SYNTHESIZABILITY-GUIDED DISCRETE DIFFUSION MODEL FOR MOLECULAR GENERATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Molecular generative models can propose candidates with favorable predicted properties, but are not guaranteed to generate synthesizable molecules. Candidates that are hard to synthesize cannot readily enter wet-lab experiments. Computer-aided synthesis planning (CASP) tools provide route-specific evidence towards synthesizability, but are too slow to integrate into generation. In this paper, we introduce *SynGuide*, a guidance mechanism that distills CASP outcomes into a noise-conditioned guidance signal that guides the generation of discrete diffusion models. We define a fixed-budget retrosynthetic-accessibility (RA) label indicating whether a CASP tool finds a route to purchasable building blocks within a fixed search budget. In order to build guided generative models, we then train a noise-conditioned predictor of the RA-label and use its gradient to guide a frozen discrete diffusion generator. Under RA-label guidance, SynGuide increases the measured planner solve rate by 24% over unguided backbone. Across three different protein targets, *SynGuide* achieves *state-of-the-art* performance improving synthesizability. This work provides a practical way to incorporate planner-derived synthesizability objectives into molecular generation without placing expensive retrosynthesis searches in the sampling loop.

1 INTRODUCTION

De novo molecular design aims to discover new compounds that satisfy the biological and physicochemical requirements of drug discovery. Recent generative models have demonstrated progress in designing candidate molecules with favorable predicted binding affinity (Eckmann et al., 2022; 2025; Guan et al., 2023; Peng et al., 2022), drug-likeness (You et al., 2018; Jin et al., 2018; Zhou et al., 2019; Shi* et al., 2020; Xie et al., 2021; Bengio et al., 2021), selectivity (Liu et al., 2021; Isigkeit et al., 2024), and ADMET properties (Lim et al., 2018; Kang & Cho, 2019; Kyro et al., 2025). However, favorable property predictions do not establish synthetic feasibility. Synthesis depends on the availability of starting materials, applicable reactions, and route-specific factors such as operational complexity and yield. Structure-based heuristics provide inexpensive approximations of synthetic accessibility but do not explicitly establish a route. Computer-aided synthesis planning (CASP) tools offer more direct evidence by searching for candidate routes to available starting materials (Genheden et al., 2020; Segler et al., 2018; Tu et al., 2025), although their computational cost can limit their use within generative optimization.

Recent work has made synthesis more explicit in generation. SynNet optimizes synthesis trees (Gao et al., 2022b), RxnFlow generates reaction pathways over purchasable building blocks (Seo et al., 2025), and SynFormer generates within a synthesis-defined chemical space (Gao et al., 2025). These studies demonstrate the value of synthesis-aware learning, but expose a tradeoff: route-generative methods restrict generation to their reaction actions and building-block catalogues, while complete-structure predictors do not score intermediate molecular states. Planner-in-the-loop optimization retains a molecular generator but repeatedly pay the CASP cost (Guo & Schwaller, 2025; Chen et al., 2026). Our goal is to preserve a molecular generator that is not bounded by a reaction-template action space while making a planner-derived signal available throughout sampling.

We address this mismatch with planner-supervised discrete guidance *SynGuide* (Figure 4). We first define a fixed-budget retrosynthetic-accessibility (RA) label indicating whether AiZynthFinder finds a route to available building blocks under a specified planner configuration. This label captures a

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.

computational planning outcome rather than an unconditional claim of experimental synthesizability. We then train a noise-conditioned predictor on partially masked SAFE sequences, using the planner outcome of each associated complete molecule as supervision. Following the predictor-guidance formulation of discrete diffusion, we use this predictor to adjust the denoising transitions of a frozen GenMol backbone. Planner evaluations are performed offline to construct the training labels; guided sampling subsequently uses the learned predictor without further planner calls. Combining this signal with a target-specific docking predictor enables the relative emphasis on accessibility and docking to be adjusted at sampling time.

We evaluate both synthesizability-only guidance and its combination with target-specific docking objectives. Synthesizability-only guidance increases the measured AiZynthFinder solve rate by 22.4%. Across three protein targets ACAA1, c-MET, and BRD4(2), intermediate objective weights improve both planner solve rates and docking scores relative to unguided sampling. Comparisons with existing de novo design methods assess the performance of the complete approach, while guidance-mechanism comparisons examine how the predictive signal is used during sampling. We further investigate sensitivity to guidance strength, objective weighting, and planner budget, and evaluate generated molecules under an alternative retrosynthesis configuration to assess transfer beyond the labeling planner. Together, these experiments examine the effectiveness, trade-offs, and configuration dependence of planner-supervised generation.

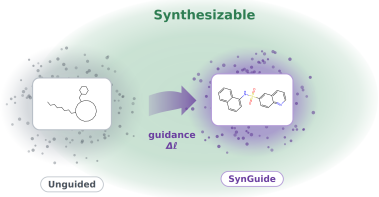

Figure 1: *SynGuide* guides generation into the synthesizable region. The green shaded area marks molecules with a retrosynthetic route found by CASP tools. Gray is the unguided generation, violet is generation with SynGuide.

We summarize the contributions of this work as follows:

- We formulate a fixed-budget retrosynthetic-accessibility prediction on partially masked molecular sequences and train a noise-conditioned predictor using CASP outcomes as supervision.
- We introduce *SynGuide*, a guidance mechanism that carries the predictor into the sampling process, enabling planner-supervised generation without planner calls or backbone updates during sampling.
- We demonstrate a 24% improvement in AiZynthFinder solve rate over unguided backbone and show that, across three protein targets, intermediate combinations of retrosynthetic-accessibility and docking guidance improve both objectives over unguided sampling.

2 Related Work

Discrete diffusion and guidance. Diffusion models are a class of generative models defined by a denoising network that is trained to remove noise from latent variables (Sohl-Dickstein et al., 2015; Ho et al., 2020). Absorbing-state discrete diffusion (Austin et al., 2021) uses a mask token to corrupt categorical sequences, and GenMol (Lee et al., 2025) applies masked discrete diffusion to molecular generation. Its reverse dynamics can be described by continuous-time transition rates (Nisonoff et al., 2025) or discrete-time categorical kernels (Schiff et al., 2025). Predictor guidance biases these transitions toward a desired outcome, with Taylor approximations reducing the cost of evaluating candidate token changes (Nisonoff et al., 2025). Other approaches include derivative-free value-based guidance (Li et al., 2025) and autoguidance, which uses a weaker generative model instead of a property predictor (Karras et al., 2024). We study guidance with a fixed GenMol backbone and a predictor trained to estimate retrosynthesis outcomes from partially masked sequences.

Synthesizability in molecular generation. Synthesizability lies as the central objective of de novo drug design. Despite advances in molecular generation, generated candidates remain challenging to synthesize (Gao & Coley, 2020; Papidocha et al., 2026). Existing synthesizability scores differ in their underlying evidence and computational requirements. Structure-based approaches such as SA (Ertl & Schuffenhauer, 2009) and SYBA (Voršilák et al., 2020) estimate synthetic ac-

cessibility from fragment statistics and structural features without explicitly identifying a synthetic route. Reaction-informed approaches instead learn from recorded chemistry (Coley et al., 2018; Chen & Jung, 2024), providing estimates informed by reaction data but not necessarily establishing a route to available starting materials. Planner-derived approaches incorporate more explicit retrosynthetic evidence. RAscore (Thakkar et al., 2021) trains a classifier to predict the outcomes of a computer-aided synthesis planning (CASP) tool for complete molecules, while Liu et al. (2024) combine retrosynthetic planning with forward reaction prediction.

Retrosynthesis planning and synthesis-oriented generation. Retrosynthesis planners search for routes from a target molecule to available starting materials (Segler et al., 2018). Their outcomes depend on the reaction model, the search algorithm, the building-block stock, and the computational budget. Planning systems such as AiZynthFinder (Genheden et al., 2020), ASKCOS (Tu et al., 2025), and SYNTHIA (Szymkuć et al., 2016; Grzybowski et al., 2018) differ in their reaction models, search procedures, and knowledge sources. Syntheseus (Maziarz et al., 2024) supports controlled comparisons of reaction models and search algorithms. A complementary generative approach constructs molecules along synthetic pathways: SynNet generates synthesis trees from building blocks (Gao et al., 2022b), RxnFlow learns a GFlowNet over reaction actions (Seo et al., 2025), and SynFormer generates within a synthesis-defined chemical space (Gao et al., 2025). These approaches depend on their specified reactions and building-block stocks. SynGuide instead learns from a planner's fixed-budget outcomes to guide molecular generation; we evaluate whether its gains persist under a different planning configuration in §5.6.

3 Background

3.1 Discrete Diffusion Models

A diffusion model generates data by learning to reverse a fixed noising process that gradually maps clean samples to a simple prior. Let $\mathcal{V}$ be a finite vocabulary and L a sequence length. A clean molecule is a token sequence $x_0 \in \mathcal{V}^L$. Masked discrete diffusion (Austin et al., 2021; Lee et al., 2025) corrupts a clean token sequence x_0 toward an absorbing mask prior π,

$$q(x_t \mid x_0) \;=\; \mathrm{Cat}\big(x_t;\, \alpha_t x_0 + (1-\alpha_t)\pi\big), \tag{1}$$

where x_t is the partially masked state of x_0 at noise time $t \in [0,1]$, and α_t decreases monotonically in t. A token that is unmasked at time t remains unmasked for all $s < t$. Nisonoff et al. (2025) write the reverse process as a continuous-time Markov chain (CTMC) with rates R_t,

$$p(x_{t+\mathrm{d}t} = \tilde{x} \mid x_t = x) \;=\; \delta_{x,\tilde{x}} + R_t(x,\tilde{x})\,\mathrm{d}t, \tag{2}$$

so that in continuous time only one token can change at a given instant. Here $R_t(x,\tilde{x})$ is the instantaneous rate from a state x to a candidate $\tilde{x}$, and $\delta_{x,\tilde{x}}$ is the Kronecker delta. Schiff et al. (2025) write the corresponding discrete-time reverse step as a categorical $p(x_s \mid x_t)$ induced by the same denoiser.

3.2 Predictor guidance for Discrete Diffusion

Predictor guidance reweights the reverse process of diffusion without retraining the generator (Nisonoff et al., 2025; Schiff et al., 2025). In the CTMC formulation, the guided rates are the unguided rates multiplied by time-dependent likelihood ratio,

$$R_t(x,\tilde{x} \mid y) \;=\; \frac{p(y \mid \tilde{x}, t)}{p(y \mid x, t)}\, R_t(x,\tilde{x}), \qquad \tilde{x} \neq x, \tag{3}$$

where y is the conditioning event or property defined on complete sequences, and $R_t(x, x \mid y) = -\sum_{\tilde{x}\neq x} R_t(x,\tilde{x} \mid y)$. Guidance strength γ raises the ratio,

$$R_t^{(\gamma)}(x,\tilde{x} \mid y) \;=\; \left[\frac{p(y \mid \tilde{x}, t)}{p(y \mid x, t)}\right]^{\gamma} R_t(x,\tilde{x}). \tag{4}$$

Taylor-approximated guidance (TAG) replaces the log-ratio with a first-order expansion around a differentiable one-hot state,

$$\log \frac{p(y \mid \tilde{x}, t)}{p(y \mid x, t)} \;\approx\; (\tilde{x} - x)^\top \nabla_x \log p(y \mid x, t), \tag{5}$$

Hence one backward pass yields a bias for every token. Schiff et al. (2025) gives the discrete-time form (D-CBG),

$$p^\gamma(x_s \mid x_t, y) \;\propto\; p_\phi(y \mid x_s, x_t)^\gamma \, p(x_s \mid x_t), \tag{6}$$

and the same first-order expansion of $\log p_\phi$. Predictor-free guidance (D-CFG) instead blends a conditional denoiser with an unconditional one. Li et al. (2025) replaces the gradient with per-step candidate selection under a value function, and Karras et al. (2024) extrapolates away from a deliberately degraded copy of the model. All of these mechanisms therefore require a score $p(y \mid x_t, t)$ at every denoising step. A CASP planner returns a verdict only for a complete molecule x_0, and is undefined on the intermediate state x_t. A noise-conditioned predictor, trained on partially masked sequences, is therefore needed to turn the planner label at x_0 into the score $p(y \mid x_t, t)$ that these mechanisms require.

4 METHOD

We use a frozen GenMol backbone as our generator. GenMol is a masked discrete diffusion model that generates Sequential Attachment-based Fragment Embedding (Noutahi et al., 2024, ABFE) sequences, with reverse dynamics described by the CTMC formulation in Section 3 (Equations 1–2). The base sampling process does not explicitly optimize retrosynthetic accessibility. We therefore reweight its transitions to favor molecules for which a retrosynthesis planner can identify routes to purchasable building blocks, while keeping the backbone parameters fixed. We define the target property on complete molecules through a binary function $G(x_0) \in \{0, 1\}$, which indicates whether the planner finds such a route under a specified configuration and search budget.

Guidance operates on partially masked states, for which G cannot be directly evaluated. At diffusion time t, guidance requires to estimate $p(y_{\mathrm{RA}} = 1 \mid x_t, t)$: the conditional probability that a complete molecule associated with the current state satisfies the planner criterion. This is the probability required by predictor guidance (Equation 3), which we estimate using a noise-conditioned predictor. Our framework comprises three components: the terminal property G, a predictor of its conditional probability on masked states, and an efficient approximation to the resulting transition likelihood ratio. The predictor avoids planner calls during sampling, while the ratio approximation reduces the cost of evaluating candidate transitions. We describe these components in the following three sections. Figure 2 illustrates the workflow of the framework.

A fixed-budget retrosynthetic-accessibility label Let $\mathcal{A}$ denote a computer-aided synthesis planner and B its search budget. We define a molecule as *solved* if $\mathcal{A}$ identifies a route to purchasable building blocks within B, with all other planner settings held fixed. The terminal property is the binary retrosynthetic-accessibility (RA) label:.

$$G(x_0) = y_{\mathrm{RA}}(x_0) = \begin{cases} 1, & \text{if } x_0 \text{ is solved,} \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

We use AiZynthFinder with a budget of $B = 500$ Monte Carlo tree search (MCTS) iterations. Appendix C.1 presents the budget sweep supporting this choice and shows that the relative ranking of the evaluated generative distributions remains stable at larger budgets.

Noise-conditioned prediction on masked sequences At diffusion time t, the sampler operates on a partially masked sequence x_t, which cannot be evaluated directly by the retrosynthesis planner. We therefore train a noise-conditioned predictor $p_\phi(y_{\mathrm{RA}} \mid x_t, t)$ to estimate the planner outcome from this intermediate state. Training inputs are obtained by corrupting complete molecular sequences x_0 using the backbone's forward process (Equation 1). Each corrupted sequence retains the label $y_{\mathrm{RA}}(x_0)$ assigned to its corresponding complete molecule, allowing the predictor to learn from planner supervision across noise levels.

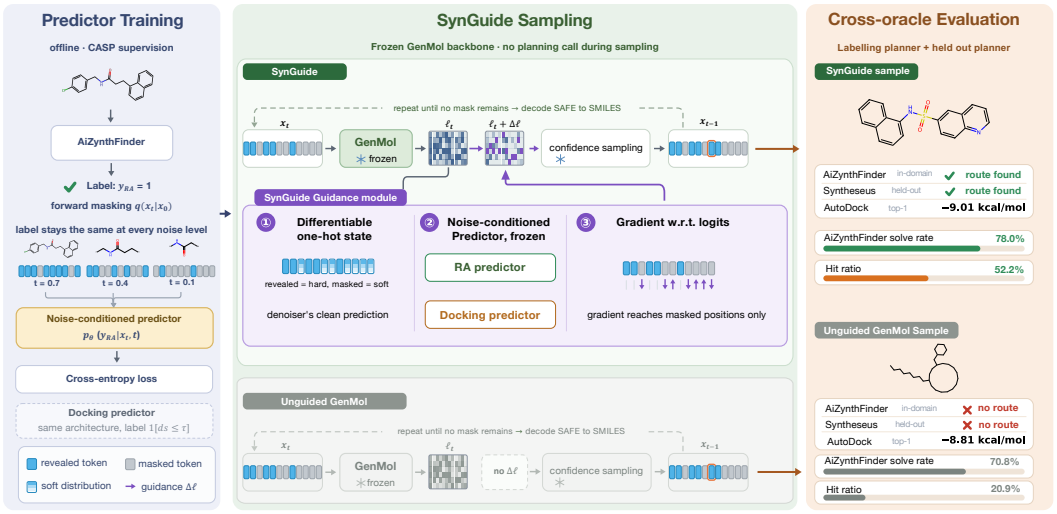

Figure 2: Workflow of *SynGuide*. In the training phase, the predictor is trained on partially masked SAFE sequences. In the sampling phase, the frozen backbone's prediction $\hat{x}_0$ is scored by the predictors at every denoising step, and the weighted gradients are added to the logits at masked positions only. In the evaluation phase, the same molecules are re-scored by the planner that supplied the labels and by a planner held out from training.

4.1 SynGuide

We introduce *SynGuide*, a planner-supervised guidance method for masked discrete diffusion with a frozen GenMol backbone. Inspired by the Taylor approximation used in predictor guidance (Equations 3 and 5), SynGuide updates the denoiser logits using gradients of the predicted target probability. The predictor evaluates a soft clean-sequence estimate constructed from the denoiser output, and gradients are propagated through this estimate to the logits.

Let $\ell_t \in \mathbb{R}^{L\times S}$ denote the denoiser logits for a sequence of length L and vocabulary size S, and let $q_t = \mathrm{softmax}(\ell_t)$, with softmax applied over the vocabulary dimension. Let $M_t \in \{0,1\}^L$ indicate masked positions, with broadcasting over the vocabulary dimension where needed. We construct the predictor input as

$$\hat{x}_0 \;=\; M_t \odot q_t + (1 - M_t) \odot \mathrm{onehot}(x_t), \tag{8}$$

where $\odot$ denotes element-wise multiplication. Thus, masked positions are represented by the denoiser's predicted token distributions, while unmasked positions retain their observed token. The RA predictor assigns the score $p_\phi(y_{\mathrm{RA}} = 1 \mid \hat{x}_0)$. For position i, let $u_i \in \mathbb{R}^S$ denote the gradient of its log-probability with respect to $\hat{x}_{0,i}$. Applying the chain rule through the softmax gives:

$$\nabla_{\ell_{t,i}} \log p_\phi(y_{\mathrm{RA}}{=}1 \mid \hat{x}_0) \;=\; M_{t,i}\, q_{t,i} \odot \Big(u_i - \langle q_{t,i}, u_i\rangle \mathbf{1}\Big), \tag{9}$$

where $\mathbf{1}$ is the all-ones vector. This gradient quantifies how changes to the denoiser logits affect the predicted accessibility of the soft clean-sequence estimate.

For multi-objective guidance, we combine the predictors through a weighted sum of log-probabilities, corresponding to a product of predicted probabilities raised to their respective weights. For each objective k, we compute $g_k = \nabla_{\ell_t} \log p_{\phi_k}(y_k = 1 \mid \hat{x}_0)$. To account for differences in gradient scale across predictors, we normalize each gradient by its root-mean-square magnitude over masked position–vocabulary coordinates:

$$\bar{g}_k = \frac{g_k}{\mathrm{RMS}_{M_t}(g_k) + \epsilon}. \tag{10}$$

The guided logits are then

$$\ell_t^{\mathrm{guided}} \;=\; \ell_t + \gamma s\, M_t \odot \sum_k \lambda_k \bar{g}_k, \tag{11}$$

The two objectives are retrosynthetic accessibility, estimated by the predictor, and target-specific docking, estimated by a binary predictor with the same architecture and training corruption process. We use nonnegative weights satisfying $\lambda_{\text{syn}}+\lambda_{\text{bind}} = 1$. Unless otherwise stated, we set $\lambda_{\text{syn}} = 0.4$, $\gamma = 2$, $s = 0.028$, and $\epsilon = 10^{-12}$, where γ controls guidance strength and s sets the base update scale. The mask restricts logit updates to masked positions, and revealed tokens remain unchanged during sampling. The backbone parameters remain frozen, and no planner calls are required during generation. Additional implementation details are provided in Appendix A.

5 EXPERIMENTS

We evaluate whether planner-supervised guidance improves retrosynthetic accessibility while retaining useful docking scores. We first compare SynGuide with molecular generators across three protein targets (§5.3). We then compare guidance mechanisms on the same frozen backbone (§5.4), test whether the gains persist under an independent planner (§5.5), and examine the objective weights and planner budget (§5.6).

5.1 EXPERIMENTAL SETUP

Protein targets. We evaluate on three protein targets: the peroxisomal thiolase ACAA1, the c-MET kinase domain, and the second bromodomain of BRD4, denoted BRD4(2). Each target has its own fixed docking grid and score threshold.

Retrosynthesis planner. We define a molecule as solved when AiZynthFinder finds a route within $B{=}500$ Monte Carlo tree search iterations. The evaluation uses the USPTO and RingBreaker expansion policies (Thakkar et al., 2019), the USPTO filter, and a 17.4-million-molecule ZINC stock (Irwin et al., 2020). All methods are evaluated with this configuration. A solve is therefore a planner outcome under a fixed budget and stock, rather than experimental evidence of synthesis.

5.2 EVALUATION PROTOCOL

Docking evaluation. We dock each generated molecule with AutoDock-GPU using fixed target-specific grids (Morris et al., 2009; Santos-Martins et al., 2021). Each evaluation uses one run with a fixed seed. We report the AutoDock4 energy of the best-ranked pose in kcal/mol; lower scores are preferred, and positive energies are clipped to zero. For each target, we specify a docking-score threshold τ to operationally define a "good binder", which gives $\tau = -6.61$, -7.58, and -8.06 kcal/mol for ACAA1, c-MET, and BRD4(2), respectively. These computational scores are proxies for binding, not measured affinities.

5.3 SYNTHESIZABILITY AND BINDING AFFINITY

Setup. Each method generates three seeds of 1000 molecules. After deduplicating, the number evaluated may be fewer than 1,000. All reported fractions and mean docking scores are computed over the resulting unique molecules for each group. We use the same docking grids and AiZynthFinder configuration for every method. Unguided GenMol is the backbone control; SAFE-GPT (Noutahi et al., 2024) uses the same SAFE representation with autoregressive generation. TAGMol (Dorna et al., 2024) uses property predictors to guide a protein-conditioned 3D diffusion model during sampling. RxnFlow (Seo et al., 2025) uses a GFlowNet to assemble molecules through reaction templates and purchasable building blocks. ToolMol (Zhou et al., 2026) combines a multi-objective genetic algorithm with an LLM agent that modifies molecules using chemical tools. We report two SynGuide settings: RA-only guidance ($\lambda_{\text{syn}}{=}1$) and combined RA and docking guidance ($\lambda_{\text{syn}}{=}0.3$). Dock. is the mean score over evaluated molecules; Synth. is the fraction solved by AiZynthFinder; and Hit is the fraction both solved and below the target-specific docking threshold. Table 1 reports the mean and standard deviation across generation seeds for each target. Section C.2 sweeps τ and shows that the ordering of the methods is stable across the range considered.

Results. Table 1 shows the main results. With RA-only guidance, SynGuide increases the solve rate from 70.7% to 87.9% on BRD4(2), from 70.8% to 87.5% on ACAA1, and from 70.7% to 87.9%

Table 1: De novo molecular design results across three protein targets. Dock. is the mean AutoDock4 score (kcal/mol, ↓); Synth. is the AiZynthFinder solve rate (%, ↑); and Hit is the percentage both solved and below the target-specific docking threshold (↑). Best results are **Bold** and second-best are underlined.

	BRD4(2)			ACAA1			c-MET		
Method	**Dock. (↓)**	**Synth. (↑)**	**Hit (↑)**	**Dock. (↓)**	**Synth. (↑)**	**Hit (↑)**	**Dock. (↓)**	**Synth. (↑)**	**Hit (↑)**
GenMol	$-10.50_{\pm 0.19}$	$\underline{70.7}_{\pm 0.5}$	$18.6_{\pm 1.1}$	$-8.81_{\pm 0.29}$	$70.8_{\pm 0.6}$	$20.9_{\pm 0.6}$	$-9.85_{\pm 0.24}$	$70.7_{\pm 0.5}$	$18.8_{\pm 0.4}$
TAGMol	$-12.50_{\pm 0.25}$	$19.3_{\pm 0.4}$	$0.1_{\pm 0.1}$	$-10.49_{\pm 0.52}$	$4.0_{\pm 0.6}$	$0.5_{\pm 0.5}$	$-11.82_{\pm 0.24}$	$1.3_{\pm 0.1}$	$0.2_{\pm 0.0}$
SAFE-GPT	$-12.48_{\pm 0.35}$	$58.0_{\pm 2.2}$	$\underline{40.4}_{\pm 2.4}$	$-9.45_{\pm 0.18}$	$57.8_{\pm 2.4}$	$\underline{32.0}_{\pm 2.0}$	$-11.31_{\pm 0.38}$	$58.0_{\pm 2.2}$	$\underline{36.8}_{\pm 1.3}$
RxnFlow	$\underline{-13.03}_{\pm 0.23}$	$17.9_{\pm 1.5}$	$16.0_{\pm 1.5}$	$\underline{-10.31}_{\pm 0.45}$	$13.7_{\pm 2.0}$	$11.5_{\pm 1.2}$	$\underline{-12.09}_{\pm 0.06}$	$12.1_{\pm 0.6}$	$11.0_{\pm 0.4}$
ToolMol	$\mathbf{-13.44}_{\pm 0.15}$	$34.4_{\pm 7.1}$	$19.0_{\pm 5.2}$	$\mathbf{-10.89}_{\pm 0.37}$	$27.6_{\pm 5.4}$	$13.5_{\pm 2.6}$	$\mathbf{-12.96}_{\pm 0.27}$	$32.3_{\pm 2.4}$	$18.6_{\pm 2.2}$
SynGuide (λ_{syn}=1)	$-10.25_{\pm 0.25}$	$\mathbf{87.9}_{\pm 1.5}$	$19.3_{\pm 0.6}$	$-8.65_{\pm 0.32}$	$\mathbf{87.5}_{\pm 1.6}$	$24.1_{\pm 1.4}$	$-9.68_{\pm 0.76}$	$\mathbf{87.9}_{\pm 1.5}$	$20.9_{\pm 2.5}$
SynGuide (λ_{syn}=0.3)	$-10.47_{\pm 0.13}$	$70.4_{\pm 1.5}$	$\mathbf{45.6}_{\pm 1.7}$	$-9.01_{\pm 0.21}$	$\underline{78.0}_{\pm 0.7}$	$\mathbf{52.2}_{\pm 1.9}$	$-9.99_{\pm 0.18}$	$\underline{72.7}_{\pm 0.6}$	$\mathbf{47.4}_{\pm 0.9}$

on c-MET. Its mean docking score becomes less favorable on all three targets, exposing the cost of optimizing only for retrosynthetic accessibility.

With combined guidance, the hit rate rises from 18.6% to 45.6% on BRD4(2), from 20.9% to 52.2% on ACAA1, and from 18.8% to 47.4% on c-MET. These gains of 27.0%, 31.3%, and 28.6% exceed the hit rates of the other listed methods. The combined setting improves solve rate over unguided GenMol on ACAA1 and c-MET, while its BRD4(2) solve rate (70.4%) is similar to the unguided value (70.7%).

The baselines illustrate different trade-offs. ToolMol has the lowest mean docking score on each target, whereas SAFE-GPT achieves the next-highest hit rate despite solve rates near 58%. TAGMol and RxnFlow obtain favorable docking scores but substantially lower AiZynthFinder solve rates under our evaluation. The planner evaluation measures routes found by the specified AiZynthFinder configuration; it does not establish that a route constructed or proposed by another method is experimentally infeasible.

Table 2: Guidance mechanisms on frozen GenMol backbone with RA as the sole target (λ_{syn}=1). Time is wall-clock seconds (↓), Validity and Synth. are percentages (↑), and the final two columns report the mean score of the ten strongest docked molecules and the mean score over all molecules for c-MET (kcal/mol, ↓). Values are mean ± standard deviation over three seeds.

Method	**Time (↓)**	**Validity (↑)**	**Synth. (↑)**	**Top-10 (↓)**	**Mean (↓)**
Unguided	$53.0_{\pm 1.0}$	$99.9_{\pm 0.1}$	$71.3_{\pm 0.7}$	$-9.46_{\pm 0.21}$	$-7.05_{\pm 0.05}$
Autoguidance	$62.6_{\pm 1.0}$	$99.0_{\pm 0.4}$	$65.1_{\pm 1.9}$	$-9.31_{\pm 0.07}$	$-6.17_{\pm 0.06}$
SVDD	$103.9_{\pm 1.0}$	$99.7_{\pm 0.2}$	$\mathbf{91.8}_{\pm 0.4}$	$\mathbf{-9.36}_{\pm 0.04}$	$\underline{-6.98}_{\pm 0.03}$
D-CBG	$\mathbf{67.0}_{\pm 1.7}$	$\underline{99.9}_{\pm 0.1}$	$74.4_{\pm 1.2}$	$\underline{-9.32}_{\pm 0.19}$	$\mathbf{-7.05}_{\pm 0.04}$
SynGuide	$\underline{68.9}_{\pm 3.8}$	$99.7_{\pm 0.1}$	$\underline{87.3}_{\pm 1.0}$	$-9.13_{\pm 0.12}$	$-6.86_{\pm 0.03}$

5.4 Guidance mechanisms

Setup. We next isolate how the distilled signal is used during sampling. we use the same frozen GenMol backbone and sampling schedule throughout and guide only toward RA (λ_{syn}=1). SVDD (Li et al., 2025), D-CBG (Schiff et al., 2025), and SynGuide use the same planner-supervised predictor. Autoguidance (Karras et al., 2024) does not use this predictor and tests whether modifying the backbone's own predictions improves the same outcomes. We compare:

- **Autoguidance**. Extrapolate away from a degraded copy of the backbone without consulting the RA predictor (more detailed description in Appendix §E).
- **SVDD**. At each denoising step, draw M candidate continuations and retain the one with the highest predicted value, without using gradients. We follow the reference implementation with M=10 and argmax selection.
- **D-CBG**. A first-order Taylor expansion of $\log p(y \mid z_t)$ converts one backward pass into an additive, per-token logit bias.

Results. Table 2 compares wall-clock cost, validity, solve rate, and docking under RA-only guidance. SynGuide raises the solve rate from 71.3% for unguided sampling to 87.3%, at 68.9 seconds versus 53.0 seconds. SVDD achieves the highest solve rate (91.8%) but takes 103.9 seconds. D-CBG uses the same predictor and takes a similar 67.0 seconds, yet reaches 74.4%. Thus, differentiating through the denoiser's soft clean prediction yields a 12.9-point higher solve rate than guidance on the current masked state at comparable wall-clock cost. Validity remains at least 99.7% for all three predictor-based methods.

Autoguidance reaches 65.1% solved molecules, below the unguided rate, and also lowers validity to 99.0%. This control does not support a gain from modifying the backbone without planner supervision in this setting. The c-MET docking scores are reported to expose collateral changes; none of the methods in this comparison receives docking guidance. SynGuide's mean score changes from -7.05 to -6.86 kcal/mol, so its greater solve rate does not imply improved docking.

5.5 Cross-planner Evaluation

Setup. Since the RA predictor learns from AiZynthFinder outcomes, an improvement under AiZynthFinder alone could reflect planner-specific optimization. We therefore evaluate the same generated molecules with an independent configuration of Syntheseus (Maziarz et al., 2024) and the template-free RootAligned single-step model (Zhong et al., 2022). Both evaluations use the same 17.4-million-molecule ZINC stock and match building blocks by InChIKey. We rank molecules by their docking scores and compute the fraction solved within each top-k subset.

Results. Table 3 reports means and standard deviations across nine target–seed groups (3 targets $\times$ 3 seeds). Among the top 50 docked molecules, RA-only SynGuide increases the AiZynthFinder solve rate from 45.8% to 81.6% relative to unguided generation. The corresponding Syntheseus rates are 96.0% and 99.8%. The same direction of change appears among the top 10: 35.6% to 74.4% under AiZynthFinder and 93.3% to 100.0% under Syntheseus. Docking-only guidance lowers AiZynthFinder Top-50 solve rate to 28.2%.

Table 3: Cross-planner solve rates (%, ↑) among the k molecules with the lowest docking scores. Values are mean ± standard deviation over three target $times$ three seeds. "RA guidance" indicates whether retrosynthetic accessibility guided generation. **Bold** marks the best filled row per column.

		AiZynthFinder		Syntheseus	
Generation method	**RA guidance**	**Top-10**	**Top-50**	**Top-10**	**Top-50**
Unguided	✗	$35.6_{\pm 14.2}$	$45.8_{\pm 6.0}$	$93.3_{\pm 7.1}$	$96.0_{\pm 5.3}$
Docking-only	✗	$25.6_{\pm 14.2}$	$28.2_{\pm 6.7}$	$96.7_{\pm 5.0}$	$94.0_{\pm 5.7}$
SynGuide ($\lambda_{\text{syn}} = 1$)	✓	$\mathbf{74.4}_{\pm 12.4}$	$\mathbf{81.6}_{\pm 5.8}$	$\mathbf{100.0}_{\pm 0.0}$	$\mathbf{99.8}_{\pm 0.7}$

The held-out planner solves most molecules in every row, leaving limited headroom to detect a difference. Its high baseline and the variation across target–seed groups preclude a strong claim about the size or statistical reliability of the transferred gain. The results nevertheless show that the RA-guided samples retain high route-found rates under a different planning configuration.

5.6 Ablation Study

The objective weight λ_{syn}. We vary λ_{syn} while keeping $\lambda_{\text{bind}} = 1 - \lambda_{\text{syn}}$ and all other sampling settings fixed. Figure 3 plots solve rate against mean docking score. Increasing the RA weight generally improves solve rate while reducing docking performance. Intermediate weights permit a more balanced outcome. This suggests that the learned generator achieves favorable joint performance beyond simple linear interpolation between the two metrics of interest. The $\lambda_{\text{syn}} = 0.4$ operating point for ACAA1 is a notable exception, exhibiting a local inward deviation and indicating less favorable joint performance at this weighting.

The docking-only endpoint ($\lambda_{\text{syn}} = 0$) also differs by target. Approximately 60% of its ACAA1 molecules are solved, compared with approximately 40% for c-MET and BRD4(2); adding RA weight can therefore recover a larger fraction of solved molecules for the latter two targets. The frontier summarizes a trade-off in computational proxies and does not establish experimental synthesis or affinity. Appendix F gives representative molecules.

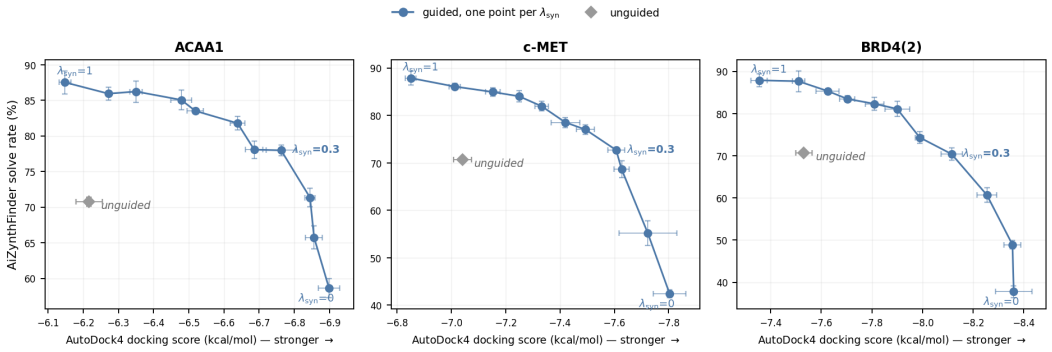

Figure 3: AiZynthFinder solve rate versus mean AutoDock4 score as the RA weight λ_{syn} varies, with one panel per target. Blue points denote guided settings and the gray diamond denotes the unguided model. Error bars are Wilson 95% intervals on the solve rate and standard errors on the mean AutoDock score. Each setting generates 1,000 molecules per seed before deduplication.

The planner's search budget B. We assess whether the relative solve rates depend on the planner budget used for evaluation. For each c-MET generation setting, we evaluate the same molecules at budgets up to $B = 500$, using the first iteration at which AiZynthFinder found a route to recover the outcome at each smaller budget. Figure 4 shows the solve rate for different RA guidance weights, alongside unguided GenMol; the vertical line marks the $B = 500$ budget used in our main results. Solve rates rise with B for every setting, but stronger RA guidance generally maintains a higher solve rate than unguided generation throughout the plotted range. The gain at $B = 500$ therefore does not arise solely from the choice of that evaluation budget. These comparisons hold the planner's reaction model and building-block stock fixed. Full results for two other targets can be found in Appendix C.1.

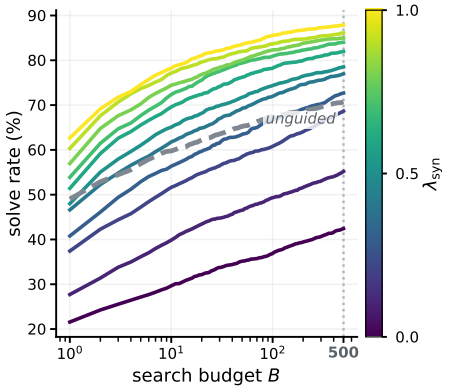

Figure 4: AiZynthFinder solve rate for the same c-MET molecules evaluated at different search budgets B.

6 Conclusion

We introduce *SynGuide*, a planner-supervised discrete diffusion guidance that trains a noise-conditioned predictor on fixed-budget retrosynthesis outcomes, and uses it to steer a frozen masked discrete diffusion model. Across three protein targets, RA-only increases the measured planner solve rate by 24% and more than doubles the hit ratio. Guiding through the denoiser's predicted clean sequence also yields a higher solve rate than guiding on the current masked state at comparable computational cost. These results show that planner outcomes collected offline can inform each sampling step without additional planner calls or backbone retraining , providing an efficient and modular route to synthesizability-aware molecular design.

Docking performance remains a limitation in the multi-objective setting. A pocket-aware backbone or stronger affinity predictor may improve target-specific guidance, while absolute binding free energy calculations (Feng et al., 2022, ABFE) and experimental validation could provide stronger evidence for selected candidates. Applying planner-supervised guidance to other generative frameworks, including flow matching, would further test its generality and require guidance rules suited to their intermediate states and generation dynamics.

AI USE STATEMENT

In this work, we used generative AI tools for minor editing of academic writing, such as drafting and refining sections. We have reviewed all AI-assisted work. We have not used generative AI tools for scientific claims, and methodological contributions. We take responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This work adheres to the ICLR Code of Ethics. We recognize that the methods presented here, like most advances in machine learning, could potentially be applied in ways that cause harm, for example, if deployed without adequate evaluation on the generated molecule samples. We encourage practitioners to carefully assess potential misuse before deploying systems built on this work.

REPRODUCIBILITY STATEMENT

We have taken several steps to ensure the reproducibility of our work. We use public available GenMol V2 checkpoint as the frozen backbone. The implementation details of our proposed framework, including model configurations, training protocols, and evaluation metrics, are described in Section 5. The CASP planner configuration (budget, policies, stock) and the docking protocol are also specified in Section 5. To facilitate replication, we release the source code and experiment scripts in supplementary material. Additional information is also included in the appendix. Together, these resources ensure that our results can be independently verified and extended.

REFERENCES

Jacob Austin, Daniel D Johnson, Jonathan Ho, Daniel Tarlow, and Rianne Van Den Berg. Structured denoising diffusion models in discrete state-spaces. *Advances in neural information processing systems*, 34:17981–17993, 2021.

Emmanuel Bengio, Moksh Jain, Maksym Korablyov, Doina Precup, and Yoshua Bengio. Flow network based generative models for non-iterative diverse candidate generation. *Advances in neural information processing systems*, 34:27381–27394, 2021.

Shuan Chen and Yousung Jung. Estimating the synthetic accessibility of molecules with building block and reaction-aware sascore. *Journal of cheminformatics*, 16(1):83, 2024.

Trevor Chen, Ariel Dai, Jason Yang, Riccardo De Santi, Daniel Khalil, Wenda Chu, Nate Gruver, Pranav Murugan, Alexander FG Goldberg, Maruan Al-Shedivat, et al. On the design space of discrete diffusion online adaptation for molecular optimization. *arXiv preprint arXiv:2607.02834*, 2026.

Oh-Hyeon Choung, Riccardo Vianello, Marwin Segler, Nikolaus Stiefl, and José Jiménez-Luna. Extracting medicinal chemistry intuition via preference machine learning. *Nature Communications*, 14(1):6651, 2023.

Connor W Coley, Luke Rogers, William H Green, and Klavs F Jensen. Scscore: synthetic complexity learned from a reaction corpus. *Journal of chemical information and modeling*, 58(2): 252–261, 2018.

Vineeth Dorna, D Subhalingam, Keshav Kolluru, Shreshth Tuli, Mrityunjay Singh, Saurabh Singal, NM Krishnan, and Sayan Ranu. Tagmol: target-aware gradient-guided molecule generation. *arXiv preprint arXiv:2406.01650*, 2024.

Peter Eckmann, Kunyang Sun, Bo Zhao, Mudong Feng, Michael K Gilson, and Rose Yu. Limo: Latent inceptionism for targeted molecule generation. *Proceedings of machine learning research*, 162:5777, 2022.

Peter Eckmann, Dongxia Wu, Germano Heinzelmann, Michael K Gilson, and Rose Yu. MF-LAL: Drug compound generation using multi-fidelity latent space active learning. In *Forty-second International Conference on Machine Learning*, 2025. URL `https://openreview.net/forum?id=3xdFvqsVnM`.

Peter Ertl and Ansgar Schuffenhauer. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *Journal of cheminformatics*, 1(1):8, 2009.

Mudong Feng, Germano Heinzelmann, and Michael K Gilson. Absolute binding free energy calculations improve enrichment of actives in virtual compound screening. *Scientific Reports*, 12(1): 13640, 2022.

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.

Wenhao Gao, Tianfan Fu, Jimeng Sun, and Connor Coley. Sample efficiency matters: a benchmark for practical molecular optimization. *Advances in neural information processing systems*, 35: 21342–21357, 2022a.

Wenhao Gao, Rocío Mercado, and Connor W Coley. Amortized tree generation for bottom-up synthesis planning and synthesizable molecular design. *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.

Samuel Genheden, Amol Thakkar, Veronika Chadimová, Jean-Louis Reymond, Ola Engkvist, and Esben Bjerrum. Aizynthfinder: a fast, robust and flexible open-source software for retrosynthetic planning. *Journal of cheminformatics*, 12(1):70, 2020.

Bartosz A Grzybowski, Sara Szymkuć, Ewa P Gajewska, Karol Molga, Piotr Dittwald, Agnieszka Wołos, and Tomasz Klucznik. Chematica: a story of computer code that started to think like a chemist. *Chem*, 4(3):390–398, 2018.

Jiaqi Guan, Wesley Wei Qian, Xingang Peng, Yufeng Su, Jian Peng, and Jianzhu Ma. 3d equivariant diffusion for target-aware molecule generation and affinity prediction. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=kJqXEPXMsE0.

Jeff Guo and Philippe Schwaller. Directly optimizing for synthesizability in generative molecular design using retrosynthesis models. *Chemical Science*, 16:6943–6956, 2025. doi: 10.1039/D5SC01476J.

Alan P Hill and Robert J Young. Getting physical in drug discovery: a contemporary perspective on solubility and hydrophobicity. *Drug discovery today*, 15(15-16):648–655, 2010.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

John J Irwin, Khanh G Tang, Jennifer Young, Chinzorig Dandarchuluun, et al. ZINC20 – a free ultralarge-scale chemical database for ligand discovery. *Journal of Chemical Information and Modeling*, 60(12):6065–6073, 2020.

Laura Isigkeit, Tim Hörmann, Espen Schallmayer, Katharina Scholz, Felix F Lillich, Johanna HM Ehrler, Benedikt Hufnagel, Jasmin Büchner, Julian A Marschner, Jörg Pabel, et al. Automated design of multi-target ligands by generative deep learning. *Nature Communications*, 15(1):7946, 2024.

Wengong Jin, Regina Barzilay, and Tommi Jaakkola. Junction tree variational autoencoder for molecular graph generation. In *International conference on machine learning*, pp. 2323–2332. PMLR, 2018.

Seokho Kang and Kyunghyun Cho. Conditional molecular design with deep generative models. *Journal of chemical information and modeling*, 59(1):43–52, 2019.

Tero Karras, Miika Aittala, Tuomas Kynkäänniemi, Jaakko Lehtinen, Timo Aila, and Samuli Laine. Guiding a diffusion model with a bad version of itself. *Advances in Neural Information Processing Systems*, 37:52996–53021, 2024.

Gregory W Kyro, Matthew T Martin, Eric D Watt, and Victor S Batista. Cardiogenai: a machine learning-based framework for re-engineering drugs for reduced herg liability. *Journal of Cheminformatics*, 17(1):30, 2025.

Seul Lee, Karsten Kreis, Srimukh Prasad Veccham, Meng Liu, Danny Reidenbach, Yuxing Peng, Saee Paliwal, Weili Nie, and Arash Vahdat. Genmol: A drug discovery generalist with discrete diffusion. *arXiv preprint arXiv:2501.06158*, 2025.

Xiner Li, Yulai Zhao, Chenyu Wang, Gabriele Scalia, Gokcen Eraslan, Surag Nair, Tommaso Biancalani, Shuiwang Ji, Aviv Regev, Sergey Levine, and Masatoshi Uehara. Derivative-free guidance in continuous and discrete diffusion models with soft value-based decoding. In D. Belgrave, C. Zhang, H. Lin, R. Pascanu, P. Koniusz, M. Ghassemi, and N. Chen (eds.), *Advances in Neural Information Processing Systems*, volume 38, Main Conference, pp. 95507–95545. Curran Associates, Inc., 2025. doi: 10.52202/085713-3194. URL https://proceedings.neurips.cc/paper_files/paper/2025/file/899af0d66d8850318a20781484416152-Paper-Conference.pdf.

Jaechang Lim, Seongok Ryu, Jin Woo Kim, and Woo Youn Kim. Molecular generative model based on conditional variational autoencoder for de novo molecular design. *Journal of cheminformatics*, 10(1):31, 2018.

Songtao Liu, Dandan Zhang, Zhengkai Tu, Hanjun Dai, and Peng Liu. Evaluating molecule synthesizability via retrosynthetic planning and reaction prediction. *arXiv preprint arXiv:2411.08306*, 2024.

Xuhan Liu, Kai Ye, Herman WT van Vlijmen, Michael TM Emmerich, Adriaan P IJzerman, and Gerard JP van Westen. Drugex v2: de novo design of drug molecules by pareto-based multi-objective reinforcement learning in polypharmacology. *Journal of cheminformatics*, 13(1):85, 2021.

Krzysztof Maziarz, Austin Tripp, Guoqing Liu, Megan Stanley, Shufang Xie, Piotr Gainski, Philipp Seidl, and Marwin Segler. Re-evaluating retrosynthesis algorithms with syntheseus. *Faraday Discussions*, 2024.

Garrett M. Morris, Ruth Huey, William Lindstrom, Michel F. Sanner, Richard K. Belew, David S. Goodsell, and Arthur J. Olson. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, 30(16):2785–2791, 2009.

Hunter Nisonoff, Junhao Xiong, Stephan Allenspach, and Jennifer Listgarten. Unlocking guidance for discrete state-space diffusion and flow models. In *International Conference on Learning Representations*, volume 2025, pp. 36052–36106, 2025.

Emmanuel Noutahi, Cristian Gabellini, Michael Craig, Jonathan SC Lim, and Prudencio Tossou. Gotta be safe: a new framework for molecular design. *Digital Discovery*, 3(4):796–804, 2024.

Sven M. Papidocha, Andreas Burger, Varinia Bernales, and Alán Aspuru-Guzik. The elephant in the lab: synthesizability in generative small-molecule design. *Current Opinion in Chemical Engineering*, 51:101217, 2026.

Xingang Peng, Shitong Luo, Jiaqi Guan, Qi Xie, Jian Peng, and Jianzhu Ma. Pocket2mol: Efficient molecular sampling based on 3d protein pockets. In *International conference on machine learning*, pp. 17644–17655. PMLR, 2022.

Diogo Santos-Martins, Leonardo Solis-Vasquez, Andreas F Tillack, Michel F Sanner, Andreas Koch, and Stefano Forli. Accelerating autodock4 with gpus and gradient-based local search. *Journal of chemical theory and computation*, 17(2):1060, 2021.

Yair Schiff, Subham Sahoo, Hao Phung, Guanghan Wang, Sam Boshar, Hugo Dalla-Torre, Bernardo Almeida, Alexander Rush, Thomas Pierrot, and Volodymyr Kuleshov. Simple guidance mechanisms for discrete diffusion models. In *International Conference on Learning Representations*, volume 2025, pp. 43776–43821, 2025.

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

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 *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=pB1XSj2y4X.

Chence Shi*, Minkai Xu*, Zhaocheng Zhu, Weinan Zhang, Ming Zhang, and Jian Tang. Graphaf: a flow-based autoregressive model for molecular graph generation. In *International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=S1esMkHYPr.

Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *International conference on machine learning*, pp. 2256–2265. pmlr, 2015.

Sara Szymkuć, Ewa P Gajewska, Tomasz Klucznik, Karol Molga, Piotr Dittwald, Michał Startek, Michał Bajczyk, and Bartosz A Grzybowski. Computer-assisted synthetic planning: the end of the beginning. *Angewandte Chemie International Edition*, 55(20):5904–5937, 2016.

Amol Thakkar, Thierry Kogej, Jean-Louis Reymond, Ola Engkvist, and Esben Jannik Bjerrum. Datasets and their influence on the development of computer assisted synthesis planning tools in the pharmaceutical domain. *Chemical science*, 11(1):154–168, 2019.

Amol Thakkar, Veronika Chadimová, Esben Jannik Bjerrum, Ola Engkvist, and Jean-Louis Reymond. Retrosynthetic accessibility score (rascore)–rapid machine learned synthesizability classification from ai driven retrosynthetic planning. *Chemical science*, 12(9):3339–3349, 2021.

Zhengkai Tu, Sourabh J. Choure, Mun Hong Fong, Jihye Roh, Itai Levin, Kevin Yu, Joonyoung F. Joung, Nathan Morgan, Shih-Cheng Li, Xiaoqi Sun, Huiqian Lin, Mark Murnin, Jordan P. Liles, Thomas J. Struble, Michael E. Fortunato, Mengjie Liu, William H. Green, Klavs F. Jensen, and Connor W. Coley. Askcos: an open source software suite for synthesis planning, 2025. URL https://arxiv.org/abs/2501.01835.

Milan Voršilák, Michal Kolář, Ivan Čmelo, and Daniel Svozil. Syba: Bayesian estimation of synthetic accessibility of organic compounds. *Journal of cheminformatics*, 12(1):35, 2020.

W Patrick Walters and Mark A Murcko. Prediction of 'drug-likeness'. *Advanced drug delivery reviews*, 54(3):255–271, 2002.

Yutong Xie, Chence Shi, Hao Zhou, Yuwei Yang, Weinan Zhang, Yong Yu, and Lei Li. {MARS}: Markov molecular sampling for multi-objective drug discovery. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=kHSu4ebxFXY.

Jiaxuan You, Bowen Liu, Zhitao Ying, Vijay Pande, and Jure Leskovec. Graph convolutional policy network for goal-directed molecular graph generation. *Advances in neural information processing systems*, 31, 2018.

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

Andrew Y Zhou, Sharvaree Vadgama, Sumanth Varambally, Peter Eckmann, Michael K Gilson, and Rose Yu. Toolmol: Evolutionary agentic framework for multi-objective drug discovery. *arXiv preprint arXiv:2605.12784*, 2026.

Zhenpeng Zhou, Steven Kearnes, Li Li, Richard N Zare, and Patrick Riley. Optimization of molecules via deep reinforcement learning. *Scientific reports*, 9(1):10752, 2019.

A GUIDANCE FORMULATION AND IMPLEMENTATION

A.1 FROM RATE REWEIGHTING TO A TOKEN-LOGIT UPDATE

For a target event y and a candidate jump from x_t to $\tilde{x}$, discrete predictor guidance reweights the unconditional transition rate by a likelihood ratio (Nisonoff et al., 2025; Schiff et al., 2025):

$$R_t^{(y)}(x_t, \tilde{x}) = R_t(x_t, \tilde{x}) \exp\left\{ \gamma \sum_k \lambda_k \left[\log p_k(y_k \mid \tilde{x}, t) - \log p_k(y_k \mid x_t, t) \right] \right\}. \tag{12}$$

Here $\gamma = 1/T$ is the inverse guidance temperature and λ_k is the weight assigned to objective k. We use two Bernoulli events: a route is found within $B = 500$ planner iterations, and the target-specific docking threshold is met. Both therefore contribute log probabilities in nats rather than mixing a binary likelihood with a regression score.

Evaluating the likelihood difference for every token substitution would require one predictor call per candidate. Taylor-approximated guidance (TAG) replaces these evaluations with a first-order expansion around a differentiable one-hot state, so one backward pass yields a bias for every token. Our implementation uses the denoiser's clean prediction rather than the current masked state. If ℓ_t are the denoiser logits, $q_t = \mathrm{softmax}(\ell_t)$, and M_t marks positions that remain masked, the predictor input is

$$\hat{x}_0 = M_t \odot q_t + (1 - M_t) \odot \mathrm{onehot}(x_t). \tag{13}$$

For objective k, we compute $g_k = \nabla_{\ell_t} \log p_k(y_k = 1 \mid \hat{x}_0, t = 1)$ and normalize it by its root-mean-square magnitude over masked token–vocabulary coordinates, $\bar{g}_k = g_k / (\mathrm{RMS}_{M_t}(g_k) + \epsilon)$. The update used in all main experiments is

$$\ell_t^{\mathrm{guided}} = \ell_t + \gamma s \, M_t \odot \sum_k \lambda_k \bar{g}_k, \tag{14}$$

with $\gamma = 2$, $s = 0.028$, and $\epsilon = 10^{-12}$. The mask ensures that an absorbing-state diffusion process never revises a token after it has been unmasked. Normalizing each objective before summation prevents the sharper classifier from dominating solely because its raw gradients are larger.

A.2 WEIGHTS AND SAMPLING-TIME COMPOSITION

We constrain $\lambda_{\mathrm{syn}} + \lambda_{\mathrm{dock}} = 1$. The main method uses $(\lambda_{\mathrm{syn}}, \lambda_{\mathrm{dock}}) = (0.3, 0.7)$. Synthesis-only mechanism comparisons set $\lambda_{\mathrm{syn}} = 1$, while docking-only generation sets $\lambda_{\mathrm{syn}} = 0$. The GenMol parameters are frozen in every case. Adding an objective therefore changes neither the backbone nor its training data: it adds one predictor, one normalized gradient term in Equation 14, and one scalar weight.

B NOISE-CONDITIONED PREDICTORS

Architecture. The RA predictor is a Bernoulli classifier $p_\theta(y_{\mathrm{RA}} = 1 \mid x_t, t)$. Token and positional embeddings are summed with a two-layer SiLU time embedding and passed through a three-layer, four-head Transformer encoder with hidden width 192, feed-forward width 768, dropout 0.1, and maximum sequence length 128. Mean pooling over non-padding tokens is followed by a two-layer MLP that outputs one logit. The target-specific docking predictors use the same architecture and classify whether $\mathrm{ds}_{\min} \leq \tau$.

Training data and corruption. The RA training set contains 9,321 GenMol-V2 molecules labelled by AiZynthFinder, of which 72.4% satisfy $y_{\mathrm{RA}} = 1$. We use a stratified 90/10 train–validation split. Each training example retains the label of its complete molecule but presents a partially masked bracket-SAFE sequence to the network: 20% of draws are clean and the remainder sample a continuous noise level uniformly from $[0, 1]$ under GenMol's masking process. Half of the batches are represented as one-hot tensors so that training also covers the differentiable inputs used by TAG. We optimize class-weighted binary cross-entropy for 20 epochs with AdamW, batch size 64, learning rate 3×10^{-4}, weight decay 10^{-4}, cosine decay, and gradient clipping at 1.0.

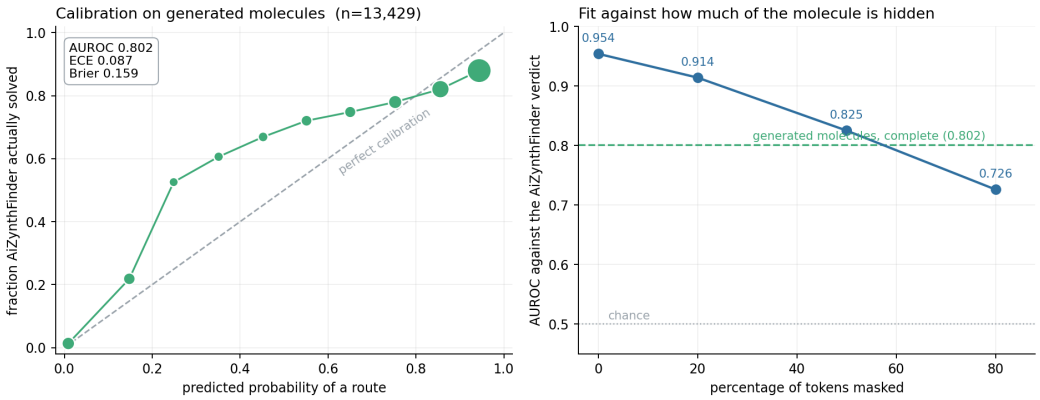

Figure 5: RA predictor diagnostics. Left: calibration on 13,429 generated molecules spanning unguided and guided conditions, all labelled by AiZynthFinder. Marker area is proportional to the number of molecules in each probability bin. Right: held-out AUROC as an increasing fraction of the SAFE tokens is masked. Complete-sequence discrimination is high (AUROC 0.954), and the predictor remains informative with 50% and 80% masking (0.825 and 0.726). On generated complete molecules the distribution-shifted AUROC is 0.802, with ECE 0.087 and Brier score 0.159.

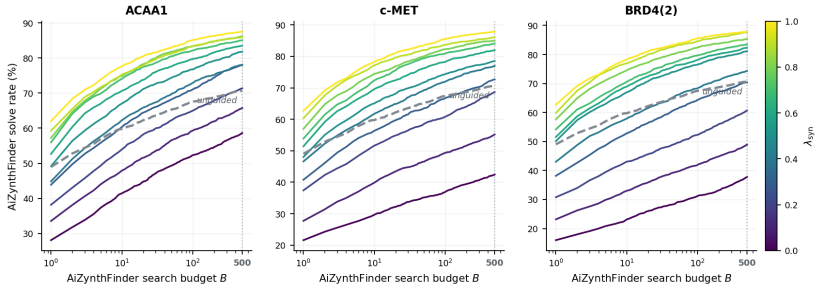

Figure 6: AiZynthFinder solve rate as a function of evaluation search budget for ACAA1, c-MET, and BRD4(2). Each curve evaluates a fixed set of generated molecules; only the planner budget B varies. Because searches run to $B = 500$ record when the first route is found, outcomes at smaller budgets can be recovered without rerunning the planner. The dotted line marks the $B = 500$ budget used to define the training label.

Diagnostic interpretation. The predictor is most accurate near the end of denoising, when most molecular structure is visible, but remains above chance early enough to provide a directional signal. The calibration curve also shows that the predictor is not a replacement for final planner evaluation: on newly generated molecules its probabilities are informative but imperfect. We therefore use it only for sampling and compute every reported Solve value with AiZynthFinder.

C Additional sensitivity analyses

C.1 Planner search budget

Figure 6 tests whether the observed accessibility gains depend on evaluating molecules at the same search budget used to define the RA label. Solve rates increase with B for every generation condition, as additional search allows the planner to find more routes. Nevertheless, conditions with $\lambda_{\text{syn}} \geq 0.7$ remain above unguided generation across the displayed budget range for all three targets. Smaller synthesis weights show a more variable relationship with unguided generation, including curve crossings. Thus, the advantage of stronger RA guidance is visible under both smaller and larger evaluation budgets within the tested range, although the measured solve rate depends on the planner budget.

Table 4: Extended k for the three methods of Table 1, mean $\pm$ sd over three seeds, all rows at 100% planner-label coverage. Top-10 and Top-100 are the order statistics the synthesizable-design literature reports (Gao et al., 2022a), and are stable enough across seeds to resolve effects that the spread at $k \leq 5$ hides. Top-k in kcal/mol (↓); Solve and Joint in % (↑).

Target	Method	Top-1	Top-5	Top-10	Top-100	Solve	Joint
ACAA1	GenMol	$-8.81_{\pm.29}$	$-8.50_{\pm.15}$	$-8.29_{\pm.11}$	$-7.53_{\pm.02}$	$70.8_{\pm0.6}$	$20.9_{\pm0.6}$
	SAFE-GPT	$\mathbf{-9.45}_{\pm.18}$	$\mathbf{-9.20}_{\pm.18}$	$\mathbf{-8.99}_{\pm.12}$	$\mathbf{-8.14}_{\pm.05}$	$57.8_{\pm2.4}$	$32.0_{\pm2.0}$
	Ours	$-8.78_{\pm.26}$	$-8.52_{\pm.09}$	$-8.38_{\pm.08}$	$-7.79_{\pm.05}$	$\mathbf{78.1}_{\pm1.2}$	$\mathbf{50.1}_{\pm1.9}$
c-MET	GenMol	$-9.85_{\pm.24}$	$-9.63_{\pm.23}$	$-9.39_{\pm.15}$	$-8.60_{\pm.05}$	$70.7_{\pm0.5}$	$18.8_{\pm0.4}$
	SAFE-GPT	$\mathbf{-11.31}_{\pm.38}$	$\mathbf{-10.90}_{\pm.23}$	$\mathbf{-10.65}_{\pm.14}$	$\mathbf{-9.62}_{\pm.07}$	$58.0_{\pm2.2}$	$36.8_{\pm1.3}$
	Ours	$-9.90_{\pm.61}$	$-9.58_{\pm.20}$	$-9.44_{\pm.13}$	$-8.85_{\pm.01}$	$\mathbf{77.0}_{\pm1.0}$	$\mathbf{47.6}_{\pm3.0}$
BRD4(2)	GenMol	$-10.50_{\pm.19}$	$-10.07_{\pm.19}$	$-9.88_{\pm.13}$	$-9.06_{\pm.04}$	$70.7_{\pm0.5}$	$18.6_{\pm1.1}$
	SAFE-GPT	$\mathbf{-12.48}_{\pm.35}$	$\mathbf{-11.83}_{\pm.13}$	$\mathbf{-11.53}_{\pm.12}$	$\mathbf{-10.49}_{\pm.04}$	$58.0_{\pm2.2}$	$40.4_{\pm2.4}$
	Ours	$-10.57_{\pm.35}$	$-10.19_{\pm.19}$	$-10.06_{\pm.14}$	$-9.42_{\pm.04}$	$\mathbf{74.3}_{\pm1.5}$	$\mathbf{44.8}_{\pm0.5}$

C.2 Docking threshold

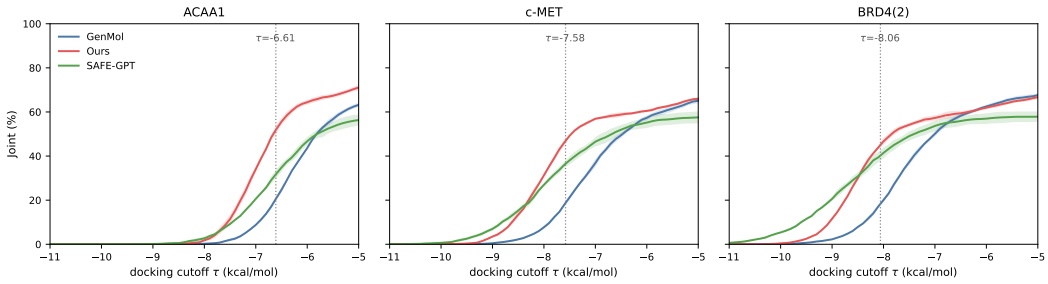

Figure 7: Sensitivity of Joint to the target-specific docking cutoff. The vertical dotted line marks the 42.15th-percentile threshold used in the main tables. Curves are recomputed from molecule-level docking scores and planner labels; no model is rerun when the cutoff changes.

As the cutoff is relaxed, all Joint rates rise because more solved molecules also qualify as docking hits. Across the displayed range, our curve remains above unguided GenMol on all three targets. The comparison with SAFE-GPT depends more strongly on the cutoff because SAFE-GPT combines stronger extreme docking scores with a lower overall solve rate. The main conclusion relative to the matched GenMol backbone is therefore not an artefact of the selected percentile.

D Extended k for the main comparison

Table 1 reports one docking statistic per target. This appendix reports four. SAFE-GPT holds every docking column on every target, and the margin widens rather than closes as k grows, so its advantage is a property of its population and not of a few outliers. Our method is ahead of unguided GenMol at $k = 100$ on all three targets, by 0.46, 0.37, and 0.42 kcal/mol on BRD4(2), ACAA1, and c-MET; at $k = 1$ the same comparison gives $+0.20$, $+0.14$, and -0.03, the last of which is a loss. Guidance therefore moves the bulk of the sample rather than its best member, which is the behaviour a population-level objective should produce and the reason the main table reports the mean.

The seed spread is what makes the small-k columns hard to use. Across the nine rows it falls from 0.13–0.38 kcal/mol at $k = 1$ to 0.00–0.07 at $k = 100$, so a $k = 1$ gap of two tenths of a kcal/mol sits inside the noise while the same gap at $k = 100$ does not. Top-10 and Top-100 are also the order statistics the synthesizable-design literature reports, which is why the comparison in §5.3 rests on them. Synth. and Joint do not vary with k and are repeated here only so each row can be read on its own.

Table 5: Comparison of Guidance mechanisms under synthesisability guidance (λ_{syn}=1). Every guidance mechanism reads the same planner-supervised predictor, with autoguidance as an exception. The last two columns are the docking score to protein target c-MET. Docking in kcal/mol (↓); Valid, Uniq and Solve in % (↑); Div is the mean pairwise Tanimoto distance over Morgan fingerprints (↑).

Mechanism	Time	Validity	Uniqueness	Diversity	Solve	Top-10	Mean
Unguided	53.0±1.0	99.9±0.1	94.9±1.1	0.868±0.011	71.3±0.7	−9.46±0.21	−7.05±0.05
Autoguidance	62.6±1.0	99.0±0.4	93.5±0.5	0.896±0.006	65.1±1.9	−9.31±0.07	−6.17±0.06
SVDD	103.9±1.0	99.7±0.2	88.9±1.6	0.857±0.004	**91.8**±0.4	**−9.36**±0.04	−6.98±0.03
D-CBG	**67.0**±1.7	99.9±0.1	**95.3**±0.7	**0.877**±0.019	74.4±1.2	−9.32±0.19	**−7.05**±0.04
Ours	68.9±3.8	99.7±0.1	94.0±0.3	0.846±0.008	87.3±1.0	−9.13±0.12	−6.86±0.03

E AUTOGUIDANCE AS A QUALITY-ONLY REFERENCE

Autoguidance replaces the unconditional model of CFG with a degraded version of the model itself, $D_w = wD_1 + (1-w)D_0$, both under identical conditioning. Our D_0 uses the same bracket-SAFE tokenizer and training corpus as D_1 but has one ninth as many parameters and is trained for one sixteenth as many steps. Matching tokenization is necessary because autoguidance combines the two models' token logits directly. We use $w = 1.5$ in Table 2.

Autoguidance carries no synthesizability signal, so in Table 2 it should be read against the unguided backbone rather than against the guided mechanisms: it tests whether better generation alone improves routability, and so bounds how much of any mechanism's gain could be explained without the planner. Its measured Solve rate is 65.1%, below the unguided value of 71.3%; the autoguidance row is marked as incomplete in Table 2, but the remaining coverage can change it by at most 3.3 points and cannot reverse that ordering. This negative result shows that planner guidance is not reproduced simply by extrapolating away from a smaller, undertrained copy of the generator.

F QUALITATIVE TOP-RANKED MOLECULES

For each target and synthesis weight λ_{syn}, we pool the generations from three seeds and deduplicate on canonical SMILES ($\sim$ 2,600 molecules per cell), then apply a six-stage cascade in which each stage targets a distinct failure mode of the generative model. We first remove structural liabilities with three complementary medicinal-chemistry filters: REOS (Walters & Murcko, 2002) to reject reactive and PAINS-like substructures (Glaxo rule set); a RingSystemLookup screen against a ChEMBL-derived reference to discard molecules containing unprecedented ring topologies; and MolSkill (Choung et al., 2023), a learned medicinal-chemistry preference model, thresholded at the 90th percentile of the training distribution so that 90% of reference drugs are retained by construction. We then impose two physicochemical constraints motivated by the oversized, lipophilic outliers observed among the raw top-ranked molecules: a longest-aliphatic-chain length $\leq$ 7 to suppress long greasy tails, and a Solubility Forecast Index (Hill & Young, 2010), $\text{SFI} = \text{clogD}_{7.4} + N_{\text{aromatic rings}} < 5$, where $\text{clogD}_{7.4}$ is predicted by a LightGBM model trained on ChEMBL, to enforce a favorable solubility profile. Finally, we require ground-truth synthetic accessibility: a molecule must admit a full retrosynthetic route to purchasable ZINC building blocks under AiZynthFinder within a budget of 500 MCTS iterations (USPTO + RingBreaker expansion policies, USPTO filter; the solved_{B500} criterion). Molecules passing all six stages constitute the filtered pool, from which we select the top-3 representatives by AutoDock docking score for display; this yields chemically sensible, soluble, and synthesizable exemplars in place of the raw docking-ranked top-3. Figures 8–13 visualize the results.

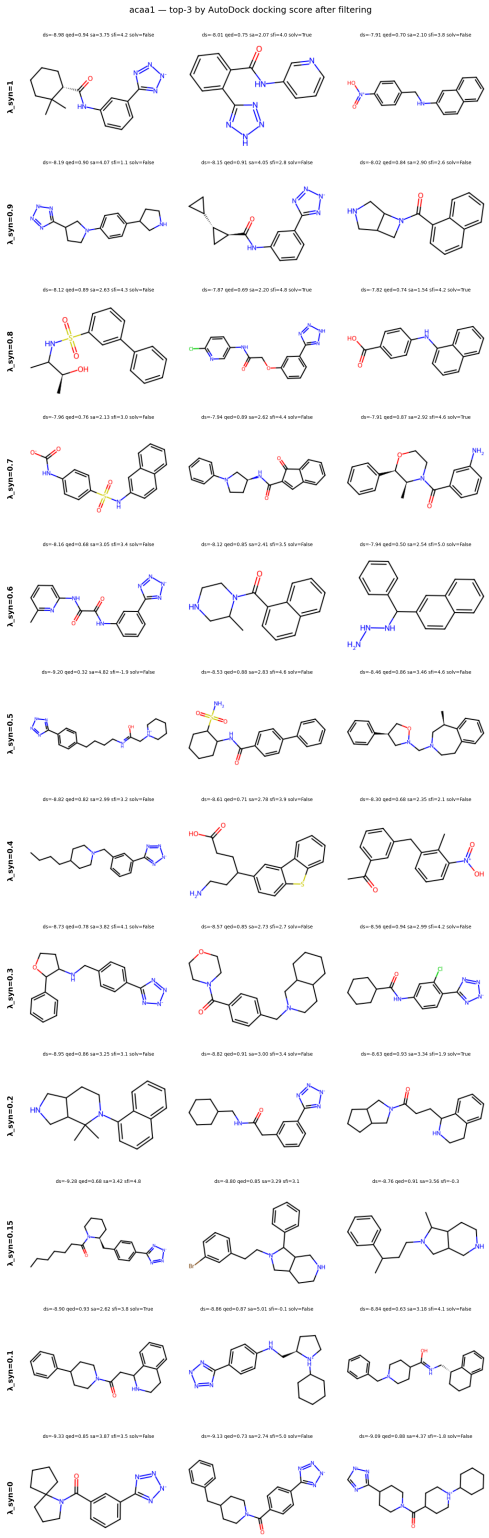

Figure 8: Top-3 strongest binders per condition, ACAA1.

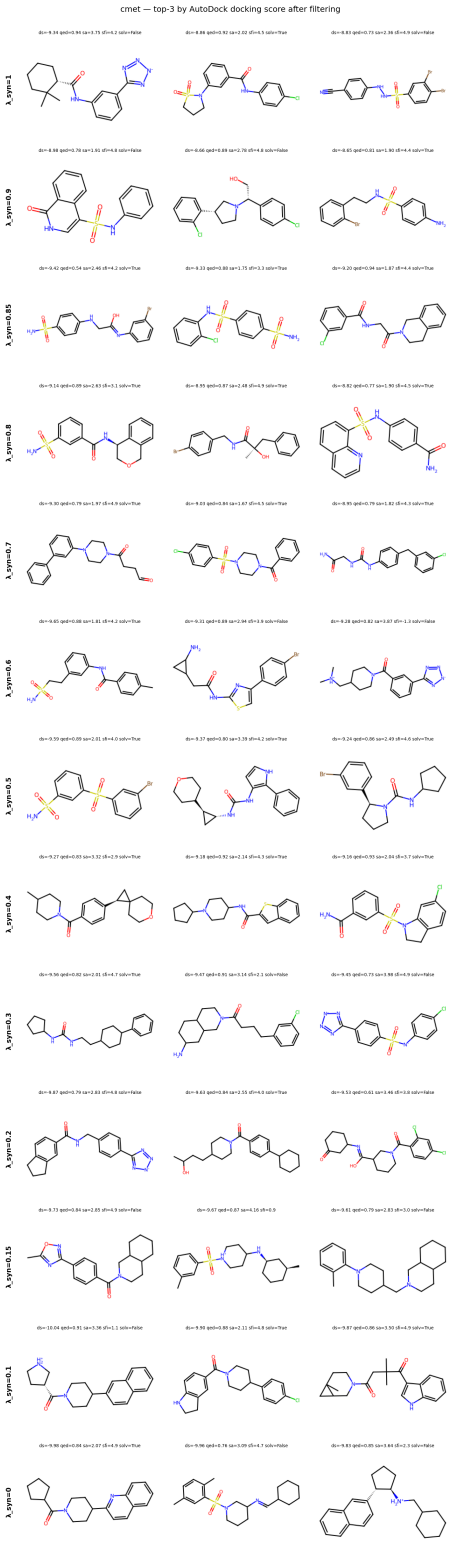

Figure 9: Top-3 strongest binders per condition, c-MET.

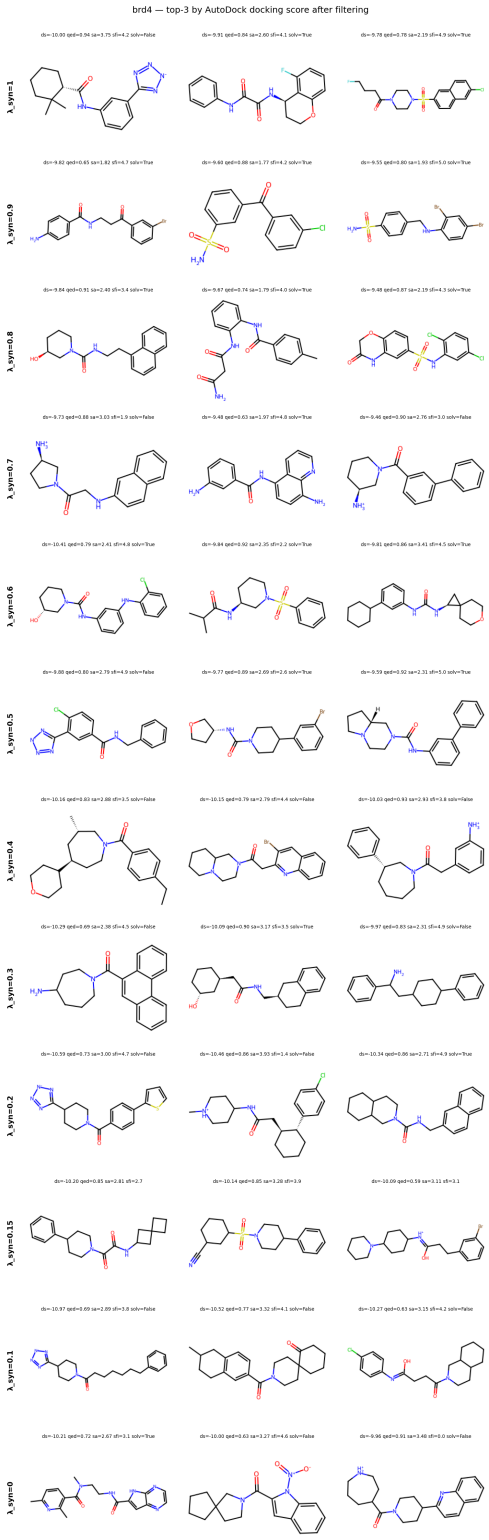

Figure 10: Top-3 strongest binders per condition, BRD4(2).

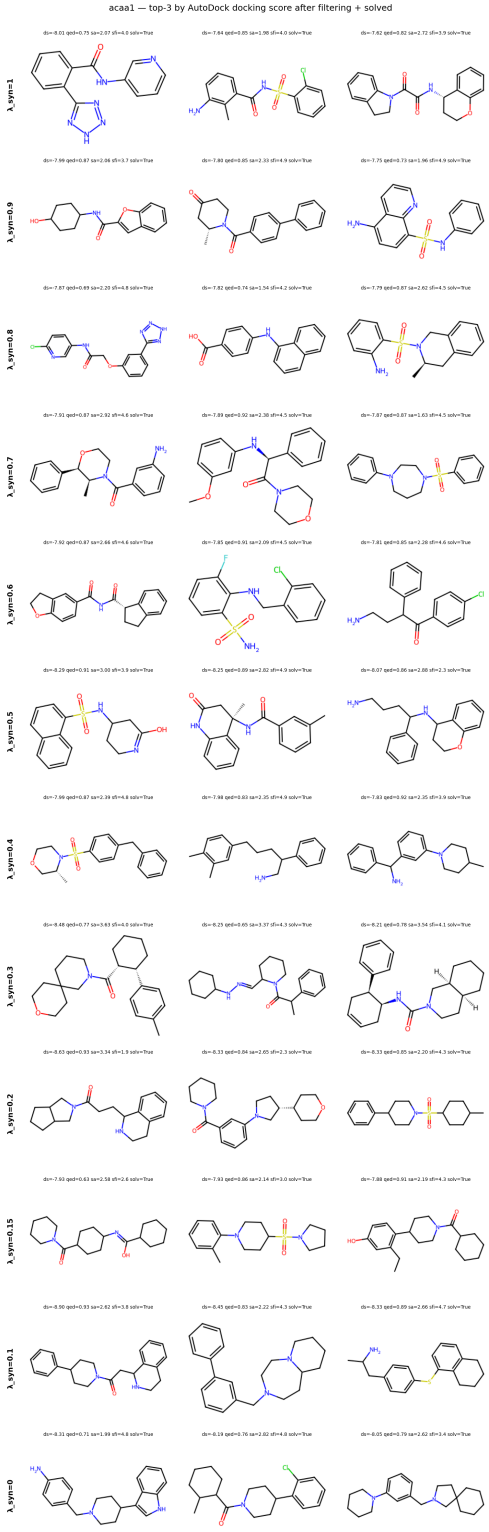

Figure 11: Top-3 binders with a measured AiZynthFinder route, ACAA1.

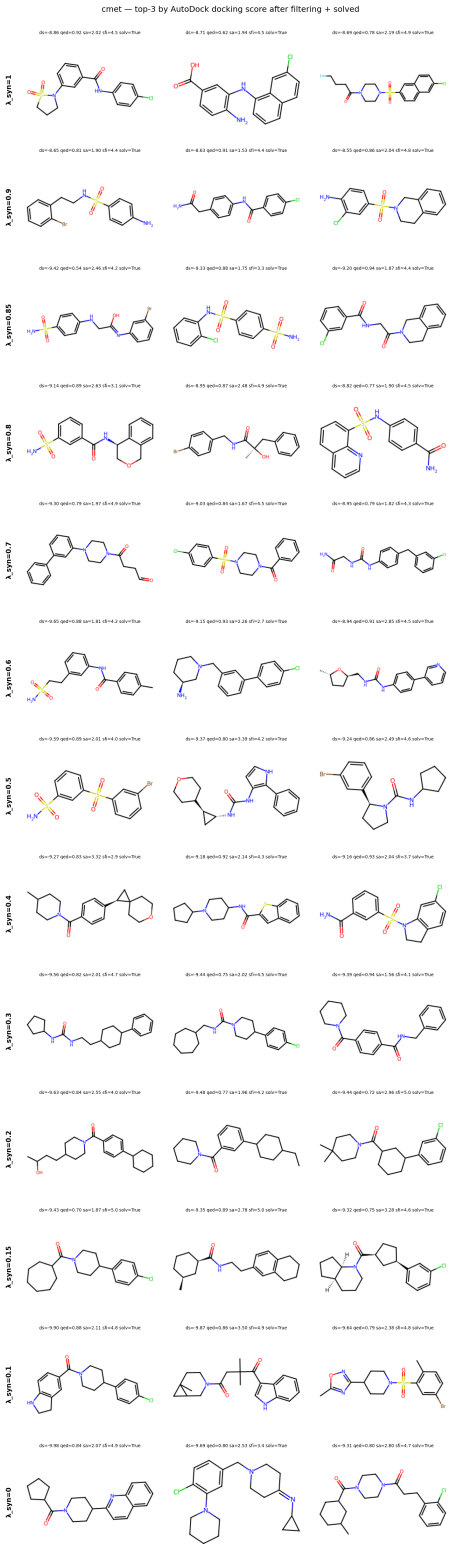

Figure 12: Top-3 binders with a measured AiZynthFinder route, c-MET.

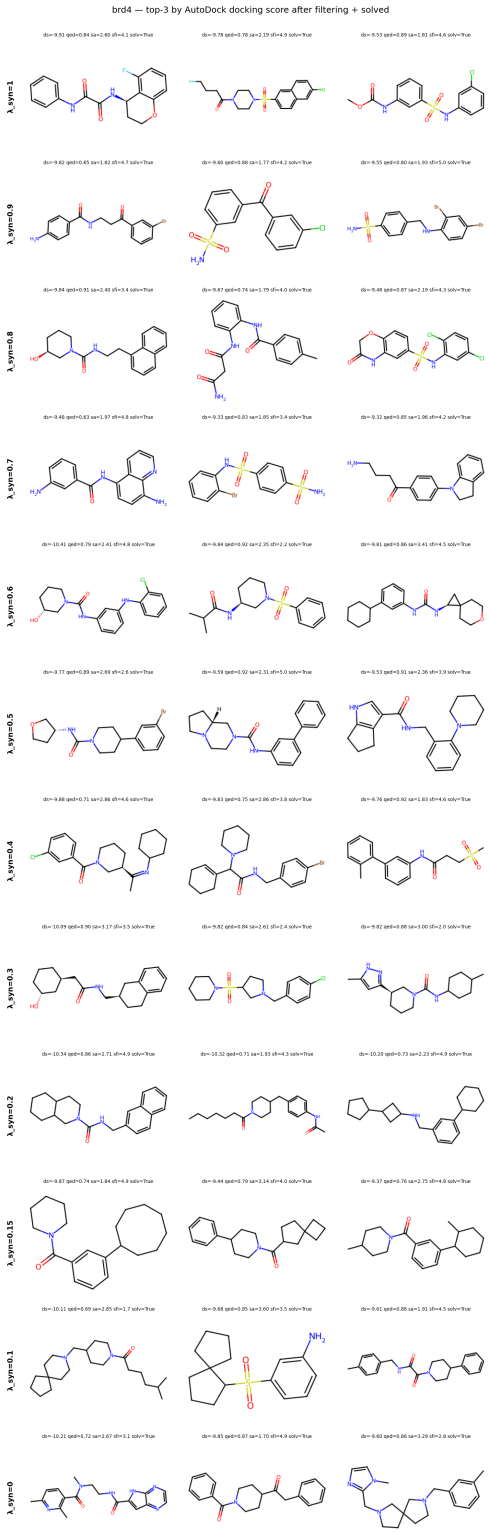

Figure 13: Top-3 binders with a measured AiZynthFinder route, BRD4(2).