COUPLING SYNTHESIS AND 3D GEOMETRY FOR POCKET-CONDITIONED MOLECULAR GENERATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Structure-based molecular generation can propose ligands for a target protein, but connecting their synthetic construction with their placement in the binding pocket remains challenging. We present SynTrace, a supervised model that couples synthesis decisions to the evolving three-dimensional state of a molecule in its pocket. The central idea is to learn reaction choices and coordinate evolution from a shared geometric representation, so that each synthesis decision can use the spatial state produced by earlier generation steps. We train this model by aligning synthesis routes with docked product poses, which provide paired action and coordinate targets without reward-based policy optimization. On 100 CrossDocked2020 test pockets, SynTrace generates 10,000 molecules with a median redocking score of -8.45 kcal/mol, a 45.49% AiZynthFinder solve rate using Enamine stock, and diversity of 0.889. Interventions show that synthesis decisions respond to relative ligand–pocket geometry and identify spatial interactions as the main source of pocket sensitivity on reference decision states.

1 INTRODUCTION

Structure-based molecular generation uses a protein pocket to guide the design of candidate ligands (Peng et al., 2022; Guan et al., 2023a; Qu et al., 2024). To guide experimental follow-up, molecular design should consider both how a molecule can be synthesized and how it fits within the pocket. A favorable computational pose alone does not provide a synthesis route, while an available route alone does not establish spatial compatibility. Joint generation can address both aspects within the same design process, linking chemical growth to the spatial environment in which it occurs.

Existing approaches address different parts of this goal. Pocket-conditioned 3D generators learn molecular structures in the receptor environment, but methods such as TargetDiff and MolCRAFT do not generate explicit construction routes (Guan et al., 2023a; Qu et al., 2024). Reaction-based models instead build molecules from available building blocks and reaction templates, making a construction route part of the output (Gao et al., 2022; Seo et al., 2025). This constrains chemical construction, but does not by itself model the intermediate's placement during growth. Combining these approaches requires synthesis decisions and coordinate generation to use the same spatial context as the molecule grows.

Recent joint models move in this direction, with different learning signals and conditioning information. The pocket-conditioned experiments of 3DSynthFlow, the molecular version of CGFlow, train the construction policy using predicted rewards; the authors identify reward-prediction accuracy as a limitation on further docking-score improvement (Shen et al., 2025, Section 6.2). SynCoGen learns joint generation without an external reward model, conditioned on pharmacophore types and positions (Rekesh et al., 2026). These interaction features do not explicitly specify the surrounding protein atoms or their occupied volume. This motivates a complementary learning strategy: use route–pose examples to supervise joint generation directly, while representing the surrounding pocket atoms so that local geometry can inform both chemical construction and spatial placement.

We introduce SynTrace, which learns synthesis actions and coordinate evolution through a shared geometric representation of the current intermediate and protein pocket. Figure 1 summarizes the training-data construction and generation process. Our design makes local ligand–pocket interactions available to both predictions: the action head selects how to extend the molecule, while the coordinate output guides its placement. Contact-stratified distance supervision emphasizes local

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.

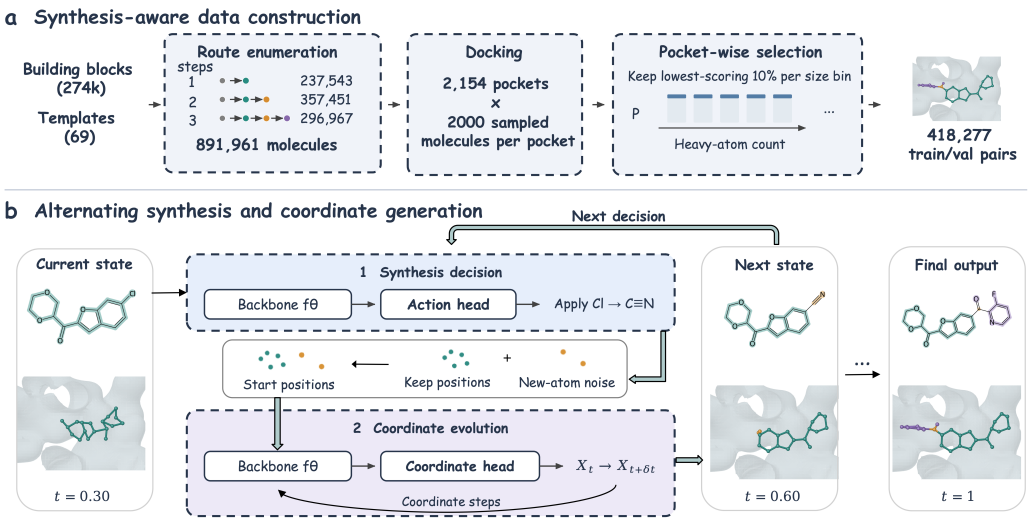

Figure 1: **Data construction and alternating generation.** (a) Enumerated routes are paired with docked product poses; filtering retains approximately the best-scoring 10% per pocket and size bin. Atom tracking aligns actions with final-pose-derived targets. (b) Both stages use the same backbone. Reactions retain mapped coordinates and initialize new atoms from noise. Repeated endpoint prediction and integration evolve the state before the next action. Colors identify introduction steps; retained coordinates can continue evolving. The $t = 0.60$ state precedes a possible third reaction; after STOP or the step limit, only coordinates evolve to $t = 1$.

ligand–pocket geometry, and a one-sided steric penalty discourages excessive overlap. During generation, reaction execution alternates with iterative coordinate integration, allowing later synthesis decisions to use the updated spatial state.

To train this model, we construct route–pose pairs through reaction enumeration and offline docking. Atom tracking aligns stepwise actions with coordinate targets derived from the final product pose, enabling joint supervision of action probabilities and coordinates. Reaction-slot scoring evaluates the exact normalizer over state-dependent subsets of a 2.1-million-entry action vocabulary. Docking supplies the training examples and poses, while policy learning uses supervised action targets. The resulting model generates a molecule, its construction route, and its pocket pose without online docking or reward-based policy optimization.

Our experiments assess both the quality of joint outputs and the role of geometric coupling. On 100 CrossDocked2020 test pockets, SynTrace achieves a median redocking score of -8.45 kcal/mol and a 45.49% retrosynthesis solve rate using Enamine stock, with pooled diversity of 0.889. Ablations support the roles of pocket conditioning, contact stratification, and steric exclusion in direct-pose scoring. Pose interventions further show that synthesis decisions respond to relative placement, while pathway interventions identify spatial interactions as the main source of changes in reference-action probabilities when the pocket is replaced.

2 Related Work

Pocket-conditioned 3D generation. Early models generate ligands through atom-wise construction or receptor-conditioned density grids (Luo et al., 2021; Ragoza et al., 2022); GraphBP uses an autoregressive flow to model atoms in a local coordinate system (Liu et al., 2022). Equivariant diffusion subsequently enabled joint generation of atom types and coordinates (Hoogeboom et al., 2022). Pocket2Mol uses an equivariant network to construct atoms and their connections incrementally (Peng et al., 2022). TargetDiff and DiffSBDD model atom types and coordinates with diffusion (Guan et al., 2023a; Schneuing et al., 2024); DecompDiff uses decomposed arm/scaffold priors (Guan et al., 2023b), while MolCRAFT performs generation in continuous parameter space (Qu et al., 2024). These models learn spatially conditioned molecular structures, but their generation

variables do not directly encode construction through reaction templates and building blocks. Synthesis routes must therefore be determined separately, for example through learned retrosynthetic search (Segler et al., 2018). We connect spatial modeling to an explicit reaction-based construction process.

Synthesis-based molecular generation. SynNet learns to construct synthesis trees from programmatically generated routes (Gao et al., 2022). SynFormer combines sequence modeling with building-block selection to navigate synthesizable chemical space (Gao et al., 2024). SynFlowNet and RxnFlow learn reward-guided reaction policies using GFlowNets (Bengio et al., 2021); RxnFlow also uses action subsampling to address large reaction and building-block spaces (Cretu et al., 2024; Seo et al., 2025). These approaches encode synthetic construction without jointly evolving intermediate coordinates. Three-dimensional information can also condition route prediction: SynthFormer generates synthesis trees from pharmacophore types and positions (Jocys et al., 2024). This differs from jointly generating a molecule's coordinates inside the pocket.

Joint route and structure generation. CGFlow interleaves a construction policy with coordinate flow, allowing the policy to read the current spatial state (Shen et al., 2025). Its molecular version, 3DSynthFlow, first trains a coordinate model and then freezes its parameters during reward-based policy learning; the CGFlow framework also describes cross-entropy policy training. SynCoGen jointly predicts building-block reaction graphs and atomic coordinates using a shared equivariant network, combining discrete denoising with continuous generation and pharmacophore conditioning (Rekesh et al., 2026). SynLaD learns a shared latent space with separate 3D-structure and synthesis-route decoders, followed by pharmacophore-conditioned latent diffusion (Cretu et al., 2026). SynTrace focuses on joint training on intermediates obtained by executing reactions, conditioned directly on protein-pocket atoms. Aligned route–pose examples train shared action and coordinate predictions, with additional supervision for local pocket geometry.

3 METHOD

3.1 JOINT GENERATION IN PROTEIN POCKETS

Our design connects chemical growth and spatial placement through a shared representation of the current intermediate and its pocket environment (Figure 2). Given a pocket $\mathcal{P} = \{(p_m, f_m)\}_{m=1}^{N_P}$, with atomic coordinates p_m and features f_m, we generate a synthesis-action sequence $\mathcal{R}$, the resulting molecular graph G, and final coordinates X. Actions select reactions and building blocks, apply available unimolecular reactions, or terminate construction. Route validity is defined relative to the supplied building-block library, reaction templates, and execution rules.

At generation time t, the state is $s_t = (G_t, X_t, \boldsymbol{\kappa}_t, t)$, where κ_i records the reaction at which atom i was introduced. The shared model predicts an action distribution $\pi_\theta(a \mid s_t, \mathcal{P})$ and coordinate endpoint $\widehat{X}_1(s_t, \mathcal{P})$. Reaction execution changes the graph and atom set; integration evolves their coordinates before the next decision. Figure 1 summarizes how aligned training states support this alternating process. Here t measures generation progress rather than physical time.

3.2 ALIGNED SUPERVISION FROM SYNTHESIS ROUTES AND POSES

To train both predictions on the same intermediate, we need coordinate labels aligned with its construction history. We obtain route–pose pairs by enumerating forward reactions and docking their final products into pockets. Within each pocket, size-stratified score filtering selects training examples. Docking therefore supplies poses and selects examples, while action learning uses the recorded synthesis decisions as targets.

Replaying each route recovers its intermediate graphs and atom identities. For final product coordinates Y, let $\phi_k(i)$ map atom i after reaction step k to its final-product identity. Surviving atoms receive targets $y_i = Y_{\phi_k(i)}$; atoms removed by later reactions remain in the input graph but are excluded from final-coordinate supervision. Thus, a single product pose supplies aligned targets across the route. These are generation targets, not claims about the binding conformations of isolated intermediates.

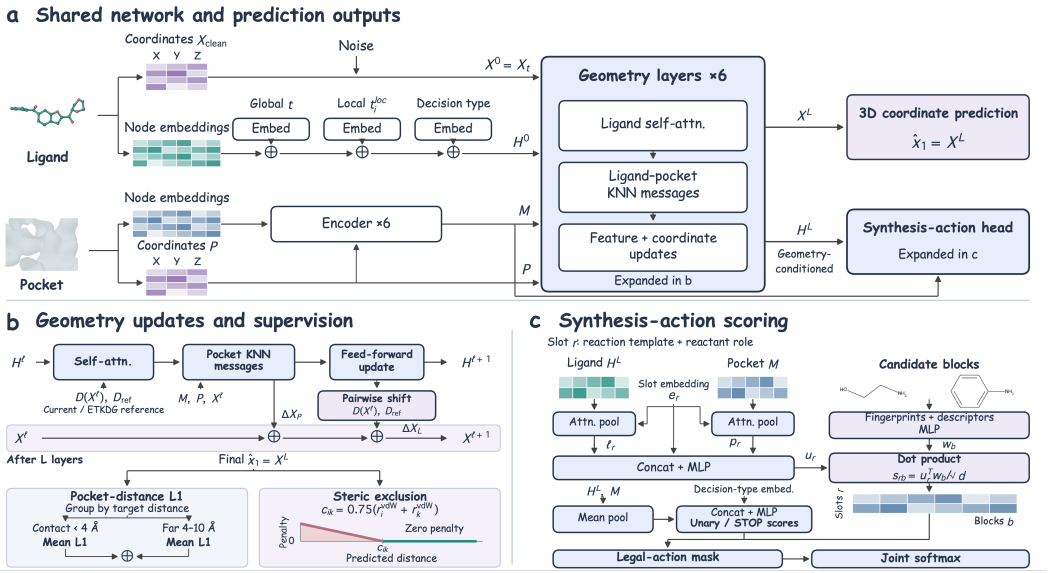

Figure 2: **Model architecture of SynTrace.** (a) The clean-coordinate/noise branch constructs training inputs; sampling supplies the current generated X_t. A shared geometric network produces ligand features H^L and endpoint coordinates $\widehat{X}_1 = X^L$. (b) Pocket and ligand displacements update coordinates within each layer. Contact-stratified and steric losses supervise the final endpoint. (c) Slot queries combine ligand and pocket features to score candidate blocks. Block logits, unimolecular logits, and STOP share one masked softmax. Figure notation and initialization rules are detailed in Appendix A.2.

New atoms enter while existing atoms may still be evolving. To represent this difference in their spatial progress, we use the overlapping local-time schedule of CGFlow (Shen et al., 2025):

$$\tau_i(t) = \mathrm{clip}\left(\frac{t - \lambda\kappa_i}{w}, 0, 1\right), \qquad \lambda = 0.30, \quad w = 0.40, \tag{1}$$

where introduction indices start at zero and routes contain at most three reactions. For supervised atoms, training coordinates interpolate between Gaussian noise ϵ_i and the aligned target:

$$x_i(t) = (1 - \tau_i(t))\epsilon_i + \tau_i(t)y_i. \tag{2}$$

A reference conformer supplies the noise scale and intramolecular distances without using the target binding pose. Action supervision uses decision times, whereas coordinate supervision samples times across the generation interval. This exposes the shared network to both synthesis decisions and spatial evolution along reference routes. Input construction for atoms without final counterparts is detailed in Appendix A.2.

3.3 Geometry-conditioned synthesis decisions

The same intermediate can occupy different regions of a pocket, providing different spatial contexts for further growth. We therefore condition action prediction on ligand features updated through local pocket interactions. A pocket encoder produces features M, and ligand features incorporate global time, per-atom local time, and the known decision stage. Six geometric layers jointly update ligand features H^ℓ and internal coordinates X^ℓ while keeping pocket coordinates fixed (Figure 2). Both action and coordinate objectives train this representation.

Local pocket interactions. To couple the representation to current placement, each layer combines ligand self-attention with messages from nearby pocket atoms. Self-attention uses current and reference intramolecular distances. The pocket neighborhood of atom i at layer ℓ is

$$\mathcal{N}_i^\ell = \mathrm{KNN}_K\left(x_i^\ell, \{p_m\}_{m=1}^{N_P}\right). \tag{3}$$

Here $K = 64$. Pairwise messages combine ligand features, neighboring pocket features, and their distances. Gated aggregation updates features; relative position vectors weighted by learned scalars update coordinates, following the equivariant construction of Satorras et al. (2021). Recomputing neighborhoods at every layer lets the model's local pocket context change with its spatial state.

The final coordinates give the endpoint prediction $\widehat{X}_1 = X^L$, with a variance readout for coordinate training. The action head reads both pocket features M and the geometrically updated ligand features H^L. These updated ligand features allow relative placement to affect synthesis decisions. Each sampling integration step calls this entire network again, separating layer updates from evolution over generation time.

Exact probabilities over applicable actions. Supervised action learning requires comparing the recorded reaction against all alternatives permitted by the state. We organize these alternatives by reaction slots so that queries can be reused across compatible building blocks. An action (r, b) pairs a slot r, specifying a template and reactant role, with a building block b. Slot-specific attention pools ligand and pocket features into $u_r \in \mathbb{R}^d$; a separate tower maps block fingerprints and descriptors to $w_b \in \mathbb{R}^d$. Their score and the action probability are

$$z_{r,b} = \frac{u_r^\top w_b}{\sqrt{d}}, \qquad \pi_\theta(a \mid s_t, \mathcal{P}) = \frac{\exp z_a}{\sum_{a' \in \mathcal{A}(s_t)} \exp z_{a'}}. \tag{4}$$

Here $\mathcal{A}(s_t)$ is the training candidate set defined by reaction applicability and the action mask. Unimolecular reactions and STOP receive scores from pooled ligand and pocket features and share the same denominator. Slot-based computation evaluates the categorical likelihood over the full masked subset of 2,146,119 action entries, without sampled negatives. Initialization-specific masks are specified in Appendix A.2.

3.4 Joint training with contact and steric constraints

Joint training must learn both which reaction to take and where the resulting structure should lie. Action likelihood supervises the discrete choice, while coordinate and intramolecular objectives fit the endpoint and preserve internal geometry. We add two pocket-relative constraints: contact-stratified distance fitting to emphasize nearby interactions, and one-sided steric exclusion to penalize overlap. The joint objective is

$$\begin{aligned}\mathcal{L} = \mathcal{L}_{\text{act}} + \lambda_{\text{coord}}\mathcal{L}_{\text{coord}} + \lambda_{\text{pair}}\mathcal{L}_{\text{pair}} + \lambda_{\text{short}}\mathcal{L}_{\text{short}} \\ + \lambda_{\text{pocket}}(\mathcal{L}_{\text{contact}} + \mathcal{L}_{\text{far}}) + \lambda_{\text{clash}}\mathcal{L}_{\text{clash}}.\end{aligned} \tag{5}$$

Here, $\mathcal{L}_{\text{coord}}$ is a variance-weighted Gaussian negative log-likelihood. Intramolecular losses fit pairwise distances, with an additional term for target distances below 1.8 Å. Appendix A.1 gives their definitions and action weighting. We focus below on contact geometry and steric exclusion.

Contact-stratified distance supervision. A single average of ligand–pocket distance errors can underweight contacts when distant pairs are more numerous. We give nearby and distant pairs separate normalization. Let $\mathcal{K}$ denote ligand atoms with final-coordinate labels, $\widehat{d}_{im} = \|\widehat{x}_{1,i} - p_m\|$ predicted pocket distances, and $d^*_{im} = \|y_i - p_m\|$ target distances. The target pose defines two groups:

$$\begin{aligned}\mathcal{C} &= \{(i,m) : i \in \mathcal{K},\ d^*_{im} < 4\,\text{Å}\}, \\ \mathcal{F} &= \{(i,m) : i \in \mathcal{K},\ 4\,\text{Å} \le d^*_{im} < 10\,\text{Å}\}.\end{aligned} \tag{6}$$

Each group receives its own mean absolute distance error:

$$\mathcal{L}_g = \frac{1}{|\mathcal{G}_g|} \sum_{(i,m)\in\mathcal{G}_g} |\widehat{d}_{im} - d^*_{im}|, \quad (\mathcal{G}_{\text{contact}}, \mathcal{G}_{\text{far}}) = (\mathcal{C}, \mathcal{F}). \tag{7}$$

Both group means receive the same coefficient in Eq.(5), so their relative counts no longer set their relative weights. Grouping by the target pose also prevents predictions from changing their own supervision category. The 4 Å threshold defines these loss groups; forward-pass neighbors are selected separately by KNN. Empty terms and batch reductions are specified in Appendix A.1.

One-sided steric exclusion. Distance fitting measures agreement with a reference pose, but does not explicitly impose a lower bound on predicted separations. To discourage overlap without continually repelling well-separated atoms, we penalize only distances below a radius-dependent bound. With ligand and pocket van der Waals radii r_i and r_m, we use

$$\mathcal{L}_{\text{clash}} = \frac{1}{|\mathcal{K}|} \sum_{i \in \mathcal{K}} \sum_{m} \max\Big(0,\, 0.75(r_i + r_m) - \widehat{d}_{im}\Big)\,. \tag{8}$$

The penalty becomes zero once the bound is satisfied. It sums over all valid pocket atoms and averages over supervised ligand atoms, independently of the KNN neighborhoods. Contact fitting and steric exclusion thus provide complementary constraints on relative placement. Both losses train the geometric representation shared with action prediction. Section 4.3 evaluates their effects on generated poses, and Section 4.4 examines how action prediction uses spatial information.

3.5 Alternating synthesis and coordinate generation

The learned endpoint provides a direction for evolving coordinates between synthesis decisions, relating the noise-to-target interpolation to continuous-flow generation (Lipman et al., 2023). Sampling begins with building-block selection. Each reaction preserves mapped coordinates, removes leaving atoms, and introduces new atoms with noisy coordinates and their introduction indices. Multiple integration steps then update the spatial state before the next action. For a step from t_a to t_b, we use

$$x_i(t_b) = x_i(t_a) + \frac{\tau_i(t_b) - \tau_i(t_a)}{\max(1 - \tau_i(t_a), 10^{-3})} \left[\widehat{x}_{1,i}(s_{t_a}, \mathcal{P}) - x_i(t_a)\right]. \tag{9}$$

Each step recomputes the endpoint from the updated input. All atoms whose local clocks are advancing can move, including atoms introduced at earlier reactions. Atoms whose local time has reached one remain fixed in the sampling state.

The initial reaction occurs at $t = 0$, with subsequent reaction decisions at $t = 0.30$ and $t = 0.60$. Each decision uses the current ligand–pocket representation. After STOP or the reaction-step limit, the graph remains fixed and coordinate integration continues to $t = 1$. The output is the executed construction route and final pose, without online docking or docking-score-guided search.

4 Experiments

We assess joint generation quality, the effects of geometric conditioning and supervision on placement, and the use of spatial information in synthesis decisions. A DYRK2 case illustrates the route and coordinate trajectory.

4.1 Experimental setup

We evaluate on 100 test pockets from the official CrossDocked2020 split (Francoeur et al., 2020), sampling 100 molecules per pocket. We evaluate Pocket2Mol, TargetDiff, DecompDiff, MolCRAFT, and RxnFlow ourselves (Peng et al., 2022; Guan et al., 2023a;b; Qu et al., 2024; Seo et al., 2025), with a sampling budget of 10,000 molecules per method. RxnFlow uses our Enamine building-block library. 3DSynthFlow is the exception: its high-β pocket-conditioned results are taken from Shen et al. (2025) and explicitly marked as literature values in Table 1. Our final model trains for 30 epochs with seed 0 and is selected by validation performance.

Scores and molecular properties. Our scoring pipeline uses Uni-Dock in balance mode with exhaustiveness 384 and Meeko ligand preparation (Yu et al., 2023). Vina scoring (Trott & Olson, 2010) is used throughout. Score evaluates generated coordinates directly; Min performs local optimization; Dock re-embeds molecules with ETKDG (Riniker & Landrum, 2015) and MMFF (Halgren, 1996) before docking into the assigned pocket. Dock therefore includes an additional pose search. Table 1 reports score means, intramolecular PoseBusters validity (PB mol) (Buttenschoen et al., 2024), heavy-atom count, QED, SA, diversity, and uniqueness. PB mol does not assess placement relative to the receptor. QED measures drug-likeness (Bickerton et al., 2012); SA estimates synthetic accessibility from fragment contributions and molecular complexity (Ertl & Schuffenhauer,

Table 1: **Docking, molecular properties, and retrosynthesis.** Docking and heavy-atom count (HA) are means; QED and SA are medians; Div. is pooled diversity. Arrows indicate preferred directions. We evaluate all methods except 3DSynthFlow with 100 samples per pocket (aggregation: Appendix B). The 3DSynthFlow row contains published high-β results (Shen et al., 2025) under the original protocol; its solve rate spans both stocks because the stock is unspecified in this comparison. Bold/underline indicate best/second-best listed model values, excluding Reference and HA, using unrounded values except QED and diversity ties at two decimals. Markers compare reported values; dashes indicate missing or inapplicable values.

	Docking			Molecular properties				Distribution		AiZynth	
Method	Score ↓ (kcal/mol)	Min ↓ (kcal/mol)	Dock ↓ (kcal/mol)	PB mol ↑ (%)	HA (atoms)	QED ↑	SA ↓	Div. ↑	Uniq. ↑ (%)	Ena. ↑ (%)	ZINC ↑ (%)
Reference	−6.61	−6.68	−7.15	93.00	22.80	0.47	3.32	0.91	96.00	37.00	42.00
Pocket2Mol	−4.78	−5.45	−6.61	7.10	16.30	0.39	4.15	0.88	97.70	8.20	9.40
TargetDiff	−5.83	−6.43	−7.42	66.10	22.70	0.49	4.58	0.89	99.20	7.20	11.00
DecompDiff	−5.42	−6.10	−7.09	84.60	21.90	0.49	4.14	0.85	94.20	16.10	19.70
MolCRAFT	**−6.25**	−6.99	−7.89	87.30	22.80	0.49	3.97	**0.90**	97.00	14.00	18.10
RxnFlow	–	–	−8.28	–	31.24	**0.73**	4.30	0.85	**100.00**	36.40	34.80
3DSynthFlow	–	–	**−9.42**	–	–	**0.73**	–	0.69	–	36.10	
SynTrace	−6.18	**−7.26**	−8.38	**87.80**	24.40	0.66	**3.15**	0.89	99.41	**45.49**	**42.41**

2009). Diversity is the pooled Morgan-fingerprint Tanimoto distance (radius 2, 2,048 bits) (Rogers & Hahn, 2010). Metric definitions and score aggregation are detailed in Appendix B.

Retrosynthesis. AiZynthFinder 4.4.1 (Genheden et al., 2020) uses MCTS with at most 100 iterations, 120 seconds, and six transformations, USPTO and RingBreaker expansion policies, and the USPTO filter. The target is excluded from stock and search continues after the first solution. Enamine 274k and ZINC 17.4M stocks are evaluated separately. Solve rates assess route recovery under each stock, complementing the construction routes generated by the model.

4.2 Generation quality and synthetic accessibility

Redocking and direct placement. We report both Dock and direct Score to distinguish performance after pose search from the quality of the generated coordinates. Table 1 reports means; the following comparisons use medians. SynTrace obtains a median Dock score of -8.45 kcal/mol, compared with -7.74 for MolCRAFT, the lowest among the four SBDD baselines. RxnFlow achieves a lower median of -8.89. For generated coordinates, MolCRAFT has a lower direct Score median than SynTrace (-6.79 versus -6.44); our median improves to -7.38 after local optimization. This difference shows why redocking and generated-pose quality should be assessed separately. After redocking, 73.52% of our molecules have lower Vina scores than the reference ligand of their respective pocket. The intramolecular PB mol pass rate is 87.8%.

Synthetic accessibility and molecular properties. AiZynthFinder solves 4,549 of our 10,000 molecules with Enamine stock (45.49%) and 4,241 with ZINC (42.41%), above the four SBDD baselines under each stock. Compared with RxnFlow, SynTrace has higher solve rates with both Enamine (45.49% versus 36.4%) and ZINC (42.41% versus 34.8%), a lower SA median (3.15 versus 4.30), and higher diversity (0.889 versus 0.853), alongside a less favorable redocking median. The corresponding QED median is 0.664, between RxnFlow (0.729) and reference ligands (0.468). All outputs can be parsed by RDKit and contain a single connected component, with 99.41% global uniqueness without deduplication. Thus, the model combines a high retrosynthesis solve rate and molecular diversity with favorable redocking scores.

Published joint-generation comparison. We were unable to reproduce 3DSynthFlow, the molecular model in CGFlow, because of a checkpoint/environment incompatibility. We therefore report its high-β pocket-conditioned results directly from the original paper (Shen et al., 2025): median Vina score -9.61, QED 0.73, diversity 0.69, and AiZynthFinder solve rate 36.1%. This literature

Table 2: **Ablations of geometric conditioning and supervision.** Scores are medians in kcal/mol. PB mol and positive direct-score fractions (Pos.) are percentages. Score medians use unmodified scores. All four rows use the 100-pocket test split. Each model selects its own best validation checkpoint; these are single-seed results.

Model	Score ↓	Min ↓	Dock ↓	PB mol (%)	Pos. (%)	HA
SynTrace	−6.44	−7.38	−8.45	87.80	2.34	24.40
A0: no pocket	+82.98	−3.33	−8.21	62.50	98.00	23.90
Pflat: no stratification	−5.82	−7.16	−8.48	82.10	5.70	24.70
L30: no hinge	−5.88	−6.99	−8.28	88.37	5.87	24.57

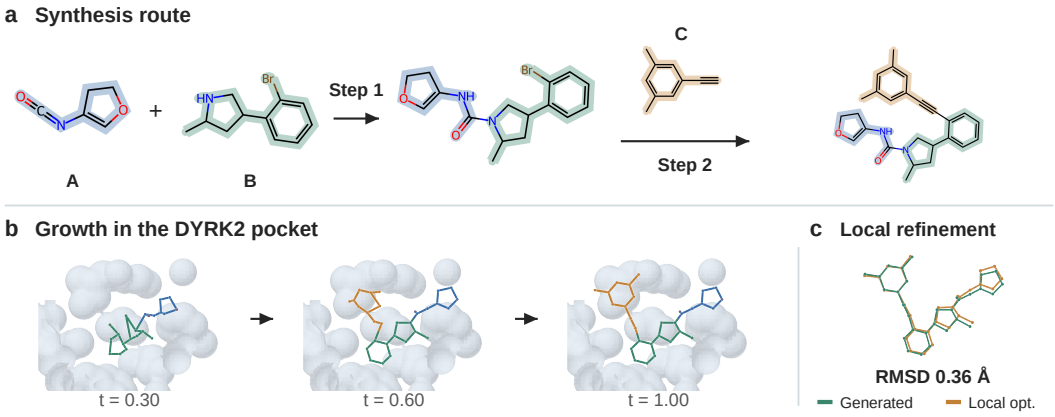

Figure 3: **A two-step route and pose in DYRK2.** (a) RDKit structures with building-block highlights matching (b); a returned AiZynthFinder route matches the depicted construction. (b) Coordinate snapshots with a common view and scale; the graph is fixed from $t = 0.60$ onward. Gray denotes the cutaway pocket envelope. (c) Generated and locally optimized poses at a separate scale, with 0.36 Å RMSD before any rigid alignment. Route matching omits stereochemistry.

comparison is separate from our evaluations of the other baselines: we did not rerun its sampling, docking, or retrosynthesis under our protocol.

4.3 Geometric conditioning and training objectives

We next test how each geometric component affects the generated poses. We ablate pocket conditioning (A0), contact/far distance stratification (Pflat), and one-sided steric exclusion (L30). Models use the same 30-epoch, 188,700-update training budget and select their own best validation checkpoints. All four evaluations use the 100 test pockets (Table 2).

Pocket conditioning and contact stratification. A0 gives positive direct scores for 98.0% of outputs, with a raw median of +82.98, despite a redocked median of −8.21. Redocking can therefore mask large errors in the generated poses. Contact stratification improves the direct-score median from Pflat's −5.82 to −6.44, the locally optimized median from −7.16 to −7.38, and PB mol from 82.1% to 87.8%; redocked medians remain numerically close. Action validation NLL is 5.5051 for SynTrace, 5.6698 for A0, and 5.4987 for Pflat. Thus, contact stratification primarily improves the measured pose quality, while action likelihood does not improve uniformly across these variants.

One-sided steric exclusion. Relative to L30, the full model improves the direct-score median from −5.88 to −6.44 and the locally optimized median from −6.99 to −7.38, while reducing the positive-score fraction from 5.87% to 2.34%. The redocking median is −8.45 versus −8.28. PB mol decreases slightly, from 88.37% to 87.8%. In this single-seed comparison, steric exclusion improves direct-pose scoring with a small reduction in intramolecular validity.

4.4 SPATIAL CONDITIONING OF SYNTHESIS DECISIONS

Having evaluated placement, we examine whether synthesis decisions respond to spatial information and which pocket pathways contribute to that response.

Do actions respond to relative placement? On 695 generated-state decisions from 92 pockets, a 1 Å ligand translation with 15°, 30°, or 60° rotation changes 30.6%, 42.6%, or 62.6% of near-greedy actions without changing the graph. Transforming ligand and pocket together changes at most 0.12%, supporting sensitivity to relative placement (Appendix C.1).

Which pocket pathways drive the response? To isolate the pathways underlying pocket sensitivity, we intervene on the final checkpoint at 18,769 validation reference decisions from 92 pockets. We hold the graph, action mask, and time fixed and compare pocket replacement with resampling training-input coordinate noise. Table 3 measures changes in the reference action's log probability. Spatial pathways include pocket-to-ligand feature messages and coordinate updates; direct pathways pass pocket features directly to the action head.

Table 3: **Pocket pathways in reference-action prediction.** Mean absolute log-probability changes (nats) over 18,769 decisions under inference-time interventions on one checkpoint. The all-disabled response is near numerical precision. Protocol and reaction-only results: Appendix C.2.

Enabled pocket pathways	Pocket replacement	Noise resampling
All (full model)	3.83	0.15
Direct pocket inputs only	0.29	0.47
Spatial interactions only	3.36	0.18
None	< 0.01	0.51

Disabling spatial interactions leaves 7.5% of the full pocket-replacement response; retaining spatial interactions alone preserves 87.7%. Among 5,136 reaction decisions excluding initialization and STOP, the full-model response is 4.80 nats to pocket replacement versus 0.41 to noise resampling. The response therefore extends to reaction choices rather than stopping alone, with spatial interactions accounting for most measured pocket sensitivity on reference states. These response ratios are not additive contributions. Candidate-continuation and mutation analyses leave the effect on pocket-specific chemical quality unresolved (Appendices C.1 and C.3).

4.5 A GENERATED ROUTE AND POSE IN THE DYRK2 POCKET

Figure 3 shows a selected DYRK2 example (PDB 4AZF, chain A) outside the main evaluation. Two reactions yield a 30-heavy-atom molecule; its graph is complete at $t = 0.60$, followed by coordinate integration to $t = 1$. Direct and locally optimized Vina scores are -11.47 and -11.63 kcal/mol, with 0.36 Å heavy-atom RMSD without rigid alignment. One of seven AiZynthFinder trees matches the starting blocks, intermediate, and reaction order, verified by template replay. This example shows a generated pocket pose together with a synthesis route also found by an independent planner; route matching omits stereochemistry.

5 CONCLUSION

SynTrace jointly learns synthesis decisions and coordinate generation from aligned route–pose examples. The generated molecules achieve favorable redocking scores and high diversity, with synthesis routes found by AiZynthFinder for 45.49% of molecules using Enamine stock. Ablations show improved direct-pose scores with geometric supervision, while interventions show that action predictions respond to relative geometry through spatial pocket interactions. Improved pocket-specific chemical choices, successful synthesis, and target binding still require validation.

AI USE STATEMENT

We used generative AI tools to improve the clarity and readability of the manuscript. All AI-assisted edits were reviewed by the authors, who take full responsibility for the final content.

REFERENCES

Emmanuel Bengio, Moksh Jain, Maksym Korablyov, Doina Precup, and Yoshua Bengio. Flow network based generative models for non-iterative diverse candidate generation. In *Advances in Neural Information Processing Systems*, volume 34, pp. 27381–27394, 2021. URL https://arxiv.org/abs/2106.04399.

G. Richard Bickerton, Gaia V. Paolini, Jérémy Besnard, Sorel Muresan, and Andrew L. Hopkins. Quantifying the chemical beauty of drugs. *Nature Chemistry*, 4:90–98, 2012. URL https://www.nature.com/articles/nchem.1243.

Martin Buttenschoen, Garrett M. Morris, and Charlotte M. Deane. PoseBusters: AI-based docking methods fail to generate physically valid poses or generalise to novel sequences. *Chemical Science*, 15:3130–3139, 2024. doi: 10.1039/D3SC04185A. URL https://doi.org/10.1039/D3SC04185A.

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. *arXiv preprint arXiv:2405.01155*, 2024. URL https://arxiv.org/abs/2405.01155.

Miruna Cretu, John Bradshaw, Patricia Suriana, Saeed Saremi, Omar Mahmood, Kirill Shmilovich, Kangway Chuang, Vishnu Sresht, and Colin Grambow. SynLaD: Latent diffusion for generating synthesizable molecules conditioned on 3D pharmacophore profiles. *arXiv preprint arXiv:2607.01105*, 2026. URL https://arxiv.org/abs/2607.01105.

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

Paul G. Francoeur, Tomohide Masuda, Jocelyn Sunseri, Andrew Jia, Richard B. Iovanisci, Ian Snyder, and David R. Koes. Three-dimensional convolutional neural networks and a cross-docked data set for structure-based drug design. *Journal of Chemical Information and Modeling*, 60: 4200–4215, 2020. doi: 10.1021/acs.jcim.0c00411. URL https://pubmed.ncbi.nlm.nih.gov/32865404/.

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*, 2022. URL https://openreview.net/forum?id=FRxhHdnxt1.

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

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*, 12:70, 2020. doi: 10.1186/s13321-020-00472-1. URL https://doi.org/10.1186/s13321-020-00472-1.

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 *International Conference on Learning Representations*, 2023a. URL https://arxiv.org/abs/2303.03543.

Jiaqi Guan, Xiangxin Zhou, Yuwei Yang, Yu Bao, Jian Peng, Jianzhu Ma, Qiang Liu, Liang Wang, and Quanquan Gu. DecompDiff: Diffusion models with decomposed priors for structure-based drug design. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202, pp. 11827–11846. PMLR, 2023b. URL https://proceedings.mlr.press/v202/guan23a.html.

Thomas A. Halgren. Merck molecular force field. I. basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry*, 17:490–519, 1996. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291096-987X%28199604%2917%3A5/6%3C490%3A%3AAID-JCC1%3E3.0.CO%3B2-P.

Emiel Hoogeboom, Víctor Garcia Satorras, Clément Vignac, and Max Welling. Equivariant diffusion for molecule generation in 3D. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162, pp. 8867–8887, 2022. URL https://proceedings.mlr.press/v162/hoogeboom22a.html.

Zygimantas Jocys, Zhanxing Zhu, Henriette M. G. Willems, and Katayoun Farrahi. SynthFormer: Equivariant pharmacophore-based generation of synthesizable molecules for ligand-based drug design. *arXiv preprint arXiv:2410.02718*, 2024. URL https://arxiv.org/abs/2410.02718.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. In *International Conference on Learning Representations*, 2023. URL https://arxiv.org/abs/2210.02747.

Meng Liu, Youzhi Luo, Kanji Uchino, Koji Maruhashi, and Shuiwang Ji. Generating 3D molecules for target protein binding. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162, pp. 13912–13924, 2022. URL https://proceedings.mlr.press/v162/liu22m.html.

Shitong Luo, Jiaqi Guan, Jianzhu Ma, and Jian Peng. A 3D generative model for structure-based drug design. In *Advances in Neural Information Processing Systems*, volume 34, pp. 6229–6239, 2021. URL https://arxiv.org/abs/2203.10446.

Xingang Peng, Shitong Luo, Jiaqi Guan, Qi Xie, Jian Peng, and Jianzhu Ma. Pocket2Mol: Efficient molecular sampling based on 3D protein pockets. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162, pp. 17644–17655. PMLR, 2022. URL https://proceedings.mlr.press/v162/peng22b.html.

Daniil Polykovskiy, Alexander Zhebrak, Benjamin Sanchez-Lengeling, Sergey Golovanov, Oktai Tatanov, Stanislav Belyaev, Rauf Kurbanov, Aleksey Artamonov, Vladimir Aladinskiy, Mark Veselov, Artur Kadurin, Simon Johansson, Hongming Chen, Sergey Nikolenko, Alán Aspuru-Guzik, and Alex Zhavoronkov. Molecular sets (MOSES): A benchmarking platform for molecular generation models. *Frontiers in Pharmacology*, 11:565644, 2020. URL https://doi.org/10.3389/fphar.2020.565644.

Yanru Qu, Keyue Qiu, Yuxuan Song, Jingjing Gong, Jiawei Han, Mingyue Zheng, Hao Zhou, and Wei-Ying Ma. MolCRAFT: Structure-based drug design in continuous parameter space. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235, pp. 41749–41768. PMLR, 2024. URL https://proceedings.mlr.press/v235/qu24a.html.

Matthew Ragoza, Tomohide Masuda, and David Ryan Koes. Generating 3D molecules conditional on receptor binding sites with deep generative models. *Chemical Science*, 13:2701–2713, 2022. URL https://doi.org/10.1039/D1SC05976A.

Andrei Rekesh, Miruna Cretu, Dmytro Shevchuk, Vignesh Ram Somnath, Pietro Liò, Robert A. Batey, Mike Tyers, Michał Koziarski, and Cheng-Hao Liu. SynCoGen: Synthesizable 3D molecule generation via joint reaction and coordinate modeling. In *International Conference on Learning Representations*, 2026. URL https://arxiv.org/abs/2507.11818.

Sereina Riniker and Gregory A. Landrum. Better informed distance geometry: Using what we know to improve conformation generation. *Journal of Chemical Information and Modeling*, 55:2562–2574, 2015. URL https://pubmed.ncbi.nlm.nih.gov/26575315/.

David Rogers and Mathew Hahn. Extended-connectivity fingerprints. *Journal of Chemical Information and Modeling*, 50:742–754, 2010. URL https://pubmed.ncbi.nlm.nih.gov/20426451/.

Víctor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E(n) equivariant graph neural networks. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139, pp. 9323–9332, 2021. URL https://proceedings.mlr.press/v139/satorras21a.html.

Arne Schneuing, Charles Harris, Yuanqi Du, Kieran Didi, Arian Jamasb, Ilia Igashov, Weitao Du, Carla Gomes, Tom L. Blundell, Pietro Liò, Max Welling, Michael Bronstein, and Bruno Correia. Structure-based drug design with equivariant diffusion models. *Nature Computational Science*, 4:899–909, 2024. doi: 10.1038/s43588-024-00737-x. URL https://www.nature.com/articles/s43588-024-00737-x.

Marwin H. S. Segler, Mike Preuss, and Mark P. Waller. Planning chemical syntheses with deep neural networks and symbolic AI. *Nature*, 555:604–610, 2018. URL https://www.nature.com/articles/nature25978.

Maximilian Seitzer, Arash Tavakoli, Dimitrije Antić, and Georg Martius. On the pitfalls of heteroscedastic uncertainty estimation with probabilistic neural networks. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=aPOpXlnV1T.

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. URL https://arxiv.org/abs/2410.04542.

Tony Shen, Seonghwan Seo, Ross Irwin, Kieran Didi, Simon Olsson, Woo Youn Kim, and Martin Ester. Compositional flows for 3D molecule and synthesis pathway co-design. In *Proceedings of the International Conference on Machine Learning*, 2025. URL https://arxiv.org/abs/2504.08051.

Oleg Trott and Arthur J. Olson. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31:455–461, 2010. URL https://doi.org/10.1002/jcc.21334.

Yuejiang Yu, Chun Cai, Jiayue Wang, Zonghua Bo, Zhengdan Zhu, and Hang Zheng. Uni-Dock: GPU-accelerated docking enables ultralarge virtual screening. *Journal of Chemical Theory and Computation*, 19(11):3336–3345, 2023. doi: 10.1021/acs.jctc.2c01145. URL https://pubmed.ncbi.nlm.nih.gov/37125970/.

A TRAINING DETAILS

A.1 SUPERVISION AND LOSS REDUCTIONS

Action supervision. For supervised action rows j, the reference action a_j^* receives the weighted negative log-likelihood

$$\mathcal{L}_{\text{act}} = -\frac{\sum_j \omega_j \log \pi_\theta(a_j^* \mid s_j, \mathcal{P}_j)}{\sum_j \omega_j}, \qquad \omega_j = \frac{1}{T_j + 1}, \tag{10}$$

where T_j is the reaction count of the route containing row j. Because one reaction may involve multiple policy calls, this weighting is not equivalent to normalization by the number of action rows in each route. Coordinate-only rows do not contribute to this loss.

Coordinate supervision. The model predicts an isotropic log variance v_i for each atom, clamped to $[-2.4, 4.0]$. Using β-NLL weighting (Seitzer et al., 2022), the coordinate loss is

$$\mathcal{L}_{\text{coord}} = \frac{1}{|\mathcal{K}|} \sum_{i \in \mathcal{K}} \text{sg}\left(e^{\beta v_i}\right) \left[\frac{3}{2} v_i + \frac{\|\widehat{x}_{1,i} - y_i\|^2}{2e^{v_i}}\right], \qquad \beta = 0.5, \tag{11}$$

where sg stops gradients through the multiplicative weight. We omit the constant Gaussian normalization term. Coordinate warmup instead minimizes mean squared Euclidean error. Variances weight training; deterministic coordinate integration does not use them to inject sampling noise.

Intramolecular geometry. For supervised atoms, let $\mathcal{Q} = \{(i, j) : i, j \in \mathcal{K},\ i \neq j\}$ and $\mathcal{Q}_{\text{short}} = \{(i, j) \in \mathcal{Q} : \|y_i - y_j\| < 1.8\,\text{Å}\}$. The intramolecular objectives are

$$\mathcal{L}_{\text{pair}} = \frac{1}{|\mathcal{Q}|} \sum_{(i,j) \in \mathcal{Q}} \left| \|\widehat{x}_{1,i} - \widehat{x}_{1,j}\| - \|y_i - y_j\| \right|, \tag{12}$$

with $\mathcal{L}_{\text{short}}$ defined identically over $\mathcal{Q}_{\text{short}}$. The latter selects pairs by distance, not an explicit bond list. Empty terms are zero. All reductions use valid entries across the batch and exclude examples without coordinate supervision; atom pairs always belong to the same example.

Pocket-loss reductions. Contact and far-group means in Eq.(7) pool valid atom pairs across the batch, always pairing ligand and pocket atoms from the same example. Empty groups contribute zero. The steric loss sums over all valid receptor atoms and divides by the total number of supervised ligand atoms; an empty supervised set contributes zero. Its penalty uses predicted distances and atomic radii rather than target distances, while retaining the coordinate-supervision mask.

A.2 INPUTS AND OPTIMIZATION

Decision stages and coordinate inputs. The input stage distinguishes empty-state selection, initial partner selection, and subsequent decisions. Selecting the first two building blocks requires two policy calls but one reaction. Subsequent unimolecular, bimolecular, and STOP targets share the same stage embedding; the target action class is not supplied as an input. KNN interactions use up to 64 valid pocket atoms.

Intermediate atoms removed by later reactions have no final-pose labels. Their input interpolation anchors use the mean target position of surviving neighbors, or of all surviving atoms if no such neighbors exist. These anchors are not supervision targets. Coordinates are predicted for all current atoms, while losses apply only to atoms with reference labels. Reference conformers supply intramolecular distances and a per-axis noise standard deviation of $\max(R_g, 0.9)/\sqrt{3}$ Å, where R_g is their radius of gyration in angstroms.

In Figure 2, $t_i^{\text{loc}} = \tau_i(t)$, P and M denote pocket coordinates and features, and X_{clean} denotes interpolation anchors. Block scores s_{rb} correspond to $z_{r,b}$ in Eq.(4), and the steric bound c_{ik} to $0.75(r_i + r_m)$ in Eq.(8). Internal layer updates $X^{\ell+1} = X^\ell + \Delta X_P + \Delta X_L$ are distinct from the outer integration state X_t, where atoms that have completed their local windows remain fixed.

Action masks. Likelihood evaluation includes STOP in the normalizer even during initial building-block selection. Sampling masks STOP until the first block and any required partner have been selected, then uses Gumbel-max on temperature-scaled legal logits. This difference applies only to initialization.

Training configuration. The final model is trained for 30 epochs with seed 0 and learning rate 3×10^{-4}. Batches contain up to 16 pockets with four molecules per pocket. In Eq.(5), the final weights for coordinate, pair, short-range, pocket, and clash terms are 1, 3, 10, 3, and 12, respectively. The first 4,000 updates use coordinate MSE with weight 0.02; coordinate NLL and clash weights then increase linearly over 800 updates. The final checkpoint minimizes validation action NLL and occurs at epoch 28 of 30. The test ablations use the same 188,700-update budget and independently select their best validation checkpoints. These are single-training-seed comparisons.

B Evaluation details

Scoring and aggregation. All methods except 3DSynthFlow use a sampling budget of 100 molecules for each of 100 test pockets. This budget is distinct from the number of successfully scored molecules; per-metric valid-sample counts were not available for all baseline summaries.

For SynTrace, its ablations, RxnFlow, and Reference, score means use $N^{-1} \sum_i \min(s_i, 0)$; the other re-evaluated baselines retain their exported aggregation, whose equivalence has not been established. All reported score medians use unmodified scores. Mean-based comparisons should therefore be read with these aggregation differences in mind.

Molecular and retrosynthesis metrics. Heavy-atom counts are means; the RxnFlow mean is 31.24. Diversity is calculated over the pooled sample, complementing per-molecule properties (Polykovskiy et al., 2020). All 10,000 SynTrace outputs parse and are connected; 59 duplicate entries give 99.41% global uniqueness without deduplication.

AiZynthFinder solve rates are computed separately for each stock. Route lengths are conditional on each method's solved subset and are not used to rank methods across different subsets. Planner success and native template execution provide computational route support, not experimental synthesis validation. The 3DSynthFlow results follow the original publication, as described in Section 4.

B.1 DYRK2 route comparison

The selected DYRK2 example comes from an additional sampling batch and is excluded from the main evaluation. One of seven returned AiZynthFinder trees matches the generated starting blocks, intermediate, and reaction order. The first returned tree uses the same blocks and reaction count but a different intermediate and order; we therefore claim a matching returned route, not a match to the first route. Route comparison omits stereochemistry, and the coordinates do not establish stereoselective synthesis. This example illustrates route–pose consistency rather than estimating a population-level route-matching rate.

C Additional analyses of synthesis decisions

C.1 Pose interventions and candidate continuations

Relative-pose sensitivity. A separate generated-state collection contains 1,107 action decisions from 92 pockets, including 695 subsequent decisions after an initial product exists. Ligand-only rigid transformations change its placement relative to the pocket; joint transformations preserve their relative geometry. Each setting combines the stated rotation with a 1 Å translation. Action-change rates first average within pockets and then equally across pockets; 95% intervals use 4,000 pocket-bootstrap samples. Readout uses temperature 10^{-6} and is near-greedy, rather than an exact deterministic argmax.

Replacing the pocket without size matching changes 76.70% [73.50, 79.80] of subsequent actions. Restricting to decisions that continue synthesis before and after ligand-only intervention gives

Table 4: **Action changes on 695 subsequent decisions.** Percentages with 95% pocket-bootstrap intervals. Every transformation includes a 1 Å translation.

Rotation	Ligand only	Joint ligand–pocket
15°	30.60 [27.40, 33.90]	0.12 [0.00, 0.36]
30°	42.60 [39.50, 45.90]	0.12 [0.00, 0.36]
60°	62.60 [59.10, 66.00]	0.00 [0.00, 0.00]

change rates of 34.80%, 49.50%, and 60.10% for the three rotations, respectively. These subsets contain 227, 225, and 230 states from 85, 84, and 84 pockets; because these subsets differ, the rates do not describe the same set of decisions at increasing perturbation strengths. Small changes under joint transformations may include near-greedy sampling variability.

Candidate choice and pocket specificity. For 351 decisions from 91 pockets, 2,106 candidate continuations are matched by reaction slot and approximately by building-block heavy-atom count. Complete products are evaluated with direct Vina scoring under a common continuation protocol. Selected products score -5.21 versus -4.94 for matched candidates, with a paired difference of -0.27 kcal/mol (95% CI $[-0.45, -0.09]$). Intervals for contact and close-overlap differences include zero.

With candidate geometry and the scoring pocket held fixed, policy preferences correlate with negative Vina score by 0.188 under the original pocket condition and 0.171 under a replaced pocket. The difference, 0.018 (95% CI $[-0.031, 0.067]$), does not establish an additional ranking advantage from the original pocket. Thus, pose sensitivity and candidate-score associations do not by themselves demonstrate improved pocket-specific chemical choices.

C.2 Pathway interventions on reference states

Design and response measure. The final checkpoint is evaluated on 150 validation batches with self-conditioning and bidirectional updates disabled. There are 18,769 unique decision states from 92 pockets and 4,763 pocket–molecule records, containing 4,714 unique molecule IDs. These comprise 4,763 initial selections, 4,107 initial-partner selections, 3,063 bimolecular reactions, 2,073 unimolecular reactions, and 4,763 STOP decisions. These reference states are distinct from the generated-state collection above.

Pocket replacement cyclically assigns each decision to a different pocket in its batch while holding the ligand input fixed. Noise resampling reconstructs the same examples using dataset noise epochs 0 and 7, preserving graphs, reference targets, interpolation anchors, decision times, and legal actions. It changes the noise component of interpolated coordinates, not the entire current conformation. Paired structural inputs are checked for equality.

For reference-action log probability $\ell_{j,c}$ in configuration c and its perturbed value $\ell_{j,c}^{(q)}$, the response is $N^{-1}\sum_j |\ell_{j,c}^{(q)} - \ell_{j,c}|$. Each configuration has its own unperturbed baseline, and no model is retrained. The measure concerns a fixed reference action rather than the full action distribution.

Spatial interactions include both KNN feature messages and pocket-induced coordinate displacements. Direct pocket inputs include slot attention over pocket memory and pooled pocket features. Disabling spatial interactions zeros both messages and displacements: suppressing messages alone can retain a coordinate response through the displacement readout's bias. Table 5 includes these partial interventions.

Uncertainty and reaction-only results. For uncertainty estimates, we average responses within each original pocket, subtract the noise response from the replacement response, and bootstrap 92 equally weighted pockets with 10,000 replicates (seed 0). The full-model mean difference is 3.63 nats (95% CI [3.48, 3.77]), positive in all 92 pockets. This pocket-weighted difference is distinct from subtracting the row-weighted means in Table 3. Intervals condition on the checkpoint and paired perturbations, without accounting for training-seed or all replacement-pocket variability.

Table 5: **Pathway responses and reaction-only analysis.** Absolute log-probability changes in nats. All-state columns use 18,769 decisions; reaction-only columns use 5,136 bimolecular/unimolecular decisions. NLL is the unperturbed all-state reference-action negative log likelihood. Values below 0.01 are shown as bounds rather than exact zeros.

	All states		Reaction only		
Enabled pathways	Pocket	Noise	Pocket	Noise	NLL
All	3.83	0.15	4.80	0.41	5.52
Direct pocket inputs only	0.29	0.47	0.30	0.62	7.28
Spatial interactions only	3.36	0.18	4.69	0.50	6.23
None	< 0.01	0.51	< 0.01	0.69	7.75
All except feature messages	0.29	0.47	0.30	0.63	7.28
Pocket displacement only	0.02	0.51	0.03	0.69	7.75

STOP contributes 59.92% of the total absolute replacement response. A post-hoc analysis restricted to the 5,136 reaction decisions retains all 92 pockets: direct pocket inputs alone preserve 6.23% of the full response, spatial interactions alone preserve 97.70%, and the noise/replacement response ratio is 8.48%. The pocket-weighted difference is 4.34 nats (95% CI [4.10, 4.58]), positive in every pocket. The pattern therefore extends beyond stopping. At empty-state initialization, spatial interactions cannot transmit a pocket signal because no ligand atoms exist, so direct pocket inputs remain necessary. Response ratios are computed before rounding and need not add to one. Differences across decision times do not isolate coordinate maturity, because action types, graphs, and legal actions also change.

Numerical limits. Evaluation uses CUDA bfloat16 autocast. With all pocket pathways disabled, the mean and maximum replacement responses are 6.47×10^{-8} and 1.91×10^{-6} nats. However, 41 states in the full-model pocket comparison have a positive returned log probability on at least one side; the largest positive value across configurations is 0.0333. This indicates a numerical normalization discrepancy in the recorded probabilities. All rows are retained; excluding these states as a sensitivity check changes the mean from 3.827 to 3.833. We therefore interpret the large differences between configurations, while recognizing this uncertainty in the exact response values.

C.3 Exploratory pocket mutation

An incomplete Asp-to-Asn sampling run yields 15 paired wild-type/mutant pockets with 20 molecules per arm per pocket. Predefined SMARTS patterns identify potentially protonatable groups rather than predicting physiological-pH charge. The hypothesis was that neutralizing negatively charged side chains would reduce the fraction of generated molecules matching these patterns.

The fraction instead increases from 19.00% to 24.33%, a difference of $+5.33$ percentage points (paired t-test $p = 0.1038$). Three pockets decrease, seven increase, and five are unchanged. The mean pattern count per molecule also increases, from 0.2100 to 0.2833 ($p = 0.04953$), opposite to the hypothesized direction. Only the 15 paired pockets are included from the available 59 wild-type and 15 mutant records. The incomplete sample does not support the proposed charge-specific response, but it also cannot establish an absence of chemical specificity. It is not used as positive evidence for pocket-specific chemical selection.