# Towards Reliable Chemical Retrosynthesis: A Self-Reflection Approach for Robust Retrosynthesis Planning

Anonymous ACL submission

## Abstract

Retrosynthetic analysis is a critical task in organic chemistry. However, the application of Large Language Models (LLMs) in retrosynthesis is often hindered by hallucinations and the high demand for internalizing vast chemical knowledge. To address these challenges, we propose ChemValiSynth, a framework designed for retrosynthetic prediction. It constructs high-quality training data via a self-reflective loop. Furthermore, we introduce a decoupled dual-model architecture that separates molecular analysis from strategy prediction, thereby enhancing the model's capacity to incorporate extensive chemical information. Empirical results on the USPTO-50K benchmark report a Top-1 accuracy of 58.7%, validating the effectiveness of our proposed approach.

## 1 Introduction

Retrosynthetic analysis is a cornerstone of organic chemistry. It is essential for designing feasible synthetic routes for target molecules. The process involves identifying reactive centers, predicting leaving groups, and ensuring that the resulting synthons are chemically reasonable (Zhu et al., 2023). However, the chemical space is vast, and reaction rules are highly complex and interdependent. Consequently, manual exploration is often inefficient and relies heavily on expert intuition. This complexity has driven the adoption of data-driven approaches to automate and accelerate the search for retrosynthetic predictions.

Computer-aided retrosynthesis has evolved rapidly, but existing models still face limitations. Early template-based approaches (Chen and Jung, 2021; Dai et al., 2019) are easy to interpret, yet they rely on rigid, predefined rules. As a result, they often struggle to generalize to new reactions. In contrast, template-free sequence methods (Tetko et al., 2020; Zhong et al., 2022) offer greater flexibility. However, without explicit chemical constraints, they frequently generate chemically invalid structures. Semi-template strategies (Yan et al., 2020; Shi et al., 2020; Somnath et al., 2021) attempt to balance these two methods. Unfortunately, they typically require complex intermediate steps like atom-mapping, which can introduce errors. Recently, Large Language Models (LLMs) and their domain-specific variants (Zhao et al., 2024b; Bai et al., 2025) have emerged as a promising paradigm (Roumeliotis and Tselikas, 2023), leveraging robust reasoning capabilities and demonstrating strong alignment with professional knowledge. Notwithstanding these advancements, this new approach introduces two distinct challenges that remain to be addressed.

First, LLMs exhibit chemistry-specific hallucinations (Zhao et al., 2024a). They often generate reactions that are syntactically valid but chemically infeasible. While these outputs adhere to notation rules, they frequently violate thermodynamic or kinetic principles. This issue originates from the probabilistic nature of LLMs rather than causal reasoning. Additionally, the "survivorship bias" in literature exacerbates the problem. The scarcity of negative data prevents models from distinguishing theoretical syntax from experimental feasibility (Maloney et al., 2023). However, this deficiency offers a unique opportunity for optimization. The discrepancy between generation and reality reveals specific reasoning gaps. By analyzing these errors to extract targeted knowledge, hallucinations serve as a valuable feedback mechanism (Madaan et al., 2023). This approach refines feasibility boundaries and enhances retrosynthetic performance.

Second, LLMs rely on linear strings (e.g., SMILES) to represent molecules, but grounding these text-based representations in chemical reality requires massive corpora. This imposes high demands on the model's capacity to internalize such vast information. Furthermore, diverse fields like drug discovery and materials science require distinct tactical preferences, making domain-specific fine-tuning essential. However, relying on a single LLM creates a conflict: fine-tuning for specific domains often causes the model to "forget" its general retrosynthetic reasoning capabilities (Guo et al., 2023). This phenomenon is formally known as catastrophic forgetting (Shi et al., 2025).

To address the above challenges, we propose ChemValiSynth. This framework centers on a **Generate–Verify–Reflect–Optimize** loop to transform raw model outputs into high-quality, verified instruction-tuning data. Through this iterative cycle, the system autonomously generates retrosynthetic plans, validates them against chemical constraints, and performs self-reflection to refine the results. To address the high demand for molecular structural information and domain-

specific fine-tuning, we utilize a **decoupled dual-model architecture**. By separating molecular analysis from strategy prediction, this design integrates extensive chemical knowledge without causing the catastrophic forgetting typical of single-model fine-tuning. By training on the resulting optimized corpus, ChemValiSynth bridges the gap between textual syntax and chemical reality. The main contributions of this paper are summarized as follows:

- We introduce ChemValiSynth, a novel retrosynthetic prediction framework that significantly enhances the retrosynthetic reasoning capabilities of LLMs.
- We propose a Generate–Verify–Reflect–Optimize pipeline that uses chemical verifiers to convert LLM-generated retrosynthetic rationales into verified training samples. This is complemented by a decoupled architecture separating molecular analysis from strategy prediction. The design provides explicit chemical cues to guide reasoning and mitigates catastrophic forgetting during domain adaptation.
- ChemValiSynth attains 58.7% top-1 accuracy on USPTO-50K, establishing a competitive baseline that approaches the current state-of-the-art methods. Ablation studies demonstrate that both the verification loop and the decoupled dual-model design are critical to achieving these performance gains.

## 2 Related Work

The retrosynthetic prediction has evolved significantly, transitioning from traditional rule-based systems to advanced deep learning architectures. Early template-based approaches (Dai et al., 2019; Chen and Jung, 2021; Zhong et al., 2024) offered high interpretability by employing encoded chemical rules; however, their applicability was fundamentally constrained to reactions within predefined rule sets, limiting generalization to novel or unexplored chemical spaces. Contemporary computational strategies have since bifurcated into two predominant paradigms. The first employs graph-based models (Sun et al., 2025; Zhong et al., 2023), such as G2GS (Shi et al., 2020), which explicitly represent molecules as graphs to capture topological relationships and reaction center transformations. While effective, these models often encounter scalability challenges when processing large or complex molecular systems. The second paradigm comprises sequence-based models (Schwaller et al., 2019; Wan et al., 2022; Maziarz et al., 2024), such as R-SMILES (Zhong et al., 2022), which formalize retrosynthesis as a neural machine translation task between molecular string representations. Although these models benefit from streamlined end-to-end training and efficient inference, they are frequently criticized for their opaque "black-box" nature, which obscures underlying chemical reasoning and impedes interpretability.

Recently, the emergence of LLMs has introduced a third paradigm, leveraging extensive pre-training and chain-of-thought reasoning to simulate logical retrosynthetic analysis (Xia et al., 2025). Subsequent efforts have produced chemically fine-tuned LLMs, such as ChemDFM (Zhao et al., 2024b), trained on specialized corpora. Despite these advancements, the performance of LLMs in retrosynthesis planning still lags behind domain-specific models. This shortfall is primarily caused by persistent issues like severe hallucination. Specifically, models often generate invalid structures or implausible reactions, while also lacking deep chemical expertise (Nicholas et al., 2025). While techniques like knowledge distillation from expert models have been explored to bridge these gaps, they typically demand substantial volumes of high-quality data and significant computational resources, resulting in prohibitive development costs.

In general domains, LLMs improve through self-reflection (Shinn et al., 2023), but this requires a reliable verifier. Fortunately, retrosynthesis offers a natural solution: forward reaction predictors (Chen and Jung, 2022) serve as these verifiers. Based on this insight, we propose a comprehensive framework. First, we design a data pipeline that uses these chemical verifiers to fix hallucinations and generate high-quality training data. Second, we introduce a dual-model retrosynthetic architecture to learn from this data. This architecture decouples the reasoning process into two specialized parts: molecular analysis and strategy prediction. By separating these tasks, our model achieves accurate and reliable retrosynthesis planning.

## 3 Method

### 3.1 Overview

We propose ChemValiSynth, a framework that improves retrosynthetic reliability via a **decoupled dual-model architecture**. It separates explicit molecular analysis from implicit strategic reasoning. The framework operates in two phases. (1) **Self–Reflective Data Construction**: A "Generate–Verify–Reflect–Optimize" loop leverages external chemical verifiers to iteratively correct hallucinations, producing a high-quality training dataset. (2) **Dual-Model Collaboration**: A reasoning agent ($LLM_{retro}$) is optimized via supervised fine-tuning and reinforcement learning to align with chemical constraints. An auxiliary agent ($LLM_{info}$) predicts structural contexts to guide reasoning. This integration ensures retrosynthetic predictions remain grounded in validated chemical principles.

### 3.2 Problem Formulation

The goal of retrosynthetic analysis is, for a given product molecule $P$, to identify a set of feasible reactants $R = \{R_1, R_2, \ldots, R_k\}$ and the corresponding reaction conditions $C$ such that $P$ can be synthesized from $R$ under $C$ with high efficiency and high selectivity.

To formalize the retrosynthetic task, the chemical space $\mathcal{M}$ is defined as the set of all possible molecules. A retrosynthetic plan can be represented as a tuple $\mathcal{S} = (R, C)$, where $R \subset \mathcal{M}$ is a non-empty set of

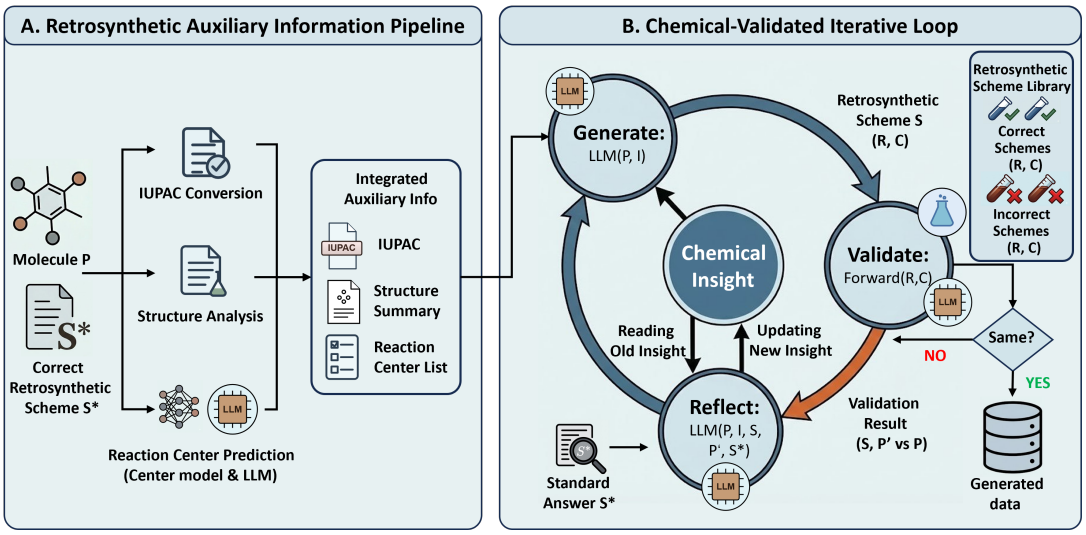



Figure 1: The data construction process of ChemValiSynth. Figure A illustrates the process of constructing auxiliary information for retrosynthesis; Figure B illustrates the self-reflection data construction pipeline, where the model explores reverse synthesis space through a Generate–Verify–Reflect–Optimize process.

reactants, and $C$ is the set of reaction conditions. Typically, multiple feasible plans exist for a target molecule. Based on the evaluation criteria, the plans can be categorized into the following two types:

1. Strictly Optimal Plan $\mathcal{S}^*$: This aligns with expert-designed standard protocols. It is optimal in metrics such as synthetic efficiency, atom economy, step simplicity, and precursor availability. Here, we treat the ground truth in the dataset as the strictly optimal plan.

2. Chemically Feasible Plan $\mathcal{S}^+$: This satisfies basic chemical correctness, meaning there exists a well-validated forward reaction prediction model $f_{\text{forward}}$ such that $f_{\text{forward}}(R, C) = P$. This plan is not required to be identical to the standard answer, but must ensure the reaction principle is plausible.

### 3.3 Self-Reflective Data Construction

The ChemValiSynth framework utilizes an automated data construction pipeline to address chemical hallucinations. This iterative process follows a "Generate–Verify–Reflect–Optimize" closed loop to produce high-quality training data without extensive manual annotation. The workflow starts with the auxiliary information construction pipeline, which provides **factual grounding** to prevent the model from generating chemically impossible structures. This stage uses three distinct components to analyze the structural data and potential retrosynthetic reaction centers of target molecules. First, the system converts molecular SMILES strings into International Union of Pure and Applied Chemistry (IUPAC) names. This conversion allows the model to utilize nomenclature priors from scientific corpora because IUPAC names represent structural information in a readable format. Next, the framework employs the RDKit library to perform pattern matching based on SMARTS strings (Bento et al., 2020). This step extracts essential functional groups and bond types to generate a concise summary of the molecular structure. Finally, a pretrained reaction center predictor, G2Retro (Chen et al., 2023), identifies potential reaction sites to encourage diverse retrosynthetic proposals. An LLM, designated as $LLM_{info}$, organizes and ranks these identified centers. This model integrates all collected data into a cohesive initial context known as auxiliary information.

This integrated auxiliary information forms the initial state ($I_0$) of the **chemical insight**. This dynamic context performs two critical functions: guiding the model toward correct plans through structural priors, and summarizing previous attempts to prevent repetitive errors. By maintaining this evolving context, the system ensures the model learns from both successes and failures throughout the construction process.

After establishing the initial chemical insights, the system begins the iterative generation phase. The core of this process is $LLM_{retro}$, which drives the self-reflection data loop through retrosynthetic reasoning. During the generation step, $LLM_{retro}$ processes the target molecule and the existing chemical insights to propose a retrosynthetic plan $\mathcal{S}_t$:

$$\mathcal{S}_t = (R_t, C_t) = LLM(P, I_{t-1}). \quad (1)$$

The system then subjects each generated plan to objective verification using ReactionT5, a pretrained forward reaction prediction model. This verifier calculates the actual product $P'_t$ that would result from the proposed reactants and conditions:

$$P'_t = f_{\text{forward}}(R_t, C_t). \quad (2)$$

If $P'_t$ does not match the original target $P$, the proposal is identified as a **chemical hallucination**—a plan that is syntactically valid but chemically infeasible.

If the generated plan is a hallucination, the system initiates a reflection phase. $LLM_{retro}$ analyzes the discrepancy between the predicted product $P'_t$ and the tar-

get molecule $P$. By integrating the previous context, the current attempt, and verification feedback, it generates an updated insight $I_t$:

$$I_t = LLM\left(P, I_{t-1}, \mathcal{S}_t, P'_t, \mathcal{S}^*\right). \quad (3)$$

This insight summarizes past failures to avoid repetitive mistakes. It effectively steers the next retrosynthetic plan toward valid pathways.

To address the limited initial capabilities of $LLM_{retro}$, we implement a **progressive termination strategy**. In the first 5,000 samples, the loop terminates once a chemically valid plan ($S^+$) is generated. In the subsequent phase, the criterion tightens to require the strictly optimal gold standard solution ($S^*$). This two-stage approach prevents model stagnation while ensuring the final dataset converges toward high-fidelity retrosynthetic prediction by learning from both successful plans and corrected hallucinations.

The system integrates three components: (1) **verified retrosynthetic plans**; (2) **reflective insights** containing corrected chemical knowledge; and (3) **auxiliary information** for molecular comprehension. Training includes both correct pathways and incorrect "hallucinations," enabling the model to critique errors and reduce future mistakes. Auxiliary information enhances molecular understanding. An LLM distills these into a fine-tuning dataset to improve retrosynthetic performance. Prompt details are in Appendix C.

### 3.4 Dual-model collaboration

To prevent catastrophic forgetting during further fine-tuning, we propose a decoupled dual-model architecture designed for retrosynthetic prediction. This framework separates structural analysis from core retrosynthetic reasoning to reduce the cognitive load on a single model. The architecture consists of $LLM_{info}$ and $LLM_{retro}$ performing collaborative reasoning. In this collaborative process, $LLM_{info}$ generates accurate chemical contexts while $LLM_{retro}$ critically utilizes this information to plan synthetic routes. We apply distinct optimization strategies for each model to enhance both the quality of auxiliary data and the robustness of the reasoning logic. This design ensures high extensibility because adapting to new domains only requires updating $LLM_{info}$, which preserves the core reasoning capabilities of $LLM_{retro}$.

#### 3.4.1 Training of Retrosynthetic Reasoning Model

As illustrated in Figure 2 (B), the training of $LLM_{retro}$ focuses on establishing robust reasoning capabilities through a sequential two-stage pipeline: supervised fine-tuning followed by reinforcement learning.

In the fine-tuning stage, the model is trained in two phases: (1) First, we adopt a standard strategy widely used in retrosynthesis research (Zhang et al., 2025; Lin et al., 2025), fine-tuning the model on QA pairs constructed from reaction data in USPTO-FULL to enhance its fundamental understanding of chemical reactions (Lin et al., 2025; Zhang et al., 2025); (2) Then, the model is further aligned using the structured data generated by the self-reflection pipeline.

As illustrated in Figure 3, the training data is organized into two distinct formats: (1) **Retrosynthetic QA** (Template 1), which directly teaches the model to generate viable synthetic routes; and (2) **Chemical Verification QA** (Template 2), which trains the model to assess the chemical feasibility of proposed plans for a target molecule. This dual-task approach enables the model to standardize its planning format while internalizing fundamental verification logic.

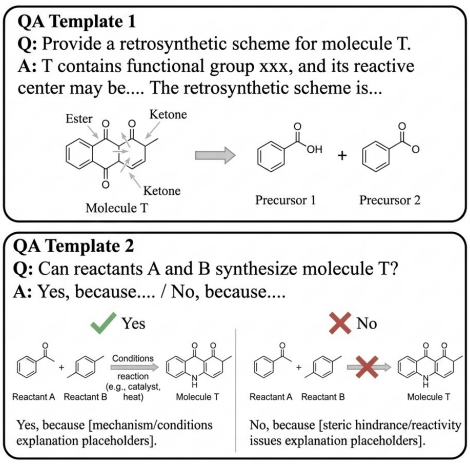



Figure 3: Two types of data

To improve plan quality, we use Group-Relative Policy Optimization (GRPO) (Guo et al., 2025). This method works well for the one-to-many retrosynthesis task. Since a single molecule can have many functional group interconversions, GRPO creates a group of candidate solutions each iteration. The model then evaluates these candidates using group-normalized rewards. This comparison helps the model better tell feasible pathways from poor ones. The objective function maximizes the relative advantage of the generated plans:

$$\begin{aligned}\mathcal{L}(\theta) =& \mathbb{E}_{x\sim\pi_\theta}\left[\min\left(r(x)\cdot A^{\text{rel}}(x),\right.\right.\\ &\left.\left.\text{clip}(r(x), 1-\varepsilon, 1+\varepsilon)\cdot A^{\text{rel}}(x)\right)\right]\\ &-\beta\cdot \text{KL}[\pi_\theta \| \pi_{\text{old}}] \quad (4)\end{aligned}$$

Here, $r(x)$ denotes the policy ratio, and $A^{\text{rel}}(x)$ represents the relative advantage calculated by normalizing scores within a group of outputs. The reward function $R(x, y)$ is hierarchically designed to guide the model toward higher-quality solutions:

$$R(x,y) = \Phi_{\text{format}}(y) + 2\Phi_{\text{valid}}(y) + 2\Phi_{\text{match}}(y, y^*). \quad (5)$$

This function decomposes the feedback into three additive components: (1) **Syntactic Validity ($\Phi_{\text{format}}$)**: Ensures the generated sequence conforms to the required format. (2) **Chemical Consistency ($\Phi_{\text{valid}}$)**: Validates the retrosynthetic step via forward reaction simulation. (3) **Canonical Alignment ($\Phi_{\text{match}}$)**: Granted only when the predicted pathway matches the ground-truth prediction exactly.

We ensure robustness and correct optimization dur-

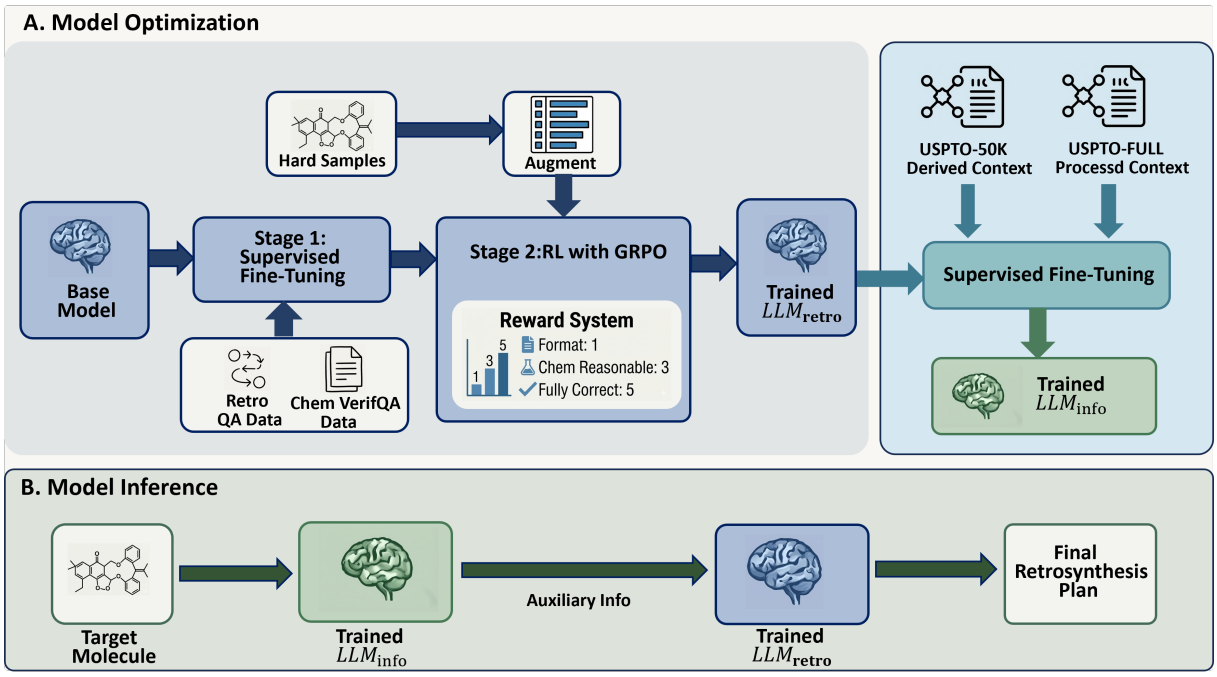



Figure 2: The model training and inference process of ChemValiSynth. Figure A illustrates the training process of ChemValiSynth. Figure B illustrates the inference process.

ing reinforcement learning through a mixed-auxiliary strategy. We employ a context mixture with a 2:2:6 ratio across three settings:

- **Complete information**: Full auxiliary data for leveraging external knowledge.
- **Partial information**: Randomly omitted data for supplementing missing information.
- **Autonomous information**: No auxiliary data, requiring internal reasoning.

This distribution encourages the model to transition from relying on internal parametric knowledge to critically evaluating and synthesizing contextual information.

#### 3.4.2 Training of Auxiliary Information Model

We developed $LLM_{info}$ based on the pre-trained $LLM_{retro}$ as shown in Figure 2 (B). The model underwent targeted fine-tuning to enhance its ability to provide auxiliary retrosynthetic information. To empower $LLM_{info}$ with robust guidance capabilities, we expanded the training dataset by integrating 200,000 new records sampled from USPTO-FULL with 40,000 auxiliary information data points previously constructed during the self-reflection loop. The entire training corpus was assembled following the pipeline depicted in Figure 1 (A). This comprehensive dataset enabled the successful optimization of $LLM_{info}$ for downstream tasks. Detailed specifications regarding the training data and experimental configurations are documented in Appendix A.

## 4 Experiments

### 4.1 Experiment Setup

We evaluate ChemValiSynth on standard retrosynthetic benchmarks against representative baselines. To assess data efficiency, we also train a variant, ChemValiSynth(50K), where the auxiliary model $LLM_{retro}$ is trained only train set of on USPTO-50K, making the entire model rely solely on that dataset.

**Benchmark.** The study is primarily based on two standard datasets: USPTO-50K and USPTO-FULL. USPTO-50K contains approximately 50,000 carefully curated single-step organic reactions spanning 10 major reaction types. USPTO-FULL, a larger dataset with about 1 million reactions, offers greater chemical diversity. We keep the same data split as GraphRetro. Please refer to Appendix A for dataset details.

To further illustrate the enhancement of our method on the model's chemical capabilities and demonstrate its generalization potential beyond retrosynthesis, we also conducted a multi-faceted evaluation of the model's performance using the Scientific Knowledge Deduction module of the ChemEval(Huang et al., 2025) benchmark.

**Baseline Models.** We evaluate our method against a diverse set of baselines, representing the main technical approaches in the field:

- **General-purpose LLMs**: Including Gemini3-Pro (Gemini Team, 2025), GPT-5.2 (OpenAI, 2025b), o4-mini (OpenAI, 2025a), DeepSeek-V3.2 (Liu et al., 2025) which provide benchmarks for general reasoning capabilities.

- **Chemistry-specific LLMs**: Including ChemDFM, ChemDual, RSGPT(Deng et al., 2025), LlaSMol-Mistral (Luo et al., 2024), and Intern-S1-mini (Bai et al., 2025), representing the chemically adapted language models.

- **Non-LLM-based Methods**:The mainstream specialized models, covering template-based models: LocalRetro (Chen and Jung, 2021), GLN (Dai et al., 2019), RetroComposer (Yan et al., 2022), graph-based models: RetroXpert, G2Gs, and GraphRetro, sequence-based models: Aug.Transformer (Tetko et al., 2020), R-SMILES and Retroformer.

### 4.2 Experimental Results on Retrosynthesis task

We conducted systematic experiments on the USPTO-50K and USPTO-FULL datasets. To clearly demonstrate the advantages of our method compared to different methods, we present the results in two parts: comparison with LLM methods, and comparison with traditional (non-LLM) state-of-the-art methods.

Table 1: Comparison with LLM-based models on USPTO-50K.

| Model | Top-1 (%) |
|---|---|
| *General-Purpose LLMs* | |
| GPT-5.2 | 5.1 |
| o4-mini | 13.5 |
| DeepSeek-V3.2 | 21.7 |
| Gemini3-pro | 26.1 |
| *Chemistry-Specific LLMs* | |
| LlaSMol-Mistral | 2.0 |
| Intern-S1-mini | 2.1 |
| ChemDFM | 16.5 |
| ChemDual | 49.95 |
| RSGPT | **63.4** |
| *Our Work* | |
| Qwen3-8B | 0.8 |
| **ChemValiSynth(50K)** | 51.9 |
| **ChemValiSynth(FULL)** | **58.7** |

**Comparison with LLM-based Methods:** As shown in Table 1, our framework, built on Qwen3-8B-Instruct, achieves a top-1 accuracy of 58.7%. This significantly outperforms recent chemistry-specific LLMs, including ChemDual (49.95%) and ChemDFM-v1.5 (16.5%). Our data-efficient variant, ChemValiSynth(50K), trained only on the USPTO-50K dataset, attains 51.9%, underscoring the method's robustness. While RSGPT (63.4%) achieves a higher score, it is trained on a billion-scale dataset. In contrast, ChemValiSynth is trained on millions of samples, demonstrating superior data efficiency while maintaining competitive performance.

**Comparison with specialized models:** We evaluate our method against a series of traditional retrosynthesis models. We summarize the Top-k accuracy results on the USPTO-50K and USPTO-FULL datasets in Tables 2 and 3, respectively.

On the widely used USPTO-50K benchmark, our method achieves the best Top-1 accuracy of 58.7% among all compared models. As shown in Table 2, ChemValiSynth outperforms the strong sequence-based baseline, R-SMILES, by a margin of 2.4% in Top-1 accuracy (58.7% vs. 56.3%). In terms of Top-5 accuracy, our method also secures the leading position with 86.7%. Notably, ChemValiSynth(50K), trained on the smaller USPTO-50K subset, also demonstrates strong competitiveness, achieving a Top-1 accuracy of 51.9% and a Top-5 accuracy of 83.3%, which is still comparable to many existing methods despite using less training data.

Table 2: Comparison with non-LLM-based models on USPTO-50K.

| Method | Top-K accuracy (%) | | | |
|---|---|---|---|---|
| | 1 | 3 | 5 | 10 |
| *Template-Based Methods* | | | | |
| GLN | 52.5 | 74.6 | 80.5 | 86.9 |
| LocalRetro | 53.4 | 77.5 | 85.9 | 92.4 |
| RetroComposer | 54.5 | 77.2 | 83.2 | 87.7 |
| *Graph-Based Methods* | | | | |
| G2Gs | 48.9 | 67.6 | 72.5 | 75.5 |
| RetroXpert | 50.4 | 61.1 | 62.3 | 63.4 |
| GraphRetro | 53.7 | 68.3 | 72.2 | 75.5 |
| *Sequence-Based Methods* | | | | |
| Aug.Transformer | 53.2 | – | 80.5 | 85.2 |
| Retroformer | 53.2 | 71.1 | 76.6 | 82.1 |
| R-SMILES | 56.3 | **79.2** | 86.2 | **91.0** |
| *Our work* | | | | |
| **ChemValiSynth(50K)** | 51.9 | 72.4 | 83.3 | 87.9 |
| **ChemValiSynth(FULL)** | **58.7** | 76.6 | **86.7** | 89.5 |

On the larger and more challenging USPTO-Full benchmark, our method demonstrates considerable performance comparable to the leading baselines. As presented in Table 3, ChemValiSynth achieves a Top-1 accuracy of 48.2%. While this result is slightly lower than RetroXpert (49.4%) and R-SMILES (48.9%), our model consistently outperforms all template-based methods. Notably, our method achieves the best performance in Top-5 accuracy at 73.1%, surpassing both R-SMILES (72.0%) and RetroXpert (67.6%), indicating its superior capability in retrieving correct reactants within the top predictions.

### 4.3 Quantitative Analysis of Guidance and Robustness

We evaluated the impact of $LLM_{\text{info}}$ guidance and the performance of $LLM_{\text{retro}}$ within the ChemValiSynth framework. Statistical analysis on the USPTO-50K test set highlights the model's predictive accuracy. To assess the contribution of specialized training, we compared the trained $LLM_{\text{info}}$ against its base version. The results are summarized in Figures 4 and 5.

Table 3: Comparison with non-LLM-based models on USPTO-Full.

| Method | Top-K accuracy (%) | | | |
|---|---|---|---|---|
| | 1 | 3 | 5 | 10 |
| *Template-Based Methods* | | | | |
| LocalRetro | 39.1 | 53.3 | 58.4 | 63.7 |
| GLN | 39.3 | – | – | 63.7 |
| *Graph-Based Methods* | | | | |
| RetroXpert | **49.4** | 63.6 | 67.6 | 71.6 |
| *Sequence-Based Methods* | | | | |
| Aug.Transformer | 46.2 | – | – | 73.3 |
| R-SMILES | 48.9 | **66.6** | 72.0 | **76.4** |
| **ChemValiSynth(FULL)** | 48.2 | 63.4 | **73.1** | 75.5 |

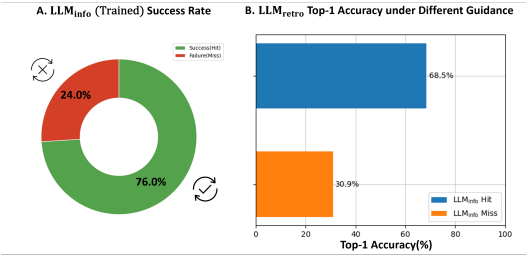



Figure 4: The success rate of $LLM_{info}$ (Trained) and the accuracy of $LLM_{retro}$ under different conditions. When $LLM_{info}$ fails, $LLM_{retro}$ can still find the correct solution.

$LLM_{\text{info}}$ identifies the correct reaction center for 76.0% of test molecules, providing high-quality guidance for retrosynthetic planning. When supplied with these correct centers, $LLM_{\text{retro}}$ achieves 68.5% Top-1 accuracy, demonstrating its ability to effectively utilize chemical context. Furthermore, GRPO training enhances the model's robustness; $LLM_{\text{retro}}$ can leverage its internal chemical knowledge to maintain accurate predictions even when guidance is noisy or incomplete.

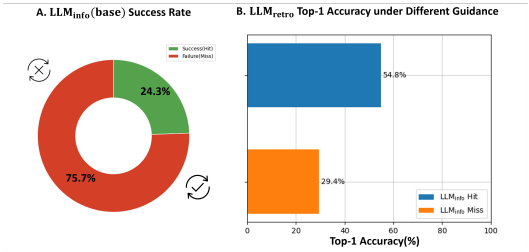



Figure 5: The success rate of $LLM_{info}$ (Base) and the accuracy of $LLM_{retro}$ under different conditions. The performance of untrained $LLM_{info}$ is substantially inferior to that of its trained counterpart.

In contrast, the base version of $LLM_{\text{info}}$ identifies correct reaction centers for only 24.3% of test molecules. The significantly higher failure rate of 75.7% limits the guidance's effectiveness compared to the trained version. The substantial performance gap highlights that specialized training is essential for $LLM_{\text{info}}$ to provide reliable chemical priors.

### 4.4 Evaluation on Enhancing Chemical Understanding

To systematically evaluate the enhancement of the model's chemical understanding by our proposed method, we conducted a comprehensive assessment of ChemvaliSynth's foundational chemical abilities using the Scientific Knowledge Deduction module of ChemEval. The experiment used Qwen3-8B as the base model and included strong baselines such as Gemini3-Pro, DeepSeek v3.2, and GPT 5.2, covering key chemical tasks including substrate recommendation and synthetic pathway recommendation. The evaluation results are presented in Table 4.

Compared to the base model Qwen3-8B, ChemvaliSynth shows significant improvement on most tasks, confirming our methods effectiveness in enhancing chemical domain expertise. Against advanced models such as Gemini3-Pro, DeepSeek v3.2, and GPT-5.2, ChemvaliSynth achieves competitive or better performance on multiple metrics. It achieves the best results in Substrate Recommendation and Reagent Recommendation. The model also shows strong gains in other tasks, including Ligand Recommendation and Solvent Recommendation. Performance on Synthetic Pathway Recommendation declines, likely due to overfitting that causes oversimplified route suggestions. Overall, ChemvaliSynth delivers broad and solid improvements, particularly in core retrosynthesis subtasks.

### 4.5 Ablation Study

We evaluated the contribution of key components in the ChemValiSynth framework. We conducted two systematic ablation studies on the USPTO-50K dataset. These studies ensure controlled and comparable results.

**Ablation Study on Fine-Tuning Data Composition:** This experiment investigates the impact of chemically validated reflection data during supervised fine-tuning. We compare two specific data configurations against the full framework:

1. **Retrosynthesis QA Data Only:** The training process uses only conventional retrosynthesis question-answer pairs.

2. **Retrosynthesis QA Data with Reflection Data:** The training process incorporates chemical verification QA reflection data alongside QA data. This includes successful and failed reasoning cases.

Table 5 presents the results. The model using Mixed Data (QA + Reflection) significantly outperforms the QA-only baseline in Top-1 accuracy. This is because analyzing both validated solutions and failed attempts allows the model to capture the underlying principles of chemical plausibility. Consequently, the accuracy and internal consistency of the predicted routes improve substantially. The full ChemValiSynth framework includes training data with molecular auxiliary information. Consequently, it achieves the best performance.

Table 4: Performance on the scientific knowledge Deduction module of ChemEval

| Task | Metric | Our Method | | Baselines | | | |
|---|---|---|---|---|---|---|---|
| | | Base | ChemvaliSynth | GPT5.2 | DeepSeek v3.2 | Gemini3 Pro | ChemDFM |
| Substrate Recommendation | F1 | 0.33 | **16.67**↑ | 1.33 | 3.00 | 4.67 | 5.00 |
| Synthetic Pathway Recommendation | F1 | 26.25 | 12.25 | 72.50 | 53.13 | 64.38 | 24.38 |
| Ligand Recommendation | Score | 4.00 | 10.00↑ | 10.00 | 6.00 | 6.00 | 25.33 |
| Reagent Recommendation | F1 | 25.50 | **48.80**↑ | 14.80 | 22.67 | 48.80 | 11.16 |
| Solvent Recommendation | F1 | 15.45 | 20.6 ↑ | 10.00 | 10.00 | 28.00 | 2.01 |
| Reaction Product Prediction | Accuracy | 5.88 | 35.92↑ | 11.33 | 37.13 | 62.00 | 17.57 |
| Product Yield Prediction | F1 | 12.00 | 14.00↑ | 46.00 | 48.00 | 24.00 | 8.00 |
| Reaction Rate Prediction | Overlap | 15.68 | 16.44↑ | 11.62 | 19.67 | 36.89 | 6.41 |
| Intermediate Derivation | Score | 42.50 | 45.80↑ | 66.25 | 52.50 | 60.25 | 24.75 |

Table 5: Impact of Training Data Composition on Model Performance (Top-1 Accuracy, %).

| Model Configuration | Top-1 (%) |
|---|---|
| Baseline (QA Only) | 11.4 |
| Mixed Data (QA + Reflection) | 35.8 |
| **ChemValiSynth (Full Framework)** | **58.7** |

Table 6: Impact of Different GRPO Reward Function Designs (Top-1 Accuracy, %).

| Reward Scheme | Top-1 (%) |
|---|---|
| Baseline | 35.8 |
| Syntax-Match Reward | 45.1 |
| Chemically-Aware Reward | **50.2** |

**Ablation on GRPO Reward Function Design:** This study investigates the effectiveness of the multi-level reward function. We compare two reward schemes for the GRPO-based reinforcement learning stage:

- **Syntax-Match Reward:** The model receives a +1 reward for generating a correctly formatted output. A +5 reward is granted only if the predicted route exactly matches the ground-truth answer ($y^*$). The reward function is defined as:

$$R(x, y) = \Phi_{\text{format}}(y) + 4 \cdot \Phi_{\text{match}}(y, y^*) \quad (6)$$

- **Chemically-Aware Reward:** Our proposed reward scheme, as shown in Equation (5), introduces an intermediate chemical plausibility reward of +3 for predictions that are correctly formatted and chemically feasible, even if they do not match the standard answer.

We present the experimental results in Table 6, using the fine-tuned $LLM_{\text{retro}}$ as the baseline. The results demonstrate that the Chemically-Aware Reward outperforms the Syntax-Match Reward in Top-1 accuracy. This performance gain suggests that intermediate rewards effectively guide the policy search. By providing a denser signal, this mechanism encourages the model to explore diverse and chemically plausible routes. Conversely, the Syntax-Match Reward lacks explicit incentives for feasible yet non-standard routes, which may lead to insufficient exploration or premature convergence. Thus, our multi-level reward design significantly enhances the robustness of the learning process.

**Ablation on the auxiliary information model:** To evaluate the effectiveness of each component in our proposed framework, we conduct an ablation study on the auxiliary information model. To verify the contribution of the $LLM_{info}$ module, we compare $LLM_{info}$ models trained on different datasets, and replace it with the rection center prediction model: G2Retro. As shown in Table 7, the $LLM_{info}$ trained solely on the train set of USPTO-50K yields only a slight improvement over using no auxiliary information and lags behind using G2Retro for auxiliary information. However, with the addition of USPTO-FULL data, the model performance surpasses that of G2Retro. This indicates that incorporating more diverse data sources could further enhance the capabilities of ChemValiSynth.

Table 7: Ablation study on the information extraction module (Substrate recommendation, top1)

| Auxiliary info module | Accuracy (%) |
|---|---|
| Base | 50.2 |
| G2Retro | 54.8 |
| $\text{LLM}_{\text{info}}$(50K) | 51.9 |
| $\text{LLM}_{\text{info}}$(FULL) | 58.7 |

## 5 Conclusion

We propose ChemValiSynth, a retrosynthesis framework integrating a reflection mechanism. Our method enhances LLM reliability and reduces hallucinations. We utilize domain-specific lightweight models for chemical verification to synthesize high-quality data. Additionally, the framework employs a dual-model architecture. It separates structural analysis from retrosynthetic reasoning to generate final predictions. Experiments show outstanding performance on multiple datasets, advancing the application of LLMs in chemistry.

## 6 Limitations

While ChemValiSynth improves the reliability of retrosynthesis, it has the following limitations:

- **Reliance on Forward Prediction Accuracy**: The performance of our framework depends heavily on the product prediction model used for verification. If this model makes an incorrect prediction, it can limit the overall effectiveness of the self-reflection loop.
- **Challenges in Comparing Multiple Valid Routes**: Our method involves comparing model-generated plans with ground-truth solutions to optimize reasoning. However, this assumes the standard answer is always superior. Currently, the framework lacks a robust way to evaluate the quality of two different but chemically feasible pathways.

## References


Lei Bai, Zhongrui Cai, Yuhang Cao, Maosong Cao, Weihan Cao, Chiyu Chen, Haojiong Chen, Kai Chen, Pengcheng Chen, Ying Chen, and 1 others. 2025. Intern-s1: A scientific multimodal foundation model. *arXiv preprint arXiv:2508.15763*.

A Patrícia Bento, Anne Hersey, Eloy Félix, Greg Landrum, Anna Gaulton, Francis Atkinson, Louisa J Bellis, Marleen De Veij, and Andrew R Leach. 2020. An open source chemical structure curation pipeline using rdkit. *Journal of Cheminformatics*, 12(1):51.

Shuan Chen and Yousung Jung. 2021. Deep retrosynthetic reaction prediction using local reactivity and global attention. *JACS Au*, 1(10):1612–1620.

Shuan Chen and Yousung Jung. 2022. A generalized-template-based graph neural network for accurate organic reactivity prediction. *Nature Machine Intelligence*, 4(9):772–780.

Ziqi Chen, Oluwatosin R Ayinde, James R Fuchs, Huan Sun, and Xia Ning. 2023. G 2 retro as a two-step graph generative models for retrosynthesis prediction. *Communications Chemistry*, 6(1):102.

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

Yafeng Deng, Xinda Zhao, Hanyu Sun, Yu Chen, Xiaorui Wang, Xi Xue, Liangning Li, Jianfei Song, Chang-Yu Hsieh, Tingjun Hou, and 1 others. 2025. Rsgpt: a generative transformer model for retrosynthesis planning pre-trained on ten billion datapoints. *Nature communications*, 16(1):7012.

Google Gemini Team. 2025. Gemini 3 pro: State-of-the-art reasoning and multimodal understanding. Technical report, Google DeepMind.

Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Peiyi Wang, Qihao Zhu, Runxin Xu, Ruoyu Zhang, Shirong Ma, Xiao Bi, and 1 others. 2025. Deepseek-r1 incentivizes reasoning in llms through reinforcement learning. *Nature*, 645(8081):633–638.

Taicheng Guo, Bozhao Nan, Zhenwen Liang, Zhichun Guo, Nitesh Chawla, Olaf Wiest, Xiangliang Zhang, and 1 others. 2023. What can large language models do in chemistry? a comprehensive benchmark on eight tasks. *Advances in Neural Information Processing Systems*, 36:59662–59688.

Yuqing Huang, Rongyang Zhang, Xuesong He, Xuyang Zhi, Hao Wang, Nuo Chen, Xin Li, Feiyang Xu, Deguang Liu, Huadong Liang, and 1 others. 2025. Chemeval: A multi-level and fine-grained chemical capability evaluation for large language models.

Xuan Lin, Qingrui Liu, Hongxin Xiang, Daojian Zeng, and Xiangxiang Zeng. 2025. Enhancing chemical reaction and retrosynthesis prediction with large language model and dual-task learning. *arXiv preprint arXiv:2505.02639*.

Aixin Liu, Aoxue Mei, Bangcai Lin, Bing Xue, Bingxuan Wang, Bingzheng Xu, Bochao Wu, Bowei Zhang, Chaofan Lin, Chen Dong, and 1 others. 2025. Deepseek-v3. 2: Pushing the frontier of open large language models. *arXiv preprint arXiv:2512.02556*.

Shunyang Luo, Yuqi Tang, Mingyuan Jiang, Kehua Feng, Qiang Zhang, and Keyan Ding. 2024. Generating multiple choice questions from scientific literature via large language models. In *2024 IEEE International Conference on Knowledge Graph (ICKG)*, pages 219–226. IEEE.

Aman Madaan, Niket Tandon, Prakhar Gupta, Skyler Hallinan, Luyu Gao, Sarah Wiegreffe, Uri Alon, Nouha Dziri, Shrimai Prabhumoye, Yiming Yang, and 1 others. 2023. Self-refine: Iterative refinement with self-feedback. *Advances in Neural Information Processing Systems*, 36:46534–46594.

Michael P Maloney, Connor W Coley, Samuel Genheden, Nessa Carson, Paul Helquist, Per-Ola Norrby, and Olaf Wiest. 2023. Negative data in data sets for machine learning training.

Krzysztof Maziarz, Guoqing Liu, Hubert Misztela, Aleksei Kornev, Piotr Gaiński, Holger Hoefling, Mike Fortunato, Rishi Gupta, and Marwin Segler. 2024. Chimera: Accurate retrosynthesis prediction by ensembling models with diverse inductive biases. *arXiv preprint arXiv:2412.05269*.

T Runcie Nicholas, Charlotte M Deane, and Fergus Imrie. 2025. Assessing the chemical intelligence of large language models. *arXiv preprint arXiv*, 2505.

OpenAI. 2025a. Openai o4-mini system card. Technical Documentation.

OpenAI. 2025b. Update to gpt-5 system card: Gpt-5.2. Technical Documentation.

Konstantinos I Roumeliotis and Nikolaos D Tselikas. 2023. Chatgpt and open-ai models: A preliminary review. *Future Internet*, 15(6):192.

Philippe Schwaller, Teodoro Laino, Théophile Gaudin, Peter Bolgar, Christopher A Hunter, Costas Bekas, and Alpha A Lee. 2019. Molecular transformer: a model for uncertainty-calibrated chemical reaction prediction. *ACS central science*, 5(9):1572–1583.

Chence Shi, Minkai Xu, Hongyu Guo, Ming Zhang, and Jian Tang. 2020. A graph to graphs framework for retrosynthesis prediction. In *International conference on machine learning*, pages 8818–8827. PMLR.

Haizhou Shi, Zihao Xu, Hengyi Wang, Weiyi Qin, Wenyuan Wang, Yibin Wang, Zifeng Wang, Sayna Ebrahimi, and Hao Wang. 2025. Continual learning of large language models: A comprehensive survey. *ACM Computing Surveys*, 58(5):1–42.

Noah Shinn, Federico Cassano, Ashwin Gopinath, Karthik Narasimhan, and Shunyu Yao. 2023. Reflexion: Language agents with verbal reinforcement learning. *Advances in Neural Information Processing Systems*, 36:8634–8652.

Vignesh Ram Somnath, Charlotte Bunne, Connor Coley, Andreas Krause, and Regina Barzilay. 2021. Learning graph models for retrosynthesis prediction. *Advances in Neural Information Processing Systems*, 34:9405–9415.

Shengyin Sun, Wenhao Yu, Yuxiang Ren, Weitao Du, Liwei Liu, Xuecang Zhang, Ying Hu, and Chen Ma. 2025. Gdiffretro: Retrosynthesis prediction with dual graph enhanced molecular representation and diffusion generation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 12595–12603.

Igor V Tetko, Pavel Karpov, Ruud Van Deursen, and Guillaume Godin. 2020. State-of-the-art augmented nlp transformer models for direct and single-step retrosynthesis. *Nature communications*, 11(1):5575.

Yue Wan, Chang-Yu Hsieh, Ben Liao, and Shengyu Zhang. 2022. Retroformer: Pushing the limits of end-to-end retrosynthesis transformer. In *International Conference on Machine Learning*, pages 22475–22490. PMLR.

Yu Xia, Rui Wang, Xu Liu, Mingyan Li, Tong Yu, Xiang Chen, Julian McAuley, and Shuai Li. 2025. Beyond chain-of-thought: A survey of chain-of-x paradigms for llms. In *Proceedings of the 31st International Conference on Computational Linguistics*, pages 10795–10809.

Chaochao Yan, Qianggang Ding, Peilin Zhao, Shuangjia Zheng, Jinyu Yang, Yang Yu, and Junzhou Huang. 2020. Retroxpert: Decompose retrosynthesis prediction like a chemist. *Advances in Neural Information Processing Systems*, 33:11248–11258.

Chaochao Yan, Peilin Zhao, Chan Lu, Yang Yu, and Junzhou Huang. 2022. Retrocomposer: composing templates for template-based retrosynthesis prediction. *Biomolecules*, 12(9):1325.

Situo Zhang, Hanqi Li, Lu Chen, Zihan Zhao, Xuanze Lin, Zichen Zhu, Bo Chen, Xin Chen, and Kai Yu. 2025. Reasoning-driven retrosynthesis prediction with large language models via reinforcement learning. *arXiv preprint arXiv:2507.17448*.

Haochen Zhao, Xiangru Tang, Ziran Yang, Xiao Han, Xuanzhi Feng, Yueqing Fan, Senhao Cheng, Di Jin, Yilun Zhao, Arman Cohan, and 1 others. 2024a. Chemsafetybench: benchmarking llm safety on chemistry domain. *arXiv preprint arXiv:2411.16736*.

Zihan Zhao, Da Ma, Lu Chen, Liangtai Sun, Zihao Li, Yi Xia, Bo Chen, Hongshen Xu, Zichen Zhu, Su Zhu, and 1 others. 2024b. Chemdfm: a large language foundation model for chemistry. *arXiv preprint arXiv:2401.14818*.

Weihe Zhong, Ziduo Yang, and Calvin Yu-Chian Chen. 2023. Retrosynthesis prediction using an end-to-end graph generative architecture for molecular graph editing. *Nature Communications*, 14(1):3009.

Zipeng Zhong, Jie Song, Zunlei Feng, Tiantao Liu, Lingxiang Jia, Shaolun Yao, Tingjun Hou, and Mingli Song. 2024. Recent advances in deep learning for retrosynthesis. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 14(1):e1694.

Zipeng Zhong, Jie Song, Zunlei Feng, Tiantao Liu, Lingxiang Jia, Shaolun Yao, Min Wu, Tingjun Hou, and Mingli Song. 2022. Root-aligned smiles: a tight representation for chemical reaction prediction. *Chemical Science*, 13(31):9023–9034.

Jiajun Zhu, Binjie Hong, Zixun Lan, and Fei Ma. 2023. Single-step retrosynthesis via reaction center and leaving groups prediction. In *2023 16th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, pages 1–5. IEEE.

## A Implementation Details

Our method employs the Qwen3-8b as the foundational model, and the training and test data are sourced from USPTO-50K and USPTO-FULL. Detailed data information is presented in Table 9.

The training of ChemValisynth comprised four distinct phases. The first three phases focused on optimizing $LLM_{retro}$. Initially, the Basic Retrosynthesis stage utilized 810,496 data points from USPTO-FULL for supervised fine-tuning. This step established the models foundational prediction capabilities. Subsequently, the Self-Reflection Training stage leveraged data derived from self-reflection loops to enable critical retrosynthesis planning. The third stage employed GRPO for reinforcement learning, which further refined the model's

Table 8: Experimental Settings for ChemValisynth Training Stages

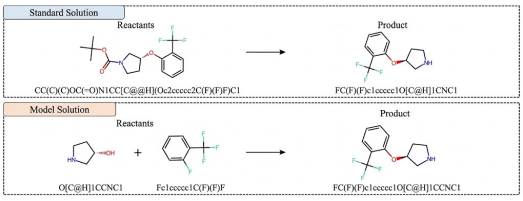
| **Stages** | **Basic Retrosynthesis** ($LLM_{retro}$) | **Self-Reflection Training** ($LLM_{retro}$) | **Reinforcement Learning** ($LLM_{retro}$) | **Auxiliary Info Training** ($LLM_{info}$) |
|---|---|---|---|---|
| Training Type | SFT | SFT | GRPO | SFT |
| Parameters initialized from | Qwen3-8b | Basic Retro Model | Self-Reflection Model | GRPO RL Model |
| Dataset | USPTO-FULL | Self-Reflection Data | Prompted RL Data | Auxiliary Info Data |
| Dataset Size | 810,496 | 23,876 | 8,400 | 200,000 |
| LoRA Rank | 64 | 64 | 64 | 64 |
| LoRA Alpha | 128 | 128 | 128 | 128 |
| Learning rate | $1 \times 10^{-4}$ | $2 \times 10^{-4}$ | $3 \times 10^{-5}$ | $2 \times 10^{-4}$ |
| Epochs | 3 | 3 | 4 | 3 |

Table 9: Dataset Information

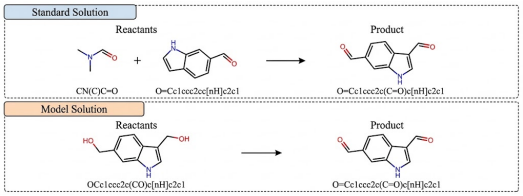
| Dataset | Train Size | Validation Size | Test Size |
|---|---|---|---|
| USPTO-50k | 40,008 | 5,001 | 5,007 |
| USPTO-full | 810,496 | 101,311 | 101,311 |

accuracy. Finally, we conducted an additional round of fine-tuning using processed USPTO-FULL data to enhance the generation of auxiliary retrosynthesis information. Table 8 details the specific training configurations.

## B Case Study




(a) Retrosynthetic plan for C1




(b) Retrosynthetic plan for C2

Figure 6: Model-generated plan vs. standard plan comparison. Subfigure (a) shows the retrosynthetic plan for molecule FC(F)(F)c1ccccc1O[C@H]1CCNC1. Subfigure (b) shows the retrosynthetic plan for molecule O=Cc1ccc2c(C=O)c[nH]c2c1.

To assess the practical utility and chemical validity of our proposed framework, we conducted a qualitative comparison between the model-generated retrosynthetic plans and standard solutions recorded in the USPTO-50K dataset. Figure 6 illustrates two representative instances where the model-generated pathways exhibit superior synthetic logic compared to the standard references.

Figure 6 (a) displays the retrosynthesis of the target molecule FC(F)(F)c1ccccc1O[C@H]1CCNC1. The standard method relies on a Boc-protected intermediate, which undergoes a deprotection step to yield the target product. However, the implementation of this strategy requires prior modification of the substrate using a corresponding Boc donor reagent, such as $(Boc)_2O$. In contrast, the model proposes a more direct method involving the nucleophilic substitution of the fluorinated precursor with O[C@H]1CCNC1. This streamlined method effectively bypasses the protection-deprotection sequence, offering a significant advantage in terms of step efficiency while maintaining feasibility through controlled reaction selectivity.

Furthermore, Figure 6 (b) demonstrates the models capability to optimize reaction conditions for the synthesis of the indole-dicarbaldehyde target O=Cc1ccc2c(C=O)c[nH]c2c1. The standard method employs the Vilsmeier-Haack formylation strategy, which achieves the desired transformation through secondary modification of substrates containing aldehyde groups. However, this process relies on highly corrosive reagents (such as $POCl_3$) and generates a vast amount of waste salts, posing severe challenges to equipment maintenance and environmental protection. Conversely, the model generates a robust alternative via the direct oxidation of the corresponding diol OCc1ccc2c(CO)c[nH]c2c1. This "one-pot" oxidation strategy mitigates the chemoselectivity issues associated with the standard route, resulting in a cleaner reaction profile, fewer byproducts, and potentially higher yields.

# C Prompts For LLMs

**stage 1 train data format**

**Q:**

> *Please provide a single-step retrosynthetic pathway for {target_mol}.*

**A:**

A single-step retrosynthetic pathway for **{target_mol}** is:

```
{retro_methods}
```

Figure 7: Prompt used to construct fine-tune data from USPTO-FULL

**Chemical insight update prompt**

You are an expert in single-step retrosynthetic analysis. Your goal is to update retrosynthetic guiding insights based on the provided context.

**You will receive:**

- Target molecule
- Expert-designed retrosynthetic route
- Previous insights (positive & negative)
- Structural & physicochemical analysis
- Past experiments & validation results

**Your task:**

Analyze inputs and output updated guiding insights in two categories:

**1. Positive insights (if expert route is given):**

- Extract design principles from the expert route.
- Interpret experimental outcomes and diagnose failure causes.
- Refine prior positive insights (focus on reasoning, not substrate names).

**2. Negative insights:**

- Refine prior negative insights based on analysis and results.
- Identify reasoning errors; cite at least one concrete failed case (goal, design, discrepancy).
- Be specific and reusable (avoid "the target molecule" phrasing).

**Requirements:**

1. Output in valid JSON list format wrapped with delimiters:
<positive_insight>```json [...] ```</positive_insight> & <negative_insight>```json [...] ```</negative_insight>

2. Insights must be complementary, self-contained, and devoid of vague references.

**Inputs:**

Target: **{target_molecule}**
Expert scheme: **{ground_truth}**
Prev Positive: **{positive_insight}**
Analysis: **{context}**
Past Exps: **{past_result}**
Prev Negative: **{negative_insight}**

**Now perform your analysis.**

Figure 8: Prompt used to extract chemical insight from validation result.

# D Constructed Data Cases

**Retrosynthetic QA**

**Q** Please provide a single-step retrosynthetic pathway for 2-O-tert-butyl 8-O-[(1-pyridin-4-ylpiperidin-4-yl)methyl] 1,3,4,5-tetrahydro-2-benzazepine-2,8-dicarboxylate.

**A** **Retrosynthetic Analysis**

The target molecule is a complex heterocyclic compound featuring a 1,3,4,5-tetrahydro-2-benzazepine core with two distinct carboxylate ester functionalities. Analysis begins with identification of strategic disconnections.

**1. Key Features and Disconnections**

The target contains a carbamate at position 2 and a benzylic ester at position 8. The **8-position ester linkage** is the most strategic disconnection point, splitting the molecule into a benzazepine carboxylic acid and a piperidine-pyridine alcohol fragment.

**2. Synthetic Equivalents**

- **2-[(2-Methylpropan-2-yl)oxycarbonyl]-1,3,4,5-tetrahydro-2-benzazepine-8-carboxylic acid:** Core fragment with a pre-installed Boc-protecting group.
- **(1-Pyridin-4-ylpiperidin-4-yl)methanol:** The alcohol component required for ester formation.

**3. Reactivity and Selectivity Assessment**

The forward esterification using **DCC** or **EDCI** is highly feasible. **Selectivity:** The Boc group is stable; pyridine and tertiary amine nitrogens are non-nucleophilic under these conditions. **Regioselectivity:** Single carboxylic acid and hydroxyl groups ensure mono-esterification without bis-esterification risks.

*Expert-designed pathway involves a single disconnection at the 8-position ester for maximum reliability.*

```
// Final Retrosynthetic Pathways (JSON)
[
    {
        "reactants": ["2-[(2-Methylpropan-2-yl)oxycarbonyl]-1,3,4,5-tetrahydro-2-benzazepine-8-carboxylic acid", "(1-Pyridin-4-ylpiperidin-4-yl)methanol"],
        "reagents": ["DCC", "DMAP"],
        "reasoning": "Standard expert pathway disconnecting the 8-position ester. Simplifies the target into stable, pre-protected precursors for high-yielding esterification."
    },
    {
        "reactants": ["1,3,4,5-Tetrahydro-2-benzazepine-2,8-dicarboxylic acid", "tert-Butanol", "(1-Pyridin-4-ylpiperidin-4-yl)methanol"],
        "reagents": ["DCC", "DMAP"],
        "reasoning": "One-pot dual esterification. Less reliable due to risks of non-regioselective bis-esterification and cross-reactivity."
    }
]
```

Figure 9: An example Retrosynthesis QA data.

## Chemical Verification QA

**Q** In designing a retrosynthetic route for N-[2-[4-(3-aminopropyl)phenyl]propyl]propane-2-sulfonamide, using reactants 4-(3-aminopropyl)benzyl bromide and propane-2-sulfonamide under conditions with potassium carbonate as a base and DMF as solvent, is this pathway feasible?

**A** **Feasibility: Infeasible**

The proposed pathway is **infeasible**. The actual product formed is **CC(C)S(=O)(=O)NCc1ccc(CCCN)cc1**, which is N-[4-(3-aminopropyl)benzyl]propane-2-sulfonamide.

**1. Regiochemistry Discrepancy**

The sulfonamide group attaches to a **benzyl carbon** instead of the intended propyl carbon. Since the sulfonamide nitrogen is a weak nucleophile (due to the electron-withdrawing sulfonyl group), under **$S_N2$ conditions**, it preferentially attacks the more accessible and reactive benzylic bromide.

**2. Reactivity & Activation Constraints**

- **Base Efficiency:** $K_2CO_3$ in DMF may not sufficiently activate the sulfonamide nitrogen.
- **Competing Sites:** The benzylic position's higher reactivity leads to incorrect regiochemistry.

**3. Recommended Improvements**

Prioritize disconnection at the **sulfonamide N–S bond**. A successful strategy uses a primary amine and sulfonyl chloride. Additionally, **benzyl carbamate (Cbz) protection** is advised for reliable amine handling.

```
// Verification Summary
{
    "conclusion": "infeasible",
    "major_error": "incorrect regiochemistry",
    "observed_site": "benzylic carbon attack",
    "target_site": "propyl carbon",
    "optimization": "use sulfonyl chloride + protected primary amine"
}
```

Figure 10: An example Chemical Verification QA data.