FORESIGHTMOL: CHEMISTRY-GUIDED RETRIEVAL-AUGMENTED GENERATION FOR STRUCTURE-BASED DRUG DESIGN

Anonymous authors
Paper under double-blind review

ABSTRACT

Retrieval-augmented molecular generation uses existing molecules to guide ligand construction, but reference compatibility does not guarantee favorable downstream properties or chemically feasible growth. Ranking the resulting candidates solely by docking score may further limit the diversity of chemical alternatives retained under fixed experimental budgets. We propose ForesightMol, a chemistry-guided framework that integrates outcome-aware reference assessment, feasible molecular growth, and complementary candidate selection. A scientific agent leverages pretrained scientific knowledge to interpret retrieved evidence, prioritize references, and guide molecular growth, while paired short-horizon rollouts assess their downstream effects. Evidence from accepted references guides local growth decisions, while shared valence and connectivity constraints exclude infeasible extensions during both reference assessment and generation. A multi-objective set selection procedure then balances molecular quality, predicted binding compatibility, and structural diversity within a fixed delivery budget. Experiments on 100 CrossDocked2020 pockets show improvements of 29.4% in mean QED and 14.8% in scaffold diversity over the best baseline. Controlled analyses identify contributions from both generation and selection, supporting the framework's effectiveness in delivering diverse, high-quality molecular candidates.

1 INTRODUCTION

Structure-based drug design (SBDD) uses the three-dimensional environment of a protein binding pocket to identify and optimize small-molecule ligands for therapeutic targets (Anderson, 2003). Growing structural datasets and advances in geometric deep learning have enabled models to generate ligand structures directly within a pocket (Francoeur et al., 2020; Peng et al., 2022; Guan et al., 2023a). These models extend molecular exploration beyond predefined compound libraries and support the design of new compounds with shapes and chemical features tailored to the target pocket. Translating these advances into experimental candidates requires chemically plausible and synthetically accessible molecules (Gao & Coley, 2020), together with diverse alternatives within a limited validation budget.

Pocket-conditioned generative models produce ligand candidates for subsequent evaluation and selection. Retrieval augmentation guides generation using molecular examples and interaction priors (Wang et al., 2023; Zhang et al., 2025; Huang et al., 2024a). However, structural similarity does not consistently imply similar biological activity (Hu & Bajorath, 2012), motivating assessment of references through the growth they induce. Reference-guided extensions must also respect the valence and connectivity of the partial ligand. After generation, docking-based ranking alone does not explicitly account for structural redundancy, while prior work highlights the value of maintaining diversity during molecular optimization (Blaschke et al., 2020). The resulting challenge is to connect useful reference guidance with chemically feasible growth and a complementary shortlist under a limited experimental budget.

Existing work addresses parts of this bottleneck through different routes. Autoregressive models make molecular construction explicit, while diffusion models capture global spatial dependencies, but improved generative modeling leaves open how intermediate choices should be judged by their

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.

downstream consequences (Luo et al., 2021; Guan et al., 2023a; Schneuing et al., 2024). Chemical constraints and property-oriented optimization improve validity and individual molecular objectives, while the composition of the final shortlist requires a separate decision (Shi et al., 2020; Zhou et al., 2024). Retrieval-based methods supply structural or interaction priors through learned matching, but matching alone does not evaluate the continuation induced by a reference (Zhang et al., 2025; Huang et al., 2024a). Diversity-aware subset selection makes complementarity explicit, although its attainable solutions depend on the supplied pool (Nakamura et al., 2022). These complementary directions motivate a framework that connects reference assessment, chemically feasible growth, and the composition of the delivered set.

In this paper, we investigate a kind of retrieval-augmented ligand generation under a fixed delivery budget of K molecules per pocket. We observed that reference value depends on the chemically feasible growth it enables, while candidate value also depends on the alternatives retained in the final shortlist. This motivates combining retrieval-grounded chemical evidence with a scientific agent that leverages pretrained knowledge to form heuristic assessments, and using short-horizon rollouts to examine their downstream utility, and multi-objective selection to account for relationships among candidates. These tools offer a way to evaluate guidance before committing to it and to consider candidate sets beyond individual molecular scores. Adapting them to molecular design, however, raises three challenges: 1) How to estimate the reference utility from incomplete ligands with limited computation, 2) how to translate chemical evidence into molecular metadata guidance that respects valence and connectivity, and 3) how to balance molecular properties with structural complementarity while making selection decisions traceable.

To address the above challenges, we propose ForesightMol, a chemistry-guided framework that integrates reference assessment, feasible molecular growth, and fixed-budget delivery. A scientific agent interprets retrieved chemical evidence using pretrained scientific knowledge to prioritize reference alternatives and formulate local growth guidance. These heuristic assessments are then evaluated through short-horizon rollouts, and accepted evidence is translated into bounded potentials that guide molecular metadata prediction. Shared valence and connectivity constraints filter infeasible extensions during both reference assessment and generation. Multi-objective set selection balances molecular quality, predicted binding compatibility, chemical and scaffold diversity, and pocket coverage, with contribution records explaining each selection. Experiments on 100 CrossDocked2020 test pockets show improved molecular quality and diversity over Rag2Mol-G using conventional candidate selection. Crossed controls support complementary improvements from generation and selection, while a feasibility diagnostic shows fewer recorded failures and more completed molecules. Our main contributions are as follows.

- We discover a fundamental mismatch between reference compatibility and downstream generation utility in retrieval-augmented ligand design, and propose ForesightMol, a chemistry-guided framework that shifts retrieved references from similarity-based retrieval to chemically interpretable evidence that activates pretrained scientific knowledge for outcome-aware generation.
- We develop a coordinated generation and delivery procedure that translates rollout-verified reference evidence into chemically constrained autoregressive growth and optimizes a fixed-budget candidate set for molecular quality, predicted binding compatibility, and structural complementarity.
- Experiments on 100 CrossDocked2020 pockets demonstrate consistent improvements, increasing mean QED from 0.494 to 0.639 and scaffold diversity from 0.860 to 0.987, while 72% of accepted interventions improve downstream molecular quality.

2 Background and Related Work

Pocket-conditioned molecular generation. Learning-based structure-based drug design has progressed from voxel-based ligand density generation to explicit modeling of three-dimensional molecular structures (Ragoza et al., 2022). Autoregressive approaches construct ligands through sequential atom placement and bond prediction under pocket conditioning (Luo et al., 2021; Peng et al., 2022). DrugGPS guides motif-based growth through learned interactions between subpocket prototypes and molecular motifs (Zhang & Liu, 2023). Diffusion-based methods, including TargetDiff and DiffSBDD, generate ligand coordinates and atom types through equivariant denoising

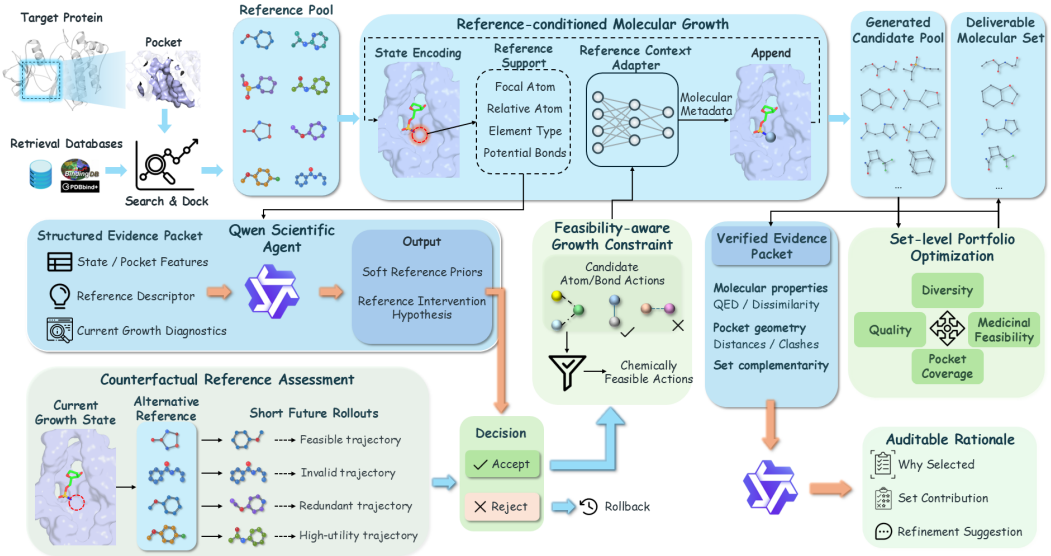

Figure 1: The overall schematic diagram of ForesightMol.

(Guan et al., 2023a; Schneuing et al., 2024). IPDiff further incorporates protein–ligand interaction priors into diffusion and sampling (Huang et al., 2024b). MolCRAFT models discrete atom types and continuous coordinates through continuous parameter distributions (Qu et al., 2024). Despite these advances, PoseBusters shows that accurate predicted binding poses can contain chemical and physical inconsistencies (Buttenschoen et al., 2024), highlighting the importance of explicit plausibility checks alongside geometric modeling.

Retrieval-guided molecular generation. Retrieval augmentation complements learned generative models with molecular examples that provide structural and interaction priors (Wang et al., 2023). At the fragment level, f-RAG combines retrieved building blocks with reference fragments to support reuse and novel generation (Lee et al., 2024). For pocket-conditioned design, Rag2Mol integrates a two-level retriever with autoregressive ligand generation (Zhang et al., 2025). IRDiff further uses interaction representations learned with binding-affinity supervision to retrieve ligands and guide diffusion (Huang et al., 2024a). These approaches extend reference matching from molecular structure to interaction-informed compatibility. Although retrieved references provide useful guidance, their compatibility with a target or an intermediate structure does not directly establish whether they will improve the properties or feasibility of subsequent molecular growth.

Chemical feasibility and candidate selection. Explicit constraints and optimization procedures provide additional control over generated molecules. GraphAF incorporates valence checks during graph construction (Shi et al., 2020), while DecompDiff introduces validity guidance and DecompOpt enables controllable substructure optimization in pocket-conditioned design (Guan et al., 2023b; Zhou et al., 2024). These approaches address the feasibility and quality of individual candidates. At the set level, SubMo-GNN combines learned molecular representations with submodular maximization to select subsets with structural and property diversity (Nakamura et al., 2022). Although these studies advance molecular validity, property optimization, and subset diversity, their coordination remains an open challenge: ranking by docking score alone does not explicitly account for structural redundancy, while diversity-aware selection is limited by the chemical alternatives available in the generated pool.

3 METHODS

3.1 PROBLEM DEFINITION

A protein pocket is represented as $\mathcal{P} = \{(\mathbf{p}_i, \mathbf{h}_i)\}_{i=1}^{N_P}$, where $\mathbf{p}_i \in \mathbb{R}^3$ and $\mathbf{h}_i$ denote atom coordinates and features, including element and residue types. A ligand $x = (\mathbf{X}, \mathbf{E}, \mathbf{B})$ comprises coordinates, element types, and a bond graph. Given $\mathcal{P}$, a retrieved reference set $\mathcal{R}(\mathcal{P})$, and a delivery budget K, our goal is to generate chemically valid ligands and select at most K molecules balancing molecular quality, predicted binding favorability, and structural complementarity.

3.2 OVERVIEW OF FORESIGHTMOL

As discussed in previous sections, we aim to improve reference utility, chemical feasibility, and set-level complementarity under limited computational and delivery budgets. To this end, we propose ForesightMol, a chemistry-guided framework for retrieval-augmented autoregressive molecular generation. As illustrated in Fig. 1, a scientific agent first prioritizes alternative references, which are evaluated through paired short-horizon rollouts from the same generation state. Chemical evidence associated with accepted references then guides subsequent growth, while shared valence and connectivity constraints filter infeasible extensions during both assessment and generation. Finally, set-level selection balances molecular quality, predicted binding compatibility, and structural diversity within a fixed delivery budget.

3.3 REFERENCE-CONDITIONED AUTOREGRESSIVE GROWTH

We build ForesightMol on a reference-conditioned autoregressive generator. Let x_t denote the partial ligand at generation step t, and define the current generation state as

$$s_t = (\mathcal{P}, x_t). \tag{1}$$

The retrieved set $\mathcal{R}(\mathcal{P})$ remains fixed for the target pocket. At each state, the generator uses a default reference

$$r_t^0 \in \arg\max_{r \in \mathcal{R}(\mathcal{P})} \phi(s_t, r), \tag{2}$$

where ϕ measures compatibility between the current state and a retrieved ligand.

Given a reference r, the encoder and cross-context interaction module produce the reference-conditioned representation

$$H_t(r) = E_\psi(s_t, r) = \big(H_t^{\mathrm{s}}(r), H_t^{\mathrm{v}}(r)\big), \tag{3}$$

where the scalar and vector features encode chemical and geometric context for subsequent equivariant growth.

A growth action $a_t = (f_t, u_t, e_t, b_t)$ specifies the focal site, relative position, element type, and bonds of the newly added atom. Conditioned on $H_t(r)$, the four prediction heads factorize the action distribution as

$$\begin{aligned} p_\theta(a_t \mid s_t, H_t(r)) = {} & p_\theta(f_t \mid s_t, H_t(r)) p_\theta(u_t \mid f_t, s_t, H_t(r)) \\ & \times p_\theta(e_t \mid u_t, f_t, s_t, H_t(r)) p_\theta(b_t \mid e_t, u_t, f_t, s_t, H_t(r)). \end{aligned} \tag{4}$$

An accepted action updates the partial ligand as $x_{t+1} = T(x_t, a_t)$. Generation proceeds until the model predicts completion, no admissible extension remains, or the maximum trajectory length is reached. This reference-conditioned process provides the common growth mechanism on which the reference intervention and feasibility constraints operate.

3.4 ROLLOUT-BASED REFERENCE INTERVENTION

At selected intermediate states of autoregressive generation, we reassess which references can best support further ligand growth. Reference utility depends on the evolving ligand structure, reference chemistry, and pocket environment, while exhaustive rollout evaluation is costly. We therefore use a frozen LLM-based scientific agent built on Qwen to bring broad scientific knowledge from large-scale pretraining into reference assessment. By combining this knowledge with current molecular

evidence, the agent jointly considers chemical growth opportunities, pocket compatibility, and potential risks beyond a single similarity criterion. This contextual assessment prioritizes promising reference alternatives under a limited rollout budget.

Using structured molecular and pocket evidence, the agent prioritizes a challenger subset

$$\mathcal{R}_t^{\mathcal{A}} \subseteq \mathcal{R}(\mathcal{P}) \setminus \{r_t^0\}, \qquad \left|\mathcal{R}_t^{\mathcal{A}}\right| \le L_A, \tag{5}$$

where L_A limits the number of challengers evaluated. For each challenger r, the agent also provides evidence for growth guidance and a confidence score $c_t(r)$. The input specification, prompt, and output schema are provided in Appendix D.

We compare each challenger with the default reference through matched short-horizon rollouts from the same generation state. All alternatives share the inherited decoding context η_t, initialization, and rules for beam search, termination, deduplication, and feasibility filtering. Each reference remains fixed during its rollout, while its encoding evolves with the partial ligand. New guidance is activated only after acceptance. For horizon H and random input ξ, we write

$$\tau_{t,H}(r;\xi) = \mathrm{Rollout}_H\big(s_t; \pi_\theta^{r,\eta_t}, \xi\big), \tag{6}$$

where π_θ^{r,η_t} is the growth policy conditioned on r under η_t.

The utility U_H rewards productive growth and valid, non-duplicate completions, with penalties for failed and duplicated branches. Using shared random inputs for each challenger–default pair, we estimate the utility gain over M replications as

$$\widehat{\Delta}_H(s_t, r) = \frac{1}{M}\sum_{m=1}^{M}\left[U_H\big(\tau_{t,H}(r;\xi_m)\big) - U_H\big(\tau_{t,H}(r_t^0;\xi_m)\big)\right]. \tag{7}$$

The default rollout is reused across challengers within each replication. Failure and duplication events also define the risk score $\widehat{R}_H(s_t, r)$.

Utility gain, risk change, and agent confidence define the eligible reference set

$$\mathcal{V}_t = \left\{ r \in \mathcal{R}_t^{\mathcal{A}} \;\middle|\; \begin{array}{r} \widehat{\Delta}_H(s_t,r) \ge \varepsilon_U, \\ \widehat{R}_H(s_t,r) - \widehat{R}_H(s_t,r_t^0) \le \varepsilon_R, \\ c_t(r) \ge \varepsilon_A \end{array} \right\}, \tag{8}$$

where $\varepsilon_U > 0$ and $\varepsilon_R \ge 0$. If $\mathcal{V}_t \neq \varnothing$, we accept

$$r_t^\star \in \arg\max_{r \in \mathcal{V}_t} \widehat{\Delta}_H(s_t, r), \tag{9}$$

breaking ties using agent assessments. Otherwise, the queried branch uses the default reference and clears its stored action prior. Unqueried branches retain their reference and guidance.

After acceptance, agent evidence guides focal-site, spatial-placement, element, and bond prediction through bounded, confidence-weighted adjustments. Reweighting the baseline distribution $\bar{p}_t^r$ over candidate actions $a \in \mathcal{B}_t(r)$ gives

$$p_t^{\mathcal{A}}(a \mid r) = \frac{\bar{p}_t^r(a)\exp(\lambda_{\mathcal{A}}\Gamma_t^r(a))}{Z_t^r}, \tag{10}$$

where $\lambda_{\mathcal{A}} > 0$ controls guidance strength, Z_t^r normalizes the distribution, and $\Gamma_t^r(a)$ combines head-wise evidence and conditional normalization terms. This reweighting preserves the generator's action support. The head-wise construction and KL-regularized interpretation appear in Appendix B.

3.5 Feasibility-Aware Growth Filtering

ForesightMol filters candidate growth actions according to local connectivity and valence constraints. Given the candidate set $\mathcal{B}_t(r)$, the feasible action set is

$$\mathcal{F}_t(r) = \{a \in \mathcal{B}_t(r) \mid \chi(s_t, a) = 1\}, \tag{11}$$

where $\chi(s_t, a)$ indicates whether the proposed extension is chemically admissible.

For an action that adds an atom of element type e_{new}, let $J(a)$ denote the existing atoms connected to the new atom, b_j the corresponding bond orders, and $v_j(x_t)$ the current bond-order sum of atom j. With $\nu(e)$ denoting the configured valence upper bound, an action is feasible when

$$J(a) \neq \varnothing, \qquad \sum_{j \in J(a)} b_j \leq \nu(e_{\text{new}}), \qquad v_j(x_t) + b_j \leq \nu(e_j),\ \forall j \in J(a). \tag{12}$$

The connectivity condition is omitted for the initial atom. These constraints are evaluated before child-state construction, and beam pruning and expansion are performed over $\mathcal{F}_t(r)$.

The same filtering rule is used during reference assessment and full molecular growth. Completed molecules are sanitized and canonically deduplicated before entering the delivery pool.

3.6 Fixed-Budget Set Delivery

The delivery stage jointly selects molecules to balance individual quality with structural diversity and pocket coverage. Let $\mathcal{C} = \{x_1, \ldots, x_N\}$ denote the sanitized, canonically deduplicated candidate pool with valid Vina Score and Vina Dock evaluations. We characterize each subset $\mathcal{S} \subseteq \mathcal{C}$ by

$$\mathbf{m}(\mathcal{S}; \mathcal{P}) = [Q_{\text{med}}, A_{\text{bind}}, D_{\text{chem}}, C_{\text{scaf}}, C_{\text{pocket}}]^\top, \tag{13}$$

where the components measure molecular quality, predicted binding favorability, molecular diversity, scaffold diversity, and contacted-residue coverage, respectively.

We combine these statistics into a weighted objective and optimize it under the delivery budget K,

$$\mathcal{S}^\star \in \arg \max_{\substack{\mathcal{S} \subseteq \mathcal{C} \\ |\mathcal{S}| = K'}} F_{\boldsymbol{\lambda}}(\mathcal{S}; \mathcal{P}), \qquad K' = \min(K, N), \tag{14}$$

where

$$F_{\boldsymbol{\lambda}}(\mathcal{S}; \mathcal{P}) = \boldsymbol{\lambda}^\top \mathbf{m}(\mathcal{S}; \mathcal{P}), \qquad \boldsymbol{\lambda} \geq 0. \tag{15}$$

The coefficients are fixed throughout selection. Because diversity and coverage depend on the selected molecules, each candidate's marginal value reflects its complementarity with the current set.

We approximate this optimization by greedy marginal addition followed by improving one-for-one exchanges. For the resulting set, we quantify each molecule's contribution by the change in set statistics upon its removal, keeping normalization scales and coverage denominators unchanged. Detailed metric definitions, coefficients, and search settings appear in Appendix C.

4 Experiments

4.1 Experimental Setup

Data and baselines We evaluate on 100 CrossDocked2020 test pockets (Francoeur et al., 2020), following the processed split and retrieval database organization of Rag2Mol (Zhang et al., 2025). Baselines include AR (Luo et al., 2021), GraphBP (Liu et al., 2022), Pocket2Mol (Peng et al., 2022), TargetDiff (Guan et al., 2023a), DiffSBDD (Schneuing et al., 2024), and Rag2Mol-G (Zhang et al., 2025). ForesightMol and Rag2Mol-G share the generator checkpoint, beam width of 128, target of 64 completed molecules, and maximum of 50 growth steps.

Evaluation We evaluate generated ligands in terms of structure, predicted binding affinity, and molecular properties. Structural metrics include Jensen–Shannon divergence (JSD) (Lee, 1999) of heavy-atom pair-distance and bond-length distributions against reference ligands, ring-size statistics, and rigid-fragment RMSD under force-field relaxation. AutoDock Vina evaluates generated poses (Vina Score) and redocked poses (Vina Dock) (Eberhardt et al., 2021). Molecular properties include mean and best QED (Bickerton et al., 2012), raw SA (Ertl & Schuffenhauer, 2009), Lipinski compliance (Lipinski et al., 2001), Morgan-fingerprint Tanimoto diversity (Rogers & Hahn, 2010), and Bemis–Murcko scaffold diversity (Bemis & Murcko, 1996). Lower JSD, RMSD, SA, and Vina values are preferred.

We assess candidate pools and delivered sets of up to $K = 10$ molecules per pocket. Details appear in Appendix E.

Method	N	Mean QED↑	Best QED↑	SA↓	Lipinski↑	Mol. div.↑	Scaf. div.↑
AR	97.22	0.507	0.785	4.370	0.886	0.843	0.610
GraphBP	99.60	0.416	0.812	5.828	0.641	**0.907**	0.920
Pocket2Mol	45.23	0.516	0.813	3.540	**0.903**	0.873	0.675
TargetDiff	78.78	0.469	0.827	4.768	0.718	0.889	0.894
DiffSBDD	100.00	0.464	**0.875**	4.809	0.772	0.899	**0.925**
Rag2Mol-G	67.91	0.518	0.790	3.570	0.823	0.823	0.753
ForesightMol	69.21	**0.577**	0.795	**3.246**	0.830	0.828	0.776

Table 1: **Unselected candidate-pool quality.** N is the mean number of evaluated molecules per pocket. All metrics are recomputed under a common evaluation protocol. Bold marks the best column mean.

		Mean QED ↑		Best QED ↑		SA ↓		Mol. div. ↑		Scaf. div. ↑		Vina Score ↓		Vina Dock ↓	
Generator	Selector	Avg.	Med.	Avg.	Med.	Avg.	Med.	Avg.	Med.	Avg.	Med.	Avg.	Med.	Avg.	Med.
Rag2Mol-G	Dock	0.494	0.518	0.641	0.6858	3.755	3.6525	0.723	0.7515	0.860	0.900	−6.835	−6.6619	−9.833	−9.5468
Rag2Mol-G	S–D	0.533	0.552	0.667	0.6856	3.662	3.6446	0.758	0.7844	0.922	**1.000**	−6.624	−6.3498	−8.880	−8.8361
ForesightMol	Dock	0.513	0.554	0.687	0.7297	3.633	**3.4279**	0.763	0.7859	0.874	0.921	**−6.938**	−6.7411	**−9.910**	**−9.6334**
ForesightMol	S–D	**0.639**	**0.660**	**0.782**	**0.8032**	**3.560**	3.5494	**0.812**	0.8220	**0.987**	**1.000**	−6.863	**−6.8664**	−9.006	−8.8256

Table 2: The effect of generation and selection strategies on fixed-budget molecular delivery. Avg. and Med. denote the mean and median across pockets. S–D denotes Score–Dock-aware selection. Bold indicates the best result in each column.

4.2 Main Results

Molecular generation. ForesightMol improves the quality of generated candidates before delivery-stage selection. As shown in Table 1, it achieves the highest mean QED (0.577) and lowest SA (3.246) among the evaluated methods. Compared with the matched baseline, it also improves best QED, Lipinski compliance, and both molecular and scaffold diversity, while producing a similar number of evaluated molecules per pocket. The improvements therefore extend beyond average property scores to the quality and variety of available candidates, providing a stronger starting pool for fixed-budget delivery.

Fixed-budget delivery. The complete framework translates these candidates into molecular sets with improved properties and structural diversity (Table 2). Compared with the matched baseline using Dock ranking, ForesightMol with S–D selection increases mean QED from 0.494 to 0.639 and scaffold diversity from 0.860 to 0.987, alongside higher best QED and molecular diversity and lower SA. The pocket-level distributions in Figure 2 show corresponding shifts toward higher QED and diversity. Mean Vina Score remains similar, while mean Vina Dock increases by 0.827 kcal/mol.

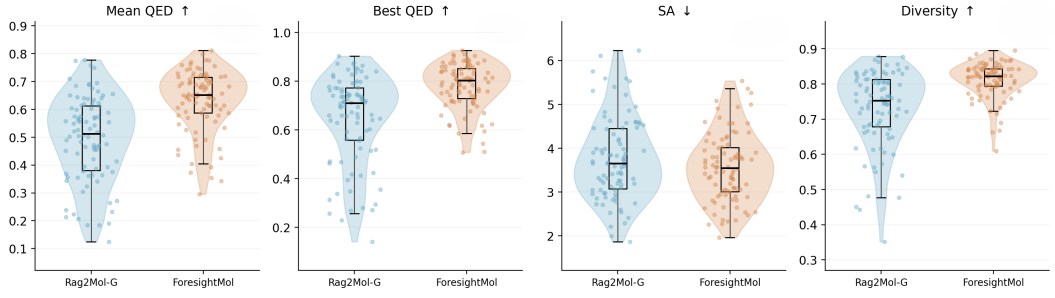

Figure 2: Distributions of per-pocket molecular quality and diversity for Rag2Mol-G and ForesightMol across test pockets. Each point represents one pocket. Violin plots show the distributions, with overlaid box plots indicating medians and interquartile ranges.

Table 3: The effect of reference intervention and chemical constraints on molecular generation. Avg. and Med. denote the mean and median across pockets. Bold indicates the best result in each column.

Method	Mean QED ↑		Best QED ↑		SA ↓		Mol. div. ↑		Scaf. div. ↑		Vina Score ↓		Vina Dock ↓	
	Avg.	Med.	Avg.	Med.	Avg.	Med.	Avg.	Med.	Avg.	Med.	Avg.	Med.	Avg.	Med.
Baseline	0.507	0.509	0.719	0.774	3.858	3.789	0.743	0.752	0.686	0.684	−5.1597	−4.1652	−7.8409	−6.9813
+ Reference intervention	0.516	**0.566**	0.736	0.782	3.797	3.673	0.745	0.766	0.752	0.793	−5.3915	−4.2990	−8.1320	−7.3484
+ Chemical constraints	0.510	0.547	0.742	0.776	3.578	3.560	0.758	0.773	0.700	0.719	−5.5602	−4.4022	−7.9188	−6.9883
ForesightMol	**0.525**	0.563	**0.762**	**0.792**	**3.461**	**3.492**	**0.773**	**0.780**	**0.753**	**0.794**	**−5.8675**	**−4.6020**	**−8.4024**	**−7.4731**

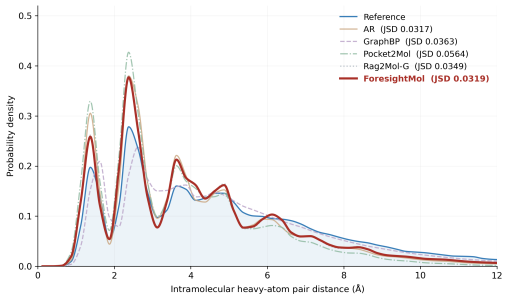

Figure 3: Distributions of intramolecular heavy-atom pair distances for reference ligands (blue) and generated molecules across test pockets. Jensen–Shannon divergence (JSD) from the reference distribution is reported for each method.

Bond	Rag2Mol-G	ForesightMol
C–C	0.4633	**0.4074**
C=C	**0.3984**	0.4343
C–N	0.3064	**0.2792**
C=N	0.6331	**0.5468**
C–O	0.3009	**0.2966**
C=O	0.4329	**0.3217**
C:C	**0.2912**	0.2946
C:N	0.3500	**0.3202**

Table 4: Jensen–Shannon divergence (JSD, in bits) between the bond-length distributions of reference ligands and delivered molecules. Lower values indicate closer agreement. "–", "=", and ":" denote single, double, and aromatic bonds, respectively.

These results demonstrate improved molecular quality and chemical variety within the same delivery budget, with a trade-off in redocking energy.

4.3 Analysis of Framework Components

Generation and selection. The crossed comparison in Table 2 shows complementary benefits from generation and selection. On both fixed pools, S–D selection improves mean and best QED, SA, and molecular and scaffold diversity over Dock ranking. With the selector held constant, ForesightMol improves all seven metric means under Dock ranking and all seven under S–D selection. These results support coordinating generation and selection to improve candidate quality and structural complementarity under a fixed delivery budget.

Reference guidance and chemical constraints. Table 3 shows complementary contributions from the two growth components. Reference intervention yields larger gains in mean QED and scaffold diversity, while chemical constraints contribute more to lower SA and higher molecular diversity. Their combination achieves the best mean value for every reported metric. Under a fixed reference policy without agent intervention, a paired diagnostic shows that filtering eliminates recorded post-beam failures and increases mean completions from 12.9 to 14.4. Although QED decreases in this diagnostic, it improves with the combined configuration, supporting the joint use of feasible continuation and reference guidance.

Reference decisions. To examine whether short-horizon assessment identifies useful reference changes, we analyze the decisions and subsequent generation outcomes. Using paired two-step rollouts with beam width eight, ForesightMol evaluates 248 query states across 98 pockets and accepts 143 interventions across 72 pockets, giving an acceptance rate of 57.7%. The remaining 105 queries retain the default reference. Further comparative testing finds that 72% of accepted interventions improve molecular quality during subsequent generation. This connects the local reference assessments to downstream molecular outcomes and supports their use in directing growth with a limited rollout budget.

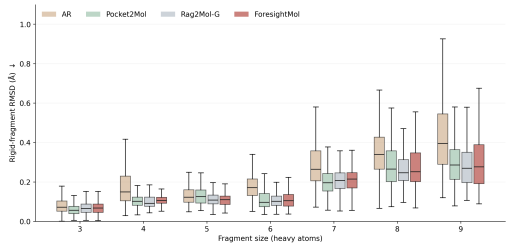

Figure 4: Median RMSD of rigid fragments before and after force-field optimization.

Ring size	Ref.	AR	Pocket2Mol	Rag2Mol-G	ForesightMol
3	3.33%	54.39%	0.34%	0.13%	0.18%
4	0.00%	0.46%	0.04%	0.03%	0.05%
5	51.11%	33.78%	34.73%	38.85%	42.12%
6	86.67%	74.54%	84.77%	96.04%	87.90%
7	2.22%	4.12%	11.05%	10.47%	11.03%
8	0.00%	2.38%	2.20%	0.25%	0.79%
9	0.00%	0.78%	0.56%	0.36%	0.36%

Table 5: Percentages of different ring sizes in reference and model-generated molecules.

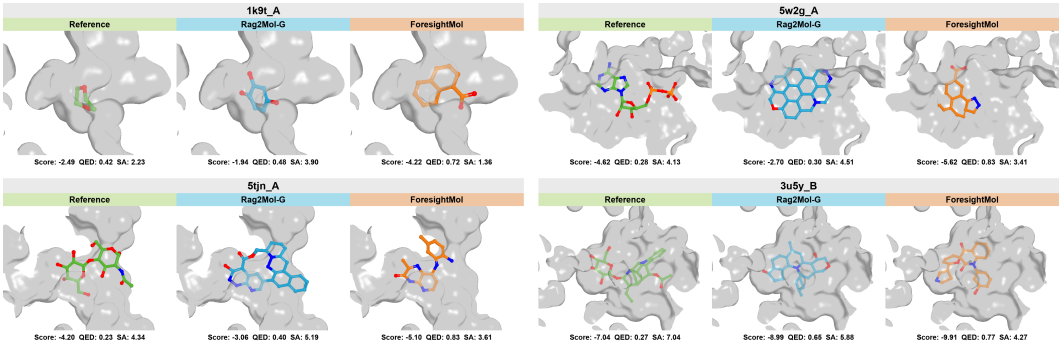

Figure 5: Illustrative examples of generated ligands for protein pockets 1k9t_A, 5w2g_A, 5tjn_A, and 3u5y_B. Carbon atoms in native reference ligands, ligands generated by Rag2Mol-G, and ForesightMol are shown in green, cyan, and orange, respectively. Protein pockets are shown as gray surfaces, using identical views and scales within each pocket. Individual-molecule Vina Score (kcal/mol, ↓), QED (↑), and SA (↓) are reported.

4.4 Structural and Qualitative Analysis

ForesightMol's property improvements are accompanied by closer agreement with reference ligands in molecular geometry and ring composition. It reduces heavy-atom pair-distance JSD relative to Rag2Mol-G, GraphBP, and Pocket2Mol (Figure 3) and improves bond-length agreement over the matched baseline in six of eight bond types (Table 4). It also more closely reproduces reference frequencies of five- and six-membered rings (Table 5). Under force-field relaxation, rigid-fragment RMSD distributions remain comparable to the matched baseline, with lower medians than AR across evaluated fragment sizes (Figure 4).

Figure 5 illustrates these results through pocket examples shown at matched views and scales. The displayed ForesightMol ligands achieve higher QED, lower SA, and more favorable Vina Score than their baseline counterparts. Their distinct ring systems and substituent arrangements illustrate how favorable molecular properties and predicted binding scores are realized in concrete chemical structures within the target pockets.

5 Conclusion

We presented ForesightMol, a chemistry-guided framework for protein-specific molecular generation that connects retrieval, generation, and fixed-budget delivery. Retrieved molecules provide chemical evidence for a scientific agent to draw on pretrained knowledge and propose reference interventions, which are tested through paired short-horizon rollouts. Verified guidance supports molecular growth under valence and connectivity constraints, followed by set-level selection that balances molecular quality, predicted binding compatibility, and structural diversity. Experiments on CrossDocked2020 show consistent gains in drug-likeness, synthetic accessibility, and delivered-set diversity.

AI USE STATEMENT

We used large language models to assist with manuscript drafting, language editing, organization, and refinement of technical explanations. We also consulted these tools for suggestions on related literature and presentation of experimental results. Separately, ForesightMol incorporates an LLM-based agent for reference assessment as part of the proposed method, with its role and configuration described in the paper.

The authors are responsible for reviewing AI-assisted material and verifying its consistency with the implementation, experimental evidence, and cited sources. We take full responsibility for the final manuscript, including all claims, analyses, and artifacts produced with AI assistance.

ETHICS STATEMENT

This work adheres to the ICLR Code of Ethics and involves no human participants, animal experiments, or personally identifiable information. We use the CrossDocked2020 dataset in accordance with its applicable license and usage guidelines. We have considered the potential biases and ethical implications of this research and are committed to maintaining transparency and research integrity.

REPRODUCIBILITY STATEMENT

We provide our source code in the supplementary materials to support reproduction of the reported experiments. Mathematical derivations, proofs, and the set-selection procedure are provided in Appendices B and C. Implementation details and experimental protocols are documented in Appendices D and E.

REFERENCES

Amy C Anderson. The process of structure-based drug design. *Chemistry & biology*, 10(9):787–797, 2003.

Guy W Bemis and Mark A Murcko. The properties of known drugs. 1. molecular frameworks. *Journal of medicinal chemistry*, 39(15):2887–2893, 1996.

G Richard Bickerton, Gaia V Paolini, Jérémy Besnard, Sorel Muresan, and Andrew L Hopkins. Quantifying the chemical beauty of drugs. *Nature chemistry*, 4(2):90–98, 2012.

Thomas Blaschke, Ola Engkvist, Jürgen Bajorath, and Hongming Chen. Memory-assisted reinforcement learning for diverse molecular de novo design. *Journal of cheminformatics*, 12(1):68, 2020.

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(9):3130–3139, 2024.

Jerome Eberhardt, Diogo Santos-Martins, Andreas F Tillack, and Stefano Forli. Autodock vina 1.2. 0: new docking methods, expanded force field, and python bindings. *Journal of chemical information and modeling*, 61(8):3891, 2021.

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.

Paul G Francoeur, Tomohide Masuda, Jocelyn Sunseri, Andrew Jia, Richard B Iovanisci, Ian Snyder, and David R Koes. 3d convolutional neural networks and a crossdocked dataset for structure-based drug design. *Journal of chemical information and modeling*, 60(9):4200, 2020.

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.

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://mlanthology.org/iclr/2023/guan2023iclr-3d/.

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 Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett (eds.), *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 11827–11846. PMLR, 23–29 Jul 2023b. URL https://proceedings.mlr.press/v202/guan23a.html.

Ye Hu and Jürgen Bajorath. Many structurally related drugs bind different targets whereas distinct drugs display significant target overlap. *RSC advances*, 2(8):3481–3489, 2012.

Zhilin Huang, Ling Yang, Xiangxin Zhou, Chujun Qin, Yijie Yu, Xiawu Zheng, Zikun Zhou, Wentao Zhang, Yu Wang, and Wenming Yang. Interaction-based retrieval-augmented diffusion models for protein-specific 3d molecule generation. In *Forty-first International Conference on Machine Learning*, 2024a.

Zhilin Huang, Ling Yang, Xiangxin Zhou, Zhilong Zhang, Wentao Zhang, Xiawu Zheng, Jie Chen, Yu Wang, Bin Cui, and Wenming Yang. Protein-ligand interaction prior for binding-aware 3d molecule diffusion models. In *International Conference on Learning Representations*, volume 2024, pp. 18453–18474, 2024b.

Lillian Lee. Measures of distributional similarity. In *Proceedings of the 37th annual meeting of the Association for Computational Linguistics*, pp. 25–32, 1999.

Seul Lee, Karsten Kreis, Srimukh P Veccham, Meng Liu, Danny Reidenbach, Saee Paliwal, Arash Vahdat, and Weili Nie. Molecule generation with fragment retrieval augmentation. *Advances in Neural Information Processing Systems*, 37:132463–132490, 2024.

Christopher A Lipinski, Franco Lombardo, Beryl W Dominy, and Paul J Feeney. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced drug delivery reviews*, 46(1-3):3–26, 2001.

Meng Liu, Youzhi Luo, Kanji Uchino, Koji Maruhashi, and Shuiwang Ji. Generating 3D molecules for target protein binding. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvari, Gang Niu, and Sivan Sabato (eds.), *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 13912–13924. PMLR, 17–23 Jul 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. *Advances in Neural Information Processing Systems*, 34:6229–6239, 2021.

Tomohiro Nakamura, Shinsaku Sakaue, Kaito Fujii, Yu Harabuchi, Satoshi Maeda, and Satoru Iwata. Selecting molecules with diverse structures and properties by maximizing submodular functions of descriptors learned with graph neural networks. *Scientific reports*, 12(1):1124, 2022.

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.

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 Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp (eds.), *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 41749–41768. PMLR, 21–27 Jul 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(9):2701–2713, 2022.

David Rogers and Mathew Hahn. Extended-connectivity fingerprints. *Journal of chemical information and modeling*, 50(5):742–754, 2010.

Arne Schneuing, Charles Harris, Yuanqi Du, Kieran Didi, Arian Jamasb, Ilia Igashov, Weitao Du, Carla Gomes, Tom L Blundell, Pietro Lio, et al. Structure-based drug design with equivariant diffusion models. *Nature Computational Science*, 4(12):899–909, 2024.

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.

Zichao Wang, Weili Nie, Zhuoran Qiao, Chaowei Xiao, Richard Baraniuk, and Anima Anandkumar. Retrieval-based controllable molecule generation. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=vDFA1tpuLvk.

Peidong Zhang, Xingang Peng, Rong Han, Ting Chen, and Jianzhu Ma. Rag2mol: structure-based drug design based on retrieval augmented generation. *Briefings in Bioinformatics*, 26(3):bbaf265, 2025.

Zaixi Zhang and Qi Liu. Learning subpocket prototypes for generalizable structure-based drug design. In *International Conference on Machine Learning*, pp. 41382–41398. PMLR, 2023.

Xiangxin Zhou, Xiwei Cheng, Yuwei Yang, Yu Bao, Liang Wang, and Quanquan Gu. Decompopt: Controllable and decomposed diffusion models for structure-based molecular optimization. In *International Conference on Learning Representations*, volume 2024, pp. 5890–5912, 2024.

A Motivation: Structural Similarity Does Not Guarantee Guidance Utility

Let S denote a molecular growth state, including the target pocket and partial ligand, and let R denote a candidate reference. Define its matching score as $M = m(S, R)$ and its expected continuation utility as

$$V_H(S, R) = \mathbb{E}_{\tau \sim P_\theta(\cdot|S,R)} \left[U(\tau) \right], \tag{16}$$

where the generator, rollout horizon H, and decoding settings are fixed.

Proposition (Information loss in similarity-based assessment). Assume that $V_H(S, R)$ is square-integrable under a given joint distribution of states and references. Among all measurable predictors that depend only on (S, M), the minimum expected squared error satisfies

$$\inf_g \mathbb{E}\Big[\big(V_H(S, R) - g(S, M) \big)^2 \Big] = \mathbb{E}[\mathrm{Var}(V_H(S, R) \mid S, M)]. \tag{17}$$

Consequently, this error is strictly positive whenever references with the same state and matching score have different expected continuation utilities with positive conditional probability.

Proof. Write $V = V_H(S, R)$ and $\mu(S, M) = \mathbb{E}[V \mid S, M]$. By the conditional expectation identity,

$$\begin{aligned} \mathbb{E}\big[(V - g(S, M))^2\big] &= \mathbb{E}[\mathrm{Var}(V \mid S, M)] \\ &\quad + \mathbb{E}\big[(\mu(S, M) - g(S, M))^2\big]. \end{aligned} \tag{18}$$

The cross term vanishes because $\mathbb{E}[V - \mu(S, M) \mid S, M] = 0$. The second term is nonnegative and equals zero for $g = \mu$, establishing Eq. equation 17. □

Corollary (Matching scores need not identify the better reference). Fix a state s and two references r_a, r_b satisfying

$$m(s, r_a) = m(s, r_b), \qquad \Delta = V_H(s, r_b) - V_H(s, r_a) > 0. \tag{19}$$

Every predictor based only on the state and matching score assigns these references the same predicted utility. Under a uniform distribution over the two references, its minimum mean squared error is $\Delta^2/4$. If the score tie is resolved uniformly, the expected selection regret relative to choosing r_b is $\Delta/2$.

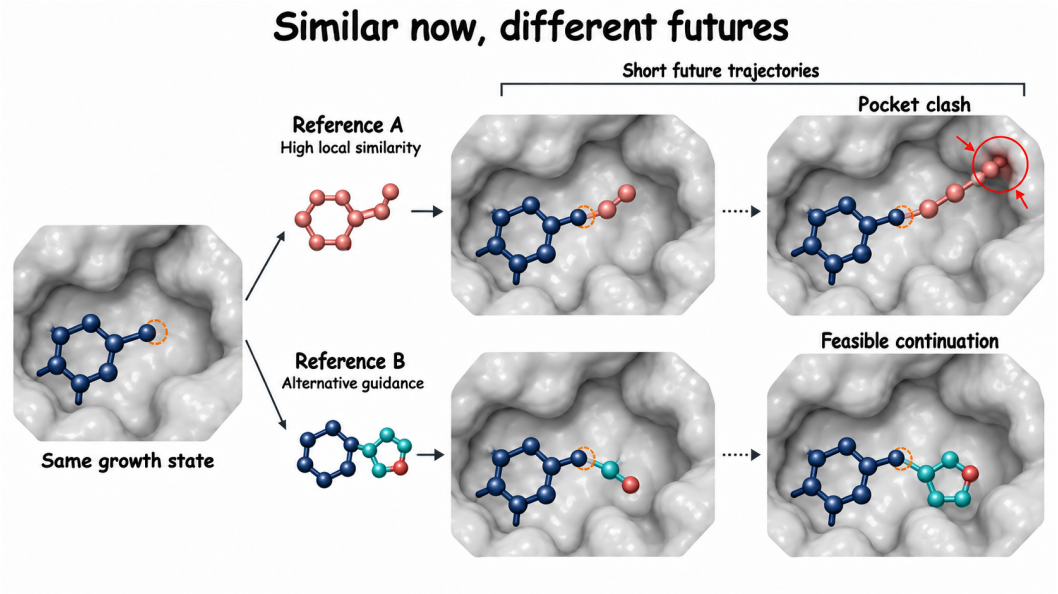

Figure S1: **Structural similarity does not guarantee favorable continuation.** From the same partial ligand and protein pocket, a reference with high local similarity may guide growth toward a sterically restricted region (top), whereas an alternative reference may support a pocket-compatible extension (bottom). Dark blue denotes the existing ligand, and colored extensions illustrate prospective growth. This conceptual illustration motivates evaluating references through their downstream consequences rather than matching scores alone.

Implication. Structural matching alone does not guarantee identification of the reference with the highest continuation utility. The relevant limitation is utility variation that remains unresolved by the matching score, rather than necessarily low overall correlation between similarity and utility. This motivates assessing reference guidance through its induced future trajectories.

B Reference-Guided Growth: Formulation and Properties

This section gives the detailed formulation of reference-guided growth and establishes its distributional and graph-level properties. We separate the mechanism into three components: structured evidence extraction, conservative policy adaptation, and chemical feasibility preservation. Implementation-specific numerical settings are collected into a single fixed configuration described in Appendix D and are shared across all experiments.

B.1 Head-Wise Evidence Guidance

Head-wise evidence potentials. Write a complete growth action as

$$a = (a^{(1)}, a^{(2)}, a^{(3)}, a^{(4)}) = (f, u, e, b),$$

corresponding to focal-site selection, spatial placement, atom-type prediction, and bond prediction. Let $a^{(<k)}$ denote the decisions preceding head k. For an accepted reference r_i, let $\mathbf{g}_{t,i}$ denote its structured assessment and let $c_{t,i} \in [0,1]$ denote the associated confidence signal. We represent the evidence adjustment for prediction head k as

$$\gamma_{t,i}^{(k)}\big(a^{(k)}; a^{(<k)}\big) = c_{t,i}\, m_{t,i}^{(k)}\big(a^{(\le k)}\big)\, h_{t,i}^{(k)}\big(a^{(\le k)}; \mathbf{g}_{t,i}\big), \tag{20}$$

where $m_{t,i}^{(k)} \in [0,1]$ measures local support supplied by the accepted reference and $h_{t,i}^{(k)}$ is a bounded head-specific evidence function.

This factorization separates three distinct quantities: the reliability of the proposal, support from the retrieved molecule, and the direction of the local adjustment. Consequently, weakly supported evidence vanishes continuously rather than introducing new actions outside the generator's support.

Conditional policy adaptation. Let $\mathcal{C}_{t,i}^{(k)}(a^{(<k)})$ denote the finite candidate support at head k, and let

$$\bar{p}_{t,i}^{(k)}(\cdot \mid a^{(<k)})$$

be the normalized baseline conditional distribution under reference r_i and the inherited decoding context. We adapt this distribution by exponential tilting:

$$p_{t,i}^{\mathcal{A},(k)}\big(a^{(k)} \mid a^{(<k)}\big) = \frac{\bar{p}_{t,i}^{(k)}\big(a^{(k)} \mid a^{(<k)}\big) \exp\Big\{\lambda_{\mathcal{A}} \gamma_{t,i}^{(k)}\big(a^{(k)}; a^{(<k)}\big)\Big\}}{Z_{t,i}^{(k)}\big(a^{(<k)}\big)}, \tag{21}$$

where

$$Z_{t,i}^{(k)}\big(a^{(<k)}\big) = \sum_{z \in \mathcal{C}_{t,i}^{(k)}(a^{(<k)})} \bar{p}_{t,i}^{(k)}\big(z \mid a^{(<k)}\big) \exp\Big\{\lambda_{\mathcal{A}} \gamma_{t,i}^{(k)}\big(z; a^{(<k)}\big)\Big\}. \tag{22}$$

The baseline and adapted joint distributions factorize as

$$\begin{aligned} \bar{p}_{t,i}(a) &= \prod_{k=1}^{4} \bar{p}_{t,i}^{(k)}\big(a^{(k)} \mid a^{(<k)}\big), \\ p_{\theta}^{\mathcal{A}}(a \mid s_t, r_i) &= \prod_{k=1}^{4} p_{t,i}^{\mathcal{A},(k)}\big(a^{(k)} \mid a^{(<k)}\big). \end{aligned} \tag{23}$$

Here, $\mathcal{B}_{t,i}$ denotes the set of complete actions induced by the conditional candidate supports. When the evidence potential vanishes, the adapted distribution reduces exactly to the baseline distribution.

Effective joint potential. For $\lambda_{\mathcal{A}} > 0$, define

$$\Gamma_{t,i}(a) = \sum_{k=1}^{4} \left[\gamma_{t,i}^{(k)}\big(a^{(k)}; a^{(<k)}\big) - \frac{\log Z_{t,i}^{(k)}\big(a^{(<k)}\big)}{\lambda_{\mathcal{A}}} \right]. \tag{24}$$

Substitution of Eq. equation 21 into Eq. equation 23 yields

$$p_{\theta}^{\mathcal{A}}(a \mid s_t, r_i) = \bar{p}_{t,i}(a) \exp\{\lambda_{\mathcal{A}} \Gamma_{t,i}(a)\}. \tag{25}$$

The conditional normalization terms account for the dependence of later prediction heads on earlier decisions and ensure that the resulting joint policy remains normalized.

Activation during generation. Reference proposals are evaluated through matched short-horizon continuations under an identical inherited decoding context. Only references satisfying the verification rule in Eq. equation 8 activate their associated evidence. Otherwise, the queried branch retains the default reference and clears newly proposed guidance. Unqueried branches retain their current reference and guidance. Thus, proposal generation, empirical verification, and graded policy adaptation remain separate stages.

B.2 KL-Regularized Policy Interpretation

Using the notation of Eq. equation 10, the corresponding joint normalizer is

$$Z_t^r = \sum_{a \in \mathcal{B}_t(r)} \bar{p}_t^r(a) \exp(\lambda_{\mathcal{A}} \Gamma_t^r(a)).$$

For finite potentials, exponential reweighting preserves baseline support. For any admissible distribution $q \ll \bar{p}_t^r$,

$$D_{\mathrm{KL}}(q \| p_t^{\mathcal{A}}) = D_{\mathrm{KL}}(q \| \bar{p}_t^r) - \lambda_{\mathcal{A}} \mathbb{E}_q[\Gamma_t^r] + \log Z_t^r. \tag{26}$$

With the effective potential in Eq. equation 24, the conditional normalizers are already absorbed into the joint representation, giving

$$D_{\mathrm{KL}}\big(q \,\|\, p_\theta^{\mathcal{A}}\big) = D_{\mathrm{KL}}(q \,\|\, \bar{p}_{t,i}) - \lambda_{\mathcal{A}} \mathbb{E}_q[\Gamma_{t,i}(a)]. \tag{27}$$

Rearranging,

$$\mathbb{E}_q[\Gamma_{t,i}(a)] - \frac{1}{\lambda_{\mathcal{A}}} D_{\mathrm{KL}}(q \,\|\, \bar{p}_{t,i}) = -\frac{1}{\lambda_{\mathcal{A}}} D_{\mathrm{KL}}\big(q \,\|\, p_\theta^{\mathcal{A}}\big)\,. \tag{28}$$

Nonnegativity of KL divergence therefore gives

$$p_t^{\mathcal{A}}(\cdot \mid r) = \arg \max_{\substack{q \in \Delta(\mathcal{B}_t(r)) \\ q \ll \bar{p}_t^r}} \left\{ \mathbb{E}_q[\Gamma_t^r(a)] - \frac{1}{\lambda_{\mathcal{A}}} D_{\mathrm{KL}}(q \| \bar{p}_t^r) \right\}. \tag{29}$$

Hence, reference-guided adaptation can be interpreted as a KL-regularized policy update: evidence-supported actions receive additional probability mass while deviation from the pretrained generator is explicitly regularized. This interpretation is independent of the downstream feasibility filter, pruning operation, and beam-search approximation.

B.3 Bounded Deviation under Confidence-Weighted Guidance

Fix a generation state and accepted reference and abbreviate $c = c_{t,i}$, $\bar{p} = \bar{p}_{t,i}$, and $p^{\mathcal{A}} = p_\theta^{\mathcal{A}}(\cdot \mid s_t, r_i)$. Suppose that

$$\left| h_{t,i}^{(k)} \right| \le B_k$$

for every supported candidate and decision prefix, with $B_k < \infty$. Then the adapted distribution preserves baseline support and satisfies

$$\left| \log \frac{p^{\mathcal{A}}(a)}{\bar{p}(a)} \right| \le 2\lambda_{\mathcal{A}} c \sum_{k=1}^{4} B_k. \tag{30}$$

Consequently,

$$D_{\mathrm{KL}}\big(p^{\mathcal{A}} \| \bar{p}\big) \le 2\lambda_{\mathcal{A}} c \sum_{k=1}^{4} B_k. \tag{31}$$

Proof. Since $m_{t,i}^{(k)} \in [0,1]$, Eq. equation 20 implies

$$\left| \gamma_{t,i}^{(k)} \right| \le c B_k.$$

For each supported prefix, the conditional normalizer is an average of exponentiated bounded potentials and therefore satisfies

$$e^{-\lambda_{\mathcal{A}} c B_k} \le Z_{t,i}^{(k)} \le e^{\lambda_{\mathcal{A}} c B_k}.$$

Applying these inequalities to Eq. equation 21 bounds the conditional log-ratio by $2\lambda_{\mathcal{A}} c B_k$. Summing over the four prediction heads proves Eq. equation 30; taking expectation under $p^{\mathcal{A}}$ gives Eq. equation 31. Finite exponential weights remain strictly positive, so support is preserved. □

The bound makes the conservative role of confidence explicit: for fixed evidence bounds, weaker evidence produces a smaller distributional perturbation, and zero evidence exactly recovers the baseline policy.

B.4 Preservation of Connectivity and Valence Bounds

Assume that the initial nonempty partial ligand is connected and satisfies the configured valence limits. Suppose that each accepted action adds one atom, forms at least one bond to the existing ligand, leaves all existing atoms and bonds unchanged, and satisfies the feasibility rule in Eq. (12). Then every retained partial ligand remains connected and satisfies the same valence limits.

Proof. Proceed by induction over growth steps. The claim holds for the initial state by assumption. Suppose it holds for x_t. Because an accepted extension connects the newly added atom to at least one atom of the existing connected graph, x_{t+1} remains connected. The feasibility operator checks both the bond-order sum of the new atom and the updated bond-order sums of all existing atoms incident to the new bonds. All other atoms retain their previous bond-order sums. The claim therefore holds for x_{t+1} and, by induction, for every subsequent retained state. □

For generation beginning from an empty ligand, the same statement applies after placement of the first atom. Because matched rollouts and full molecular generation use the same feasibility operator, the property holds in both stages.

C Set Delivery: Metrics and Selection Protocol

This section specifies the fixed-budget delivery procedure. The formulation separates candidate-level descriptors from set-level complementarity and expresses heterogeneous criteria through compact normalized vectors. All mappings and scalarizations are fixed globally and shared across generator configurations; they are not adapted to individual pockets or to method identity.

C.1 Set Statistics and Normalization

Candidate pools and normalization. Let $\mathcal{C}_0$ denote the loaded pool of parseable molecules and let $\mathcal{C} \subseteq \mathcal{C}_0$ denote the sanitized, canonically deduplicated selection pool with finite Vina evaluations. Candidate-level molecular descriptors are normalized over their corresponding source pool, while binding descriptors are normalized over the eligible selection pool.

For a scalar feature f and normalization pool $\mathcal{D}$, define

$$z_{\uparrow,\mathcal{D}}(f(x)) = \begin{cases} \dfrac{f(x) - f_{\min,\mathcal{D}}}{f_{\max,\mathcal{D}} - f_{\min,\mathcal{D}}}, & f_{\max,\mathcal{D}} > f_{\min,\mathcal{D}}, \\ \frac{1}{2}, & f_{\max,\mathcal{D}} = f_{\min,\mathcal{D}}, \end{cases} \qquad z_{\downarrow,\mathcal{D}}(f) = 1 - z_{\uparrow,\mathcal{D}}(f). \tag{32}$$

The resulting scales are frozen before greedy construction and remain unchanged during local refinement and contribution analysis.

Molecular quality and local geometry. For ligand atom ℓ, let d_ℓ denote its nearest distance to a pocket atom. We summarize local pocket geometry by

$$\mathbf{g}(x) = [\mathrm{contact}(x), \mathrm{sweet}(x), \mathrm{clash}(x), \delta(x)]^\top, \tag{33}$$

where the components respectively describe pocket contact, occupancy of a favorable interaction-distance region, short-range steric conflict, and deviation of the mean ligand–pocket distance from the nominal interaction region. The implementation uses a single fixed geometric profile Θ_{geom} to instantiate these descriptors for every candidate and every experiment.

Let $L(x)$ denote satisfaction of the standard Lipinski criteria. We collect the normalized molecular and geometric descriptors into

$$\mathbf{u}(x) = \left[z_\uparrow(\mathrm{QED}(x)), z_\downarrow(\mathrm{SA}(x)), L(x), z_\uparrow(\mathbf{g}(x))^\top\right]^\top.$$

The candidate-level molecular-quality summary is

$$q(x) = \boldsymbol{\omega}_q^\top \mathbf{u}(x), \qquad \boldsymbol{\omega}_q \in \Delta^{d_q - 1}, \tag{34}$$

where $\boldsymbol{\omega}_q$ is a fixed global scalarization contained in the released evaluation configuration.

We additionally record a coarse medicinal-chemistry indicator

$$h(x) = \mathbf{1}[\mathbf{c}_{\mathrm{med}}(x) \in \mathcal{H}_{\mathrm{med}}], \tag{35}$$

where $\mathbf{c}_{\mathrm{med}}(x)$ contains the relevant drug-likeness, rule-based, and short-range geometry descriptors and $\mathcal{H}_{\mathrm{med}}$ denotes the fixed operating region used throughout the experiments. This quantity is a descriptive component of the set objective rather than a chemical-validity filter: candidates outside $\mathcal{H}_{\mathrm{med}}$ remain eligible for selection.

For a nonempty set $\mathcal{S}$, write

$$\overline{f}_{\mathcal{S}} = \frac{1}{|\mathcal{S}|} \sum_{x \in \mathcal{S}} f(x).$$

Define the set-level quality vector

$$\mathbf{v}_{\mathrm{med}}(\mathcal{S}) = \left[\overline{q}_{\mathcal{S}}, \overline{h}_{\mathcal{S}}, \overline{\mathbf{g}}_{\mathcal{S}}^{\top}\right]^{\top}.$$

The corresponding quality component is

$$Q_{\mathrm{med}}(\mathcal{S}) = \boldsymbol{\omega}_{\mathrm{med}}^{\top} \mathbf{v}_{\mathrm{med}}(\mathcal{S}), \tag{36}$$

with one fixed coefficient vector $\boldsymbol{\omega}_{\mathrm{med}}$ shared across all candidate pools. All constituent descriptors are computed directly from the generated SDF structures.

Predicted binding favorability. Vina Score evaluates the generated pose, whereas Vina Dock evaluates the best energy returned by the common redocking protocol. Let z_{Score} and z_{Dock} denote their independently reverse-normalized values. We write their aggregate as

$$A_{\mathrm{bind}}(\mathcal{S}; \mathcal{P}) = \boldsymbol{\omega}_{\mathrm{bind}}^{\top} \begin{bmatrix} \overline{z}_{\mathrm{Score},\mathcal{S}} \\ \overline{z}_{\mathrm{Dock},\mathcal{S}} \end{bmatrix}, \tag{37}$$

where $\boldsymbol{\omega}_{\mathrm{bind}}$ is fixed once for the complete evaluation protocol. All methods use the same Vina implementation and evaluator configuration.

Chemical diversity and coverage. Structural complementarity is measured independently of the preceding scalar molecular summaries. Using the common Morgan fingerprint representation, for $n = |\mathcal{S}| \geq 2$,

$$D_{\mathrm{chem}}(\mathcal{S}) = \frac{2}{n(n-1)} \sum_{\{x,y\} \subseteq \mathcal{S}} \left[1 - \mathrm{Tanimoto}\big(\mathrm{fp}(x), \mathrm{fp}(y)\big)\right], \tag{38}$$

with $D_{\mathrm{chem}}(\mathcal{S}) = 0$ for $n < 2$.

Let $\mathrm{scaf}(x)$ denote the Bemis–Murcko scaffold of candidate x. Scaffold diversity is

$$C_{\mathrm{scaf}}(\mathcal{S}) = \frac{|\{\mathrm{scaf}(x) : x \in \mathcal{S},\ \mathrm{scaf}(x) \neq \varnothing\}|}{\max(1, n)}. \tag{39}$$

Acyclic molecules remain in the denominator but do not contribute a nonempty scaffold identity.

Let $T(x)$ denote the set of pocket residues contacted by x under the same geometric contact definition used throughout evaluation. We define

$$C_{\mathrm{pocket}}(\mathcal{S}; \mathcal{P}) = \frac{\left|\bigcup_{x \in \mathcal{S}} T(x)\right|}{\max\left(1, \left|\bigcup_{x \in \mathcal{C}} T(x)\right|\right)}. \tag{40}$$

This measures coverage relative to interactions attainable within the eligible candidate pool.

The complete set descriptor is therefore

$$\mathbf{m}(\mathcal{S}; \mathcal{P}) = \left[Q_{\mathrm{med}}(\mathcal{S}), A_{\mathrm{bind}}(\mathcal{S}; \mathcal{P}), D_{\mathrm{chem}}(\mathcal{S}), C_{\mathrm{scaf}}(\mathcal{S}), C_{\mathrm{pocket}}(\mathcal{S}; \mathcal{P})\right]^{\top}.$$

C.2 Scalarization and Set Search

The delivery objective uses one globally fixed nonnegative scalarization profile,

$$\boldsymbol{\lambda} = \boldsymbol{\lambda}_{\text{S-D}}, \qquad \boldsymbol{\lambda}_{\text{S-D}} \succeq 0, \tag{41}$$

applied identically to every pocket and generator configuration. Its machine-readable specification is part of the released experimental configuration.

For interpretability, the same objective can be viewed hierarchically as a non-binding quality–complementarity block combined with the two binding-related summaries:

$$F_{\boldsymbol{\lambda}}(\mathcal{S}; \mathcal{P}) = \alpha_{\mathrm{base}} F_{\mathrm{base}}(\mathcal{S}; \mathcal{P}) + \boldsymbol{\alpha}_{\mathrm{bind}}^{\top} \begin{bmatrix} \overline{z}_{\mathrm{Score},\mathcal{S}} \\ \overline{z}_{\mathrm{Dock},\mathcal{S}} \end{bmatrix}, \tag{42}$$

where

$$F_{\mathrm{base}}(\mathcal{S};\mathcal{P}) = Q_{\mathrm{med}}(\mathcal{S}) + \boldsymbol{\alpha}_{\mathrm{comp}}^{\top} \begin{bmatrix} D_{\mathrm{chem}}(\mathcal{S}) \\ C_{\mathrm{scaf}}(\mathcal{S}) \\ C_{\mathrm{pocket}}(\mathcal{S};\mathcal{P}) \end{bmatrix}.$$

Thus, the expanded coefficient vector in Eq. equation 41 is induced by a small number of semantic objective groups rather than being treated as a collection of independently optimized terms.

Starting from $\mathcal{S} = \varnothing$, define the marginal gain

$$\Delta_F(x \mid \mathcal{S}) = F_{\boldsymbol{\lambda}}(\mathcal{S} \cup \{x\};\mathcal{P}) - F_{\boldsymbol{\lambda}}(\mathcal{S};\mathcal{P}). \tag{43}$$

Greedy construction selects

$$x_k \in \arg\max_{x \in \mathcal{C} \setminus \mathcal{S}_k} \left[F_{\boldsymbol{\lambda}}(\mathcal{S}_k \cup \{x\};\mathcal{P}) - F_{\boldsymbol{\lambda}}(\mathcal{S}_k;\mathcal{P})\right], \tag{44}$$

and sets

$$\mathcal{S}_{k+1} = \mathcal{S}_k \cup \{x_k\}$$

until

$$|\mathcal{S}_k| = K' = \min(K, |\mathcal{C}|).$$

The greedy solution is then refined through a bounded number of cardinality-preserving one-for-one exchange passes. For $x \in \mathcal{S}$ and $y \in \mathcal{C} \setminus \mathcal{S}$, the exchange gain is

$$F_{\boldsymbol{\lambda}}\big((\mathcal{S} \setminus \{x\}) \cup \{y\};\mathcal{P}\big) - F_{\boldsymbol{\lambda}}(\mathcal{S};\mathcal{P}). \tag{45}$$

Only improving exchanges are applied, and refinement terminates when no admissible exchange changes the objective beyond the fixed numerical tolerance of the implementation. The final set is denoted by $\widehat{\mathcal{S}}$.

Scale invariance. A common positive scaling of all components of $\boldsymbol{\lambda}$ does not alter either greedy additions or exchange decisions. Indeed, for $c > 0$,

$$F_{c\boldsymbol{\lambda}}(\mathcal{S};\mathcal{P}) = cF_{\boldsymbol{\lambda}}(\mathcal{S};\mathcal{P}),$$

and hence all marginal-gain orderings are preserved. Only relative trade-offs among criterion families can therefore affect the selected set.

The delivery cardinality K is treated as the external experimental budget rather than an optimized model parameter.

C.3 Removal-Based Contribution Analysis

For each $x \in \widehat{\mathcal{S}}$, we record its contribution to the final delivered set as

$$\mathbf{d}_x = \mathbf{m}(\widehat{\mathcal{S}};\mathcal{P}) - \mathbf{m}(\widehat{\mathcal{S}} \setminus \{x\};\mathcal{P}). \tag{46}$$

Both terms use the normalization scales and pocket-coverage denominator defined by the original eligible pool $\mathcal{C}$. This provides a molecule-level account of how each retained candidate affects quality, binding compatibility, diversity, and pocket coverage.

D Implementation Details

This appendix summarizes the implementation needed to reproduce the reported experiments. To distinguish methodological structure from engineering choices, all fixed numerical settings are treated as a single experimental configuration,

$$\Theta_{\mathrm{impl}} = \{\Theta_{\mathrm{model}}, \Theta_{\mathrm{train}}, \Theta_{\mathrm{search}}, \Theta_{\mathrm{int}}, \Theta_{\mathrm{rollout}}, \Theta_{\mathrm{delivery}}\},$$

which is frozen before evaluation and shared across all reported test pockets and matched method comparisons.

Algorithm 1 Supervised training of the reference-conditioned generator

Require: Training complexes $\mathcal{D}$, associated references, parameters (θ, ψ), fixed training profile Θ_{train}

for each training minibatch $\mathcal{B}$ **do**
  for each example $i \in \mathcal{B}$ **do**
    Sample a masking pattern and construct partial-ligand state s_i
    Prepare atom features, spatial graphs, queries $\mathcal{Q}_i$, and targets $\mathcal{Y}_i$
    Encode reference-conditioned context $H_i = E_\psi(s_i, r_i)$
    Predict $\widehat{\mathcal{Y}}_i = D_\theta(s_i, H_i, \mathcal{Q}_i)$
    Compute ℓ_i using Eq. equation 47
  end for
  Compute $\widehat{\mathcal{L}}_{\text{gen}} = |\mathcal{B}|^{-1} \sum_{i \in \mathcal{B}} \ell_i$
  Update (θ, ψ)
end for
return (θ, ψ)

D.1 Input Representation and Architecture

Each example contains a protein pocket, a partial ligand, and retrieved references represented in the common pocket coordinate system. Protein-atom features encode element identity, residue identity, and backbone membership. Ligand-atom features encode element identity, local graph structure, valence-related quantities, and bond-type counts. A molecular indicator distinguishes protein and ligand atoms.

Separate equivariant geometric encoders process the current complex and reference complex. Their outputs are integrated through cross-context interaction to produce a reference-conditioned representation for molecular growth. The decoder then predicts the four components of an autoregressive growth action: focal site, relative position, element type, and bond pattern. Relative atom placement is represented by a finite mixture density, whereas element and bond prediction use the local geometric and reference-conditioned context.

All architectural dimensions, geometric-neighborhood parameters, and mixture settings are contained in Θ_{model} and remain unchanged across ForesightMol and the matched Rag2Mol-G configuration.

D.2 Supervised Training

Training examples are obtained by partially masking complete ligands and constructing the corresponding intermediate molecular-growth states. The supervised objective is

$$\mathcal{L}_{\text{gen}} = \mathcal{L}_{\text{focal}} + \mathcal{L}_{\text{pos}} + \mathcal{L}_{\text{element}} + \mathcal{L}_{\text{bond}} + \mathcal{L}_{\text{aux}}. \tag{47}$$

The focal-site, element, and bond terms are classification objectives. The position term is the negative log-likelihood under the predicted mixture distribution. Auxiliary supervision covers protein growth sites and real-versus-perturbed spatial queries.

Optimization, masking, perturbation, batching, and regularization settings are specified by the common training profile Θ_{train}. The resulting generator checkpoint is shared by ForesightMol and the matched Rag2Mol-G baseline; differences in the reported generation experiments therefore arise from the inference-time mechanisms studied in this work rather than independently trained backbones.

Algorithm 1 summarizes the common training procedure, where E_ψ denotes reference encoding and context fusion and D_θ denotes the prediction heads.

D.3 Generation and Reference Intervention

Generation uses a fixed search profile Θ_{search} that determines beam search, completion budget, maximum trajectory length, and the generation states at which reference assessment may be queried. The same profile is used for every evaluated pocket.

At a queried state s_t, the agent produces a small challenger subset

$$\mathcal{R}_t^{\mathcal{A}} \subseteq \mathcal{R}(\mathcal{P}) \setminus \{r_t^0\},$$

where r_t^0 is the generator's default reference. Each challenger is compared with the default reference using a matched short-horizon continuation. The alternatives share the initial partial-ligand state, inherited decoding context, stochastic input, feasibility rules, search policy, and stopping conditions. The default continuation is reused across challengers whenever possible, so the comparison isolates the effect of reference choice.

For challenger r, define

$$\widehat{\Delta}_H(s_t, r) = \frac{1}{M} \sum_{m=1}^{M} \Big[U_H(\tau_{t,H}(r; \xi_m)) - U_H(\tau_{t,H}(r_t^0; \xi_m)) \Big], \tag{48}$$

$$\Delta R_H(s_t, r) = \widehat{R}_H(s_t, r) - \widehat{R}_H(s_t, r_t^0). \tag{49}$$

Benefit, trajectory risk, and proposal reliability are combined only through the fixed verification operator

$$\mathcal{A}_{\text{int}} : (\widehat{\Delta}_H, \Delta R_H, c_t) \mapsto \{0, 1\},$$

whose configuration is denoted by Θ_{int}. Equivalently, the verified set is

$$\mathcal{V}_t = \Big\{ r \in \mathcal{R}_t^{\mathcal{A}} : \mathcal{A}_{\text{int}}\big(\widehat{\Delta}_H(s_t, r), \Delta R_H(s_t, r), c_t(r)\big) = 1 \Big\}.$$

If $\mathcal{V}_t \neq \varnothing$, the eligible reference with the largest paired rollout advantage is used. If no challenger is verified, the queried branch retains the default reference and newly proposed guidance is cleared. Unqueried branches retain their existing reference and guidance.

The verification operator is fixed globally and is never adapted to an individual pocket, reference, or generator. Agent confidence is not interpreted as a calibrated probability and cannot by itself activate an intervention: deployment additionally requires external rollout evidence. Valence and connectivity filtering remains active in both reference assessment and full molecular generation.

D.4 Rollout Utility

The short-horizon evaluator summarizes a continuation using a fixed feature map

$$\boldsymbol{\phi}_U(\tau) = \big[\boldsymbol{\phi}_{\text{completion}}(\tau), \boldsymbol{\phi}_{\text{progress}}(\tau), \boldsymbol{\phi}_{\text{support}}(\tau), \boldsymbol{\phi}_{\text{failure}}(\tau)\big],$$

covering successful completion, productive continuation, model support, and adverse or redundant branches. The rollout ranking functional is written compactly as

$$U_H(\tau) = \boldsymbol{\omega}_U^\top \boldsymbol{\phi}_U(\tau), \tag{50}$$

where $\boldsymbol{\omega}_U$ is the globally fixed rollout profile contained in Θ_{rollout}.

Importantly, the algorithm does not interpret U_H as an absolute estimate of molecular value. It is used only through matched differences between continuations from the same state. For challenger r and default reference r_t^0,

$$\Delta U_H = U_H(\tau_r) - U_H(\tau_0).$$

Consequently, every common additive offset cancels. Moreover, for any $a > 0$ and $b \in \mathbb{R}$,

$$U'_H(\tau) = aU_H(\tau) + b \quad \Longrightarrow \quad \Delta U'_H = a\Delta U_H.$$

Thus the absolute scale and origin of the ranking functional are immaterial; only its induced relative ordering under the fixed verification profile can affect reference decisions.

Trajectory risk is recorded separately:

$$R_H(\tau) = \rho_{\text{risk}}(\boldsymbol{\phi}_R(\tau)), \tag{51}$$

where $\boldsymbol{\phi}_R$ collects failure and redundancy events and ρ_{risk} is the normalized risk summary specified in Θ_{rollout}. Keeping this statistic separate prevents a utility improvement from being interpreted as sufficient evidence when it is accompanied by an unacceptable deterioration in continuation reliability.

The complete rollout definition is identical across all pockets, candidate references, query states, and compared generators.

D.5 Agent Configuration

We use a fixed scientific-agent configuration throughout all experiments, with Qwen instantiated as `qwen3.7-max-2026-06-08` via API. We set the temperature to 0 and disable thinking mode, with a 120-second request timeout and up to three request attempts. The agent receives anonymous molecular identifiers together with structured summaries of the current partial ligand, target pocket, and candidate references, and returns structured reference proposals and graded support assessments. Proposals are constructed before rollout evaluation, without access to the resulting rollout outcomes.

Each structured response records reference preference and evidence relevant to the four autoregressive prediction heads, together with assessments of continuation value, molecular quality, synthetic tractability, trajectory risk, and confidence. The returned scores are bounded assessment signals rather than calibrated probabilities. Malformed, missing, or otherwise unusable outputs revert to the default-reference path.

The scientific agent therefore serves as a proposal mechanism rather than a deployment authority:

$$\text{proposal} \longrightarrow \text{matched rollout verification} \longrightarrow \text{programmatic decision}.$$

Head-wise use of accepted evidence is described in Appendix B.

D.6 Agent Prompt

The prompt is organized around state-grounded reference assessment, structured chemical evidence, and separation between proposal and deployment.

> **Role and task.** Assess candidate references for pocket-conditioned molecular growth from the same partial-ligand state. Use only the supplied partial-ligand, pocket, and reference summaries to evaluate their potential usefulness for subsequent growth.
>
> **Reference assessment.** Compare candidate references with respect to their support for focal-site selection, spatial placement, element prediction, and bond prediction. Consider whether unmatched reference regions provide plausible continuation evidence and assess drug-likeness, synthetic tractability, steric conflict, invalid continuation, dead ends, and redundancy. Ground assessments in the supplied information and represent uncertainty explicitly when evidence is incomplete.
>
> **Structured output.** Return the supplied reference identifiers together with structured assessment fields for reference preference, four-head support, continuation value, molecular-quality support, synthetic-tractability support, trajectory risk, confidence, and concise reason codes. All scalar assessment fields must lie in the prescribed bounded range.
>
> **Deployment separation.** Do not determine whether a proposed reference is deployed. The external program evaluates every proposal using matched rollouts and applies the fixed verification operator.

E Experimental Details

E.1 Dataset and Comparisons

We use the processed CrossDocked2020 dataset (Francoeur et al., 2020) distributed with Rag2Mol (Zhang et al., 2025), following its preprocessing, retrieval organization, and fixed split of 100,000 training examples and 100 test pockets. All experiments, including component ablations and structural analyses, use the 100-pocket test set. Each example contains up to four associated references.

Baselines include AR (Luo et al., 2021), GraphBP (Liu et al., 2022), Pocket2Mol (Peng et al., 2022), TargetDiff (Guan et al., 2023a), DiffSBDD (Schneuing et al., 2024), and Rag2Mol-G (Zhang et al., 2025). Available outputs from external baselines are evaluated using common metric definitions.

The crossed comparison applies Dock ranking and Score–Dock-aware selection to each generator's fixed eligible pool. The component study compares the baseline, reference intervention alone, chemical constraints alone, and their combination under the same selection protocol and delivery budget.

E.2 Candidate Processing and Delivery

Generated molecules are sanitized and canonically deduplicated using RDKit. Eligible delivery candidates have finite Vina Score and Vina Dock results. Both selectors operate on the same eligible pool and return $\min(10, N)$ molecules, where N is the eligible pool size.

Dock ranking selects molecules by increasing Vina Dock energy. Score–Dock-aware selection balances molecular quality, predicted binding favorability, and set complementarity. Metric normalization, coefficients, and search settings are provided in Appendix C.

E.3 Molecular Properties and Docking

We evaluate QED (Bickerton et al., 2012), raw SA (Ertl & Schuffenhauer, 2009), and Lipinski compliance (Lipinski et al., 2001). Mean and best QED denote the within-pool or within-set mean and maximum. Lipinski compliance requires molecular weight ≤ 500, logP ≤ 5, hydrogen-bond donors ≤ 5, and acceptors ≤ 10.

Molecular diversity is the mean pairwise Tanimoto distance between Morgan fingerprints (Rogers & Hahn, 2010), using radius two and 2048 bits. Scaffold diversity is the number of distinct nonempty Bemis–Murcko scaffolds (Bemis & Murcko, 1996) divided by the number of molecules, including acyclic molecules. Molecular diversity is zero for sets containing fewer than two molecules.

Vina Score evaluates generated poses, and Vina Dock evaluates redocked poses. We use AutoDock Vina 1.2.7 (Eberhardt et al., 2021), with exhaustiveness eight, nine output modes, and seed 2026. The docking box is a 20 Å cube centered on the native ligand's heavy-atom centroid. Lower SA and Vina values are preferred.

E.4 Structural Evaluation

Structural distributions are compared with native test ligands. Bond-length JSD uses delivered molecules, with bonds pooled by element pair and bond type. Histograms use 0.005 Å bins over 1.1–1.7 Å and additional bins for out-of-range values.

Heavy-atom pair-distance JSD uses full generated pools and all unordered heavy-atom pairs within each molecule. Histograms use 0.25 Å bins up to 30 Å and an overflow bin. Molecular histograms are normalized individually, averaged within pockets, and then averaged equally across pockets. Both JSD measures use base-two logarithms and are reported in bits.

Ring-size prevalence is the fraction of molecules containing at least one ring of a given size, averaged equally across pockets. It is computed on full generated pools, including acyclic molecules in the denominator. A molecule may contribute to several ring sizes.

Fragment geometry is evaluated on fragments of three to nine heavy atoms obtained by cutting non-ring single bonds that separate multi-atom components. Fragments are relaxed in isolation using MMFF94s with torsion and out-of-plane bending terms disabled. Symmetry-aware aligned RMSD measures the change from the original geometry. Distributions and medians are computed over fragments pooled by method and fragment size.

F Additional Analyses and Examples

F.1 Reference Intervention Strategies

We compare three reference strategies under the same initial reference policy and valence filtering. The fixed-reference configuration retains the initial reference, while random replacement selects a non-default reference under the same query conditions, without agent assessment, rollout verification, or agent-derived guidance. The complete intervention combines agent proposals, rollout verification, and guidance after acceptance.

Table S1 shows that the complete intervention achieves the highest mean and median QED, Lipinski compliance, molecular and scaffold diversity, and contact ratio, together with the lowest SA. Random replacement reduces mean QED and both diversity measures relative to retaining the ini-

Method	Mean QED ↑		Best QED ↑		SA ↓		Lipinski (%) ↑		Mol. div. ↑		Scaf. div. ↑		Contact ratio (%) ↑	
	Avg.	Med.	Avg.	Med.	Avg.	Med.	Avg.	Med.	Avg.	Med.	Avg.	Med.	Avg.	Med.
Default reference	0.5158	0.5655	0.7419	0.7759	3.7965	3.6734	80.15	94.27	0.6934	0.7023	0.6861	0.6843	91.24	96.97
Random replacement	0.5114	0.5249	0.7617	0.7693	3.8378	4.0789	77.92	94.97	0.6400	0.6581	0.6224	0.6313	92.62	96.28
Agent + rollout	**0.5770**	**0.5769**	**0.7962**	**0.7920**	**3.4610**	**3.4920**	**84.69**	**98.53**	**0.7551**	**0.7658**	**0.6995**	**0.7188**	**94.28**	**98.50**

Table S1: The effect of reference intervention strategies on generated molecular properties. Avg. and Med. denote the mean and median across pockets. Agent + rollout includes reference proposals, rollout verification, and guidance after acceptance. Bold indicates the best result in each column.

Method	Pocket clash-free ↑		Clash P90 ↓		Internal clash-free ↑		Ring planarity ↑	
	Avg.	Med.	Avg.	Med.	Avg.	Med.	Avg.	Med.
Reference	100	100	0	0	93.33	100	100	100
AR	**99.52**	**100**	**0.07**	**0**	31.15	29.53	41.55	37.35
Pocket2Mol	97.42	**100**	0.52	**0**	61.12	65.45	94.55	95.67
TargetDiff	95.54	**100**	0.95	**0**	65.28	66.04	93.59	95.90
DiffSBDD	89.90	95.50	2.71	**0**	43.89	43.50	95.38	96.54
Rag2Mol-G	98.43	**100**	0.22	**0**	59.39	63.01	93.68	96.59
ForesightMol	98.93	**100**	0.09	**0**	**79.44**	**81.36**	**99.24**	**100**

Table S2: Geometric plausibility of full generated pools. All values are percentages. Avg. and Med. denote the mean and median across eligible pockets. Clash P90 is calculated within each pocket before aggregation, and ring planarity is evaluated over aromatic rings. Reference denotes native ligands. Bold indicates the best result among generators in each column, including ties.

tial reference, whereas the complete intervention improves these properties. These results support evaluating reference utility and incorporating accepted evidence into subsequent growth.

Contact ratio measures the fraction of all input atoms in a ligand whose nearest pocket atom is within 4.5 Å, counting each ligand atom once. We average this ratio across generated molecules within each pocket, then report the mean and median across pockets.

F.2 Geometric Plausibility of Generated Pools

We assess geometric plausibility in the full generated pools through pocket clashes, intramolecular clashes, and aromatic-ring planarity. These checks complement distributional comparisons by measuring specific geometric defects.

Pocket clash-free rate measures the percentage of molecules with no ligand–protein heavy-atom distance below 1.6 Å. Clash P90 denotes the within-pocket 90th percentile of the fraction of ligand heavy atoms involved in such clashes. An intramolecular clash occurs when a heavy-atom pair is closer than 0.75 times the sum of its van der Waals radii, excluding pairs separated by at most two bonds. Internal clash-free rate measures the percentage of molecules without these clashes. Ring planarity measures the percentage of aromatic rings whose atoms have an RMS distance of at most 0.1 Å from the best-fit plane. Pockets without aromatic rings are excluded from planarity aggregation. We report the mean and median of each pocket-level metric.

Table S2 shows that ForesightMol achieves the highest internal clash-free rate and ring planarity among the evaluated generators. It also maintains a mean pocket clash-free rate of 98.93% and a mean Clash P90 of 0.09%. These results indicate improved internal geometry alongside limited severe overlap with the protein environment.

F.3 Selector Sensitivity

The set-level objective provides control over the balance between docking performance, molecular quality, and structural diversity. Figure S2 compares selection configurations applied to fixed candi-

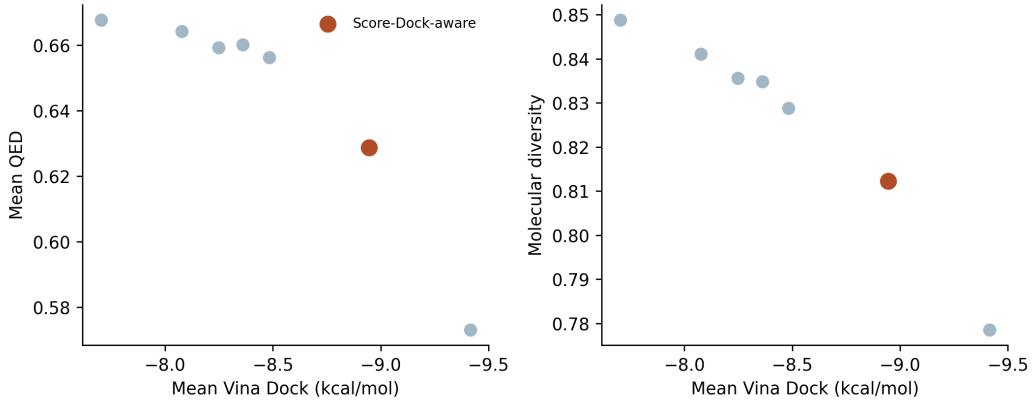

Figure S2: Effect of selector weights on delivered molecular quality and diversity. Each point represents a configuration evaluated on fixed candidate pools, with the fixed Score–Dock-aware protocol highlighted in orange. All configurations in the sweep are shown. More favorable Vina Dock energies lie toward the right, while higher QED and diversity lie toward the top.

date pools. Increasing the docking contribution shifts the delivered sets toward more favorable Vina Dock energies, while quality-oriented configurations retain higher QED and molecular diversity.

Across the two endpoints, mean Vina Dock changes from -7.703 to -9.414 kcal/mol, mean QED from 0.668 to 0.573, and diversity from 0.849 to 0.779. The fixed Score–Dock-aware protocol occupies an intermediate operating point, with Vina Dock of -8.946 kcal/mol, QED of 0.629, and diversity of 0.812 in this analysis. This balance reflects the joint consideration of individual molecular properties and set complementarity.

F.4 Effect of Feasibility Filtering

Feasibility filtering improves the reliability of molecular growth by screening extensions before they enter the beam. Under a fixed reference policy without agent intervention, the paired diagnostic reduces the mean recorded post-beam failure count from 8.7 to 0.0 per pocket. This result supports enforcing valence and connectivity constraints at the action level, preventing invalid extensions from propagating through subsequent growth steps.

F.5 Molecular Size and Synthetic Accessibility

Figure S3 relates molecular size to synthetic accessibility. ForesightMol follows the overall size–SA relationship observed in native ligands and shows a more pronounced concentration in the smaller-molecule region. Together with the aggregate SA improvements, this pattern is consistent with the generation of compact candidates with lower predicted synthetic complexity. The joint view connects the reported property gains with the molecular structures explored by the framework.

F.6 Additional Delivered-Molecule Examples

Figures S4 and S5 illustrate the chemical alternatives obtained by the two frameworks. Pocket poses are paired with two-dimensional depictions of the same molecules, connecting their spatial placement with ring systems, substituents, and reported properties.

In Figure S4, the displayed ForesightMol candidates jointly improve Vina Score, QED, and SA over their paired baseline candidates. These examples show how favorable property combinations are realized through distinct chemical structures within the target pockets. Figure S5 presents candidate pairs favoring the baseline, providing a complementary view of the molecular solutions explored by both methods.

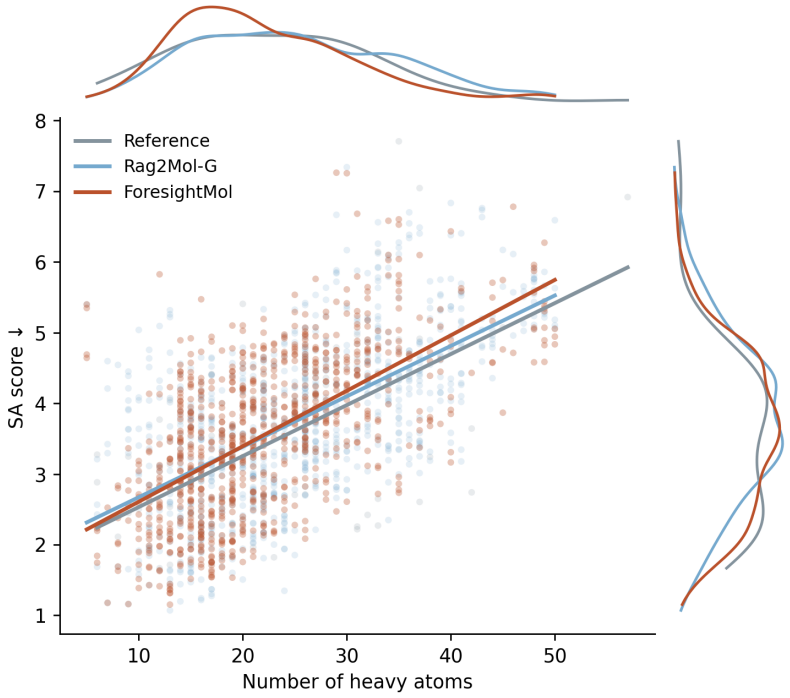

Figure S3: Joint distributions of heavy-atom count and raw SA score for native ligands and generated molecules. Points represent individual molecules, solid lines show fitted linear trends, and marginal curves summarize each property distribution. Lower SA scores indicate greater predicted synthetic accessibility.

F.7 Agent-Assisted Interpretation of Delivered Sets

We also use Qwen to improve the interpretability of delivered sets by translating molecular measurements and contribution records into candidate-specific explanations. These explanations connect individual properties with structural complementarity and suggest how candidates can be considered together for experimental follow-up.

Figure S5 illustrates this process for the ten-molecule set delivered for pocket 3gs6/A. M01 achieves the highest QED (0.875), while M04 combines a QED of 0.818 with the highest mean structural dissimilarity to the other selected molecules (0.842). Removing M04 reduces set diversity by 0.0123. M08 combines a QED of 0.810 with a mean dissimilarity of 0.803, representing another balance between drug-likeness and structural complementarity.

Using these measurements, Qwen recommends considering M04 alongside M01 to explore a distinct chemical alternative while retaining favorable drug-likeness. An excerpt from its generated rationale states

> "If the experimental budget permits only a small number of follow-up candidates, M04 should be considered as a diversity-oriented complement to M01."

This recommendation connects the set-level diversity contribution to a concrete follow-up strategy. Evaluating M01 and M04 together would probe whether favorable experimental outcomes extend across distinct structures within the delivered set.

Pocket measurements further inform these explanations. M01 and M08 have minimum pocket distances of 2.38 Å and 2.50 Å, respectively, with no detected heavy-atom clashes below 1.6 Å. Together, the molecular properties, pocket measurements, and contribution records provide an explicit basis for reviewing Qwen's recommendations and prioritizing complementary candidates.

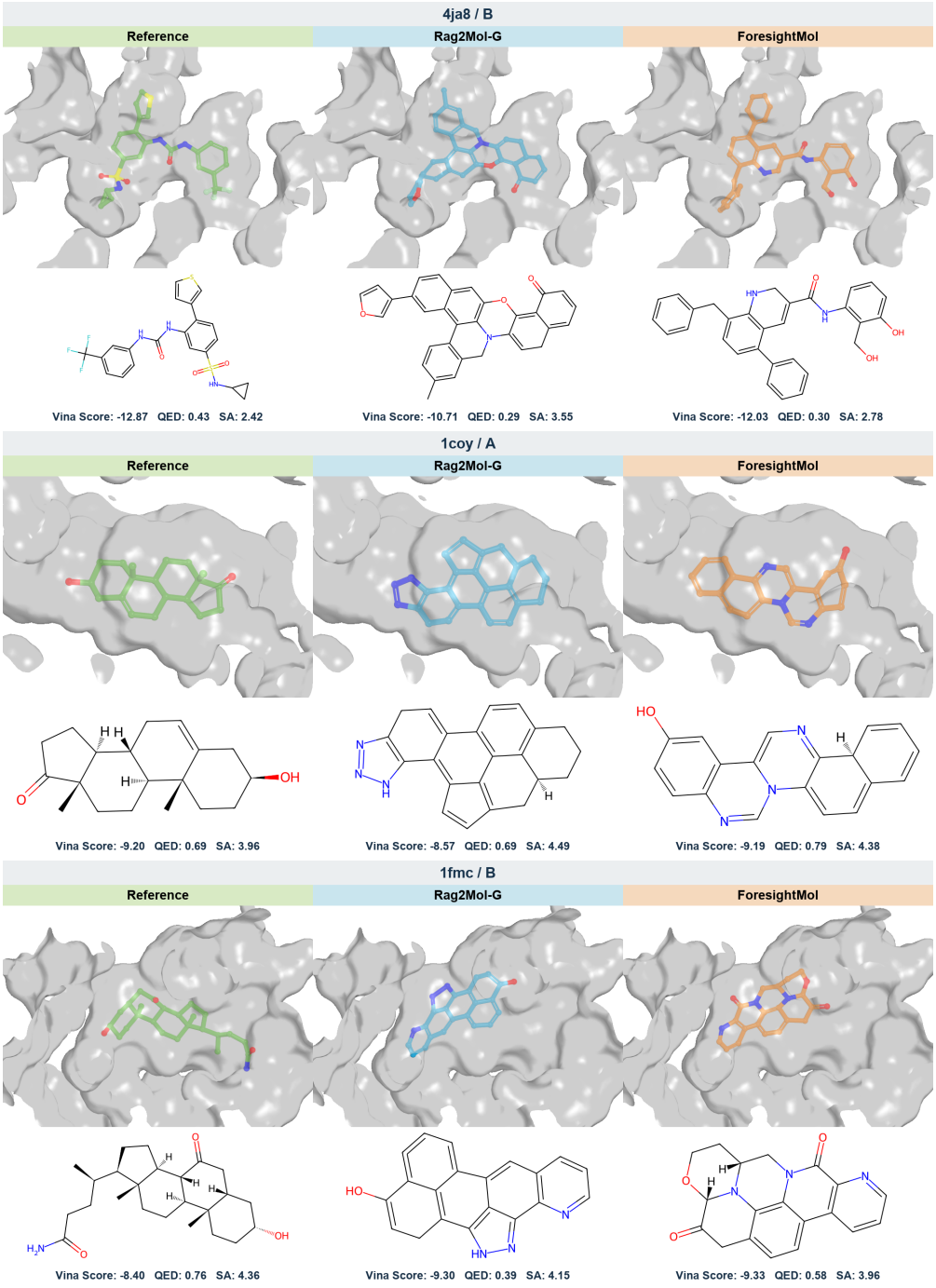

Figure S4: Additional delivered-molecule examples for pockets 4ja8/B, 1coy/A, and 1fmc/B. The displayed ForesightMol candidates achieve more favorable Vina Score, higher QED, and lower raw SA than their paired Rag2Mol-G candidates. Carbon atoms in native ligands, Rag2Mol-G candidates, and ForesightMol candidates are shown in green, cyan, and orange, respectively. Protein pockets are shown as gray surfaces. Each pose is paired with its corresponding two-dimensional structure and molecular metrics. Vina Score is reported in kcal/mol.

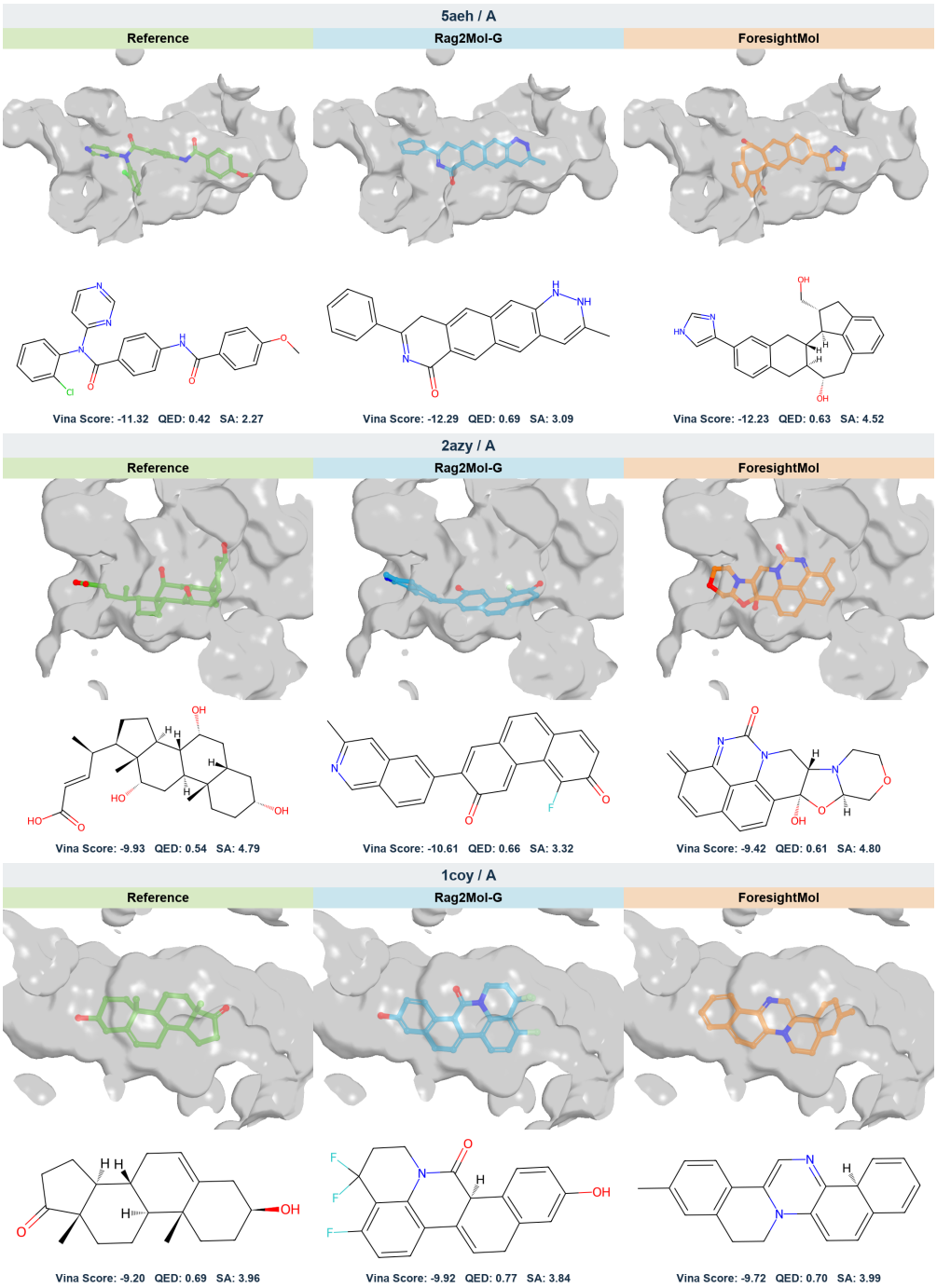

Figure S5: Additional delivered-molecule examples for pockets 5aeh/A, 2azy/A, and 1coy/A. The displayed Rag2Mol-G candidates achieve more favorable Vina Score, higher QED, and lower raw SA than their paired ForesightMol candidates. Colors and metric conventions follow Figure S4.

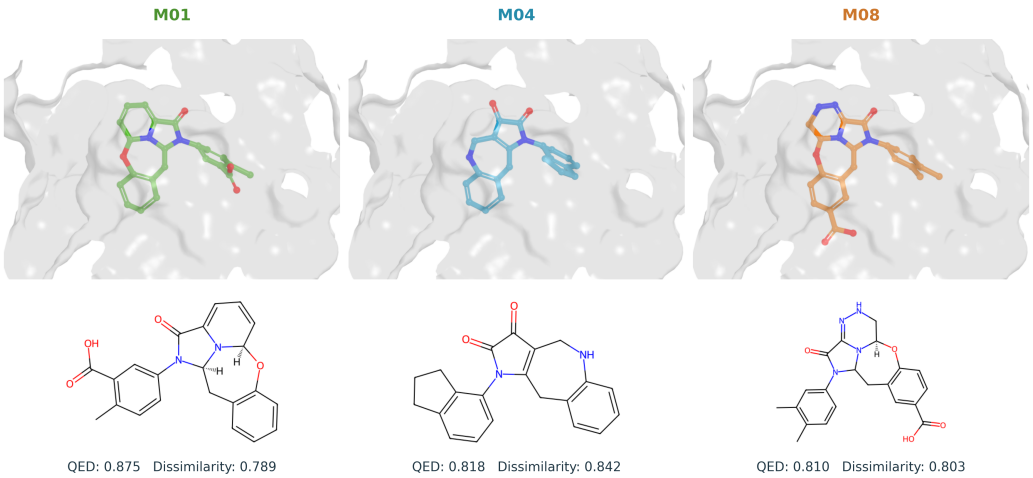

Figure S6: Qwen-assisted interpretation of candidates delivered for pocket 3gs6/A. M01 highlights molecular quality, M04 structural complementarity, and M08 a balance of the two. Dissimilarity is the mean Morgan-fingerprint Tanimoto distance to the other selected molecules. Pocket poses and corresponding two-dimensional structures accompany the measured properties used to support the candidate rationales.

G Future Work

Future work will explore adaptive reference assessment, allocating agent queries and rollout depth according to uncertainty and growth difficulty to better use the available computational budget. Integrating reaction knowledge, synthesis planning, and additional property predictors could further align reference guidance and candidate selection with practical drug discovery requirements. We also plan to investigate transfer across generation backbones and target families, and to evaluate delivered candidates through prospective synthesis and binding assays. Such experimental feedback could refine reference utility estimation and the balance between molecular quality and structural complementarity. Together, these directions aim to support more precise and effective drug discovery.