ROUTEDIFF: MOLECULAR DESIGN VIA DISCRETE DIFFUSION OVER SYNTHETIC ROUTES

Anonymous authors
Paper under double-blind review

ABSTRACT

Molecular generative models can propose molecules with desirable properties, but the resulting molecules are not necessarily synthesizable. Synthesis-aware molecular generation addresses this gap by choosing both a molecule and an executable route that constructs it. Existing sequence-based route generators typically commit route decisions autoregressively or refine completed pathways in separate stages. We introduce RouteDiff, a masked discrete diffusion model over complete postfix routes. A bidirectional Transformer predicts unresolved building-block and reaction decisions from the currently visible route context, while diffusion progressively determines when those decisions are revealed and committed. To scale to a 223,243-entry building-block catalog, we combine frozen molecular representations with learned low-rank identity components, and compose supported property conditions at inference time without task-specific retraining. On Enamine, RouteDiff combines a 92.8% reconstruction rate and 0.987 similarity with the highest building-block diversity (0.783) among the compared methods. Under a $10{,}000$-oracle-call budget, it achieves the highest synthesis-constrained AUC Top-10 on eight of ten PMO tasks, with gains of up to 13.2 points. Denoising-trajectory analysis further finds pre-commit changes in unresolved building-block or reaction predictions in $36/48$ analyzed successful trajectories.

1 INTRODUCTION

Generative molecular design seeks molecules with desirable properties, but high model scores do not ensure that proposed structures can be made. De novo generators can produce unrealistic or synthetically inaccessible molecules even when they perform well on standard benchmarks (Gao & Coley, 2020). Synthetic-accessibility heuristics can bias generation toward more tractable molecules (Gao & Coley, 2020), but a score alone does not specify an executable synthesis route. A direct approach is to generate within a reaction-defined synthesis space, assembling purchasable building blocks through predefined reaction rules so that each candidate is accompanied by an executable route (Gao et al., 2022b; Cretu et al., 2025). This changes the modeling problem from generating molecular structures alone to generating complete synthetic routes.

Generating within a reaction-defined synthesis space still leaves open how the synthesis route itself should be represented and generated. Existing methods construct trees one reaction at a time (Gao et al., 2022b), learn sequential reaction policies with GFlowNets (Cretu et al., 2025; Seo et al., 2025), or optimize route structure and molecular choices at separate levels (Sun et al., 2025). Sequence-based methods instead represent complete pathways directly: ChemProjector and SynFormer use postfix routes (Luo et al., 2024; Gao et al., 2025), while ReaSyn combines bottom-up and top-down autoregressive pathway generation with explicit editing (Lee et al., 2026). Across these formulations, building-block and reaction decisions remain coupled through the intermediates they create, while alternative choices can give rise to multiple feasible routes.

Sequence models require a linear representation of the complete synthetic route. Postfix notation linearizes a synthesis tree into a stack-executable sequence of building blocks and reactions (Luo et al., 2024; Gao et al., 2025). When this serialization also determines left-to-right generation, however, each building-block or reaction decision is predicted only from the preceding prefix and committed before later route context becomes available (Luo et al., 2024; Gao et al., 2025). Iterative

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

refinement can revisit an already generated pathway in subsequent passes (Lee et al., 2026), but does so after an autoregressive route or subtree has first been produced. This motivates our central question: can postfix order specify how a route is executed without also fixing what context is available for prediction and when its decisions are committed?

RouteDiff answers this question with masked discrete diffusion over the complete postfix route and a bidirectional Transformer denoiser (Austin et al., 2021; Sahoo et al., 2024), while retaining postfix order for route execution. Starting from a fully masked route, the bidirectional Transformer predicts unresolved building-block and reaction decisions from the route context currently visible on both sides. Diffusion progressively reveals and commits these decisions; predictions at still-masked positions may change as new route context becomes visible. Thus, postfix order specifies execution without also prescribing prediction context or commitment timing.

Beyond the route-generation mechanism itself, two practical requirements remain. First, the model must select exact building blocks from a 223,243-entry catalog while retaining useful chemical structure across the catalog. Prior route models have likewise used molecular representations to handle building-block vocabularies at this scale (Gao et al., 2025). We address this with frozen ChemBERTa representations and learned low-rank identity components, combining molecular semantics with exact catalog identity (Chithrananda et al., 2020). Second, molecular design often requires satisfying multiple desired properties simultaneously (Han et al., 2024). RouteDiff conditions its denoiser on molecular descriptors and numerical constraints, and composes supported conditions at inference time without task-specific retraining. Together, these components allow full-route diffusion to operate over a large synthesis vocabulary and support compositional molecular design.

We evaluate RouteDiff on target recovery and synthesis breadth, property optimization, compositional generation, and denoising dynamics. On Enamine reconstruction, RouteDiff reconstructs 92.8% of targets with 0.987 similarity and achieves the highest building-block diversity (0.783) among the compared methods. Under a fixed $10{,}000$-oracle-call budget, it achieves the highest synthesis-constrained AUC Top-10 on eight of ten PMO tasks. Direct compositional generation further shows that multiple supported property conditions can be combined at inference time, achieving feasible-route rates of 37.62%–99.54% and product diversity of 0.909–0.938 across five broad physicochemical regions. Finally, denoising-trajectory analysis finds pre-commit changes in unresolved building-block or reaction predictions in 36 of 48 (75.0%) feasible single-product trajectories in a diagnostic cohort, showing that this behavior recurs within the analyzed successful trajectories.

Our contributions are summarized as follows:

- **Decoupling execution, prediction context, and commitment in synthetic-route generation.** We introduce RouteDiff, which retains postfix serialization for route execution, uses a bidirectional Transformer to predict unresolved decisions from visible context on both sides, and uses masked discrete diffusion to progressively commit those decisions over the complete route.
- **Large-catalog building-block selection and compositional conditioning.** We combine frozen molecular representations with learned low-rank identity components for exact selection from a 223,243-entry building-block catalog, and compose supported molecular property conditions at inference time without task-specific retraining.
- **Strong synthesis-aware generation with multi-level diversity.** We show that RouteDiff combines high-fidelity target recovery and strong fixed-budget optimization with diverse feasible solutions across building-block, product, scaffold, and route levels. Denoising-trajectory analysis further finds recurrent pre-commit changes in unresolved building-block or reaction predictions, providing direct evidence that route hypotheses can remain revisable before commitment.

2 METHOD

2.1 PROBLEM FORMULATION AND ROUTE REPRESENTATION

Let $\mathcal{B}$ denote the catalog of purchasable building blocks and $\mathcal{R}$ the set of reaction templates defining the synthesis space. Following prior synthetic-route models (Luo et al., 2024; Gao et al., 2025), we let $\mathcal{X}$ denote the space of typed postfix candidate sequences over building-block, reaction, and

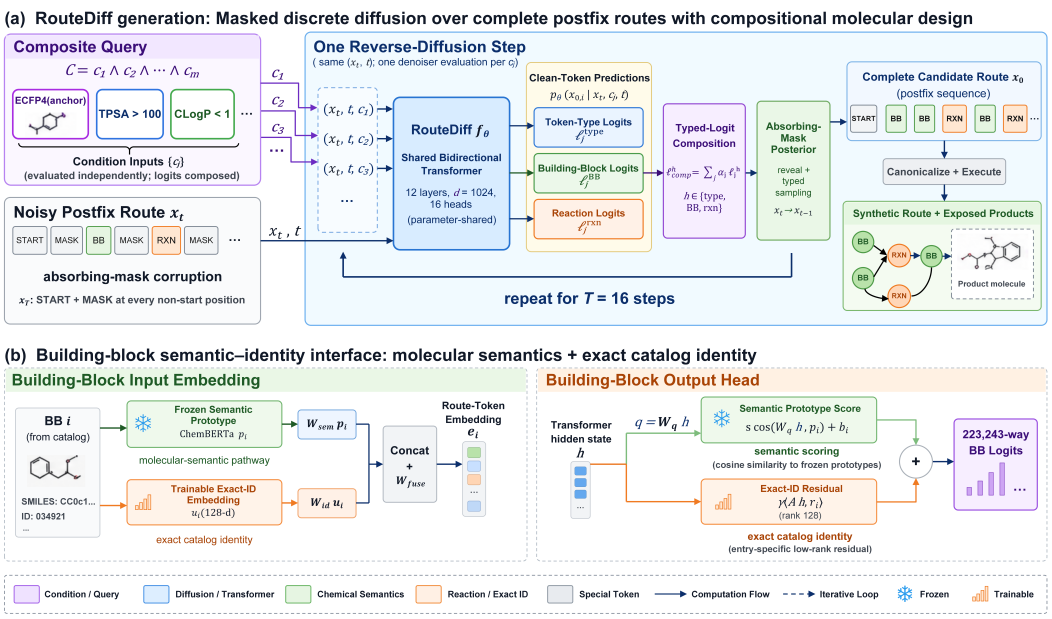

Figure 1: **Overview of RouteDiff.** (a) RouteDiff applies masked discrete diffusion to complete postfix routes using a bidirectional denoiser. For composite queries, condition-wise typed logits are composed before each reverse transition; the final route is executed by $\mathcal{E}$. (b) The building-block interface combines frozen ChemBERTa semantic prototypes with trainable low-rank exact-identity components at input and output for exact selection from the 223,243-entry catalog.

special tokens. A complete candidate route is represented as

$$x_0 = (x_{0,1}, \ldots, x_{0,L}) \in \mathcal{X}. \tag{1}$$

Each position has a route-token type—building block, reaction, or special token—and building-block and reaction positions additionally specify an identity in $\mathcal{B}$ or $\mathcal{R}$. Before stack execution, a fixed-length candidate state is deterministically canonicalized as described in Appendix B.1. The resulting postfix sequence is executed with a stack: building blocks are pushed onto the stack, while reaction tokens consume the required compatible reactants and push the resulting product. We denote this canonicalization-and-execution procedure by $\mathcal{E}$. If execution succeeds, $\mathcal{E}(x_0) = Y \neq \varnothing$ denotes the set of exposed products; otherwise, $\mathcal{E}(x_0) = \perp$. We therefore define the executable route space as

$$\mathcal{X}_{\text{exec}} = \{x \in \mathcal{X} : \mathcal{E}(x) \neq \perp\}. \tag{2}$$

Given a molecular design goal g, let $s_g(Y)$ denote a task-dependent utility measuring how well the exposed products Y satisfy the requested objective. We formulate synthesis-aware molecular design as seeking an executable route whose products maximize this utility:

$$x^\star(g) \in \arg\max_{x \in \mathcal{X}_{\text{exec}}} s_g(\mathcal{E}(x)). \tag{3}$$

A solution is therefore an executable sequence of building-block and reaction decisions whose resulting products satisfy the design objective, rather than only a high-scoring molecular structure.

Postfix notation specifies how a route is executed, but it need not dictate how candidate routes are generated. Building-block and reaction choices are coupled through the intermediates they create: each reaction requires compatible reactants, while downstream reactions constrain which upstream choices can participate in an executable route. We therefore treat the complete route as the generation state rather than forcing postfix execution order to also determine the context available for predicting unresolved decisions or when those decisions become committed. Given a supported molecular condition c, RouteDiff parameterizes $p_\theta(x_0 \mid c)$ over complete candidate sequences $x_0 \in \mathcal{X}$. Samples are executed by $\mathcal{E}$ to determine whether they belong to $\mathcal{X}_{\text{exec}}$. The task-level goal g and model condition c are distinct: supported conditions may express the goal directly or provide conditioning

within an external optimization procedure. Rather than factorizing this distribution in postfix order, the next section instantiates it with bidirectional discrete diffusion.

2.2 Bidirectional Discrete Diffusion over Complete Routes

RouteDiff instantiates $p_\theta(x_0 \mid c)$ with masked discrete diffusion over the complete candidate route. At each denoising step, a bidirectional Transformer predicts clean-token distributions for unresolved positions from the current partially masked route state. Because self-attention is non-causal, visible route context may lie on either side of an unresolved position in postfix order. The reverse diffusion process progressively reveals and commits masked positions, without requiring them to be committed in postfix execution order.

We adopt an absorbing-mask discrete diffusion process (Austin et al., 2021; Sahoo et al., 2024). Given a clean route x_0, the cumulative forward corruption independently masks each non-start position i:

$$q(x_{t,i} \mid x_{0,i}) = (1 - p_t)\delta_{x_{0,i}} + p_t \delta_{\texttt{MASK}}, \qquad p_t = \frac{t}{T}. \tag{4}$$

Here, $t \in \{1, \ldots, T\}$ indexes the diffusion step and T is the total number of steps. The start token is kept fixed, and since $p_T = 1$, every non-start position in x_T is masked, yielding a fixed initialization for reverse generation.

We parameterize the denoiser with a bidirectional Transformer. Given a corrupted route state x_t, supported molecular condition c, and timestep t, it predicts the clean-token distribution at each route position i:

$$p_\theta(x_{0,i} \mid x_t, c, t). \tag{5}$$

Because route tokens are typed, the model realizes this distribution with separate output heads for route-token type, building-block identity, and reaction identity. With non-causal self-attention, the Transformer produces these distributions for all unresolved positions in parallel. Here, bidirectional refers to context within a single diffusion state, not to switching between synthesis-tree traversal orders.

Starting from x_T, RouteDiff applies the reverse absorbing-mask process for $t = T, \ldots, 1$. At step t, each masked position is independently revealed with probability $1 - p_{t-1}/p_t = 1/t$ under the linear schedule; when revealed, its value is sampled from $p_\theta(x_{0,i} \mid x_t, c, t)$, and otherwise it remains masked. Once revealed, a position remains fixed in subsequent steps. The denoiser is reevaluated at every step, so predictions for still-masked positions can change as additional route context becomes visible before those positions are committed. Thus, predictions at still-masked positions are provisional, whereas revealed tokens in x_t constitute the committed sampler state.

2.3 Building-block Semantic-Identity Parameterization

As illustrated in Figure 1(b), RouteDiff parameterizes the 223,243-entry building-block catalog through complementary molecular-semantic and exact-identity pathways at both the input and output interfaces. A purely categorical representation treats catalog entries as unrelated classes, whereas a molecular representation captures chemical similarity without explicitly parameterizing exact catalog identity. We therefore combine frozen ChemBERTa prototypes with trainable low-rank identity components, preserving chemical similarity while retaining precise catalog identity.

For input building block i, let p_i denote its frozen ChemBERTa prototype and $u_i \in \mathbb{R}^{128}$ its trainable catalog-identity embedding. We project the two components to the model dimension and fuse them by concatenation:

$$e_i = W_{\text{fuse}}[W_{\text{id}} u_i; W_{\text{sem}} p_i]. \tag{6}$$

The ChemBERTa prototype table is fixed during training, while the identity embeddings and projection and fusion layers are learned.

For output prediction, we project the denoiser hidden state h into the ChemBERTa prototype space, $q = W_q h$. The semantic logit for catalog entry i is

$$\ell_i^{\text{sem}} = s \cos(q, p_i) + b_i, \tag{7}$$

where s is a learned logit scale and b_i is an entry-specific bias. We augment this prototype score with a low-rank exact-identity residual:

$$\ell_i = \ell_i^{\text{sem}} + \gamma \langle Ah, r_i \rangle, \tag{8}$$

where $r_i \in \mathbb{R}^{128}$ is a separate trainable residual for catalog entry i and γ is a learned residual scale. The building-block distribution is obtained by applying a softmax over $\{\ell_i\}_{i \in \mathcal{B}}$.

2.4 Constraint-Conditioned Route Generation

Figure 1(a) summarizes RouteDiff's compositional conditioning mechanism. Condition inputs c_j are constructed from the supported condition families, including ECFP4/FCFP4 vectors and scalar exact-value, upper-bound, and lower-bound specifications. A condition input may encode one or more descriptor requirements. For a composite query $C = \bigwedge_{j=1}^{m} c_j$, each condition input is evaluated independently by the same parameter-shared denoiser on the shared noisy route state (x_t, t):

$$\ell_j^h(x_t, t) = f_\theta^h(x_t, c_j, t), \qquad h \in \{\text{type}, \text{BB}, \text{rxn}\}. \tag{9}$$

Rather than combining condition embeddings before denoising, RouteDiff composes the resulting typed logits. All composite queries reported in this work use positive condition weights $\alpha_j > 0$, and their typed logits are composed as

$$\ell_{\text{comp}}^h(x_t, t) = \sum_{j=1}^{m} \alpha_j \ell_j^h(x_t, t), \qquad h \in \{\text{type}, \text{BB}, \text{rxn}\}, \tag{10}$$

where α_j weights condition input c_j. The composed logits define the clean-token distributions used by the reverse absorbing-mask transition. Because all condition inputs are evaluated by the same parameter-shared denoiser, supported conditions can be recombined at inference without training a separate generator for each composite query.

2.5 Training Objective

For each training route–condition pair (x_0, c) with $x_0 \in \mathcal{X}_{\text{exec}}$, we sample $t \sim \text{Uniform}\{1, \ldots, T\}$ and obtain x_t using the forward corruption process in Section 2.2. The denoiser is trained to recover the clean route at positions corrupted by this process, using the typed objective

$$\mathcal{L} = \lambda_{\text{type}} \mathcal{L}_{\text{type}} + \lambda_{\text{BB}} \mathcal{L}_{\text{BB}} + \lambda_{\text{rxn}} \mathcal{L}_{\text{rxn}}, \tag{11}$$

with equal loss weights in our implementation. The building-block loss is evaluated only when the clean target is a building-block token, and the reaction loss only when it is a reaction token. Each loss is normalized within each training sample before averaging across the batch, preventing diffusion steps with more corrupted positions from receiving greater weight solely due to their corruption level.

3 Experiments

RouteDiff is trained on synthetic routes generated online from a fixed synthesis space containing 223,243 Enamine building blocks and 115 reaction templates; all evaluations freeze this training-time synthesis vocabulary and use the checkpoint with the lowest validation loss observed during training. Experiment 1 draws 256 route samples per target for each of three seeds (2026, 2027, and 2028); Experiment 2 reports three independent optimization runs under a fixed 10,000-oracle-call budget; and Experiment 3 reports five-seed statistics. Published baselines follow the synthesis spaces and search procedures of their source benchmarks; task-specific protocols and metrics are introduced in the corresponding experiments.

3.1 Experiment 1: Synthesizable Molecule Reconstruction

Setup. We first ask whether RouteDiff can recover target molecules through executable routes while also sampling alternative synthesis choices. Following the SynFormer/ReaSyn benchmark

Table 1: Synthesizable molecule reconstruction results. Baseline values are the three-run mean±SD results reported in ReaSyn v2 Table 1 (Lee et al., 2026); RouteDiff reports mean±SD over three fixed sampling seeds with 256 samples per target. Best results are bold.

Dataset	Method	Recon. (%) ↑	Similarity ↑	Div. (Pathway) ↑	Div. (BB) ↑
Enamine	SynNet (Gao et al., 2022b)	25.2±0.1	.661±.000	.014±.001	.239±.002
	SynFormer (Gao et al., 2025)	66.3±0.6	.913±.001	.101±.001	.587±.002
	ReaSyn (Lee et al., 2026)	**95.0±0.0**	**.987±.001**	**.118±.002**	.753±.004
	RouteDiff	92.8±0.3	**.987±.001**	.113±.001	**.783±.003**
ChEMBL	SynNet (Gao et al., 2022b)	7.9±0.0	.542±.000	.009±.000	.090±.001
	SynFormer (Gao et al., 2025)	19.7±0.4	.668±.002	.039±.000	.192±.002
	ReaSyn (Lee et al., 2026)	**31.7±0.3**	**.751±.001**	**.050±.000**	**.321±.002**
	RouteDiff	21.0±0.2	.733±.002	.040±.001	.291±.003

(Gao et al., 2025; Lee et al., 2026), we evaluate 1,000 targets from the Enamine REAL Diversity Set (Enamine Ltd., 2023) and 1,000 targets from ChEMBL (Gaulton et al., 2012). Given a target molecule, RouteDiff samples complete route candidates through its denoising process and executes them with $\mathcal{E}$; the resulting executable routes are used to assess both target fidelity and sampling breadth.

Metrics. Following ReaSyn v2 (Lee et al., 2026), we use reconstruction rate and similarity for target fidelity, and diversity (pathway) and diversity (BB) for sampling breadth. Reconstruction rate is the fraction of targets recovered exactly by at least one sampled route, while similarity averages, across targets, the highest Morgan-fingerprint Tanimoto similarity (Rogers & Hahn, 2010) reached by the sampled products. Among analogous routes whose products have similarity at least 0.8, diversity (pathway) measures molecular diversity among their end products, and diversity (BB) measures molecular diversity among the building blocks used by those routes. Exact metric definitions are provided in Appendix F.

Results. On Enamine (Table 1), RouteDiff combines a $92.8\pm0.3\%$ reconstruction rate and best-reported 0.987 ± 0.001 similarity with the highest building-block diversity among the compared methods (0.783 ± 0.003). On ChEMBL, where targets can fall outside the fixed synthesis space, RouteDiff reaches $21.0\pm0.2\%$ reconstruction and 0.733 ± 0.002 similarity, with building-block diversity of 0.291 ± 0.003. Together, these results show that RouteDiff combines target fidelity with sampling breadth over alternative synthesis choices within a fixed reaction-defined synthesis space.

Ablation study. We use the Enamine reconstruction task to examine the building-block parameterization and generation formulation, with all variants trained and evaluated under the same protocol. Removing either the low-rank parameterization or the ChemBERTa prototype pathway reduces reconstruction rate and similarity, with the joint ablation producing the largest drop, while both diversity metrics remain close to the full model (Table 2). The matched autoregressive counterpart retains comparable target fidelity and pathway diversity but reduces building-block diversity from 0.783 ± 0.003 to 0.756 ± 0.005. Together, these results show that the semantic–identity parameterization primarily supports target recovery, while the full-route bidirectional-diffusion formulation maintains comparable fidelity and pathway-level breadth while sampling more diverse building-block choices than its autoregressive counterpart.

3.2 Experiment 2: Goal-Directed Molecular Optimization

Setup. Beyond reconstruction, we evaluate RouteDiff on goal-directed molecular optimization under a limited oracle budget. We use 10 PMO tasks reported in the synthesis-constrained ReaSyn evaluation (Gao et al., 2022a; Lee et al., 2026), with a fixed budget of 10,000 oracle calls per task.

Metrics. We evaluate optimization performance using AUC Top-10, following the PMO evaluation protocol (Gao et al., 2022a). At each oracle-call budget, Top-10 is the mean objective score of the ten highest-scoring molecules discovered so far, and AUC Top-10 measures the area under this

Table 2: **Ablation study on Enamine reconstruction.** Mean±SD over three fixed sampling seeds with 256 samples per target. All variants are reported on the same four metrics under the matched training and evaluation setup described in Appendix B.6.

Variant	Recon. (%) ↑	Similarity ↑	Div. (Pathway) ↑	Div. (BB) ↑
Full RouteDiff	92.8±0.3	.987±.001	.113±.001	.783±.003
w/o low-rank parameterization	89.9±0.4	.971±.001	.109±.002	.770±.004
w/o ChemBERTa prototypes	91.2±0.3	.979±.001	.113±.001	.772±.003
w/o both components	88.1±0.5	.964±.002	.112±.001	.766±.004
Autoregressive counterpart	92.6±0.2	.982±.003	.113±.002	.756±.005

Table 3: **Goal-directed molecular optimization on 10 PMO tasks.** AUC Top-10 under a budget of 10,000 oracle calls. Baseline results are from ReaSyn v2 Table 3 (Lee et al., 2026); RouteDiff reports the mean over three runs. Best results are bold.

Objective	SynthesisNet	Graph GA–SF	Graph GA–ReaSyn	RouteDiff
Amlodipine MPO	.608	.696	.620	**.786**
Celecoxib rediscovery	.582	.559	**.810**	.780
Fexofenadine MPO	.791	.786	.788	**.812**
Median1	.305	.308	.274	**.365**
Median2	.257	.258	.259	**.295**
Osimertinib MPO	.810	.816	.823	**.827**
Perindopril MPO	.524	.530	.561	**.693**
Ranolazine MPO	.741	.751	.752	**.764**
Sitagliptin MPO	.313	.338	.314	**.423**
Zaleplon MPO	**.528**	.478	.460	.524

trajectory over the full 10,000-call budget. Higher AUC Top-10 therefore reflects both the quality of discovered molecules and how efficiently they are identified within the oracle budget.

Results. RouteDiff achieves the best synthesis-constrained AUC Top-10 on eight of ten PMO tasks (Table 3). The largest gains over the strongest synthesis-constrained baseline are $+0.132$ on Perindopril, $+0.090$ on Amlodipine, and $+0.085$ on Sitagliptin. Overall, RouteDiff shows strong fixed-budget, synthesis-constrained optimization across diverse objectives.

3.3 Experiment 3: Compositional Generation, Diversity, and Denoising Dynamics

Setup. Experiment 2 evaluates RouteDiff through an external optimization loop; we next test whether its inference-time compositional conditioning can directly satisfy multiple molecular requirements without task-specific retraining. Beyond constraint satisfaction, we ask what diversity remains among the resulting feasible solutions. Using the frozen model trained with all supported condition families, we evaluate nine composite queries comprising five physicochemical regions and four drug-anchored objectives. The exact query definitions are provided in Appendix D.

Metrics. We evaluate each composite query using task score, feasibility, and diversity. Best task score records the highest query-specific score reached in the sampled outputs. Feasible-route rate is the fraction of attempted routes that yield at least one product satisfying all explicit hard constraints, and we additionally report the number of unique feasible products. Product diversity is the mean pairwise ECFP4 Tanimoto dissimilarity among unique feasible products. For drug-anchored queries, fingerprint similarity contributes to the task score, while feasibility is determined by the explicit scalar constraints. Exact metric definitions and additional diagnostics are provided in Appendix E.

Results. Across the five physicochemical regions, RouteDiff achieves feasible-route rates of 37.62%–99.54% while maintaining product diversity of 0.909–0.938, showing that satisfying these broad composed constraints does not collapse feasible outputs to a narrow fingerprint-level region. The four drug-anchored queries define substantially narrower target regions: best task scores range

Table 4: **Direct compositional conditioning across nine composite queries.** Mean ± sample standard deviation over five sampling runs with 1,000 attempted routes per query and run. Valid routes execute successfully; feasible routes yield at least one product satisfying all explicit hard constraints. A route may yield multiple products, so unique feasible-product counts can exceed 1,000.

Task	Best task score ↑	Valid routes (%) ↑	Feasible routes (%) ↑	Unique feasible products ↑	Product div. ↑
Physicochemical regions					
Lipinski+Veber region	1.000 ± 0.000	99.94 ± 0.05	98.82 ± 0.22	739.2 ± 10.3	0.936 ± 0.001
Lipinski rule-of-five region	1.000 ± 0.000	100.00 ± 0.00	99.54 ± 0.24	792.6 ± 13.6	0.938 ± 0.001
Veber	1.000 ± 0.000	95.30 ± 0.48	92.94 ± 0.61	1082.0 ± 23.7	0.909 ± 0.001
Congreve Rule-of-3	1.000 ± 0.000	100.00 ± 0.00	76.60 ± 0.79	571.2 ± 10.3	0.937 ± 0.001
Ghose	1.000 ± 0.000	96.98 ± 0.81	37.62 ± 1.11	326.4 ± 30.3	0.914 ± 0.002
Drug-anchored objectives					
Cobimetinib	0.913 ± 0.012	91.56 ± 0.56	1.78 ± 0.38	35.8 ± 10.8	0.800 ± 0.022
Osimertinib	0.762 ± 0.031	95.74 ± 0.64	0.42 ± 0.24	5.6 ± 3.0	$0.834 \pm 0.039^{*}$
Amlodipine	0.827 ± 0.048	98.42 ± 0.43	0.16 ± 0.09	2.2 ± 0.8	$0.463 \pm 0.098^{*}$
Perindopril	0.523 ± 0.217	99.80 ± 0.07	0.06 ± 0.05	0.8 ± 0.8	$0.230^{\dagger}$

Product diversity is averaged over runs containing at least two unique feasible products: *4/5 runs and †1/5 run; all other rows use 5/5 runs.

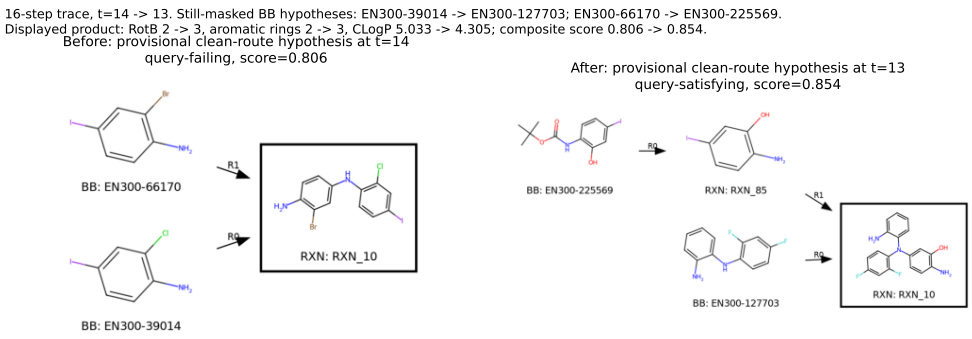

Figure 2: **Pre-commit revision of a provisional route hypothesis.** In Cobimetinib medoid 3, two still-masked building-block predictions change from $t = 14$ to $t = 13$, moving the provisional route from infeasible to feasible. Panels show argmax $\hat{x}_0$ predictions, not committed sampler states.

from 0.523 to 0.913, while feasible-route rates range from 0.06% to 1.78%. As Table 4 shows, route validity nevertheless remains high at 91.56%–99.80%, showing that the main difficulty in these queries is satisfying the composed molecular requirements rather than generating executable routes.

When the generated object is an executable synthetic route, diversity is not a single property: feasible solutions may differ in products, scaffolds, building blocks, or reaction programs. The Lipinski rule-of-five and Veber queries have similar single-product ECFP4 diversity (0.938 vs. 0.911), yet their effective-scaffold fractions are 0.020 vs. 0.304 and reaction-skeleton uniqueness is 0.009 vs. 0.264 (Appendix Table 6). Lipinski's 0.806 exact-route uniqueness despite 0.009 reaction-skeleton uniqueness indicates substantial building-block substitution within shared reaction programs, whereas Veber shows broader scaffold and reaction-program variation. Restricting the physicochemical tasks to single-product routes changes ECFP4 diversity by at most 0.007.

Route-hypothesis revision during sampling. The preceding analyses characterize diversity across final samples; we next examine pre-commit revision in six feasible single-product medoids selected before trajectory inspection (Appendix E). In Cobimetinib medoid 3 (Figure 2), two still-masked building-block predictions change between $t = 14$ and $t = 13$, moving the provisional route from infeasible to feasible under all three hard constraints. Because neither position has yet been revealed, the transition illustrates RouteDiff's separation of prediction and commitment: unresolved route hypotheses can change before diffusion commits them.

To test recurrence beyond the displayed example, we replay an independently selected cohort of 48 feasible single-product routes (32 Ghose and 16 Cobimetinib) chosen before trajectory inspection. Under the stricter final-token-aligned criterion, 36/48 (75.0%) trajectories contain a still-masked building-block or reaction-token prediction that changes toward the final emitted identity. Thus, pre-commit semantic revision recurs across the success-conditioned cohort.

4 Discussion and Limitations

Sequence-based synthetic-route models often use postfix order both to execute a route and to determine the prediction context and commitment order of autoregressive generation (Luo et al., 2024; Gao et al., 2025); ReaSyn can subsequently revisit generated pathways through explicit refinement (Lee et al., 2026). RouteDiff suggests separating three roles that are often coupled in sequence generation: serialization determines how a structured object is executed, bidirectional denoising determines what context informs unresolved decisions, and diffusion determines when those decisions become committed. Experiment 1 provides a matched formulation comparison in which the autoregressive counterpart retains comparable target fidelity and pathway diversity but shows lower building-block diversity; across final samples, Experiments 1 and 3 show diversity at building-block, product, scaffold, and reaction-program levels; and within trajectories, Experiment 3 finds final-token-aligned pre-commit revisions in 36/48 (75.0%) success-conditioned trajectories. Together, these observations motivate a broader question for structured discrete generation: should serialization order also determine when local decisions become irreversible? RouteDiff provides one concrete answer by retaining postfix execution semantics while decoupling them from prediction context and commitment timing.

Two limitations define the present scope. First, RouteDiff generates within a fixed reaction-defined synthesis space, so synthesizability is operationally defined by successful execution under the specified building-block catalog, reaction templates, and executor. Extending full-route diffusion to broader or dynamically changing synthesis spaces is an important next step. Second, inference-time composition is currently demonstrated for the supported molecular descriptors and numerical predicates. Extending the same composition mechanism to richer learned objectives and additional condition families would broaden the range of design problems that can be expressed directly within the generator.

5 Conclusion

We introduced RouteDiff for synthesis-aware molecular generation, where the generated object is an executable synthetic route rather than a molecular structure alone. RouteDiff separates three roles that are coupled in autoregressive postfix generation: postfix serialization specifies route execution, a bidirectional Transformer uses visible context on both sides to predict unresolved decisions, and masked discrete diffusion determines when those decisions become committed. Empirically, RouteDiff combines high-fidelity target recovery and strong fixed-budget optimization with diverse feasible solutions across building-block, product, scaffold, and reaction-program levels. Denoising-trajectory analysis further finds final-token-aligned pre-commit revisions in 36/48 (75.0%) success-conditioned trajectories, showing that unresolved route hypotheses can remain revisable before commitment. More broadly, these results suggest that in structured discrete generation, serialization can specify how an object is executed without necessarily prescribing the context or timing with which its decisions become irreversible.

AI use statement

In this work, we used generative AI tools to assist with interpretation of reported results and to provide feedback on experimental presentation. Additionally, we used generative AI tools to draft and edit portions of the manuscript for readability, polish writing, review and refine wording. The authors independently verified all AI-assisted text, analyses, claims, citations, and artifacts against the underlying experiments and sources. The authors take responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This work develops generative methods for small-molecule design and therefore has potential dual-use implications. Generated molecules and synthetic routes are computational proposals within a predefined reaction space and do not constitute experimental validation of synthesis, safety, or efficacy. The experiments use standard molecular-design benchmarks and do not introduce objectives targeting controlled substances or chemical warfare agents. Experimental use of generated candidates requires appropriate chemical-safety, regulatory, and domain-expert review.

REPRODUCIBILITY STATEMENT

We provide the model formulation, route representation, diffusion process, building-block parameterization, conditioning mechanism, training-data generation, training objective, and optimization configuration in Section 2 and Appendices B–C. The experimental protocols, datasets, sampling budgets, composite-query definitions, and evaluation metrics for all three experiments are specified in Section 3 and Appendices D–F. Appendix E additionally documents the multi-scale diversity and denoising-trajectory analyses, including the medoid and independent-cohort selection procedures.

REFERENCES

Jacob Austin, Daniel D. Johnson, Jonathan Ho, Daniel Tarlow, and Rianne van den Berg. Structured denoising diffusion models in discrete state-spaces. In *Advances in Neural Information Processing Systems*, volume 34, 2021.

Guy W. Bemis and Mark A. Murcko. The properties of known drugs. 1. molecular frameworks. *Journal of Medicinal Chemistry*, 39(15):2887–2893, 1996. doi: 10.1021/jm9602928.

Seyone Chithrananda, Gabriel Grand, and Bharath Ramsundar. ChemBERTa: Large-scale self-supervised pretraining for molecular property prediction. *arXiv preprint arXiv:2010.09885*, 2020.

Miles Congreve, Robin Carr, Chris Murray, and Harren Jhoti. A "rule of three" for fragment-based lead discovery? *Drug Discovery Today*, 8(19):876–877, 2003. doi: 10.1016/S1359-6446(03)02831-9.

Miruna Cretu, Charles Harris, Ilia Igashov, Arne Schneuing, Marwin Segler, Bruno Correia, Julien Roy, Emmanuel Bengio, and Pietro Liò. SynFlowNet: Design of diverse and novel molecules with synthesis constraints. In *International Conference on Learning Representations*, 2025.

Jörg Degen, Christof Wegscheid-Gerlach, Andrea Zaliani, and Matthias Rarey. On the art of compiling and using "Drug-Like" chemical fragment spaces. *ChemMedChem*, 3(10):1503–1507, 2008. doi: 10.1002/cmdc.200800178.

Enamine Ltd. REAL Database, 2023. URL `https://enamine.net/compound-collections/real-compounds/real-database`.

Wenhao Gao and Connor W. Coley. The synthesizability of molecules proposed by generative models. *Journal of Chemical Information and Modeling*, 60(12):5714–5723, 2020. doi: 10.1021/acs.jcim.0c00174.

Wenhao Gao, Tianfan Fu, Jimeng Sun, and Connor W. Coley. Sample efficiency matters: A benchmark for practical molecular optimization. In *Advances in Neural Information Processing Systems*, volume 35, pp. 21342–21357, 2022a.

Wenhao Gao, Rocío Mercado, and Connor W. Coley. Amortized tree generation for bottom-up synthesis planning and synthesizable molecular design. In *International Conference on Learning Representations*, 2022b.

Wenhao Gao, Shitong Luo, and Connor W. Coley. Generative AI for navigating synthesizable chemical space. *Proceedings of the National Academy of Sciences*, 122(41):e2415665122, 2025. doi: 10.1073/pnas.2415665122.

Anna Gaulton, Louisa J. Bellis, A. Patricia Bento, Jon Chambers, Mark Davies, Anne Hersey, Yvonne Light, Shaun McGlinchey, David Michalovich, Bissan Al-Lazikani, and John P. Overington. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Research*, 40(D1):D1100–D1107, 2012. doi: 10.1093/nar/gkr777.

Arup K. Ghose, Vellarkad N. Viswanadhan, and John J. Wendoloski. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. a qualitative and quantitative characterization of known drug databases. *Journal of Combinatorial Chemistry*, 1(1):55–68, 1999.

Xu Han, Caihua Shan, Yifei Shen, Can Xu, Han Yang, Xiang Li, and Dongsheng Li. Training-free multi-objective diffusion model for 3D molecule generation. In *The Twelfth International Conference on Learning Representations*, 2024.

Seul Lee, Karsten Kreis, Srimukh Prasad Veccham, Meng Liu, Danny Reidenbach, Yuxing Peng, Saee Gopal Paliwal, Weili Nie, and Arash Vahdat. GenMol: A drug discovery generalist with discrete diffusion. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 33205–33226. PMLR, 2025.

Seul Lee, Karsten Kreis, Srimukh Prasad Veccham, Meng Liu, Danny Reidenbach, Saee Gopal Paliwal, Weili Nie, and Arash Vahdat. Exploring synthesizable chemical space with iterative pathway refinements. In *International Conference on Learning Representations*, 2026.

Christopher A. Lipinski, Franco Lombardo, Beryl W. Dominy, and Paul J. Feeney. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, 23(1–3):3–25, 1997. doi: 10.1016/S0169-409X(96)00423-1.

Sulin Liu, Juno Nam, Andrew Campbell, Hannes Stärk, Yilun Xu, Tommi S. Jaakkola, and Rafael Gómez-Bombarelli. Think while you generate: Discrete diffusion with planned denoising. In *International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=MJNywBdSDy.

Shitong Luo, Wenhao Gao, Zuofan Wu, Jian Peng, Connor W. Coley, and Jianzhu Ma. Projecting molecules into synthesizable chemical spaces. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 33289–33304. PMLR, 2024.

David Rogers and Mathew Hahn. Extended-connectivity fingerprints. *Journal of Chemical Information and Modeling*, 50(5):742–754, 2010. doi: 10.1021/ci100050t.

Subham Sekhar Sahoo, Marianne Arriola, Yair Schiff, Aaron Gokaslan, Edgar Marroquin, Justin T. Chiu, Alexander Rush, and Volodymyr Kuleshov. Simple and effective masked diffusion language models. In *Advances in Neural Information Processing Systems*, 2024.

Seonghwan Seo, Minsu Kim, Tony Shen, Martin Ester, Jinkyoo Park, Sungsoo Ahn, and Woo Youn Kim. Generative flows on synthetic pathway for drug design. In *International Conference on Learning Representations*, 2025.

Michael Sun, Alston Lo, Minghao Guo, Jie Chen, Connor W. Coley, and Wojciech Matusik. Procedural synthesis of synthesizable molecules. In *International Conference on Learning Representations*, 2025.

Daniel F. Veber, Stephen R. Johnson, Hung-Yuan Cheng, Brian R. Smith, Keith W. Ward, and Kenneth D. Kopple. Molecular properties that influence the oral bioavailability of drug candidates. *Journal of Medicinal Chemistry*, 45(12):2615–2623, 2002.

A Related Work

Synthesis-aware molecular generation. Reaction-based molecular generators restrict generation to synthesis spaces defined by purchasable building blocks and predefined reaction rules. SynNet models bottom-up synthetic-tree construction one reaction step at a time (Gao et al., 2022b). SynFlowNet and RxnFlow formulate molecular construction as sequential reaction decisions trained with GFlowNet objectives; RxnFlow additionally uses action-space subsampling to operate over very large building-block and reaction action spaces (Cretu et al., 2025; Seo et al., 2025). Procedural Synthesis decouples the structure of a synthetic program from its molecular choices and optimizes the two levels separately (Sun et al., 2025). These methods differ in how they represent synthetic pathways and organize the decisions used to construct them. RouteDiff is closest to the line of work that represents complete synthetic pathways as sequences, which we discuss next.

Synthetic-route sequence modeling and refinement. ChemProjector introduces postfix pathway sequences that linearize synthesis trees into stack-executable representations and autoregressively decodes them with a Transformer conditioned on an input molecular graph (Luo et al., 2024). SynFormer adopts the same postfix representation for autoregressive route generation, while using diffusion specifically to generate building-block fingerprints for catalog selection (Gao et al., 2025). ReaSyn extends this line with a unified autoregressive model for bottom-up route generation and top-down subtree re-prediction, followed by an Edit Bridge that refines generated pathways through insertion, deletion, and substitution (Lee et al., 2026). ChemProjector and SynFormer therefore use postfix order both for route execution and for left-to-right prediction and commitment, whereas ReaSyn can subsequently revisit an autoregressively generated pathway through top-down and holistic refinement. RouteDiff instead retains postfix serialization for route execution while treating the complete route as the masked-diffusion state: bidirectional denoising predicts unresolved positions from visible context on both sides, and diffusion determines when those positions are revealed and committed.

Masked discrete diffusion for sequence generation. D3PM introduced diffusion processes directly over discrete state spaces, including absorbing-state transitions that correspond to masking (Austin et al., 2021). MDLM develops this masking formulation for sequence generation, training encoder-only denoisers with masked-token objectives and generating through iterative unmasking (Sahoo et al., 2024). DDPD further separates planning from denoising, using a learned planner to adaptively select positions for denoising and refinement (Liu et al., 2025). In molecular generation, GenMol applies discrete diffusion to SAFE sequences with non-autoregressive bidirectional parallel decoding (Lee et al., 2025). RouteDiff builds on masked discrete diffusion but changes the generated object from a molecular representation to a complete postfix route candidate, jointly specifying building-block identities and reaction choices. Unlike DDPD, RouteDiff does not learn an explicit planner: under its absorbing-mask process, still-masked positions may change their predicted identities across denoising steps, whereas revealed tokens remain fixed.

B Additional Method and Implementation Details

B.1 Route Serialization and Execution

RouteDiff represents each route as a fixed-length tensor with maximum length $L = 16$. Each training route is serialized as a clean state x_0 that begins with a `START` token, contains the postfix sequence of building-block and reaction tokens, terminates with `END`, and uses `PAD` for the remaining positions. Building-block and reaction positions additionally carry their exact catalog and reaction indices, respectively, while special-token positions carry no identity. The `MASK` token is introduced only by the diffusion process and is not part of a clean training state. `PAD` is treated as a clean x_0 target and can therefore be corrupted and predicted during denoising.

Operationally, $\mathcal{E}$ first canonicalizes the sampled fixed-length state after reverse diffusion. The first predicted `END` or `PAD` position terminates the route; that position is represented as `END`, and all subsequent positions are set to `PAD`. Identity indices are retained only at positions of the corresponding building-block or reaction type. The resulting postfix sequence is then evaluated by the stack executor. A successful execution may expose one or more products, yielding $\mathcal{E}(x_0) = Y \neq \varnothing$; how such

products are selected or aggregated for a particular benchmark is specified by the corresponding evaluation protocol.

B.2 Absorbing-Mask Diffusion and Typed Sampling

For completeness, the cumulative corruption schedule in Section 2.2 corresponds to an absorbing one-step transition. Let $p_0 = 0$ and $p_t = t/T$ with $T = 16$. For a currently unmasked position, the one-step masking probability is

$$\beta_t = \frac{p_t - p_{t-1}}{1 - p_{t-1}} = \frac{1}{T - t + 1}, \tag{12}$$

while `MASK` is absorbing under the forward process. Conditioned on a currently masked position and its clean value, the exact posterior is

$$q(x_{t-1,i} = \texttt{MASK} \mid x_{t,i} = \texttt{MASK}, x_{0,i}) = \frac{p_{t-1}}{p_t}, \tag{13}$$

and

$$q(x_{t-1,i} = x_{0,i} \mid x_{t,i} = \texttt{MASK}, x_{0,i}) = 1 - \frac{p_{t-1}}{p_t} = \frac{1}{t}. \tag{14}$$

The reverse sampler substitutes the denoiser distribution over $x_{0,i}$ for the unknown clean value in this posterior.

At each reverse step, the denoiser produces token-type, building-block, and reaction logits for every route position. RouteDiff first samples the clean token type. If the sampled type is a building block, its catalog identity is then sampled from the building-block distribution; if it is a reaction, its reaction identity is sampled from the reaction distribution. Special tokens carry no building-block or reaction identity. Thus, the three output heads implement the clean-route distribution in Section 2.2 through typed sampling, rather than through a single flat vocabulary over all route symbols and catalog entries.

At inference, RouteDiff initializes x_T as `START` followed by `MASK` at every non-start position and applies the stochastic ancestral transition for $t = T, \ldots, 1$. `START` and `MASK` are excluded from the clean-token sampling support at generated positions, the first post-`START` position is constrained to be a building block, and `PAD` remains a valid clean-token prediction. At step t, every currently masked position is independently revealed with probability $1/t$. A revealed position receives its sampled type and, when applicable, the corresponding building-block or reaction identity; positions not selected for revelation remain masked. Revealed positions remain fixed in all subsequent steps, and at $t = 1$ every remaining masked position is revealed. Route executability is then evaluated by $\mathcal{E}$, which applies the canonicalization in Appendix B.1 before stack execution.

B.3 Building-Block Parameterization Details

Let $N_{\mathrm{BB}} = 223{,}243$ denote the size of the Enamine building-block catalog used by RouteDiff. The model uses a frozen ChemBERTa prototype matrix

$$P \in \mathbb{R}^{N_{\mathrm{BB}} \times 768} \tag{15}$$

together with separate trainable input and output identity tables

$$U \in \mathbb{R}^{N_{\mathrm{BB}} \times 128}, \qquad R \in \mathbb{R}^{N_{\mathrm{BB}} \times 128}. \tag{16}$$

The rows of P, U, and R are aligned by exact building-block catalog index. The same frozen prototype table supplies the semantic vectors at the input and output interfaces, whereas U and R are independently learned and do not share parameters. The prototypes are 768-dimensional representations from DeepChem/ChemBERTa-100M-MLM.

At the input interface, the 768-dimensional semantic vector and the 128-dimensional identity vector are independently projected to the $d = 1024$ Transformer dimension. The two projected vectors are concatenated and mapped back to 1024 dimensions by the learned fusion projection. The ChemBERTa table remains fixed throughout training, while the identity table and all projection and fusion parameters are optimized with the rest of the model.

At the output interface, a denoiser state $h \in \mathbb{R}^{1024}$ is projected to a 768-dimensional query. The query and all prototype vectors are ℓ_2-normalized before their dot products are computed, yielding the cosine-similarity term in Section 2.3. The exact-identity residual separately projects h to 128 dimensions and scores it against the corresponding row of R. The input identity vectors and output identity vectors are initialized from $\mathcal{N}(0, 0.02^2)$. The semantic logit scale is initialized to 10, the identity-residual scale to 0.1, and the entry-specific prototype biases are initialized to zero. The resulting logits are normalized over all N_{BB} entries, so building-block sampling selects an exact catalog identity directly.

B.4 Training Data Generation and Objective

Training examples are generated online from a fixed reaction-defined synthesis space rather than drawn from a finite corpus of stored routes. We construct this space from an October 2023 Enamine building-block collection and the 115 reaction templates used by SynFormer (Gao et al., 2025); after preprocessing, the serialized synthesis space used throughout training and evaluation contains 223,243 building blocks and 115 reaction templates. Each generated synthesis row provides a clean postfix route x_0, together with molecular descriptors evaluated for the corresponding generated product. The benchmark target lists used in Experiment 1 are not used to construct the training stream. One supported condition representation c is assigned to each training row; the condition families and their training mixture are detailed in Appendix C. For each route–condition pair (x_0, c), we sample $t \sim \mathrm{Uniform}\{1, \ldots, T\}$ and obtain x_t using the absorbing-mask forward process in Section 2.2.

For training sample n, let

$$M_n = \{i : x_{t,n,i} = \texttt{MASK}\} \tag{17}$$

denote the positions corrupted by the sampled forward process, and let $\tau_{n,i}$ denote the clean token type. The target sets for the three output heads are

$$M_n^{\mathrm{type}} = M_n, \qquad M_n^{\mathrm{BB}} = \{i \in M_n : \tau_{n,i} = \mathrm{BB}\}, \qquad M_n^{\mathrm{rxn}} = \{i \in M_n : \tau_{n,i} = \mathrm{RXN}\}. \tag{18}$$

Thus, building-block and reaction identities are supervised only when the corresponding clean token type is present at a corrupted position.

Let $\mathrm{CE}_{n,i}^h$ denote the cross-entropy loss for head $h \in \{\mathrm{type}, \mathrm{BB}, \mathrm{rxn}\}$ at position i. For the token-type head, we use

$$w_{n,i}^{\mathrm{type}} = \begin{cases} 0.3, & \tau_{n,i} = \mathrm{PAD}, \\ 1, & \text{otherwise}, \end{cases} \tag{19}$$

while $w_{n,i}^{\mathrm{BB}} = w_{n,i}^{\mathrm{rxn}} = 1$. Each head is first normalized within each sample and then averaged over samples containing at least one target for that head:

$$\mathcal{L}_h = \frac{1}{|\mathcal{N}_h|} \sum_{n \in \mathcal{N}_h} \frac{\sum_{i \in M_n^h} w_{n,i}^h \, \mathrm{CE}_{n,i}^h}{\sum_{i \in M_n^h} w_{n,i}^h}, \qquad \mathcal{N}_h = \{n : M_n^h \neq \varnothing\}. \tag{20}$$

The final objective is

$$\mathcal{L} = \mathcal{L}_{\mathrm{type}} + \mathcal{L}_{\mathrm{BB}} + \mathcal{L}_{\mathrm{rxn}}, \tag{21}$$

corresponding to unit weight for all three heads. Per-sample normalization prevents examples with more corrupted positions from receiving greater weight solely because of their diffusion timestep, while the reduced PAD weight applies only to token-type prediction.

B.5 Model and Optimization Configuration

The denoiser uses 12 bidirectional Transformer layers with hidden dimension 1024, 16 attention heads, and feed-forward dimension 2048. Route positions use learned positional embeddings, and the Transformer uses ReLU activations and dropout 0.1. The diffusion timestep is represented by a learned 1024-dimensional embedding that is added to every route-position representation before the Transformer layers.

We optimize RouteDiff with AdamW using an initial learning rate of 3×10^{-4}. The learning rate follows cosine annealing with $T_{\max} = 500$ and minimum learning rate 10^{-5}, with the scheduler

updated once per epoch. Training uses one NVIDIA H800 GPU in FP32 precision, with a micro-batch of $1{,}024$ routes and gradient accumulation over two micro-batches, giving an effective batch size of $2{,}048$. We train for $640{,}000$ optimizer updates (320 epochs), processing approximately 1.31 billion online-generated route examples, with random seed 2026 and gradient clipping at 100. Training takes approximately 220 hours. Validation is performed once per epoch on a fixed cache of $10{,}240$ route examples generated at the start of training with validation seed 2026. Validation loss is computed on the full cache, while route-decoding diagnostics use 64 examples from each of the ten validation batches. All reported experiments use the checkpoint with the lowest validation loss observed during training.

B.6 Ablation Configurations

All Experiment 1 ablations use the same synthesis space, online data generation, training budget, optimizer setup, best-validation checkpoint selection, route executor, three sampling seeds, and evaluation implementation as the full model. In the *w/o low-rank parameterization* variant, the 128-dimensional catalog-identity embedding u_i is replaced by a directly learned full-dimensional building-block embedding with the same model-input dimensionality, while the low-rank output residual $\gamma\langle Ah, r_i\rangle$ is removed; the ChemBERTa prototype pathway is retained. In the *w/o ChemBERTa prototypes* variant, the frozen ChemBERTa prototype branch is removed so that the learned identity pathway provides the building-block input representation, while prototype-based output scoring is replaced by a learned fully connected building-block classifier; the low-rank identity components are retained. The *w/o both components* variant combines these changes, removing both the ChemBERTa prototype pathway and the low-rank identity parameterization.

For the autoregressive counterpart, we retain the Transformer size, conditioning inputs, route vocabulary, and complete building-block parameterization, including the ChemBERTa prototypes and low-rank identity components. The online training space and data generation, training budget, optimizer setup, and best-validation checkpoint selection are also unchanged. We change only the generation formulation: non-causal self-attention and the masked-diffusion objective and sampling procedure are replaced by causal self-attention, next-token prediction, and left-to-right generation in postfix order. All ablation results use the Enamine reconstruction benchmark with 1,000 targets and 256 samples per target. All variants are evaluated on the same four Experiment 1 metrics.

C Conditioning and Composite-Guidance Implementation

C.1 Condition Representations

Each condition input is represented by one or more $d = 1024$ prompt tokens that are concatenated with the noisy route representation before the Transformer layers. The bidirectional Transformer therefore processes the condition prompts and the complete noisy route jointly. RouteDiff supports six condition families: ECFP4, FCFP4, exact scalar descriptors, scalar upper bounds, scalar lower bounds, and BRICS fragments (Degen et al., 2008).

For ECFP4 and FCFP4 conditions, the fixed-width fingerprint vector is mapped by a learned MLP to four prompt tokens. A BRICS condition is represented by the ECFP4 fingerprints of its molecular fragments. Each retained fragment contributes one prompt token, using a 2048-bit fingerprint, with at most eight fragment tokens per condition input.

Scalar conditions pair a numerical value with a descriptor-type identity. For exact-value conditioning, each pair (v, k) is embedded by combining an MLP representation of value v with a learned embedding of descriptor type k, producing one prompt token per scalar requirement. Multiple exact-value requirements can therefore be represented jointly as a set within one condition input. Upper- and lower-bound conditions use separate learned embedders for (u, k) and (l, k), respectively, and each bound condition is expanded to four prompt tokens.

During training, bound thresholds are constructed from the descriptor value v of the generated product. For descriptor type k, an upper-bound threshold is sampled between v and its running per-descriptor upper reference, whereas a lower-bound threshold is sampled between the running per-descriptor lower reference and v. At inference, the supplied threshold is embedded directly.

Lower-bound requirements are therefore represented as a native positive condition family rather than by negating an upper-bound condition.

C.2 Training Condition Mixture

Each $1{,}024$-row training micro-batch is partitioned across the six condition families described above: 204 ECFP4 rows, 204 FCFP4 rows, 204 exact-scalar rows, 102 upper-bound rows, 102 lower-bound rows, and 208 BRICS rows. For exact-scalar conditioning, four descriptor types are evaluated for each training row. Each resulting scalar prompt token is independently dropped with probability 0.5 during training. For BRICS conditioning, each retained fragment prompt token is independently dropped with probability 0.5.

No separate whole-condition dropout is applied to the denoiser input ($p_{\mathrm{cond-drop}} = 0$). This is distinct from the descriptor-internal token dropout above: scalar-property and BRICS dropout act independently on prompt tokens within their respective condition representations, whereas whole-condition dropout would remove all condition prompts for a training row. ECFP4, FCFP4, upper-bound, and lower-bound condition inputs are used without descriptor-internal token dropout.

C.3 Typed-Logit Composition

For a conjunction of condition inputs $c_1, \ldots, c_m$, RouteDiff evaluates each c_j independently with the same parameter-shared denoiser at the same noisy route state and diffusion timestep. A condition input may itself contain multiple prompt tokens; in particular, a set of exact scalar requirements is processed jointly within one denoiser evaluation. For each condition input, the denoiser produces separate token-type, building-block, and reaction logits,

$$\ell_j^h(x_t, t) = f_\theta^h(x_t, c_j, t), \qquad h \in \{\mathrm{type}, \mathrm{BB}, \mathrm{rxn}\}. \tag{22}$$

All composite queries reported in this work use positive condition weights $\alpha_j > 0$. Their predictions are composed independently for each output head:

$$\ell_{\mathrm{comp}}^h(x_t, t) = \sum_{j=1}^{m} \alpha_j \ell_j^h(x_t, t), \qquad h \in \{\mathrm{type}, \mathrm{BB}, \mathrm{rxn}\}. \tag{23}$$

Composition therefore acts on denoiser outputs rather than on condition embeddings, and the three typed output spaces remain separate throughout composition. Upper- and lower-bound prompts both enter this expression as positive condition inputs; a two-sided scalar range is represented by combining its lower- and upper-bound conditions in the same conjunction.

At every reverse-diffusion step, the composed logits define the clean-token distributions used by the typed sampler in Appendix B.2. The sampler first predicts token type and then the corresponding building-block or reaction identity, after which positions are revealed according to the absorbing reverse transition. Unresolved positions are reevaluated from the shared route state at the next denoising step, so condition composition is applied throughout the reverse trajectory.

D Composite-Query Catalog

The nine-query catalog contains five physicochemical regions and four drug-anchored objectives. The physicochemical regions are based on the Lipinski rule of five, Veber criteria, Congreve rule of three, and Ghose property ranges (Lipinski et al., 1997; Veber et al., 2002; Congreve et al., 2003; Ghose et al., 1999); Table 5 gives the exact operational inequalities used here.

Table 5: Operational definitions for the nine composite queries used in Experiment 3. For drug-anchored queries, ECFP4 similarity contributes to the graded task score, while feasibility is determined only by the listed scalar constraints.

Task	Operational query / evaluator
Lipinski+Veber region	MW<500 AND HBD<5 AND HBA<10 AND RotB<10 AND TPSA<140 AND CLogP<5.0
Lipinski rule-of-five region	MW≤ 500 AND HBD≤ 5 AND HBA≤ 10 AND CLogP≤ 5
Veber	RotB≤ 10 AND TPSA≤ 140
Congreve Rule-of-3	MW<300 AND HBD≤ 3 AND HBA≤ 3 AND CLogP≤ 3
Ghose	$-0.4 \leq$CLogP≤ 5.6, $160 \leq$MW≤ 480, $40 \leq$MR≤ 130, $20 \leq$TotalAtoms≤ 70
Cobimetinib	Score: ECFP4 similarity to Cobimetinib; hard constraints: RotB$=3$ AND AromaticRings$=3$ AND CLogP<5.0
Osimertinib	Score: ECFP4 similarity to Osimertinib; hard constraints: TPSA>100 AND CLogP<1.0
Amlodipine	Score: ECFP4 similarity to Amlodipine; hard constraint: Rings$=3$
Perindopril	Score: ECFP4 similarity to Perindopril; hard constraint: AromaticRings$=2$

For the four drug-anchored queries, ECFP4 similarity remains a graded score and is not converted into a hard pass/fail threshold.

E Additional Experiment 3 Diagnostics

E.1 Feasibility-Conditioned Molecular and Route Diversity

For each task and sampling seed, let $\mathcal{S}$ be the 1,000 attempted route sequences, $\mathcal{S}_{\text{valid}}$ the executable subset, and $\mathcal{P}_{\text{feas}}$ the unique canonical products from valid routes that satisfy every literal hard constraint. Attempt-level feasibility is

$$\text{FeasibleRouteRate} = \frac{\#\{r \in \mathcal{S} : r \text{ yields at least one feasible product}\}}{|\mathcal{S}|}.$$

A route contributes at most once to this rate even if the executor exposes multiple products.

For $n = |\mathcal{P}_{\text{feas}}| \geq 2$, molecular diversity is mean pairwise Morgan/ECFP4 Tanimoto dissimilarity,

$$D_{\text{ECFP4}} = 1 - \frac{2}{n(n-1)} \sum_{i<j} T(fp_i, fp_j).$$

It is undefined for $n < 2$ rather than set to zero. To measure scaffold concentration, let p_s be the fraction of feasible unique products assigned to Bemis–Murcko scaffold s (Bemis & Murcko, 1996). We report both the number of observed scaffolds and the effective scaffold count

$$N_{\text{eff}} = \exp\left(-\sum_s p_s \log p_s \right).$$

The normalized effective-scaffold fraction is N_{eff}/n.

Route diversity is computed on feasible route instances. An *exact route signature* preserves every building-block and reaction identity in postfix order. A *reaction-skeleton signature* replaces each building-block identity with a generic BB symbol while retaining reaction identities and program structure. Their uniqueness rates are the number of distinct signatures divided by the number of feasible route instances. This distinction separates reactant substitution within the same reaction program from changes to the reaction program itself.

Table 6: Multi-scale diversity of feasible outputs. Statistics are computed per seed and averaged over the sampling seeds for which the corresponding per-seed statistic is defined; undefined values are excluded rather than imputed as zero. Effective scaffold fraction is the effective number of Bemis–Murcko scaffolds divided by the number of unique feasible products; exact-route and reaction-skeleton uniqueness are computed over feasible route instances. SP product diversity recomputes ECFP4 diversity using only single-product routes. For Perindopril, scaffold- and route-level statistics are defined in 3/5 seeds, while product diversity is defined in 1/5 seed. "–" denotes an undefined SP product-diversity statistic.

Task	Eff. scaffold fraction	Exact route uniq.	Rxn-skeleton uniq.	SP product div.
Lipinski+Veber region	0.020	0.756	0.016	0.936
Lipinski rule-of-five region	0.020	0.806	0.009	0.938
Veber	0.304	0.999	0.264	0.911
Congreve Rule-of-3	0.030	0.751	0.005	0.937
Ghose	0.159	0.612	0.237	0.921
Cobimetinib	0.686	1.000	0.926	0.877
Osimertinib	0.767	1.000	0.971	0.894
Amlodipine	0.926	0.933	0.933	–
Perindopril	1.000	1.000	1.000	–

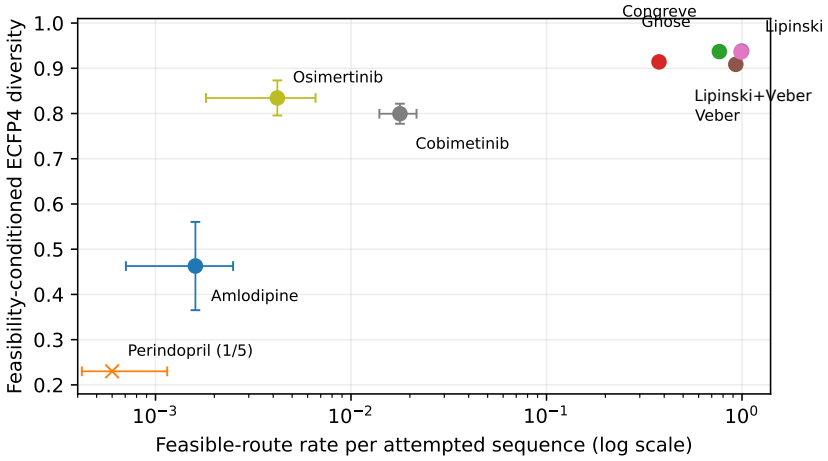

Figure 3: Constraint satisfaction and diversity among feasible outputs. Feasible-route rate is plotted against ECFP4 diversity among unique feasible products for the five-seed Experiment 3 evaluation. Feasible-route rate is averaged over all five seeds, whereas ECFP4 diversity is averaged only over seeds containing at least two unique feasible products; undefined diversity values are not imputed as zero. Diversity is defined in 4/5 seeds for Osimertinib and Amlodipine, in 1/5 seed for Perindopril, and in 5/5 seeds for all other tasks. The horizontal axis is logarithmic to resolve rare-feasible drug-anchored objectives.

Table 7: Additional executor diagnostic: fraction of valid routes exposing multiple products. Values are five-run means.

Task	Multi-product rate
Lipinski+Veber region	0.007
Lipinski rule-of-five region	0.003
Veber	0.157
Congreve Rule-of-3	0.002
Ghose	0.533
Cobimetinib	0.736
Osimertinib	0.201
Amlodipine	0.059
Perindopril	0.004

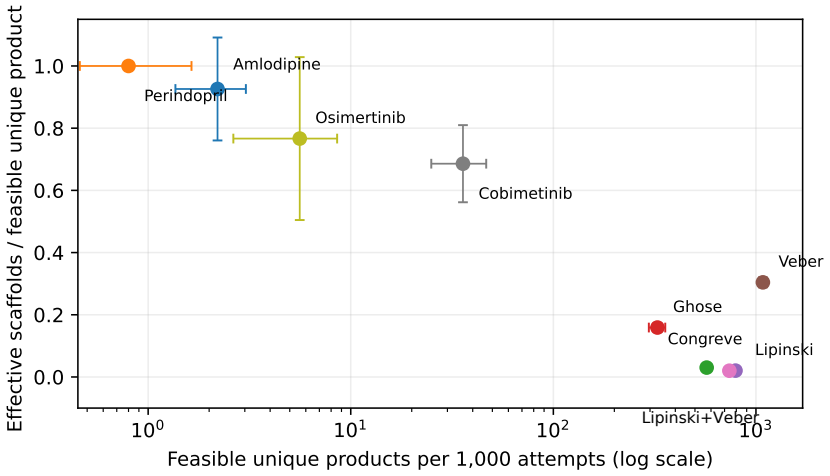

Figure 4: Scaffold occupancy among hard-feasible products. The horizontal coordinate is the five-seed mean number of unique feasible products. The vertical coordinate is the mean effective-scaffold fraction over seeds containing at least one unique feasible product; undefined values are not imputed as zero. Error bars show ± 1 sample standard deviation over the seeds contributing to the corresponding coordinate. For Perindopril, the effective-scaffold fraction is defined in 3/5 seeds. The horizontal axis is logarithmic because hard-feasible-set sizes span three orders of magnitude. High molecular-fingerprint diversity does not necessarily imply broad scaffold occupancy: for example, Lipinski has high ECFP4 diversity but a small effective-scaffold fraction.

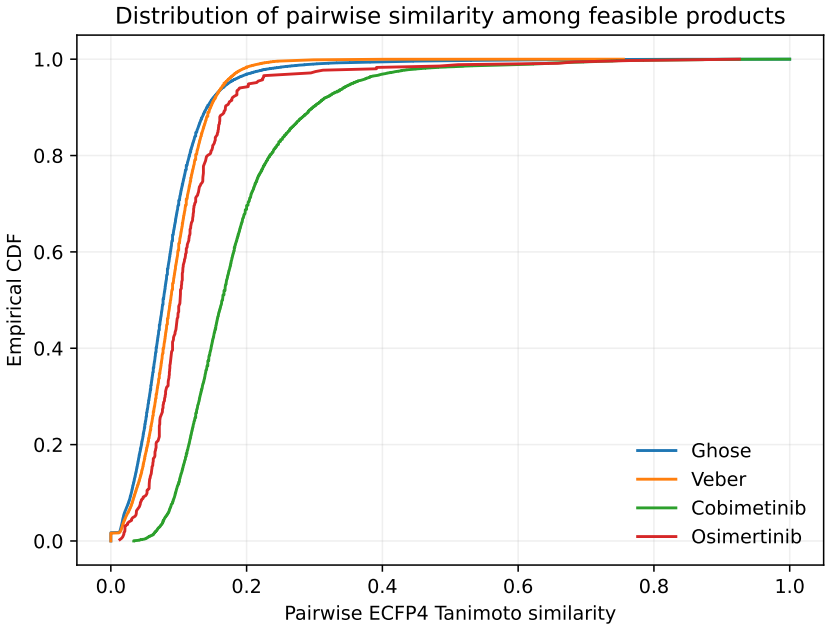

Figure 5: Descriptive ECDF of pairwise ECFP4 Tanimoto similarity among feasible products for four representative tasks. Pairs are pooled only for visualization (up to 50,000 sampled pairs per task with a fixed analysis seed); scalar diversity in Table 4 is computed within each sampling seed before aggregation.

E.2 Denoising-Trajectory Diagnostics

For qualitative route analysis, we pool feasible unique products across the five seeds using a pre-specified eligibility filter: the representative occurrence must come from a single-product route, contain at least one reaction step, and have a complete serialized route record. We then select $k = 3$ deterministic ECFP4 medoids per displayed task. Selection is fixed before trajectory inspection. The corresponding route tree is loaded from the sampling-time record, so the displayed pathway is model-emitted rather than retrospectively inferred.

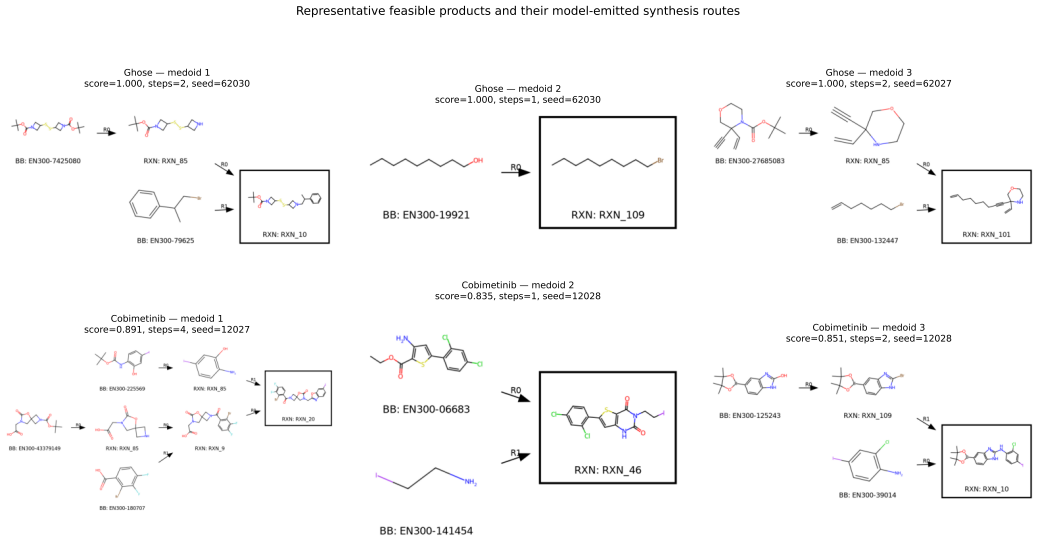

Figure 6: Complete six-medoid route case study used for Experiment 3: three Ghose medoids (top) and three Cobimetinib medoids (bottom). Ghose medoids all have score 1.000; Cobimetinib medoid scores range from 0.835 to 0.891. All satisfy every literal hard constraint and all displayed routes were emitted by the model during sampling.

Diffusion-trajectory replay. The route cases above were selected before trajectory inspection. We replayed exactly these six medoids and verified equality of the stored and replayed exact-route signatures as well as recovery of the selected product in all six cases. For the retrospective token-level diagnostic, a final-token-aligned pre-commit revision is counted only when (i) the position remains masked after the earlier reverse step and before the later denoising pass, (ii) the earlier argmax x_0 token differs from the *final emitted token*, and (iii) the later argmax token equals that final emitted token. Under this definition, all six traces contain a revision and three contain revision-and-commit; restricting to semantic BB/RXN identities gives five of six and two of six, respectively, with the remaining events involving END/PAD termination tokens.

Table 8: Trajectory diagnostics for the six pre-selected Experiment 3 medoids. Replay match requires the exact emitted route signature and selected product to match the stored sample. "BB/RXN rev." counts at least one final-token-aligned semantic pre-commit revision; "BB/RXN R&C" additionally requires that identity to be sampled and revealed at the same step.

Task	Medoid	Replay match	BB/RXN rev.	BB/RXN R&C
Ghose	1	yes	yes	no
Ghose	2	yes	yes	no
Ghose	3	yes	yes	no
Cobimetinib	1	yes	yes	yes
Cobimetinib	2	yes	yes	yes
Cobimetinib	3	yes	no	no

Figure 7 shows two complementary transitions from this fixed six-trace set. Ghose medoid 2 exhibits a final-token-aligned reaction-token revision toward its emitted route. Cobimetinib medoid 3 instead exhibits a feasibility-rescuing hypothesis revision; because the later building-block hypotheses in that transition are not the final emitted identities, it is not counted as a final-token-aligned semantic revision in Table 8.

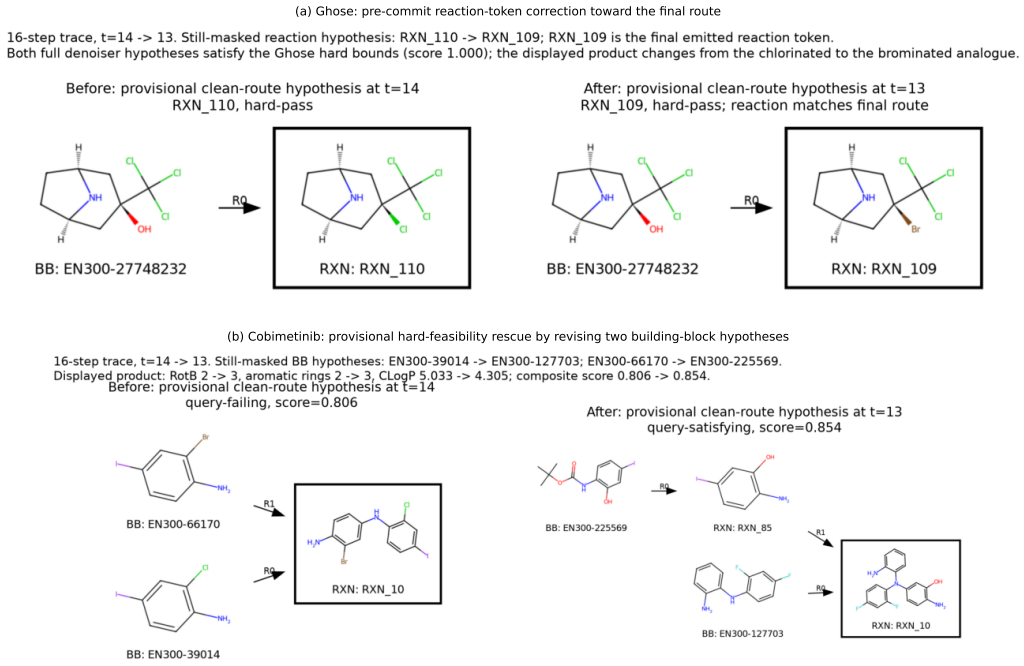

Figure 7: Two complementary denoiser-trajectory examples from the fixed six-medoid set. (a) Ghose medoid 2: at $t = 14 \rightarrow 13$, a still-masked reaction hypothesis changes from RXN_110 to RXN_109; RXN_109 is the final emitted reaction token, while both full denoiser hypotheses remain within the Ghose hard region. (b) Cobimetinib medoid 3: two still-masked building-block hypotheses change and the full provisional route moves from infeasible to feasible under the explicit scalar constraints, with task score $0.806 \rightarrow 0.854$. Panel (b) is a feasibility-rescuing hypothesis revision but not a final-token-aligned semantic revision. Both panels depict full argmax clean-route predictions for diagnostic visualization rather than sampler states; the absorbing ancestral sampler never rewrites an already revealed token.

Success-conditioned trajectory cohort. To test whether pre-commit semantic revision extends beyond the six medoids, we use an independently selected trajectory-blind cohort. Selection depends only on final Experiment 3 metadata—hard success, a single-product route, complete serialization, and deterministic hash ordering—before any trajectory is inspected. The selector requests up to 32 routes per task; 32 Ghose routes satisfy the criteria, whereas only 16 Cobimetinib routes do, so all 16 eligible Cobimetinib routes are included. All 48 replayed route signatures and selected products match their stored benchmark records. A semantic revision occurs in 22/32 Ghose and 14/16 Cobimetinib traces; revision-and-commit occurs in 8/32 and 8/16, respectively. Overall, 36/48 traces contain a final-token-aligned semantic BB/RXN revision and 16/48 contain semantic revision-and-commit. Three Cobimetinib cohort routes overlap with the six-medoid set; excluding them leaves 34/45 semantic-revision and 14/45 revision-and-commit traces. All reported cohort statistics are conditioned on final hard success.

Table 9: Success-conditioned trajectory-cohort diagnostics. "Semantic rev." means at least one final-token-aligned BB/RXN pre-commit revision; "Semantic R&C" additionally requires the revised identity to be sampled and revealed at that step. Hamming reports the mean full typed-sequence token mismatch count between the denoiser's argmax x_0 hypothesis and the final emitted route at the first and last reverse steps.

Task	n	Replay	Semantic rev.	Semantic R&C	Hamming t=16 $\rightarrow$ 1
Ghose	32	32/32	22/32 (68.8%)	8/32 (25.0%)	4.22$\rightarrow$1.53
Cobimetinib	16	16/16	14/16 (87.5%)	8/16 (50.0%)	5.50$\rightarrow$2.38
Overall	48	48/48	36/48 (75.0%)	16/48 (33.3%)	–

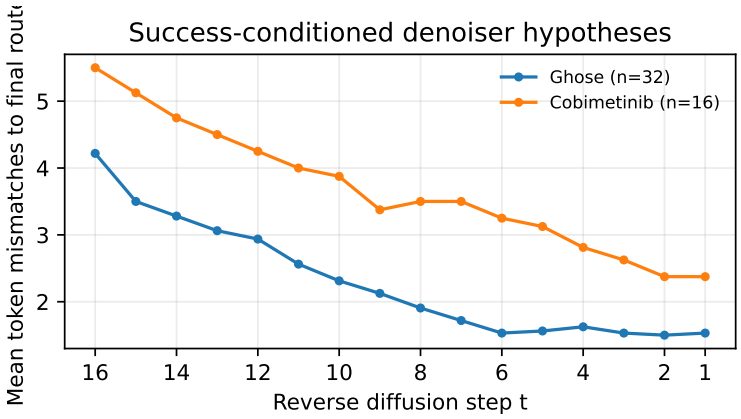

Figure 8: Mean full typed-sequence Hamming distance between the denoiser's argmax clean-route hypothesis and the final emitted route across the success-conditioned trajectory cohort. The distance decreases overall from $t = 16$ to $t = 1$ for both tasks (Ghose: $4.22 \rightarrow 1.53$; Cobimetinib: $5.50 \rightarrow 2.38$), although individual steps need not improve monotonically. This Hamming diagnostic counts all token types, whereas the semantic-revision statistics above count BB/RXN identities only. Because the final route is sampled stochastically, the argmax x_0 hypothesis need not equal the final emitted route even at $t = 1$. The cohort is conditioned on final hard success.

F Evaluation Metrics

Experiment 1: reconstruction and diversity. For each target x in the Enamine or ChEMBL reconstruction set, RouteDiff samples 256 synthesis programs per seed and executes them in the frozen RouteDiff chemical space. If a valid route r exposes more than one product, its representative product is

$$p(r) = \arg\max_{p} T\left(fp^{4096}_{\text{Morgan},r=2}(p), fp^{4096}_{\text{Morgan},r=2}(x)\right).$$

Let $\text{can}(\cdot)$ denote canonical SMILES. Target-level reconstruction is

$$R_x = \mathbf{1}[\exists r : \text{can}(p(r)) = \text{can}(x)], \qquad \text{Recon} = \frac{1}{1000}\sum_x R_x.$$

For each valid route, $s(r, x)$ is the Morgan Tanimoto similarity with radius 2 and 4096 bits. The target score is

$$S_x = \max_r s(r, x), \qquad \text{Similarity} = \frac{1}{1000}\sum_x S_x.$$

Following the ReaSyn v2 benchmark framing, analogous routes are

$$\mathcal{A}_x = \{r : s(r, x) \geq 0.8\}.$$

Table 10: **Full Experiment 2 benchmark context on 10 PMO tasks.** Baseline values are three-run means from ReaSyn v2 Table 3 (Lee et al., 2026); RouteDiff reports three-run means. Bold marks the best synthesis-constrained result in each row. †Graph GA is unconstrained and is shown only as a reference.

Objective	SynthesisNet	SynNet	DoG-Gen	DoG-AE	Graph GA–SF	Graph GA–ReaSyn	RouteDiff	Graph GA†
Amlodipine MPO	.608	.567	.537	.509	.696	.620	**.786**	.651
Celecoxib rediscovery	.582	.443	.466	.357	.559	**.810**	.780	.682
Fexofenadine MPO	.791	.764	.697	.681	.786	.788	**.812**	.785
Median1	.305	.219	.218	.172	.308	.274	**.365**	.261
Median2	.257	.237	.213	.183	.258	.259	**.295**	.257
Osimertinib MPO	.810	.797	.776	.751	.816	.823	**.827**	.829
Perindopril MPO	.524	.559	.475	.433	.530	.561	**.693**	.533
Ranolazine MPO	.741	.743	.712	.690	.751	.752	**.764**	.745
Sitagliptin MPO	.313	.026	.048	.010	.338	.314	**.423**	.524
Zaleplon MPO	**.528**	.341	.123	.050	.478	.460	.524	.458

For any canonicalized, deduplicated molecular set $M = \{m_1, \ldots, m_n\}$, we use Morgan radius 2, 2048 bits, `useChirality=False`, and

$$D(M) = \begin{cases} \frac{2}{n(n-1)} \sum_{i<j} \left(1 - T(fp_i, fp_j)\right), & n \geq 2, \\ 0, & n < 2. \end{cases}$$

Div. (Pathway) is $D(M)$ for the unique representative end products produced by $\mathcal{A}_x$; despite the inherited label, it is an end-product molecular-diversity statistic rather than exact route-sequence diversity. Div. (BB) applies the same $D(M)$ to unique canonical leaf building blocks collected from $\mathcal{A}_x$. Zero padding is applied at the target level before averaging, so targets with fewer than two unique analogous products or building blocks contribute 0 rather than being dropped. RouteDiff reports mean±sample standard deviation across sampling seeds 2026, 2027, and 2028.

Experiment 2: fixed-budget goal-directed optimization. Each objective is evaluated under a hard maximum of 10,000 oracle calls. At an evaluation step, Top-10 is the mean score of the ten best molecules observed so far (using all available molecules before ten evaluations); AUC Top-10 summarizes this trajectory over the oracle-call budget following the PMO/ReaSyn goal-directed evaluation protocol (Gao et al., 2022a; Lee et al., 2026). Each RouteDiff optimization run starts from 500 molecules; each iteration selects 50 parents and generates 50 offspring. Three independent optimization runs are evaluated under the full budget, and the reported value for each objective is their mean AUC Top-10. Experiment 2 focuses on fixed-budget optimization performance, while diversity is characterized separately in Experiments 1 and 3. Published baselines retain their original search procedures and synthesis spaces. Table 10 provides the full 10-task benchmark context, including weaker synthesis-constrained baselines and the unconstrained Graph GA reference omitted from the main-text table.

Experiment 3: composite-query generation. Each task uses 1,000 attempted route sequences per sampling seed and five sampling seeds from the same frozen checkpoint. Route validity is the fraction of attempts that execute. Feasible-route rate is the fraction of attempted routes yielding at least one product that satisfies every explicit hard constraint. For the drug-anchored tasks, ECFP4 similarity remains graded and is not converted into a hard pass/fail threshold. Product-level metrics use unique canonical feasible products and include count, pairwise ECFP4 diversity, Bemis–Murcko scaffold statistics, and the effective-scaffold fraction N_{eff}/n. SP product diversity recomputes ECFP4 diversity after restricting to single-product routes. Route-level metrics use feasible route instances and report exact-route and reaction-skeleton uniqueness. Multi-product rate is the fraction of valid routes that expose more than one product. Scalar statistics are computed per seed. Statistics defined in all five seeds are summarized as mean ± sample standard deviation over those five seeds. If a statistic is undefined for a seed because the relevant feasible set is empty or contains fewer than the required number of products, that seed is excluded only from the summary of that statistic; undefined values are not imputed as zero.

Experiment 3: denoising-trajectory diagnostics. Trajectory replay is evaluated in two success-conditioned sets: the six medoids fixed by the preceding diversity analysis and an independently se-

lected trajectory-blind cohort chosen from final hard-success, single-product, serialization-complete routes before trace inspection (32 Ghose and all 16 eligible Cobimetinib routes). Three cohort traces overlap with the six-medoid set. Replay match requires the exact final route signature and selected product to agree with the stored benchmark sample. A final-token-aligned semantic pre-commit revision requires a BB/RXN position to remain masked across two consecutive denoising passes, with the earlier argmax x_0 token differing from and the later argmax token equaling the final emitted token. A semantic revision-and-commit additionally requires the later stochastic sample to equal that final token and the position to be revealed at that step. Step-wise trajectory convergence is summarized by full typed-sequence token Hamming distance between the argmax x_0 hypothesis and the final emitted route. All reported trajectory statistics are conditioned on final hard success.