R^3: End-to-End Reasoning-based Planning for Multi-step Retrosynthesis via Reinforcement Learning

Anonymous ACL submission

Abstract

Multi-step retrosynthetic planning is a fundamental challenge in organic chemistry, traditionally modeled as a combinatorial search problem guided by single-step prediction models. However, this search-centric paradigm often disconnects from the explicit chemical reasoning processes employed by human experts. In this paper, we propose R^3 (**R**einforced **R**easoning **R**etrosynthesis), a novel framework that reformulates this task as end-to-end generative reasoning. Instead of traversing a search tree, R^3 simulates the problem-solving logic of chemists to directly generate complete synthetic pathways. To achieve this, we initialize the model with domain knowledge and employ end-to-end Reinforcement Learning (RL) to optimize the entire planning policy. Experimental results on Retrobench show that R^3 achieves a state-of-the-art Top-1 accuracy of 43.7%, demonstrating that generative reasoning offers a superior alternative to traditional search algorithms in solving complex retrosynthetic problems.

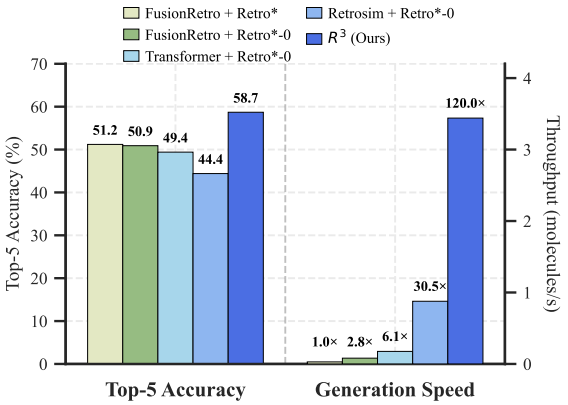

Figure 1: Performance comparison on Retrobench. The left chart displays the Top-5 accuracy (%), where R^3 achieves **58.7%**. The right chart illustrates the generation speed (throughput in molecules/s) annotated with relative speedup factors. R^3 demonstrates superior efficiency, achieving a **120.0**× speedup compared to the FusionRetro + Retro* baseline. Here, Retro*-0 denotes the variant of Retro* that does not employ a value model for search guidance; by contrast, Retro* integrates a value model to optimize its search strategy, which leads to a slight improvement in performance but a trade-off of slower generation speed.

1 Introduction

Retrosynthesis planning is a fundamental strategy in organic chemistry and drug discovery, aiming to systematically deconstruct complex target molecules into simpler, commercially available precursors (Corey, 1991; Zheng et al., 2022). This task is generally categorized into two levels: single-step retrosynthesis and multi-step retrosynthetic planning. Specifically, single-step retrosynthesis aims to identify a set of immediate reactants that can transform into a given target molecule via a one-step chemical reaction (Jiang et al., 2023). In contrast, multi-step planning seeks to discover a complete, sequential reaction pathway that recursively transforms the target into a set of commercially available starting materials (building blocks) (Zheng et al., 2022; Zhong et al., 2023).

Traditionally, AI-driven methods have treated multi-step planning as a heuristic search problem. These approaches typically focus on either enhancing the accuracy of single-step models to prune the search tree (Kim et al., 2021; Liu et al., 2023a) or training value networks to guide algorithms like Monte Carlo Tree Search (MCTS) (Segler et al., 2018; Hong et al., 2023) or A* search (Chen et al., 2020). While effective to some extent, these methods often suffer from high computational overhead and a "black-box" nature, lacking the explicit chemical reasoning that human experts employ when designing a route. Additionally, they are constrained by the performance ceiling of single-step models.

Recently, Large Language Models (LLMs) have shown remarkable capabilities in complex reason-

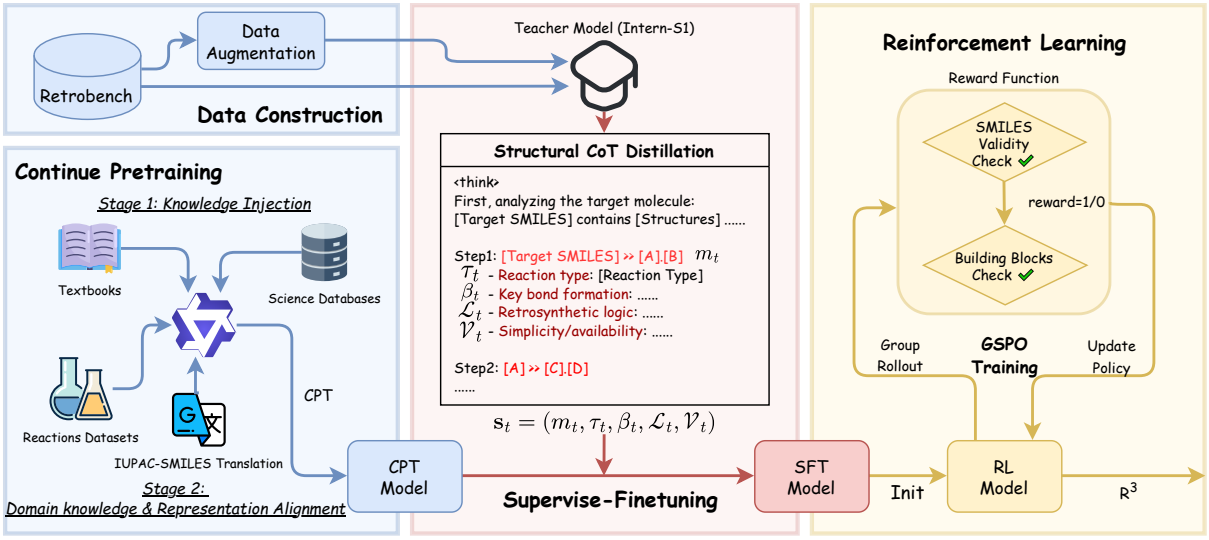

Figure 2: The overall framework of R^3. The training pipeline consists of three phases: (1) Continual Pretraining (CPT): Starting from Qwen3-8B-Base, the model undergoes a two-stage pretraining process on biomedical texts and reaction datasets to align chemical representations (SMILES/IUPAC) and inject domain knowledge. (2) Structural CoT Distillation (SFT): A teacher model (Intern-S1) distills reasoning capabilities into the student model using a structural reasoning protocol. The reasoning trajectory is structured into a 5-tuple sequence (reaction type τ, key disconnection β, rationale L, etc.) to prevent logical drift. (3) Reinforcement Learning (RL): The model is further optimized using Group Sequence Policy Optimization (GSPO). The reward function $r\,(x, y)$ strictly enforces both chemical validity of intermediate SMILES and the successful termination of the route at commercially available building blocks.

ing tasks across various domains (Guo et al., 2025; Jaech et al., 2024; Yu et al., 2025; Trinh et al., 2024; Li et al., 2022). Recent research has begun to explore the use of LLMs for retrosynthesis; however, these efforts are primarily limited to single-step retrosynthesis (Zhang et al., 2025a) or rely on integrating LLMs into search algorithms (Kang et al., 2025; Liu et al., 2025; Wang et al., 2025). The dependency on single-step models and search algorithms prevents these methods from fully exploiting the generative reasoning capabilities of LLMs for direct multi-step retrosynthetic route planning.

In this paper, we propose a novel framework to train an LLM (Qwen3-8B-Base) as a specialized agent for multi-step retrosynthetic planning. Our method shifts the paradigm from "searching" to "generative reasoning-based planning", enabling the model to directly generate complete retrosynthetic routes in an end-to-end manner. This approach mimics the problem-solving logic of human chemists, who reason through each step of the retrosynthesis process rather than exhaustively searching through all possibilities.

To achieve this, we first establish a robust chemical reasoning foundation by injecting domain knowledge and structured logic through continual pretraining and supervised fine-tuning. Building upon this, we employ end-to-end Reinforcement Learning (RL) to optimize the entire planning policy, ensuring that the model learns to generate valid retrosynthetic routes. Experimental results on Retrobench (Liu et al., 2023b) demonstrate that our model achieves state-of-the-art performance with a Top-1 accuracy of 43.7% and a Top-5 accuracy of 58.7%. Beyond accuracy, R^3 demonstrates superior test-time scaling and inference efficiency, achieving a 120x speedup compared to traditional search-based baselines (Figure 1).

2 Related Work

Search-based Multi-step Retrosynthesis Traditionally, search-based multi-step retrosynthesis formulates the planning task as a search problem over a molecular graph or tree, aiming to identify a valid pathway of chemical reactions to transform a target molecule into commercially available starting materials. Foundationally, Monte Carlo Tree Search (MCTS) (Segler et al., 2018) and A*-like search strategies, such as Retro* (Chen et al., 2020), formulate planning as a search problem guided by neural policy and value networks. Subsequent research has focused on enhancing search efficiency

and route quality: Retro*+ (Kim et al., 2021) and PDVN (Liu et al., 2023a) introduce self-improving mechanisms and dual-value networks to better estimate synthesis costs; EG-MCTS (Hong et al., 2023) and MEEA* (Zhao et al., 2024) incorporate experience guidance and look-ahead exploration into tree search. Beyond single-target planning, methods like RetroGraph (Xie et al., 2022) and DreamRetroer (Zhang et al., 2025b) enable efficient group retrosynthesis. Additionally, recent frameworks like FusionRetro (Liu et al., 2023b) and CREBM (Liu et al., 2024a) integrate advanced molecular representations and energy-based reranking to further ensure the feasibility of the predicted routes.

LLM-based Retrosynthesis Recently, LLMs are increasingly being explored for retrosynthesis tasks due to their powerful reasoning capabilities. Some studies train foundation models on chemical data, however not specifically for retrosynthesis tasks (Frey et al., 2023; Bai et al., 2025). On single-step retrosynthesis tasks, LLM-based methods have been proven to possess state-of-the-art (SOTA) capabilities. (Zhang et al., 2025a; Yang et al., 2025b; Liu et al., 2024b). In multi-step retrosynthesis tasks, some studies have explored integrating LLMs into search algorithms to assist in refining retrosynthetic routes (Wang et al., 2025) or utilizing text descriptions as a modality to guide the search process (Kang et al., 2025). Specifically, Retro-R1 (Liu et al., 2025) provides the single-step model as a callable tool to the LLM, transforming multi-step retrosynthesis into an agent tool-use scenario. While the aforementioned methods leverage LLMs for retrosynthesis tasks, most are limited by the performance of the underlying single-step models and fail to utilize the LLMs' superior reasoning capabilities to propose retrosynthetic routes directly.

3 Methodology

3.1 Problem Formulation

We formulate multi-step retrosynthetic planning as a sequential decomposition problem. Let $\mathcal{M}$ denote the chemical space and $\mathcal{B} \subset \mathcal{M}$ be the set of available building blocks. Given a target molecule $x \in \mathcal{M} \setminus \mathcal{B}$, the goal is to generate a retrosynthesis plan $\mathcal{T}$ — a directed acyclic graph where the root is x and all leaves belong to $\mathcal{B}$.

Traditionally, this problem is solved using search algorithms. The search process can be defined as finding a path in a state space graph where nodes represent molecules or sets of molecules, and edges represent chemical reactions. Let S_t be the set of precursor molecules at step t, with $S_0 = \{x\}$. At each step, a single-step retrosynthesis model $f(m)$ predicts a set of possible reactants R for a molecule $m \in S_t$. The search algorithm (e.g. A*) explores the tree of possible reactions to find a sequence such that the final set of leaves $S_T \subseteq \mathcal{B}$. The objective is often to minimize the cost of the route, defined as $C(\mathcal{T}) = \sum_{r \in \mathcal{T}} c(r)$, where $c(r)$ is the cost of an individual reaction or material, or the depth of the route $D(\mathcal{T})$, which is the maximum number of reaction steps in any branch of the retrosynthesis tree.

Unlike searching algorithms, we propose a Generative Reasoning Paradigm. We model the retrosynthesis planner as a parameterized policy π_θ. The model does not merely output the reaction pathway y; instead, it generates a joint distribution of a reasoning trajectory z and the final plan y:

$$\pi_\theta(z, y|x) = \pi_\theta(y|x, z)\pi_\theta(z|x). \quad (1)$$

In this framework, the reasoning process z acts as a latent variable that bridges the gap between the target x and the solution y. The model first generates the chain-of-thought z to deduce the synthesis strategy step-by-step, and conditioned on this reasoning, it subsequently generates the structured retrosynthetic plan y.

3.2 Continual Pretraining

To enhance the model's domain-specific foundation and specialized reasoning capabilities, we perform a two-stage continual pre-training process starting from Qwen3-8B-Base model (Yang et al., 2025a).

The first stage focuses on injecting general biomolecular and scientific knowledge. The training corpora for this stage are curated from specialized databases and literatures, including PubChem (Kim et al., 2025), PubMed (White, 2020), bioRxiv (Sever et al., 2019), and the biology and chemistry subsets of FineFineWeb (M-A-P et al., 2024). To ensure the model retains its general-purpose capabilities, we also incorporate a balanced portion of general data from FineWeb-Edu (Lozhkov et al., 2024). During this phase, we follow standard unsupervised pre-training objectives, where the loss is calculated over the entire text sequence.

The second stage is specifically designed to enhance retrosynthesis-specific domain knowledge

and representational alignment. We utilize a vast collection of chemical reactions sourced from the Open Reaction Database (ORD) (Kearnes et al., 2021a) and the USPTO (Lowe, 2012). We also integrate bidirectional translation tasks between molecular IUPAC names and SMILES. This representational alignment is essential because, during the subsequent reasoning process, molecular fragments and functional groups are frequently referenced by their IUPAC names in the textual chain-of-thought. Aligning these two modalities enables the model to bridge the gap between structural data and linguistic chemical concepts. This second stage shifts to a training paradigm closer to SFT, where the loss is calculated exclusively on the output tokens. Detailed statistics regarding dataset scales and specific data formats for each task are provided in Appendix Section A.

3.3 Structural Chain-of-Thought Distillation

A key challenge in end-to-end planning is that multi-step retrosynthesis inherently requires the model to maintain and navigate a complex search tree during its reasoning process. Although long CoT reasoning, as seen in models like DeepSeek-R1, enhances problem-solving, it often produces excessively long and unstructured text that is inefficient for representing tree-structured planning. This lack of structure makes it difficult for model to effectively learn the underlying chemical logic during the SFT process. To address this, we introduce Structural CoT Distillation.

Reasoning Structure Instead of free-form generation, we enforce a structured reasoning protocol. We define the reasoning trajectory z as a sequence of steps, $z = (\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_L)$, where L represents the depth of the retrosynthesis tree. Each step $\mathbf{s}_t$ is modeled as a 5-tuple representing a comprehensive chemical decision unit:

$$\mathbf{s}_t = (m_t, \tau_t, \beta_t, \mathcal{L}_t, \mathcal{V}_t) \tag{2}$$

m_t **(Transformation):** The specific reaction mapping (e.g., SMILES string $P \gg R_1.R_2$).

τ_t **(Reaction Type):** The categorical classification (e.g., 1,3-Dipolar cycloaddition), providing mechanistic context.

β_t **(Key Disconnection):** The identification of the strategic bond to break (e.g., N-O bond of isoxazole ring), ensuring topological awareness.

$\mathcal{L}_t$ **(Strategic Rationale):** A natural language justification explaining why this step is favorable (e.g., thermodynamic stability, functional group tolerance).

$\mathcal{V}_t$ **(Availability Check):** A verification step assessing if the generated precursors are commercially available building blocks.

Reasoning Data Distillation We utilize the training dataset of Retrobench (Liu et al., 2023b) as the initial data source. To capture the multimodal nature of retrosynthesis, we augment the dataset using retro* (Chen et al., 2020) to discover diverse valid pathways.

We employ Intern-S1 (Bai et al., 2025) as the teacher model. During the distillation phase, prompts provided to the teacher model includes the ground-truth route to facilitate the generation of high-quality reasoning traces. During the SFT phase, prompts for the student model contain only target molecules, ensuring that the model learns to plan without access to the answer.

Supervised Fine-Tuning We fine-tune the student model π_θ to minimize the negative log-likelihood of this structured sequence:

$$\mathcal{L}_{SFT} = -\sum_{(x,z,y)\in\mathcal{D}_{SFT}} \log P_\theta(z, y|x) \tag{3}$$

3.4 Reinforcement Learning

To further improve the model's performance on retrosynthesis tasks, we employ GSPO (Zheng et al., 2025), a variant of GRPO (Shao et al., 2024), as our RL algorithm.

Given a target product x, the policy π_θ generates a retrosynthetic route y. We utilize a binary outcome-supervised reward function $r(x, y)$, defined to enforce both chemical validity and precise termination at the desired starting materials. Specifically, $r(x, y) = 1$ if and only if all intermediate products in the route y are chemically valid SMILES, and the set of leaf nodes (starting materials) matches the ground-truth building blocks. Following the standard evaluation protocol on Retrobench (Liu et al., 2023b), this matching is performed by comparing the canonical InChiKeys of the molecules.

The GSPO objective is written as:

$$\mathcal{J}_{\text{GSPO}}(\theta) = \mathbb{E}_{\substack{x\sim\mathcal{D} \\ \{y_i\}\sim\pi_{\theta_{\text{old}}}}} \left[\frac{1}{G} \sum_{i=1}^{G} \mathcal{L}_i(\theta) \right], \tag{4}$$

Table 1: Main Results on Retrobench. We compare R^3 with SOTA baselines across three categories. R^3 achieves new SOTA performance on all metrics, from Top-1 to Top-5 accuracy.

Category	Method	Top-1 (%)	Top-2 (%)	Top-3 (%)	Top-4 (%)	Top-5 (%)
Template-based	Retrosim (Coley et al., 2017)	35.1	40.5	42.9	44.0	44.4
	Neuralsym (Segler and Waller, 2017)	42.0	49.3	52.0	53.6	54.3
	GLN (Dai et al., 2019)	39.6	48.9	52.7	54.6	55.7
Template-free	Transformer (Karpov et al., 2019)	31.3	40.4	44.7	47.2	49.4
	Megan (Sacha et al., 2021)	19.5	29.7	37.2	42.6	45.9
	FusionRetro (Liu et al., 2023b)	37.5	45.0	48.3	50.2	51.2
	FusionRetro+CREBM (Liu et al., 2024a)	39.6	46.7	49.5	51.0	51.7
	RetroInText (Kang et al., 2025)	42.1	49.9	53.0	54.7	55.7
Generative LLMs	Intern-S1 (Bai et al., 2025)	2.0	3.5	5.7	8.1	11.5
	GPT-5 (OpenAI, 2025)	2.4	6.3	10.7	12.0	14.9
	Gemini-3-pro (Google, 2025)	6.4	8.5	12.5	14.5	20.0
	R^3 **(Ours)**	**43.7**	**50.4**	**55.0**	**57.3**	**58.7**

where $\mathcal{L}_i$ represents the clipped objective for the i-th sample:

$$\mathcal{L}_i = \min\left(s_i\widehat{A}_i, \text{clip}\left(s_i, 1-\varepsilon, 1+\varepsilon\right)\widehat{A}_i\right). \quad (5)$$

GSPO utilize the group-based standardized advantage $\widehat{A}_i = (r(x, y_i) - \mu_r)/\sigma_r$ and the importance ratio $s_i = (\pi_\theta(y_i|x)/\pi_{\theta_{\text{old}}}(y_i|x))^{1/|y_i|}$. The importance ratio of GSPO is based on sequence-level, making it a more stable RL algorithm than GRPO (Zheng et al., 2025; Shao et al., 2024).

4 Experiments

4.1 Experimental Setup

Dataset We evaluate our method on Retrobench, a widely used benchmark for retrosynthesis prediction proposed by FusionRetro (Liu et al., 2023b). This benchmark is constructed from the USPTO reaction dataset and a set of commercially available building blocks. Detailed statistics regarding the dataset are provided in Appendix B. To ensure fair comparison and data consistency, we canonicalize all molecules in the dataset following the methodology established in FusionRetro (Liu et al., 2023b).

Baseline We compare our method against two categories of baselines: (1) Search-based methods, including Retrosim (Chen et al., 2020), Neuralsym (Kim et al., 2021), GLN (Dai et al., 2019), Transformer (Karpov et al., 2019), Megan (Sacha et al., 2021), FusionRetro (Liu et al., 2023b), CREBM (Liu et al., 2024a), and RetroInText (Kang et al., 2025); (2) Generative LLMs, including Intern-S1 (Bai et al., 2025), GPT-5 (OpenAI, 2025), and Gemini-3-pro (Google, 2025). For search-based methods, we report the results using the search algorithm (Retro* or Retro*-0) that achieves the best performance. For LLM baselines, we use the same prompt as our method during inference to ensure a fair comparison, and evaluate them on a subset of the test set due to computational constraints.

Metric We evaluate the performance using the Top-1 exact match accuracy. The match is based on the comparison of the InChiKey of the predicted building blocks and the ground-truth building blocks, following the evaluation protocol established in FusionRetro (Liu et al., 2023b).

For Generative LLMs, we adopt a specific protocol for Top-K evaluation. We oversample 8 responses for each test instance and remove duplicates. To compute Top-K accuracy, we randomly select k unique samples, extract their corresponding retrosynthetic route trees, and merge them into a unified graph. A test case is considered solved if the set of ground-truth starting materials is contained within the nodes of this merged graph. We conduct sampling across various temperatures and

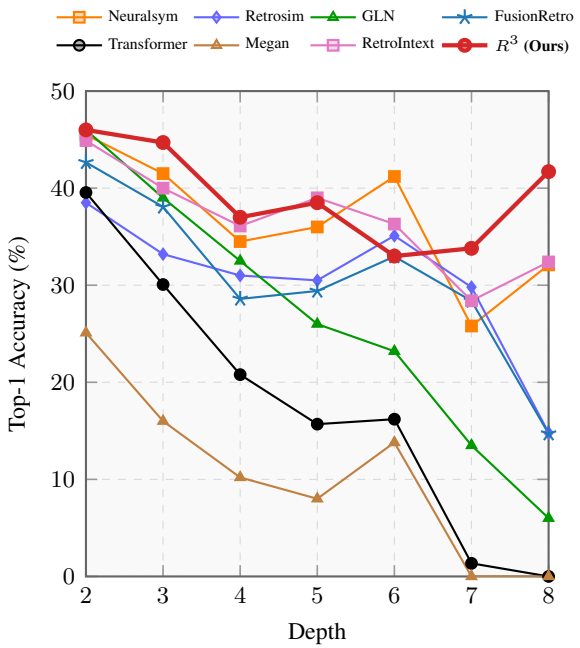

Figure 3: Comparison of Top-1 Accuracy across varying synthetic route depths. R^3 demonstrates superior performance, particularly at greater depths, compared to baselines including Neuralsym, Retrosim, and FusionRetro.

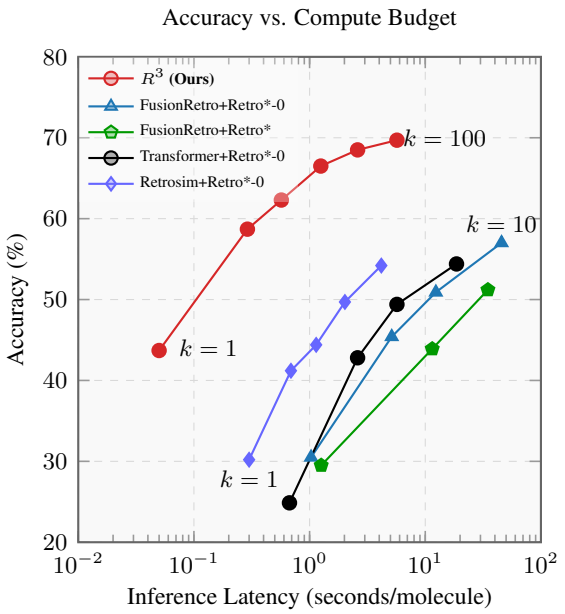

Figure 4: Inference Scaling Efficiency. We plot the cumulative Top-K accuracy against the average inference latency per molecule on a logarithmic scale. For FusionRetro, distinct points correspond to increasing beam sizes (from $k = 1$ to 10), illustrating the significant latency cost required to expand the search space. In contrast, R^3 demonstrates a superior Pareto frontier, efficiently converting test-time compute (sampling budget k) into performance gains without the exponential overhead of planning algorithms.

top-p settings and report the best performance for Top-K metrics.

4.2 Results

Comparison with Baselines The main results are summarized in Table 1. Our method achieves a Top-1 accuracy of 43.7% on Retrobench, establishing a new state-of-the-art. This performance significantly surpasses both the strongest search-based baseline, RetroInText (42.1%), and the fusion-based approach, FusionRetro (37.5%).

Beyond Top-1 accuracy, R^3 demonstrates superior capability in generating diverse valid pathways. In terms of Top-k performance, our model consistently outperforms all baselines, achieving 50.4% at Top-2 and 58.7% at Top-5. Crucially, this significant performance gain at higher k values evidences the model's **test-time scaling** capability. It demonstrates that by increasing the computational budget at inference time, R^3 can effectively tackle a wider array of challenging retrosynthesis problems that remain unsolved by a single greedy decoding pass.

Furthermore, the results highlight the necessity of domain-specific adaptation. General-purpose LLMs, such as GPT-5 and Gemini-3-pro, achieve negligible accuracy ($< 7\%$), confirming that the complex logic of retrosynthesis cannot be solved by reasoning capabilities alone without specialized chemical knowledge injection and alignment.

Performance at Different Depths To analyze the model's robustness across varying levels of retrosynthesis complexity, we evaluate R^3 against baselines across route depths ranging from 2 to 8. As illustrated in Figure 3, R^3 consistently outperforms all baselines at nearly every depth level.

Notably, traditional search-based methods often suffer from the "curse of dimensionality" as the search tree deepens. In contrast, R^3 exhibits a substantial advantage at greater depths (e.g., maintaining $> 41\%$ accuracy at depth 8), whereas methods like FusionRetro degrade significantly (dropping to $\sim 14\%$). This suggests that our generative reasoning paradigm effectively captures global structural dependencies, preventing the error propagation typical in greedy or heuristic search algorithms.

4.3 Computing Efficiency

In high-throughput scenarios, inference latency is a critical bottleneck. As shown in Table 1, R^3 achieves an average throughput of **3.44** molecules

Model	Top-1 Acc (%)
Qwen3-8B-Base	0.0
+ CPT	0.0
+ SFT (w/o Structural CoT)	4.0
+ SFT	28.3
+ SFT + RL	29.2
+ CPT + SFT	33.1
+ CPT + SFT + RL	**43.7**

Table 2: Ablation study on Retrobench Top-1 Accuracy. We illustrate the impact of CPT, SFT, and RL (200 steps) on the Qwen3-8B-Base model.

per second to generate 5 samples, representing a **120× speedup** over FusionRetro+Retro* (0.028mol/s). This dramatic acceleration stems from our shift to a generative paradigm: unlike search algorithms that iteratively navigate a combinatorial tree (often stalling GPU utilization below 35% due to CPU-GPU overhead), R^3 generates the complete route in a single autoregressive pass. This allows R^3 to leverage optimizations like FlashAttention to maintain $> 90\%$ GPU utilization, while search-based methods only achieve $\sim$35% GPU utilization.

Figure 4 further visualizes the trade-off between compute budget and performance. R^3 establishes a superior **Pareto frontier**, delivering competitive accuracy at a fraction of the latency. Crucially, our model demonstrates effective **test-time scaling**: increasing the sampling budget to $k = 100$ boosts coverage to 69.7% while maintaining a total latency (~ 5.7s per molecule) that is still orders of magnitude faster than the search-based baseline. In contrast, search algorithms like FusionRetro face a scaling wall: expanding the search width (e.g., beam size) from $k = 1$ to $k = 10$ incurs a massive latency penalty (from $\sim$1s to $\sim$45s) due to the combinatorial explosion of the search space, yielding only marginal accuracy gains due to the diminishing returns of exploring lower-probability branches.

4.4 Ablation Study

To decouple the contributions of our three-stage training pipeline, we analyze the functional role of each phase as shown in Table 2. The results highlight a synergistic dependency among the stages:

SFT as the Cold-Start Foundation SFT is the prerequisite for the model to function. The base model (0.0%) completely fails to solve any retrosynthesis tasks. SFT is essential to "cold start" the policy; without the structural constraints learned during this phase, the model cannot generate valid reasoning traces, rendering subsequent RL impossible to launch.

CPT as the Knowledge Reservoir While SFT adapts the format, the model does not acquire plentiful new domain knowledge. The necessity of CPT is evident when comparing *Base+SFT+RL* (29.2%) with our full method (43.7%). Without the domain-specific representation space constructed during CPT, the model hits an early performance ceiling. CPT provides the fundamental chemical logic that allows the model to respond to RL optimization effectively.

RL as the Performance Catalyst Finally, RL acts as the catalyst to unlock the model's latent potential. On top of the CPT-enhanced backbone, RL drives a massive **10.6% improvement** (33.1% $\rightarrow$ 43.7%). This confirms that while CPT provides the knowledge and SFT provides the form, only RL can align these capabilities with the long-horizon goal of successful multi-step planning.

Ablation on Structural CoT To assess the impact of our proposed Structural Chain-of-Thought distillation, we conduct an ablation where we directly fine-tune the model using free-form CoT traces generated by the teacher model, without enforcing the structured 5-tuple schema. This variant achieves only 4.0% accuracy. By analyzing the SFT data, we observe that the teacher model often produces excessively long and unstructured reasoning paths, and tends to reveal ground-truth answers in the prompt. This leads to logical drift and poor generalization during inference. In contrast, our structured approach compels the model to explicitly reason about each chemical decision, significantly enhancing both interpretability and performance.

4.5 Building Block Awareness Analysis

A critical concern for generative planners is the potential hallucination of unavailable starting materials. Unlike search-based methods that query a database at every step, R^3 must internalize the definition of available building blocks ($\mathcal{B}$).

To quantify this capability, we constructed a classification probe dataset from Retrobench. We labeled ground-truth starting materials as *positive* and intermediate products as *negative*. We

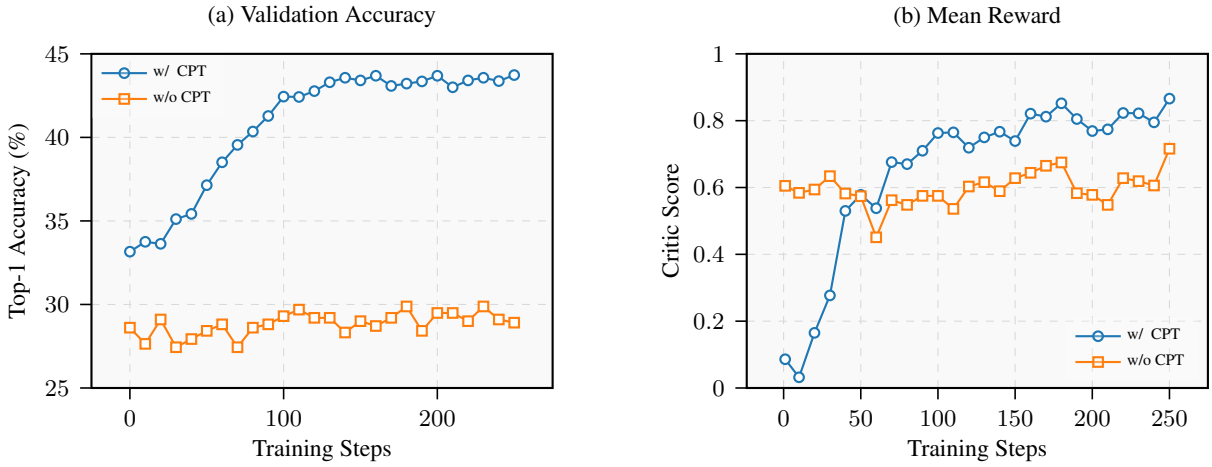

Figure 5: Training dynamics of CPT model vs. non-CPT model during RL stage.

then queried Qwen3-8B to determine if a given molecule belongs to $\mathcal{B}$. As shown in Table 3, Qwen3-8B achieves a Pass@1 accuracy of 82.0% and Pass@3 of 94.3%. This high accuracy demonstrates that the base model already learned the chemical space of commercially available materials, which is crucial for generating valid retrosynthetic routes without external database lookups.

Table 3: **Building Block Identification Accuracy.** The model is asked to classify whether a molecule is a commercially available building block.

Model	Pass@1	Pass@3
Qwen3-8B	82.0%	94.3%

4.6 RL Training Dynamics

We present the training dynamics during the RL stage in Figure 5. The model enhanced with CPT exhibits steady improvements in both validation accuracy and mean reward over 250 training steps. In contrast, the model without CPT fails to improve, stagnating at its initial SFT performance level. We observe that the CPT model starts with a lower initial reward but quickly surpasses the non-CPT model within the first 50 steps. This may be attributed to the injection of non-conversational corpora during the CPT phase, which temporarily misaligns the response style but ultimately provides a richer knowledge base for effective RL optimization.

5 Conclusion

In this work, we present R^3, a novel framework that leverages Large Language Models for multi-step retrosynthetic planning. By reformulating the task as a generative reasoning problem and employing end-to-end reinforcement learning, we enable the model to generate interpretable and chemically valid retrosynthesis routes. Our extensive experiments on Retrobench demonstrate that R^3 achieves state-of-the-art performance, significantly outperforming existing search-based and LLM-based methods. This work highlights the potential of LLMs in complex scientific reasoning tasks and opens new avenues for future research in AI-driven retrosynthetic planning.

Limitations

Despite the promising results, this work relies primarily on the exact match accuracy of starting materials to evaluate performance, which may not fully capture the holistic quality of the generated retrosynthetic routes, such as reaction conditions, yield, or atom economy. Furthermore, the evaluation is constrained by the Retrobench dataset, which contains a limited number of ground-truth pathways for each target molecule; consequently, chemically valid and efficient routes generated by our model that deviate from these annotated references may be overlooked, potentially underestimating the model's true planning capabilities due to the incomplete coverage of the vast chemical search space.

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*.

He Cao, Yanjun Shao, Zhiyuan Liu, Zijing Liu, Xiangru Tang, Yuan Yao, and Yu Li. 2024. Presto: Progressive pretraining enhances synthetic chemistry outcomes. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 10197–10224.

Binghong Chen, Chengtao Li, Hanjun Dai, and Le Song. 2020. Retro*: Learning retrosynthetic planning with neural guided a* search. In *International Conference on Machine Learning*, pages 1608–1616. PMLR.

Connor W. Coley, Luke Rogers, William H. Green, and Klavs F. Jensen. 2017. Computer-assisted retrosynthesis based on molecular similarity. *ACS Central Science*, 3(12):1237–1245.

Elias James Corey. 1991. The logic of chemical synthesis: Multistep synthesis of complex carbogenic molecules (nobel lecture). *Angewandte Chemie International Edition in English*, 30(5):455–465.

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.

Nathan C. Frey, Ryan Soklaski, Simon Axelrod, Siddharth Samsi, Rafael Gómez-Bombarelli, Connor W. Coley, and Vijay Gadepally. 2023. Neural scaling of deep chemical models. *Nature Machine Intelligence*, 5(11):1297–1305.

Google. 2025. Gemini 3 Pro Model Card. https://deepmind.google/models/gemini/pro/. Accessed: 2025-12-28.

Daya Guo and 1 others. 2025. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint*.

Siqi Hong, Hankz Hankui Zhuo, Kebing Jin, Guang Shao, and Zhanwen Zhou. 2023. Retrosynthetic planning with experience-guided monte carlo tree search. *Communications Chemistry*, 6(1).

Aaron Jaech, Adam Kalai, Adam Lerer, Adam Richardson, Ahmed El-Kishky, Aiden Low, Alec Helyar, Aleksander Madry, Alex Beutel, Alex Carney, and 1 others. 2024. Openai o1 system card. *arXiv preprint arXiv:2412.16720*.

Yinjie Jiang, Yemin Yu, Ming Kong, Yu Mei, Luotian Yuan, Zhengxing Huang, Kun Kuang, Zhihua Wang, Huaxiu Yao, James Zou, Connor W. Coley, and Ying Wei. 2023. Artificial intelligence for retrosynthesis prediction. *Engineering*, 25:32–50.

Chenglong Kang, Xiaoyi Liu, and Fei Guo. 2025. Retrointext: A multimodal large language model enhanced framework for retrosynthetic planning via in-context representation learning. In *The Thirteenth International Conference on Learning Representations*.

Pavel Karpov, Guillaume Godin, and Igor V. Tetko. 2019. *A Transformer Model for Retrosynthesis*, page 817–830. Springer International Publishing.

Steven M Kearnes, Michael R Maser, Michael Wleklinski, Anton Kast, Abigail G Doyle, Spencer D Dreher, Joel M Hawkins, Klavs F Jensen, and Connor W Coley. 2021a. The open reaction database. *Journal of the American Chemical Society*, 143(45):18820–18826.

Steven M Kearnes, Michael R Maser, Michael Wleklinski, Anton Kast, Abigail G Doyle, Spencer D Dreher, Joel M Hawkins, Klavs F Jensen, and Connor W Coley. 2021b. The open reaction database. *Journal of the American Chemical Society*, 143(45):18820–18826.

Junsu Kim, Sungsoo Ahn, Hankook Lee, and Jinwoo Shin. 2021. Self-improved retrosynthetic planning. In *International Conference on Machine Learning*, pages 5486–5495. PMLR.

Sunghwan Kim, Jie Chen, Tiejun Cheng, Asta Gindulyte, Jia He, Siqian He, Qingliang Li, Benjamin A Shoemaker, Paul A Thiessen, Bo Yu, and 1 others. 2025. Pubchem 2025 update. *Nucleic acids research*, 53(D1):D1516–D1525.

Yujia Li, David Choi, Junyoung Chung, Nate Kushman, Julian Schrittwieser, Rémi Leblond, Tom Eccles, James Keeling, Felix Gimeno, Agustin Dal Lago, Thomas Hubert, Peter Choy, Cyprien de Masson d'Autume, Igor Babuschkin, Xinyun Chen, Po-Sen Huang, Johannes Welbl, Sven Gowal, Alexey Cherepanov, and 7 others. 2022. Competition-level code generation with alphacode. *Science*, 378(6624):1092–1097.

Guoqing Liu, Di Xue, Shufang Xie, Yingce Xia, Austin Tripp, Krzysztof Maziarz, Marwin Segler, Tao Qin, Zongzhang Zhang, and Tie-Yan Liu. 2023a. Retrosynthetic planning with dual value networks. In *International conference on machine learning*, pages 22266–22276. PMLR.

Songtao Liu, Hanjun Dai, Yue Zhao, and Peng Liu. 2024a. Preference optimization for molecule synthesis with conditional residual energy-based models. *arXiv preprint arXiv:2406.02066*.

Songtao Liu, Zhengkai Tu, Minkai Xu, Zuobai Zhang, Lu Lin, Rex Ying, Jian Tang, Peilin Zhao, and Dinghao Wu. 2023b. Fusionretro: molecule representation fusion via in-context learning for retrosynthetic planning. In *International Conference on Machine Learning*, pages 22028–22041. PMLR.

Wei Liu, Jiangtao Feng, Hongli Yu, Yuxuan Song, Yuqiang Li, Shufei Zhang, Lei Bai, Wei-Ying Ma, and Hao Zhou. 2025. Retro-r1: Llm-based agentic retrosynthesis. *arXiv preprint*. Submitted to NeurIPS 2025.

Yifeng Liu, Hanwen Xu, Tangqi Fang, Haocheng Xi, Zixuan Liu, Sheng Zhang, Hoifung Poon, and Sheng Wang. 2024b. T-rex: Text-assisted retrosynthesis prediction. *arXiv preprint arXiv:2401.14637*.

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

Anton Lozhkov, Loubna Ben Allal, Leandro von Werra, and Thomas Wolf. 2024. Fineweb-edu: the finest collection of educational content.

M-A-P, Ge Zhang, Xinrun Du, Zhimiao Yu, Zili Wang, Zekun Wang, Shuyue Guo, Tianyu Zheng, Kang Zhu, Jerry Liu, Shawn Yue, Binbin Liu, Zhongyuan Peng, Yifan Yao, Jack Yang, Ziming Li, Bingni Zhang, Minghao Liu, Tianyu Liu, and 6 others. 2024. Fine-fineweb: A comprehensive study on fine-grained domain web corpus.

OpenAI. 2025. GPT-5 System Card. `https://openai.com/zh-Hans-CN/index/gpt-5-system-card/`. Accessed: 2025-12-28.

Qizhi Pei, Lijun Wu, Kaiyuan Gao, Xiaozhuan Liang, Yin Fang, Jinhua Zhu, Shufang Xie, Tao Qin, and Rui Yan. 2024. Biot5+: Towards generalized biological understanding with iupac integration and multi-task tuning. In *Findings of the Association for Computational Linguistics ACL 2024*, pages 1216–1240.

Mikołaj Sacha, Mikołaj Błaż, Piotr Byrski, Paweł Dąbrowski-Tumański, Mikołaj Chromiński, Rafał Loska, Paweł Włodarczyk-Pruszyński, and Stanisław Jastrzębski. 2021. Molecule edit graph attention network: Modeling chemical reactions as sequences of graph edits. *Journal of Chemical Information and Modeling*, 61(7):3273–3284.

Marwin H. S. Segler and Mark P. Waller. 2017. Neural-symbolic machine learning for retrosynthesis and reaction prediction. *Chemistry – A European Journal*, 23(25):5966–5971.

Marwin HS Segler, Mike Preuss, and Mark P Waller. 2018. Planning chemical syntheses with deep neural networks and symbolic ai. *Nature*, 555(7698):604–610.

Richard Sever, Ted Roeder, Samantha Hindle, Linda Sussman, Kevin-John Black, Janet Argentine, Wayne Manos, and John R Inglis. 2019. biorxiv: the preprint server for biology. *BioRxiv*, page 833400.

Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, Y. K. Li, Y. Wu, and Daya Guo. 2024. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. *Preprint*, arXiv:2402.03300.

Guangming Sheng, Chi Zhang, Zilingfeng Ye, Xibin Wu, Wang Zhang, Ru Zhang, Yanghua Peng, Haibin Lin, and Chuan Wu. 2024. Hybridflow: A flexible and efficient rlhf framework. *arXiv preprint arXiv: 2409.19256*.

Mujeen Sung, Minbyul Jeong, Yonghwa Choi, Donghyeon Kim, Jinhyuk Lee, and Jaewoo Kang. 2022. Bern2: an advanced neural biomedical named entity recognition and normalization tool. *Bioinformatics*, 38(20):4837–4839.

Trieu H. Trinh, Yuhuai Wu, Quoc V. Le, He He, and Thang Luong. 2024. Solving olympiad geometry without human demonstrations. *Nature*, 625(7995):476–482.

Haorui Wang, Jeff Guo, Lingkai Kong, Rampi Ramprasad, Philippe Schwaller, Yuanqi Du, and Chao Zhang. 2025. Llm-augmented chemical synthesis and design decision programs. *Preprint*, arXiv:2505.07027.

Jacob White. 2020. Pubmed 2.0. *Medical reference services quarterly*, 39(4):382–387.

Shufang Xie, Rui Yan, Peng Han, Yingce Xia, Lijun Wu, Chenjuan Guo, Bin Yang, and Tao Qin. 2022. Retrograph: Retrosynthetic planning with graph search. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, KDD '22, page 2120–2129, New York, NY, USA. Association for Computing Machinery.

An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, Chujie Zheng, Dayiheng Liu, Fan Zhou, Fei Huang, Feng Hu, Hao Ge, Haoran Wei, Huan Lin, Jialong Tang, and 41 others. 2025a. Qwen3 technical report. *arXiv preprint*.

Yifei Yang, Runhan Shi, Zuchao Li, Shu Jiang, Bao-Liang Lu, Qibin Zhao, Yang Yang, and Hai Zhao. 2025b. Batgpt-chem: A foundation large model for chemical engineering. *Research*, 8.

Qiying Yu, Zheng Zhang, Ruofei Zhu, Yufeng Yuan, Xiaochen Zuo, YuYue, Weinan Dai, Tiantian Fan, Gaohong Liu, Juncai Liu, LingJun Liu, Xin Liu, Haibin Lin, Zhiqi Lin, Bole Ma, Guangming Sheng, Yuxuan Tong, Chi Zhang, Mofan Zhang, and 17 others. 2025. DAPO: An open-source LLM reinforcement learning system at scale. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*.

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

Xuefeng Zhang, Haowei Lin, Muhan Zhang, Yuan Zhou, and Jianzhu Ma. 2025b. A data-driven group retrosynthesis planning model inspired by neurosymbolic programming. *Nature Communications*, 16(1).

Dengwei Zhao, Shikui Tu, and Lei Xu. 2024. Efficient retrosynthetic planning with mcts exploration enhanced a* search. *Communications Chemistry*, 7(1).

Chujie Zheng, Shixuan Liu, Mingze Li, Xiong-Hui Chen, Bowen Yu, Chang Gao, Kai Dang, Yuqiong Liu, Rui Men, An Yang, Jingren Zhou, and Junyang Lin. 2025. Group sequence policy optimization. *Preprint*, arXiv:2507.18071.

Shuangjia Zheng, Tao Zeng, Chengtao Li, Binghong Chen, Connor W. Coley, Yuedong Yang, and Ruibo Wu. 2022. Deep learning driven biosynthetic pathways navigation for natural products with bionavi-np. *Nature Communications*, 13(1).

Yaowei Zheng, Richong Zhang, Junhao Zhang, Yanhan Ye, Zheyan Luo, Zhangchi Feng, and Yongqiang Ma. 2024. Llamafactory: Unified efficient fine-tuning of 100+ language models. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations)*, Bangkok, Thailand. Association for Computational Linguistics.

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).

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.

A Details of Continual Pretraining

This section provides a comprehensive breakdown of the data sources, formatting strategies, and training objectives employed during the two-stage continual pre-training process. The detailed statistics for each dataset are summarized in Table 4, and the training hyperparameters are summarized in Table 5.

A.1 Stage 1: Biomolecular and Scientific Knowledge Injection

The primary objective of Stage 1 is to imbue the model with a broad foundation in biomolecular science and chemical structures. We categorize the data used in this stage into three distinct types based on their structural format:

Molecule+Text: This data type consists of molecular structural information paired with descriptive natural language. Specifically, we extract SMILES sequences from PubChem (Kim et al., 2025) and pair them with their corresponding textual descriptions (e.g., chemical properties, taxonomy, and pharmacological action). This allows the model to associate structural representations with scientific concepts.

Interleaved Text: To enhance the model's ability to handle multi-modal scientific literature, following Pei et al. (2024), we process corpora from bioRxiv (Sever et al., 2019) and PubMed (White, 2020) Abstracts using BERN2 (Sung et al., 2022), a high-performance biomedical entity recognition tool. We identify biological and chemical entities within the text and either insert or replace them with their corresponding SMILES sequences or IUPAC names. This creates a dense, interleaved representation that enables the model to establish a direct associative link between natural language mentions and their corresponding chemical structures.

Pure Sequence Data: This includes large-scale collections of SMILES and IUPAC names from PubChem (Kim et al., 2025), as well as scientific text from FineFineWeb-biology/chemistry (M-A-P et al., 2024) and PubMed Central (White, 2020). We also include FineWeb-Edu (Lozhkov et al., 2024) to maintain capabilities in the general domain. These datasets are treated as pure sequences.

During this stage, we employ a standard unsupervised pre-training objective, calculating the cross-entropy loss over the entire sequence to maximize the model's predictive accuracy across both textual and chemical modalities.

A.2 Stage 2: Retrosynthesis Specialization and Representational Alignment

The second stage transitions toward retrosynthesis reaction and the alignment of different molecular notations. The data types in this stage are formulated as follows:

Retrosynthesis Task: Utilizing data from the Open Reaction Database (ORD) (Kearnes et al., 2021b) and USPTO (Lowe, 2012), we formulate retrosynthesis as a sequence-to-sequence task. The input is the SMILES of product molecule, and the model is trained to predict the corresponding reactants.

Cross-Modal Alignment (IUPAC $\leftrightarrow$ SMILES): To bridge the gap between nomenclature and structure, we perform bidirectional translation tasks. For the IUPAC to SMILES task, the input is a molecule's IUPAC name and the output is its SMILES sequence; the SMILES to IUPAC task follows the reverse logic. This alignment is critical for ensuring the model can accurately reference molecular fragments by name during complex chain-of-thought reasoning.

USPTO-Application: A unique case in Stage 2 is the USPTO-Application dataset. We use its interleaved version provided by PRESTO (Cao et al., 2024). Specifically, the USPTO-Application dataset comprises chemical reaction equations and experimental procedures, where molecular entities identified by BERN2 (Sung et al., 2022) are replaced with their corresponding SMILES. To accommodate the SFT loss format, its input is set as an empty sequence, while the entire interleaved text serves as the output.

Unlike Stage 1, the training objective in Stage 2 shifts to an SFT-style loss, where the gradient is calculated exclusively on the output tokens (e.g., the reactants, the translated notation, or the interleaved content in USPTO-Application), thereby refining the model's precision in specialized chemical generation.

B Dataset Details

The Retrobench dataset consists of 46,458 training routes, 5,803 validation routes, and 5,838 test routes for evaluation. We summarize the data statistics of Retrobench in Table 6. The dataset is categorized by the depth of the ground-truth retrosynthetic routes.

Table 4: Detailed data configuration of the two-stage continual pre-training.

Data Source	Data Type	Samples	Tokens
Stage 1			
PubChem (Kim et al., 2025)	Molecule+Text	387.4K	70.7M
bioRxiv (Sever et al., 2019)	Interleaved Text	2.3M	495.3M
PubMed Abstract (White, 2020)	Interleaved Text	33.4M	10.4B
PubChem (Kim et al., 2025)	SMILES	114.8M	4.1B
PubChem (Kim et al., 2025)	IUPAC	114.8M	5.6B
FineFineWeb-biology (M-A-P et al., 2024)	Text	7.7M	8.1B
FineFineWeb-chemisty (M-A-P et al., 2024)	Text	9.0M	8.0B
PubMed Central (White, 2020)	Text	70.6M	13.0B
FineWeb-Edu (Lozhkov et al., 2024)	Text	19.7M	20.2B
Stage 2			
Open Reaction Database (Kearnes et al., 2021b)	Retrosynthesis	1.8M	135.5M
USPTO-Full (Dai et al., 2019)	Retrosynthesis	771.9K	50.1M
USPTO-MIT (Zhong et al., 2022)	Retrosynthesis	406.9K	31.2M
USPTO-Application (Cao et al., 2024)	Interleaved Text	1.6M	532.1M
PubChem (Kim et al., 2025)	IUPAC to SMILES	15M	1.2B
PubChem (Kim et al., 2025)	SMILES to IUPAC	15M	1.2B

Table 5: Detailed training configuration of the two-stage continual pre-training.

	Continual Pretraining Stage 1	Continual Pretraining Stage 2
Batch Size	1,024	256
Max Length	4,096	4,096
Training Tokens	70.1B	3.2B
Learning Rate Schedule	Cosine with min lr	Cosine
Peak Learning Rate	6.0×10^{-5}	1.0×10^{-4}
GPU Devices	$4 \times$ A100	$4 \times$ A100

C Hyperparameters

We utilize the Qwen3-8B-Base model (Yang et al., 2025a) as base model, LlamaFactory (Zheng et al., 2024) as SFT framework, and VeRL (Sheng et al., 2024) as RL framework. During SFT, we set the batch size to 64, learning rate to $3e-5$, and train for 3 epochs. During RL, we set the batch size to 64, group size to 16, and learning rate to $1e-6$. The training curves are shown in Figure 5.

D Case Study

We provide a concrete example of a retrosynthetic route generated by R^3 in Figure 6. This visualization demonstrates the model's ability to decompose a complex target into available building blocks. For brevity, the input prompts and the output reasoning traces are truncated.

Table 6: Data statistics of Retrobench. The table reports the number of molecules for each route depth across Training, Validation, and Test sets.

Dataset	**Depth**											
	2	3	4	5	6	7	8	9	10	11	12	13
Training	22,903	12,004	5,849	3,268	1,432	594	276	107	25	0	0	0
Validation	2,862	1,500	731	408	179	74	34	13	2	0	0	0
Test	2,862	1,500	731	408	179	74	34	13	2	32	2	1

MODEL INPUT (ABBREVIATED)

```
You are an expert in retrosynthesis planning. Your task is to propose a sequence of chemical
reactions that can synthesize a target molecule from available starting materials.

## Task Description
Given a target molecule (in SMILES format), work backwards to find a sequence of reactions
that breaks down the target into simpler molecules that are reasonably available as starting
materials.

## Output Format
Provide the final answer as a JSON list of reaction strings in the format
"PRODUCT>>REACTANTS".

Please propose a retrosynthesis route for the following target molecule:
Target molecule: Cn1ncc(C(=O)N2CCCC2)c1C(=O)Nc1ccc2nc(-c3cccnc3)nn2c1
```

MODEL OUTPUT (PARTIAL DISPLAY)

```
<think>
Step 1: Cn1ncc(C(=O)N2CCCC2)c1C(=O)Nc1ccc2nc(-c3cccnc3)nn2c1>>Cn1ncc(C(=O)O)c1C(=O)Nc1ccc2nc(-
c3cccnc3)nn2c1.C1CCNC1
- Reaction type: Amide Bond Disconnection (Aliphatic)
- Key bond formed: Amide bond between a pyrazole-carboxylic acid and pyrrolidine.
- Retrosynthetic Logic: The aliphatic amide is disconnected first. This is strategic because
pyrrolidine is a simple, commercially available secondary amine. Forming this bond late-stage
avoids potential side reactions involving the amine during the earlier construction of the
heteroaromatic core.
- Reactant Availability: Pyrrolidine (C1CCNC1) is a common commodity chemical.

Step 2: Cn1ncc(C(=O)O)c1C(=O)Nc1ccc2nc(-c3cccnc3)nn2c1>>CCOC(=O)c1cnn(C)c1C(=O)Nc1ccc2nc(-
c3cccnc3)nn2c1
- Reaction type: Functional Group Interconversion (Ester Hydrolysis)
- Key bond formed: Transformation of carboxylic acid back to ethyl ester.
- Retrosynthetic Logic: The free carboxylic acid is traced back to its ethyl ester precursor.
This is necessary because the ester group is typically carried through the ring-forming steps
(e.g., Knorr synthesis or condensation) used to build the pyrazole skeleton. Reverting to the
ester prepares the molecule for the *next* disconnection: breaking down the complex core into
simpler fragments.
- Reactant Availability: The ester is a key synthetic intermediate (not yet a starting
material) that facilitates further structural decomposition.

Step 3: ... [Remaining steps omitted for brevity] ...
</think>
<answer> ["Cn1ncc(C(=O)N2CCCC2)c1C(=O)Nc1ccc2nc(-
c3cccnc3)nn2c1>>Cn1ncc(C(=O)O)c1C(=O)Nc1ccc2nc(-c3cccnc3)nn2c1.C1CCNC1", ... ] </answer>
```

Figure 6: A representative case study generated by R^3.