a relatively long time, but pre-training only needs to be done once)

We run Saturn using the same hyperparameters than a previous study on synthesis-constrained molecular optimization (Guo et al., 2025).

## C ADDITIONAL DEVELOPMENT RESULTS

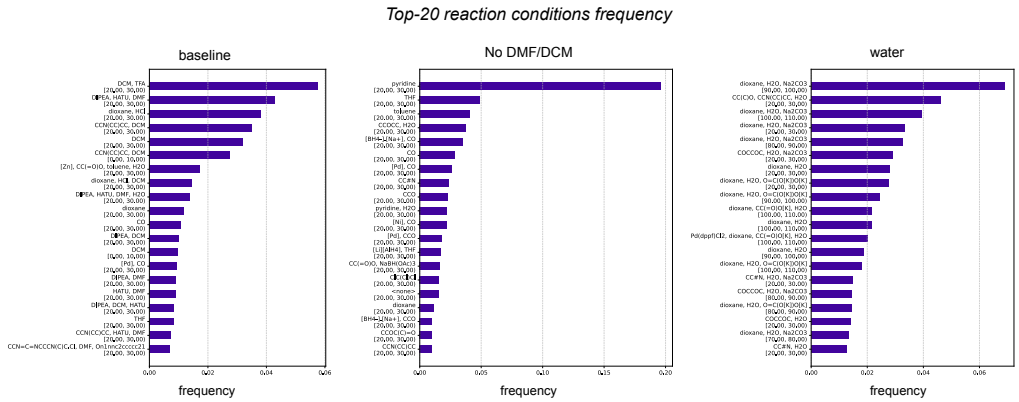


Figure 6: Top-20 reaction condition distribution for each development experiment.

## D PESTICIDE MPO SETUP DETAILS

The pesticide-derived MPO reward is defined as follows:

$$R_{pesticide}(m) = [DockingScore(m) \cdot R_{synthesizability}(m)]^{1/2} \in [0, 1] \tag{7}$$

$R_{synthesizability}$ is the synthesizability reward combining the following synthetic parameters:

1. Reaction conditions: binary reward as explained in the Methods section, avoiding DMF and DCM and enforcing the 20-30°C range).
2. Reaction types: binary reward avoiding protection/deprotection reaction types as defined in Guo et al. (2025).
3. Biomass-derived building blocks: enforce biomass-derived building blocks from a given list (see section E) using the TANGO reward (Guo & Schwaller, 2026).

$DockingScore(m)$ is the docking score of the given molecule against a fungi-related protein. To establish the docking oracle for the fungicide succinate dehydrogenase inhibitor (SDHI) case study, the crystal structure of the protein-ligand complex was retrieved from the Protein Data Bank (PDB ID: 2FBW), following the use case of previous work on molecular generative design for agrochemicals (Zhang et al., 2025). Structural preparation involved the removal of solvent molecules and metal complexes, followed by the isolation of a single subunit from the native tetrameric assembly. The protein and reference ligand (carboxin) were separated, and the protein structure was subsequently refined using PDBFixer via the Dockstream interface (Guo et al., 2021). The receptor was exported in .pdbqt format and integrated as the primary input for the QuickVina2-GPU docking oracle to evaluate binding affinities. The docking score raw values are reshaped to fall under the [0,1] range using a reverse sigmoid transformation with $k$ = 0.15 and low ($l$) and high ($h$) boundaries at -16 and 0, respectively (see equation 3).

## E ENFORCED BIOMASS-DERIVED BUILDING BLOCKS

The following enforced building blocks correspond to molecules that can be derived from lignin, a sustainable resource (Bertella & Luterbacher, 2020). These blocks are used as the enforced block set in the sustainable-by-design experiments.

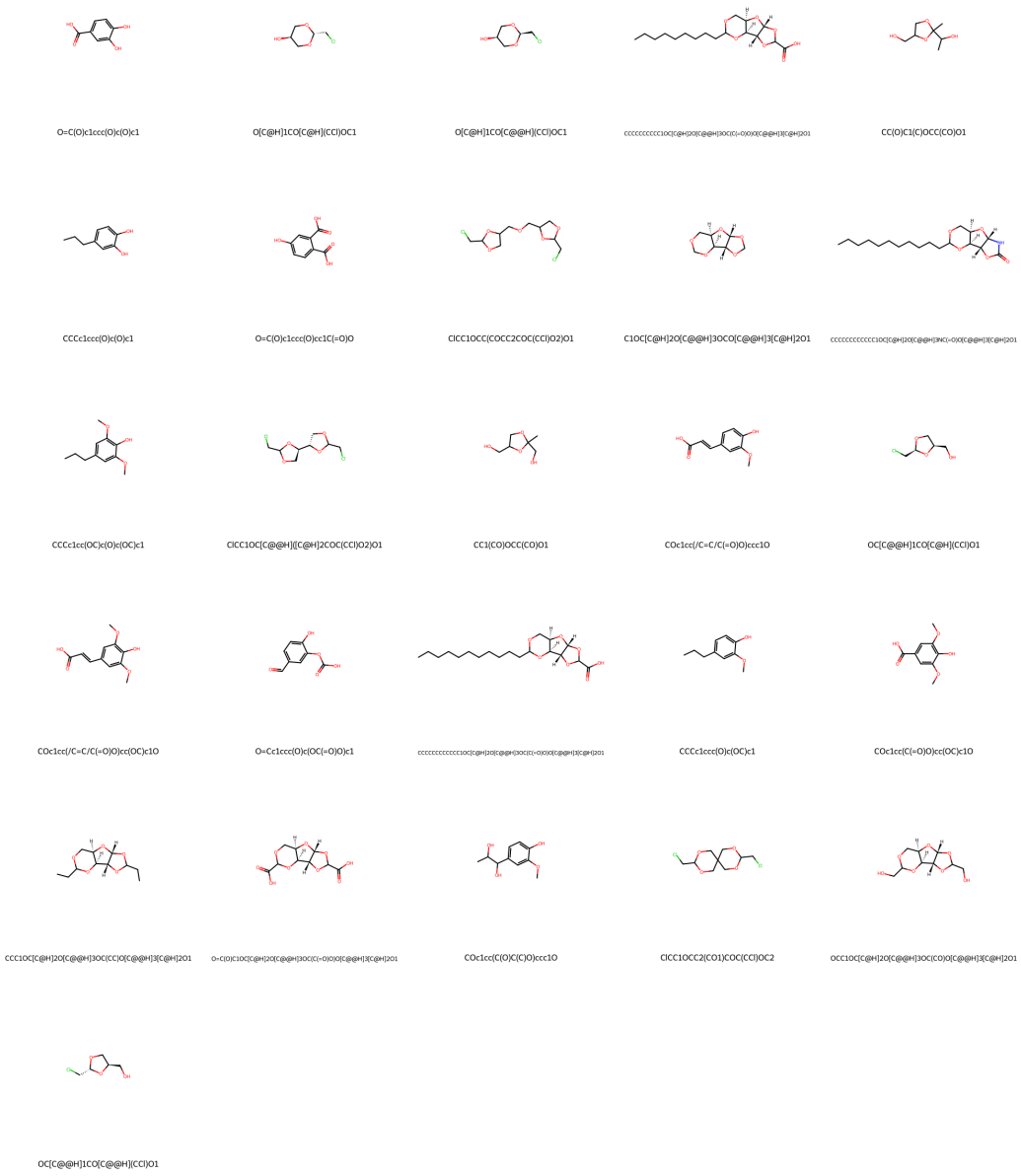


Figure 7: Sustainable building blocks derived from lignocellulosic biomass.

## F RETROSYNTHESIS BENCHMARK ON PESTICIDE-DERIVED TARGETS

To complement the generative experiments, we constructed an external retrosynthesis benchmark based on pesticide-like molecules and evaluated a set of Syntheseus (Maziarz et al., 2025b) single-step model/search algorithm combinations in this collection. The benchmark was designed to probe multi-step retrosynthetic performance on agrochemical molecular structures rather than on standard medicinal chemistry targets to tailor the experiments towards a realistic use case.

The target set was derived from the BCPC Pesticide Compendium (BCPC, 2021) by extracting pesticide entries together with their CAS number and then resolved molecular structures as SMILES using CAS-based CIR lookups with PubChem InChIKey fallback. The resulting raw collection contained 2564 molecules. We subsequently filtered it to retain molecules with $\mathrm{QED} \geq 0.5$, while excluding compounds with invalid SMILES, heavy metals, very short aliphatic structures (less than 5 carbon atoms), and single simple aromatic rings. We then removed targets already present in the purchasable inventory used by Syntheseus (Enamine EU September 2025 commercial building blocks), reducing the set to 1527 molecules.

To promote structural diversity, the remaining molecules were clustered with RDKit Morgan fingerprints (radius 2, 2048 bits) and Butina clustering Butina (1999) at a distance cutoff of 0.4. We finally selected the centroids of the 100 largest clusters to create the target molecules for the retrosynthesis benchmark. This procedure ensured that the benchmark emphasized representative and structurally diverse pesticide-like targets that were not trivially solvable by direct inventory lookup.

Retrosynthetic searches were run through Syntheseus using three different models (LocalRetro (Chen & Jung, 2021), MEGAN (Sacha et al., 2021) and RetroKNN (Xie et al., 2023)), combined with two search algorithms (Monte Carlo Tree Search (MCTS) and Retro* (Chen et al., 2020)), resulting in six model/search settings. All the single-step reaction models were trained on USPTO-50KSchneider et al. (2016a). All runs used the same Enamine building-block inventory, a wall-clock limit of 180 s per target, and the same set of 100 targets. Benchmark aggregation was performed by parsing the route files for each solved target and reporting the solving rate, mean wall-clock solution time, average route length, and average number of returned routes.

Additionally, we also modified the number of single-step reaction proposals retained at each expansion step, using 1 and 50 proposals (table 2 and table 3, respectively). Table 2 shows the benchmark results for the first case, where only one reaction proposal is retained. Under this strict setting, solving rates were low across all methods (20–23%), but MEGAN and LocalRetro gave the fastest solution times, whereas RetroKNN produced slightly higher solving rates at the cost of longer search times. Table 3 reports the second scenario in which 50 single-step proposals were retained in each expansion. In this more permissive setting, solving rates increased substantially to 71–83%, with MEGAN + Retro* providing the best overall performance on this benchmark. As expected, the larger number of retained solutions also increased average route length and wall-clock time.

Based on these benchmark results, we selected MEGAN and Retro* for the subsequent experiments, as it provided the most favorable balance between solving performance and computational cost. In the broader search setting, it achieved the highest solving rate (0.83) while maintaining a lower average wall-clock time (4.70 s) than the other methods. In the top-1 single-step expansion setting used for the main experiments, it also remained one of the fastest configurations (0.022 s) with competitive solving performance (0.21). We set the number of single-step reaction proposals returned at each expansion step to 1 to ensure feasible runtimes during optimization.

Table 2: Syntheseus benchmark on 100 pesticide-derived targets when retaining 1 reaction proposal per expansion step.

| Model + search algorithm | Solving rate | Mean time (s) | Avg. route length | Avg. # routes |
|---|---|---|---|---|
| LocalRetro + MCTS | 0.20 | 0.041 | 1.80 | 1.10 |
| LocalRetro + Retro* | 0.20 | 0.034 | 1.60 | 1.00 |
| MEGAN + MCTS | 0.22 | 0.025 | 2.14 | 1.05 |
| MEGAN + Retro* | 0.21 | 0.022 | 1.90 | 1.00 |
| RetroKNN + MCTS | 0.23 | 0.174 | 2.35 | 1.22 |
| RetroKNN + Retro* | 0.22 | 0.157 | 2.27 | 1.00 |

Table 3: Syntheseus benchmark on 100 pesticide-derived targets when retaining 50 reaction proposals per expansion step.

| Model + search algorithm | Solving rate | Mean time (s) | Avg. route length | Avg. # routes |
|---|---|---|---|---|
| LocalRetro + MCTS | 0.71 | 5.255 | 3.24 | 4.83 |
| LocalRetro + Retro* | 0.75 | 8.098 | 5.43 | 4.91 |
| MEGAN + MCTS | 0.79 | 7.223 | 4.43 | 4.82 |
| MEGAN + Retro* | 0.83 | 4.697 | 5.60 | 5.00 |
| RetroKNN + MCTS | 0.71 | 9.647 | 3.45 | 4.76 |
| RetroKNN + Retro* | 0.72 | 7.036 | 5.17 | 4.89 |