MSF-HDIFF: MOLECULAR SEMANTIC FIELD-GUIDED MULTI-AGENT HIERARCHICAL DIFFUSION FOR SMALL-MOLECULE ANTIBIOTIC DISCOVERY

Anonymous authors
Paper under double-blind review

ABSTRACT

The continued rise in antimicrobial resistance creates an urgent need to accelerate antibiotic discovery. Artificial intelligence (AI) has shown great promise in accelerating small-molecule antibiotic discovery, specifically generative models that can design de novo small molecules. However, existing generative models often struggle to simultaneously satisfy molecular requirements while making targeted chemical modifications to optimize desired properties. To address this, we introduce MSF-HDiff, a closed-loop framework that combines multi-agent reasoning with hierarchically conditioned diffusion for de novo small-molecule antibiotic discovery. Our model features a novel framework that integrates hierarchical condition search, transferable molecular semantic fields for property-directed chemical guidance, and reference-based feasibility control during diffusion sampling. This enables iterative molecular design that balances global antibiotic objectives with local chemical optimization. MSF-HDiff outperforms other generative models in generating de novo small molecules with strong antibacterial activity, favorable molecular properties, and high structural reliability as shown in Figure 1.

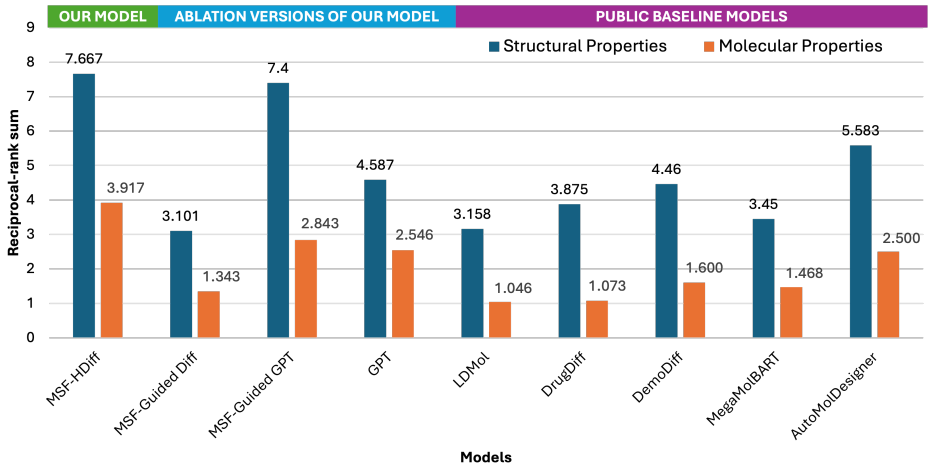

Figure 1: Reciprocal rank sum of average performance for MSF-HDiff, ablation models, and baseline models across structural properties (blue) and molecular properties (orange) of compounds generated for *Escherichia coli*.

1 INTRODUCTION

Antimicrobial resistance (AMR) is a growing global health threat that occurs when microorganisms become resistant to the medicines used to treat them, making infections more difficult to treat, longer-lasting, and potentially more serious. Bacterial AMR was estimated to be associated with

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.

more than 4.7 million deaths globally in 2021 (Collaborators, 2024), yet relatively few new antibiotics are being developed. This creates an urgent need for approaches that can accelerate the discovery of new antibiotic compounds. Small-molecule drugs offer a promising avenue because they can effectively target bacteria while offering relatively low production costs, chemical stability, and the potential for oral administration (Bergkessel et al., 2023).

Artificial intelligence (AI) offers an opportunity to accelerate antibiotic discovery by efficiently exploring chemical space and optimizing molecules for desired properties. Generative AI is particularly promising for generating novel antibiotic candidates from scratch (Stokes et al., 2020; Krishnan et al., 2025). Recent advances in molecular diffusion models enable molecular generation and modification. In contrast, tool-augmented language models and multi-agent systems connect chemical reasoning with computational analysis and iterative optimization (Oestreich et al., 2025; M. Bran et al., 2024; Kim et al., 2025). Together, these developments enable an integrated design process where scientific evidence informs generation and computational feedback guides subsequent designs.

However, applying these capabilities to antibacterial design presents several challenges. First, an antibacterial design objective, such as inhibiting a particular bacterial strain, does not directly specify the molecular structures needed to achieve that objective. Translating biological evidence into a design strategy therefore requires coordinating global compositional constraints, exploration across structural families, and refinement of local chemical patterns. Second, property targets indicate what should improve but do not specify which combinations of chemical features will produce that improvement. Although chemical language and diffusion models can learn property-associated molecular modifications (Guo et al., 2026; Xiong et al., 2026), applying these modifications to new molecular contexts requires understanding how local features interact within a complete molecule. Finally, local property optimization can compete with broader compositional and structural requirements. Value-based diffusion guidance can connect intermediate molecular states to desired outcomes (Li et al., 2025), but effectively coordinating these objectives requires controlling how local guidance influences the probability of satisfying higher-level requirements simultaneously.

We introduce MSF-HDiff, a molecular semantic field-guided multi-agent hierarchical diffusion framework for small-molecule antibiotic discovery. Our approach bridges desired molecular properties and chemical generation by representing molecules as meaningful local structural units. These units are organized into ordered molecular representations, allowing the model to capture how neighboring structural features work together within their broader molecular context. We then introduce molecular semantic fields to learn how changes in combinations of these features relate to changes in whole-molecule properties. Based on 228,539 molecular-pair records spanning ten physicochemical properties, a Semantic Field Agent learns these relationships and uses them to identify structural changes likely to achieve desired property profiles. The selected structural features are then provided as conditions to a generative model, enabling the design of complete de novo molecules with targeted chemical properties. Our contributions are threefold:

1. **Closed-loop multi-agent antibacterial design.** We formulate antibacterial molecule design as a closed-loop process in which specialized agents integrate biological evidence and computational feedback to iteratively guide molecular generation toward favorable antibacterial activity and molecular properties.
2. **Molecular semantic fields for property-directed design.** We introduce molecular semantic fields, which capture how local chemical features influence whole-molecule properties, and develop a database of 228,539 molecular pairs to learn these relationships for property-directed generation.
3. **Hierarchical diffusion with feasibility control.** We develop a hierarchical diffusion framework that integrates property guidance with structural and compositional constraints, balancing desired property changes against molecular feasibility using semantic-window-OFF generation under the same global and structural conditions as a reference.

2 Related Work

Antibiotic Discovery and Agentic Molecular Design. Generative AI has been applied to de novo antibiotic discovery through property-guided generation, chemical optimization, and agentic work-

flows. SyntheMol searches chemical building blocks and reactions for synthesizable antibacterial candidates, while related methods incorporate predicted antibacterial activity and physicochemical properties during optimization (Swanson et al., 2024a; 2026). LLM-based agents further combine language-model reasoning with domain tools and iterative feedback for chemistry analysis, synthesis planning, molecular generation, and verification (M. Bran et al., 2024; Liu et al., 2024; Kim et al., 2025). However, these approaches provide limited mechanisms for converting heterogeneous evidence and feedback into structured optimization signals that directly guide molecular generation.

Chemical Vocabulary and Structured Semantic Guidance. Structured representations have also been explored to improve interpretable and controllable molecular generation. DrugLLM introduces Functional Group Tokenization to represent molecules as chemically meaningful sequences with explicit attachment information, while text-conditioned diffusion enables multi-property molecular editing through textual conditioning (Guo et al., 2026; Xiong et al., 2026). Beyond chemistry, TOUCH demonstrates how intermediate structured representations can connect high-level intent with hierarchical diffusion (Han et al., 2025). Nevertheless, existing methods offer limited mechanisms for linking localized chemical changes to whole-molecule property responses and using these relationships as structured conditions for complete de novo generation.

Conditional Diffusion and Molecular Optimization. Conditional diffusion methods steer molecular generation through structural or property conditions, optimization, or reward-guided sampling. DiGress and GenMol perform discrete graph or sequence diffusion for molecular generation and optimization, while MolEditRL combines graph diffusion with reinforcement learning to align property changes and structural preservation (Vignac et al., 2023; Lee et al., 2025; Zhuang et al., 2026). Soft Value-Based Decoding instead uses continuation values to guide pretrained diffusion models toward downstream rewards without generator fine-tuning (Li et al., 2025). Despite these advances, existing approaches largely rely on predefined conditions or objectives and provide limited mechanisms for coordinating local semantic preferences with higher-level requirements during complete de novo generation while explicitly managing molecular feasibility trade-offs.

3 Proposed Approach

Given a target bacterium, design requirements, and user-provided examples, our framework searches for molecules with favorable predicted antibacterial activity and properties. Global conditions (C) constrain the chemical space, structural conditions (F_s) specify scaffold and topology at structural stage s, and semantic windows ($W_{T,s,r}$) guide local chemical configurations at semantic loop r. The search proceeds hierarchically across runs, stages, and inner loops, with conditions fixed at their respective levels to accumulate feedback within stable structural contexts before exploring new regions of chemical space. The overall framework is shown in Figure 2, with the complete execution procedure provided in Algorithm 1.

Notation summary and pseudocode algorithms are provided in Appendix A and B, respectively. Further details on multi-agent coordination and evaluation, molecular semantic field construction and agent learning, and HDiff training and reference-guided sampling are provided in Appendix C, D, and E, respectively.

3.1 Multi-Agent Conditioning Module

Librarian Agent. The Librarian Agent establishes the initial task context by retrieving literature evidence (G_1), active molecules from ChEMBL (G_2), and user-provided examples (G_3). These sources define $C = (C_1, C_2)$, comprising the allowed element set and molecular-weight interval. The target, design requirements, representation constraints, and admission criteria are recorded in TaskSpec. Retrieved and user-provided molecules are evaluated by the Property Evaluation Module, and the highest-ranked eligible examples initialize the global candidate set.

Discovery Agent. The Discovery Agent performs iterative exploration at two hierarchical timescales. At the beginning of structural stage s, it integrates the initial evidence, previous structural plans, and completed-batch feedback to define a structural plan F_s and Semantic Field Guidance ($SFG_{s,0}$). The structural plan specifies the scaffold and topological constraints, while $SFG_{s,0}$ identifies prioritized property directions and trade-offs for local refinement.

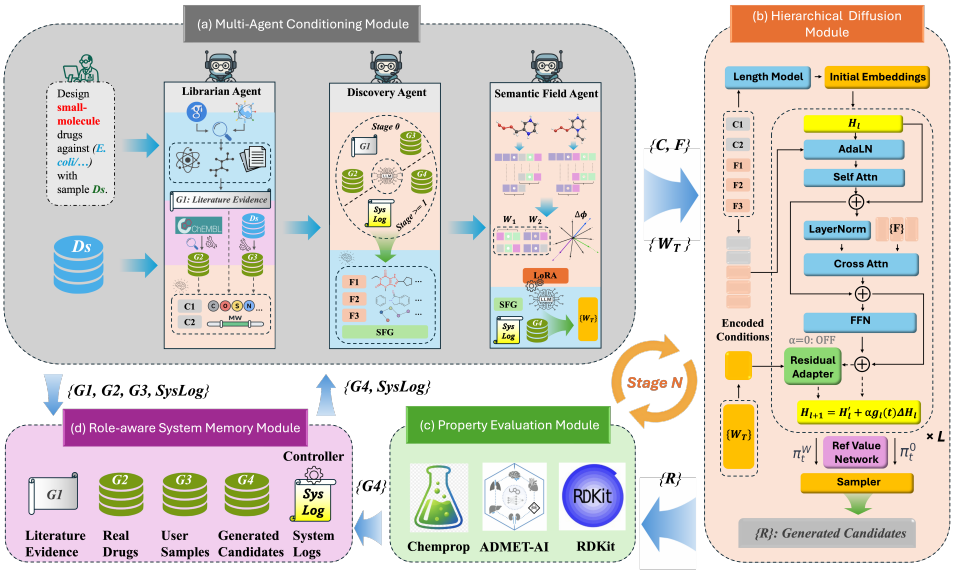

Figure 2: Overall framework. (a) Multi-Agent Conditioning Module: integrates evidence, structural planning, and semantic guidance to define generation conditions. (b) Hierarchical Diffusion Module: generates molecules under structural and semantic conditioning with reference-based feasibility control. (c) Property Evaluation Module: evaluates molecular validity, constraints, properties, and antibacterial activity. (d) Role-aware System Memory Module: records outcomes and feedback to support iterative planning and refinement.

Within a structural stage, F_s remains fixed while $SFG_{s,r}$ is updated after each semantic loop r. The reference population is initialized from G_3 and subsequently updated using the most recently completed generation batch. This separation allows semantic exploration to accumulate feedback within a fixed structural context before a new structural region is explored.

Semantic Field Agent. The Semantic Field Agent converts $SFG_{s,r}$ into local structural conditions for de novo generation. Given the current guidance, reference population, and role-specific feedback, it proposes candidate semantic windows and selects a target collection $W_{T,s,r}$ containing one to four windows:

$$W_{T,s,r} \subseteq P_{s,r} \cup W_{T,s,r-1}, \quad 1 \le |W_{T,s,r}| \le 4, \qquad W_{T,s,-1} = \varnothing. \tag{1}$$

Windows may therefore be newly proposed or retained from the preceding loop within the same structural stage. During deployment, the selected windows are passed to the frozen Hierarchical Diffusion Module (HDiff), and their evaluated outcomes provide feedback for subsequent semantic decisions. The complete Semantic Field Agent inference procedure is provided in Algorithm 4.

The Semantic Field Agent is trained in two phases using a LoRA adapter on a frozen language model (Figure 3). Phase I learns contextual property responses from molecular transformations, while Phase II adapts these representations to executable target-window selection. The complete training procedure is summarized in Algorithm 3, with training objectives and policy optimization detailed in Appendix D.

Controller and generation. The Controller supplies HDiff with C, F_s, and $W_{T,s,r}$ to generate a batch of independent attempts:

$$R_{s,r} \leftarrow HDiffBatch\,(C, F_s, W_{T,s,r}; b)\,, \tag{2}$$

where b is the batch size and $R_{s,r}$ records all attempts, including invalid and repeated molecules. Conditions and model parameters remain fixed within each batch. Each attempt begins from a fully masked SELFIES representation and is independently sampled under the specified conditions.

After evaluation, the Controller updates memory and continues semantic refinement within the current structural stage, requesting a new structural plan after the prescribed number of loops. Gener-

ation ends when the fixed attempt budget is exhausted. Invalid and duplicate outputs count toward the budget.

3.1.1 Molecular Semantic Fields

Inspired by semantic fields in language, where meaning emerges from contextual relationships between words, we define molecular semantic fields to capture how local structural units ("chemical words") collectively influence molecular properties in context. A paired window links source and target structures around an unchanged core adjacent to an edit, and windows derived from the same molecular transformation form a semantic field group. Molecular transformations are represented as directed pairs $(m_{\mathrm{src}}, m_{\mathrm{tgt}})$. For each molecule, we define a ten-dimensional property vector (created by the Property Evaluation Module) and its target-minus-source response:

$$\Phi(m) = (\phi_1(m), \ldots, \phi_{10}(m)), \qquad \Delta\Phi = \Phi(m_{\mathrm{tgt}}) - \Phi(m_{\mathrm{src}}). \tag{3}$$

The response is associated with the complete paired-window group rather than assigned additively to individual windows, allowing the model to learn context-dependent molecular transformations.

The semantic-field database contains 228,539 molecular-pair records after standardization and deduplication. Transformations are represented using Functional Group Tokenization (FGT), with localized structural changes and surrounding context forming the semantic windows used for learning. Target-side windows condition HDiff, with semantic-field database preparation summarized in Algorithm 2 and further details provided in Appendix D.

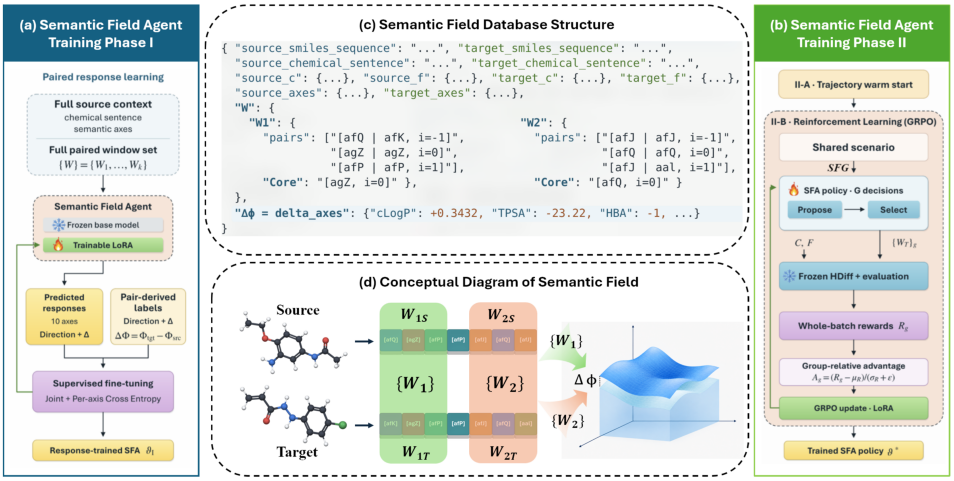

Figure 3: Semantic field agent and molecular semantic fields. (a) Phase I training for contextual property-response prediction. (b) Phase II training for target-window selection. (c) Database records linking paired structural windows to whole-molecule property responses. (d) Conceptual illustration of context-dependent property responses.

3.2 Property Evaluation Module

The Property Evaluation Module applies a common protocol to retrieved, user-provided, and generated molecules. It verifies molecular validity, representation constraints, and single-component structure, and evaluates ten semantic-field property axes capturing lipophilicity, polarity, hydrogen-bonding behavior, molecular flexibility, carbon hybridization, primary amine content, and ionization-related properties, together with task-level molecular properties. Candidate ranking uses a fixed composite score incorporating drug-likeness, synthetic accessibility, target-specific predicted antibacterial activity, and general toxicity. Global eligibility requires molecular validity, satisfaction of C, completed evaluations, and task admission, while satisfaction of F_s is retained as a structural diagnostic. The complete evaluation and ranking procedure is provided in Appendix C.

3.3 Role-aware System Memory Module

The Role-aware System Memory Module maintains information at multiple levels of the hierarchy. Literature evidence, ChEMBL molecules, and user-provided examples are retained as G_1, G_2, and G_3, respectively. For each structural stage, $G_{4,s}$ stores up to 1,000 distinct eligible generated molecules, while a persistent global set maintains the current top-ten candidates.

A system log records C, F_s, $SFG_{s,r}$, $W_{T,s,r}$, stage and loop identifiers, batch statistics, and generation outcomes. Complete archives retain failed and duplicate attempts, allowing subsequent planning to account for both molecular quality and generation efficiency. Memory update and candidate replacement rules are provided in Appendix C.

3.4 Hierarchical Diffusion Module

HDiff generates molecules under global, structural, and semantic conditioning while constraining deviations from a reference-defined feasible region. It uses absorbing-mask diffusion over SELFIES, with C and F_s defining the compositional and structural scope and $W_{T,s,r}$ providing local semantic guidance.

HDiff is trained sequentially through T1–T3 for reference generation and length modeling, semantic-window adaptation, and Reference Value Network learning, respectively, with the complete procedure provided in Algorithm 6 and Appendix E.

Molecular length is sampled from C and F_s, followed by denoising from a fully masked SELFIES sequence. Global and structural conditions enter the diffusion backbone, while semantic windows are incorporated through residual adapters:

$$H_{l+1} = H'_l + \alpha g_l(t)\Delta H_l, \qquad \alpha \in \{0, 1\}, \tag{4}$$

where H'_l is the backbone state, ΔH_l is the window-conditioned residual, and $g_l(t)$ is a learned diffusion-time gate. The binary variable α switches semantic-window conditioning off or on. A local encoder represents each window, and a set encoder aggregates the selected windows into a permutation-invariant semantic memory H_W.

3.4.1 Reference-based feasibility

A Reference Value Network estimates the probability that an intermediate diffusion state produces a valid molecule satisfying C and F_s under continuation without semantic-window conditioning:

$$V_\tau^{\mathrm{OFF}}(x_\tau; C, F_s) = \mathbb{E}_{X_0 \sim \mathrm{OFF}(\cdot|x_\tau,\tau,C,F_s)} \left[H_{\mathrm{qual}}(X_0; C, F_s) \right]. \tag{5}$$

Here, H_{qual} indicates successful decoding into a valid, single-component molecule satisfying C and F_s. The Reference Value Network V_ψ approximates V_τ^{OFF} from the partial sequence, diffusion time, and C/F_s conditions.

During reverse diffusion, HDiff combines reference and semantically conditioned candidates and computes

$$w^* = \arg\min_{w \in \Delta_N} D_{\mathrm{KL}}(w \,\|\, p) \quad \text{s.t.} \quad w^\top v \geq \max\left(0, b_0 - \epsilon_t\right). \tag{6}$$

Here, Δ_N is the probability simplex over N candidates, p contains the importance-corrected semantic weights, v contains the candidates' OFF-continuation qualification values, b_0 is the mean value among direct OFF candidates, and ϵ_t is the tolerance. The next state is sampled using w^*, with HDiff inference summarized in Algorithm 5 and further details in Appendix E.

4 Experiments

4.1 Implementation Details

Experiments were conducted on 1 x NVIDIA H100 GPU. The HDiff backbone comprised 6 Transformer layers with a hidden dimension of 256, 8 attention heads, and a feed-forward dimension of 1,024. The local window encoder and window set encoder contained 2 and 1 layers, respectively, both with a hidden dimension of 128 and 4 attention heads. The residual adapter used a bottleneck

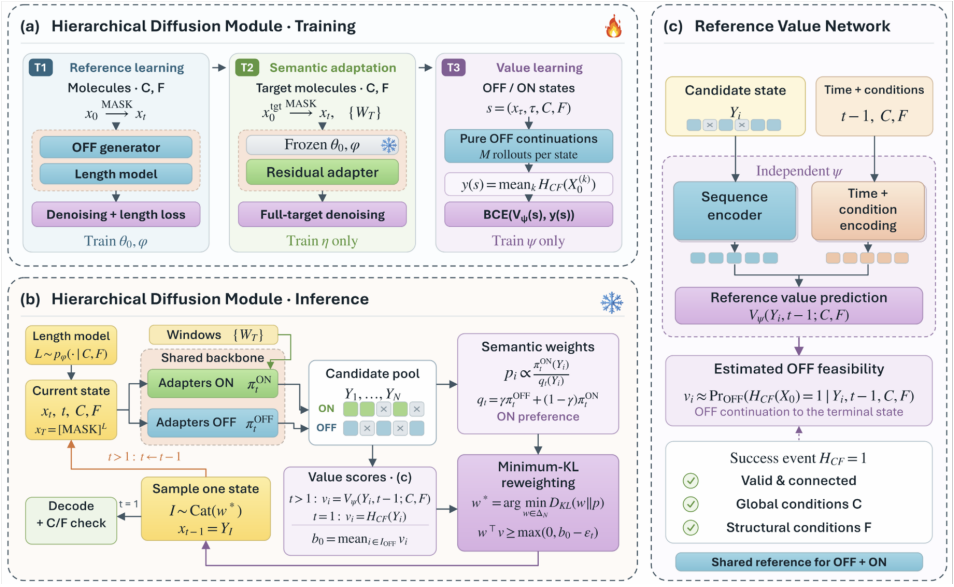

Figure 4: Hierarchical Diffusion Module. (a) Training of generation, semantic conditioning, and feasibility components. (b) Inference with hierarchical structural and semantic conditioning and reference-bounded sampling. (c) The Reference Value Network estimates terminal qualification under frozen OFF continuation, providing a common feasibility reference.

dimension of 64. The reference value network comprised 4 Transformer layers with a hidden dimension of 256 and 8 attention heads. Molecules were represented as SELFIES sequences with a maximum length of 256 tokens, and the diffusion process used 64 denoising steps.

4.2 Small Molecule Generation Testing

Bacterial Targets and Dataset. We evaluated small-molecule antibacterial design against *E. coli* and *S. aureus*, representing Gram-negative and Gram-positive bacteria, respectively. Molecular structures and antibacterial activity records from ChEMBL 36 yielded 2,746,426 unique molecules after standardization and deduplication. Of these, 497,978 molecules were used for HDiff reference learning with corresponding c/f values, and 228,539 molecular-pair records formed the molecular semantic field database. Species-specific MIC predictors were fine-tuned using 14,231 molecules active against *E. coli* and 41,273 molecules active against *S. aureus* (MIC $\leq 16\ \mu g/mL$).

Ablation and Baseline Models. We compared MSF-HDiff with five public methods: LDMol (Chang & Ye, 2025), DrugDiff (Oestreich et al., 2025), DemoDiff (Liu et al., 2025), MegaMolBART (NVIDIA Corporation, 2024), and AutoMolDesigner (Shen et al., 2024), as well as three study-specific ablations. GPT performs direct generation without multi-agent planning or hierarchical guidance; MSF-Guided GPT retains these components but replaces HDiff with GPT-5.5; and MSF-Guided Diff replaces HDiff with a standard diffusion generator without a Reference Value Network or reference-guided sampling. Further baseline model details are provided in Appendix F.

Structural Reliability. Under a fixed budget of 1,000 generation attempts per model and target, we evaluated validity, uniqueness, novelty, VUN (validity $\times$ uniqueness $\times$ novelty), 3D construction success, internal-energy validity, and end-to-end PoseBusters (PB) validity (Buttenschoen et al., 2024) (Table 1), with additional structural metrics reported in Supplementary Table A4 and Appendix G. MSF-HDiff achieves VUN rates of 100.00% for *E. coli* and 99.13% for *S. aureus*, alongside 100.0% 3D construction, bond-angle validity, and steric-clash-free rates and PB-validity of 98.20% and 98.30%. Relative to MSF-Guided Diff, this improves VUN by 56.70 and 28.23 percentage points and PB-validity by 17.33 and 10.20 points. For *E. coli*, adding the full planning and guidance framework to GPT also increases PB-validity from 87.07% to 100.00%, supporting the contribution of the proposed guidance framework to structural reliability.

Molecular Performance. We evaluated diversity, antibacterial activity (pMIC), drug-likeness (QED), synthetic accessibility (SAS), toxicity, and composite Ranking Scores (Table 2), with metric definitions and calculation details provided in Appendix H. Among the top 1,000 candidates from 10,000 generation attempts per model and target, MSF-HDiff achieves the highest Ranking Scores for *E. coli* (63.54) and *S. aureus* (65.32), with QED/SAS values of 0.71/2.87 and 0.77/2.79, respectively. Independent composite scores are 62.02 and 63.50, exceeding the strongest public baseline by 3.98 and 4.46 points, with toxicity of 0.14 for both targets. For *E. coli*, planning and guidance increase GPT's independent composite score from 53.02 to 62.01; MSF-HDiff maintains comparable performance (62.02) while improving QED (0.71 vs. 0.66), SAS (2.87 vs. 3.00), and toxicity (0.14 vs. 0.28). Additional component-level ablations are provided in Appendix I.

Table 1: Structural reliability of generated small molecules across bacterial species. Values are mean ± SD (%) over three seeds; higher is better. Best and second-best results are bolded and underlined, respectively.

Model	Validity	Uniqueness	Novelty	VUN	3D construction success	Internal energy valid	End-to-end PB-valid
			Escherichia coli (E. coli)				
MSF-HDiff	**100.00 ± 0.00**	**100.00 ± 0.00**	**100.00 ± 0.00**	**100.00 ± 0.00**	**100.00 ± 0.00**	98.20 ± 0.35	98.20 ± 0.35
MSF-Guided Diff	**100.00 ± 0.00**	43.30 ± 0.35	**100.00 ± 0.00**	43.30 ± 0.35	90.10 ± 0.35	93.45 ± 0.35	80.87 ± 0.70
MSF-Guided GPT	**100.00 ± 0.00**	95.30 ± 0.17	**100.00 ± 0.00**	95.30 ± 0.17	**100.00 ± 0.00**	**100.00 ± 0.00**	**100.00 ± 0.00**
GPT	99.90 ± 0.10	96.90 ± 0.17	**100.00 ± 0.00**	96.80 ± 0.10	99.90 ± 0.10	87.15 ± 0.60	87.07 ± 0.55
LDMol	87.40 ± 0.60	82.38 ± 0.23	**100.00 ± 0.00**	72.00 ± 0.44	86.10 ± 0.50	98.84 ± 0.20	83.90 ± 0.61
DrugDiff	**100.00 ± 0.00**	98.70 ± 0.10	**100.00 ± 0.00**	98.70 ± 0.10	97.40 ± 0.20	92.30 ± 0.35	87.97 ± 0.42
DemoDiff	93.40 ± 0.40	**100.00 ± 0.00**	**100.00 ± 0.00**	93.40 ± 0.40	92.60 ± 0.35	99.35 ± 0.11	91.90 ± 0.53
MegaMolBART	99.93 ± 0.06	98.90 ± 0.10	95.55 ± 0.20	94.43 ± 0.21	95.50 ± 0.35	99.69 ± 0.10	94.10 ± 0.35
AutoMolDesigner	98.00 ± 0.26	**100.00 ± 0.00**	**100.00 ± 0.00**	98.00 ± 0.26	97.60 ± 0.20	99.59 ± 0.10	97.20 ± 0.30
			Staphylococcus aureus (S. aureus)				
MSF-HDiff	**100.00 ± 0.00**	99.13 ± 0.12	**100.00 ± 0.00**	99.13 ± 0.12	**100.00 ± 0.00**	98.30 ± 0.17	98.30 ± 0.17
MSF-Guided Diff	**100.00 ± 0.00**	70.90 ± 0.35	**100.00 ± 0.00**	70.90 ± 0.35	93.60 ± 0.35	94.66 ± 0.35	88.10 ± 0.30
MSF-Guided GPT	**100.00 ± 0.00**	98.40 ± 0.10	**100.00 ± 0.00**	98.40 ± 0.10	99.50 ± 0.10	**99.90 ± 0.10**	**99.30 ± 0.17**
GPT	**100.00 ± 0.00**	98.10 ± 0.10	**100.00 ± 0.00**	98.10 ± 0.10	**100.00 ± 0.00**	87.50 ± 0.56	87.50 ± 0.56
LDMol	**100.00 ± 0.00**	99.80 ± 0.10	**100.00 ± 0.00**	**99.80 ± 0.10**	90.70 ± 0.35	98.24 ± 0.20	88.40 ± 0.53
DrugDiff	**100.00 ± 0.00**	99.20 ± 0.10	**100.00 ± 0.00**	99.20 ± 0.10	97.00 ± 0.20	91.24 ± 0.35	86.23 ± 0.40
DemoDiff	92.70 ± 0.40	**100.00 ± 0.00**	**100.00 ± 0.00**	92.70 ± 0.40	91.90 ± 0.35	98.91 ± 0.19	90.90 ± 0.46
MegaMolBART	99.93 ± 0.12	99.20 ± 0.61	87.40 ± 0.88	86.63 ± 0.25	98.50 ± 0.20	99.70 ± 0.10	97.90 ± 0.10
AutoMolDesigner	98.10 ± 0.26	**100.00 ± 0.00**	**100.00 ± 0.00**	98.10 ± 0.26	97.70 ± 0.20	99.69 ± 0.10	97.40 ± 0.26

Table 2: Molecular properties of generated small molecules across bacterial species. Values summarize the top 1,000 of 10,000 generated candidates. Best and second-best results are bolded and underlined, respectively.

Model	Diversity ↑	pMIC (Indep.) ↑	pMIC (Oracle) ↑	QED ↑	SAS ↓	Toxicity ↓	Ranking score ↑	Indep. composite score ↑
				Escherichia coli (E. coli)				
MSF-HDiff	0.86	4.32	4.47	**0.71**	**2.87**	0.14	**63.54**	**62.02**
MSF-Guided Diff	0.82	4.28	4.36	0.53	3.05	0.26	56.05	55.27
MSF-Guided GPT	0.87	**4.77**	4.40	0.66	3.00	0.28	59.09	62.01
GPT	**0.92**	3.85	3.42	0.49	4.08	**0.03**	50.39	53.02
LDMol	0.89	4.15	4.24	0.34	3.65	0.17	52.47	51.06
DrugDiff	0.90	4.09	4.07	0.40	4.93	0.17	49.22	48.84
DemoDiff	0.91	4.21	4.12	0.50	3.62	0.15	54.67	55.17
MegaMolBART	0.90	4.63	**4.70**	0.37	3.84	0.19	55.55	54.78
AutoMolDesigner	**0.92**	4.37	4.29	0.54	3.09	0.17	57.40	58.04
				Staphylococcus aureus (S. aureus)				
MSF-HDiff	0.85	4.43	4.60	**0.77**	**2.79**	0.14	**65.32**	**63.50**
MSF-Guided Diff	0.85	4.49	4.64	0.47	3.19	0.32	55.19	53.50
MSF-Guided GPT	**0.92**	4.13	3.65	0.52	3.91	**0.07**	51.46	54.73
GPT	0.84	4.45	4.59	0.63	2.82	0.31	58.94	57.27
LDMol	0.90	4.29	4.07	0.55	4.07	0.15	53.29	54.64
DrugDiff	0.90	4.25	3.86	0.41	4.99	0.18	46.10	48.94
DemoDiff	0.91	4.28	4.02	0.50	3.77	0.15	52.58	54.37
MegaMolBART	0.86	**5.01**	**4.96**	0.46	3.55	0.23	58.62	59.04
AutoMolDesigner	**0.92**	4.40	4.28	0.52	3.21	0.14	56.50	57.32

Aggregate Performance. For *E. coli*, we summarized performance across metrics using dense ranks, with ties receiving the same rank, followed by reciprocal-rank summation within each eval-

uation group. MSF-HDiff achieves the highest aggregate performance for both structural reliability and molecular properties, demonstrating strong overall performance (Figure 1).

Chemical-Space Analysis and Representative Molecules. MSF-HDiff samples multiple neighborhoods within the main ChEMBL reference-density region, whereas AutoMolDesigner spans a broader, ChEMBL-like distribution (Figure 5A). Descriptor PCA similarly shows concentrated sampling within the reference space, with the first three components explaining 89.8% of variance (Figure 5B) and a selected-set diversity of 0.86 (Table 2). Two representative candidates illustrate distinct architectures: Candidate 1 combines a pyrimidine carboxamide–benzenesulfonamide motif (Figure 5C), while Candidate 2 contains a quinolone-carboxylic-acid core with morpholine and cyclopropyl substituents (Figure 5D). Their QED/SAS values are 0.86/1.97 and 0.93/2.38, predicted toxicity is 0.09 and 0.21, and predicted *E. coli* MICs are 0.23 and 0.21 μg/mL, respectively. In addition, we characterized the structural and physicochemical profiles of MSF-HDiff-generated molecules and compared them with real-world drugs. This is detailed in Appendix J.

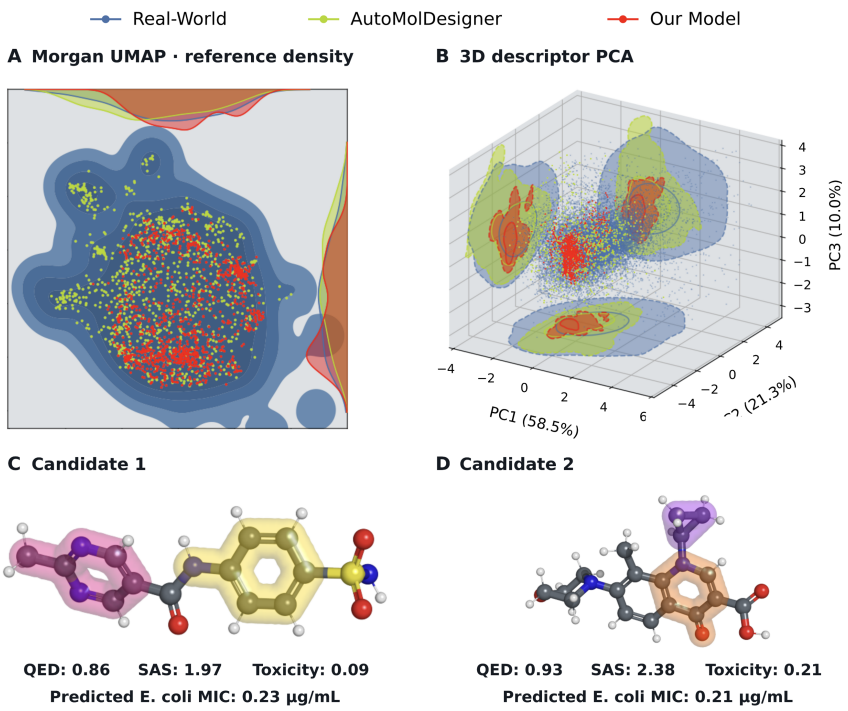

Figure 5: (A) Morgan-fingerprint UMAP and (B) descriptor PCA projections comparing MSF-HDiff, AutoMolDesigner, and ChEMBL reference molecules. (C-D) Representative MSF-HDiff candidates with distinct molecular architectures and favorable predicted properties.

5 Conclusion

We introduce MSF-HDiff, a novel molecular semantic field-guided multi-agent hierarchical diffusion framework for de novo antibacterial design that integrates closed-loop condition search, local-to-global chemical guidance, and reference-based feasibility control. Molecular semantic fields learn how combinations of local structural features relate to whole-molecule properties, enabling targeted conditions for hierarchical diffusion while balancing property objectives with molecular feasibility. Across *E. coli* and *S. aureus*, MSF-HDiff achieves high structural reliability and improved aggregate performance over ablation and public baselines, while retaining favorable drug-likeness, synthetic accessibility, and independent MIC predictions. Chemical-space analyses further demonstrate focused exploration while preserving structural diversity, highlighting the potential of MSF-HDiff for controlled de novo antibiotic discovery. Limitations and future directions are discussed in Appendix K.

AI USE STATEMENT

In this work, we used generative AI tools to assist with polishing and improving the clarity, grammar, and presentation of the manuscript. We did not use generative AI tools to generate research findings, conduct analyses, or make scientific decisions. All AI-assisted text was reviewed and edited by the authors, who verified its accuracy and appropriateness in the context of the work. We take full responsibility for the final content of this work, including all text, claims, and artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This work introduces MSF-HDiff, a framework for generating molecules with desired antibacterial activity and molecular properties. The framework is intended to support early-stage antibacterial drug discovery and does not involve human subjects or personally identifiable data. We are not aware of any immediate ethical concerns arising from its development or intended use. However, generated molecules should be treated as computational candidates rather than validated therapeutics and should undergo appropriate safety and experimental evaluation before downstream use. Users should consider the model's capabilities and limitations, apply appropriate scientific and safety practices, and use MSF-HDiff to complement rather than replace human scientific judgment.

REPRODUCIBILITY STATEMENT

We have made multiple efforts to ensure reproducibility throughout our study. We have included extensive details in the main manuscript and additional sections in our Appendix that provide algorithms and more in-depth explanations of the theoretical concepts. Upon acceptance, a public GitHub repository will be made available with all source code.

REFERENCES

Megan Bergkessel, Barbara Forte, and Ian H. Gilbert. Small-Molecule Antibiotic Drug Development: Need and Challenges. *ACS Infectious Diseases*, 9(11):2062–2071, November 2023. ISSN 2373-8227, 2373-8227. doi: 10.1021/acsinfecdis.3c00189.

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.

Jinho Chang and Jong Chul Ye. LDMol: A text-to-molecule diffusion model with structurally informative latent space surpasses AR models. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 7374–7392. PMLR, 13–19 Jul 2025.

Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance 1990–2021: a systematic analysis with forecasts to 2050. *The Lancet*, 404(10459):1199–1226, September 2024. ISSN 01406736. doi: 10.1016/S0140-6736(24)01867-1.

Yan Guo, Menglan Luo, Wenbo Zhang, Peidong Liu, Jin Liu, Shudong Huang, Jiancheng Lv, Bowen Ke, and Xianggen Liu. Few-shot molecular property optimization *via* a domain-specialized large language model. *Chemical Science*, 17(10):4928–4941, 2026. ISSN 2041-6520, 2041-6539. doi: 10.1039/D5SC08859C.

Guangyi Han, Wei Zhai, Yuhang Yang, Yang Cao, and Zheng-Jun Zha. Touch: Text-guided controllable generation of free-form hand-object interactions, 2025.

Hyomin Kim, Yunhui Jang, and Sungsoo Ahn. MT-Mol: Multi Agent System with Tool-based Reasoning for Molecular Optimization. In *Findings of the Association for Computational Linguistics: EMNLP 2025*, pp. 11544–11573. Association for Computational Linguistics, 2025. doi: 10.18653/v1/2025.findings-emnlp.619.

Aarti Krishnan, Melis N. Anahtar, Jacqueline A. Valeri, Wengong Jin, Nina M. Donghia, Leif Sieben, Andreas Luttens, Yu Zhang, Seyed Majed Modaresi, Andrew Hennes, Jenna Fromer, Parijat Bandyopadhyay, Jonathan C. Chen, Danyal Rehman, Ronak Desai, Paige Edwards, Ryan S. Lach, Marie-Stéphanie Aschtgen, Margaux Gaborieau, Massimiliano Gaetani, Samantha G. Palace, Satotaka Omori, Lutete Khonde, Yurii S. Moroz, Bruce Blough, Chunyang Jin, Edmund Loh, Yonatan H. Grad, Amir Ata Saei, Connor W. Coley, Felix Wong, and James J. Collins. A generative deep learning approach to de novo antibiotic design. *Cell*, 188(21):5962–5979.e22, October 2025. ISSN 00928674. doi: 10.1016/j.cell.2025.07.033.

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

Xiner Li, Yulai Zhao, Chenyu Wang, Gabriele Scalia, Gokcen Eraslan, Surag Nair, Tommaso Biancalani, Shuiwang Ji, Aviv Regev, Sergey Levine, and Masatoshi Uehara. Derivative-free guidance in continuous and discrete diffusion models with soft value-based decoding. In *Advances in Neural Information Processing Systems*, volume 38, Main Conference, pp. 95507–95545. Curran Associates, Inc., 2025. doi: 10.52202/085713-3194.

Gang Liu, Jie Chen, Yihan Zhu, Michael Sun, Tengfei Luo, Nitesh V Chawla, and Meng Jiang. Graph Diffusion Transformers are In-Context Molecular Designers, 2025.

Shengchao Liu, Jiongxiao Wang, Yijin Yang, Chengpeng Wang, Ling Liu, Hongyu Guo, and Chaowei Xiao. Conversational drug editing using retrieval and domain feedback. In *The Twelfth International Conference on Learning Representations*, 2024. URL `https://openreview.net/forum?id=yRrPfKyJQ2`.

Andres M. Bran, Sam Cox, Oliver Schilter, Carlo Baldassari, Andrew D. White, and Philippe Schwaller. Augmenting large language models with chemistry tools. *Nature Machine Intelligence*, 6(5):525–535, May 2024. ISSN 2522-5839. doi: 10.1038/s42256-024-00832-8.

NVIDIA Corporation. Megamolbart: A deep learning model for small molecule drug discovery and cheminformatics based on smiles, 2024. URL https://github.com/NVIDIA/MegaMolBART.

Marie Oestreich, Erinc Merdivan, Michael Lee, Joachim L. Schultze, Marie Piraud, and Matthias Becker. DrugDiff: small molecule diffusion model with flexible guidance towards molecular properties. *Journal of Cheminformatics*, 17(1):23, February 2025. ISSN 1758-2946. doi: 10.1186/s13321-025-00965-x.

Tao Shen, Jiale Guo, Zunsheng Han, Gao Zhang, Qingxin Liu, Xinxin Si, Dongmei Wang, Song Wu, and Jie Xia. AutoMolDesigner for Antibiotic Discovery: An AI-Based Open-Source Software for Automated Design of Small-Molecule Antibiotics. *Journal of Chemical Information and Modeling*, 64(3):575–583, February 2024. ISSN 1549-9596, 1549-960X. doi: 10.1021/acs.jcim.3c01562.

Jonathan M. Stokes, Kevin Yang, Kyle Swanson, Wengong Jin, Andres Cubillos-Ruiz, Nina M. Donghia, Craig R. MacNair, Shawn French, Lindsey A. Carfrae, Zohar Bloom-Ackermann, Victoria M. Tran, Anush Chiappino-Pepe, Ahmed H. Badran, Ian W. Andrews, Emma J. Chory, George M. Church, Eric D. Brown, Tommi S. Jaakkola, Regina Barzilay, and James J. Collins. A Deep Learning Approach to Antibiotic Discovery. *Cell*, 180(4):688–702.e13, February 2020. ISSN 00928674. doi: 10.1016/j.cell.2020.01.021.

Kyle Swanson, Gary Liu, Denise B. Catacutan, Autumn Arnold, James Zou, and Jonathan M. Stokes. Generative AI for designing and validating easily synthesizable and structurally novel antibiotics. *Nature Machine Intelligence*, 6(3):338–353, March 2024a. ISSN 2522-5839. doi: 10.1038/s42256-024-00809-7.

Kyle Swanson, Parker Walther, Jeremy Leitz, Souhrid Mukherjee, Joseph C Wu, Rabindra V Shivnaraine, and James Zou. ADMET-AI: a machine learning ADMET platform for evaluation of large-scale chemical libraries. *Bioinformatics*, 40(7):btae416, July 2024b. ISSN 1367-4803, 1367-4811. doi: 10.1093/bioinformatics/btae416.

Kyle Swanson, Gary Liu, Denise B Catacutan, Stewart McLellan, Autumn Arnold, Megan M Tu, Eric D Brown, James Zou, and Jonathan M Stokes. SyntheMol-RL: a flexible reinforcement learning framework for designing easily synthesizable antibiotics. *Molecular Systems Biology*, 22(6):833–867, April 2026. ISSN 1744-4292. doi: 10.1038/s44320-026-00206-9.

Clement Vignac, Igor Krawczuk, Antoine Siraudin, Bohan Wang, Volkan Cevher, and Pascal Frossard. Digress: Discrete denoising diffusion for graph generation. In *The Eleventh International Conference on Learning Representations*, 2023.

Yida Xiong, Kun Li, Jiameng Chen, Hongzhi Zhang, Di Lin, Yan Che, and Wenbin Hu. Text-guided multi-property molecular optimization with a diffusion language model. *Information Fusion*, 127:103907, March 2026. ISSN 15662535. doi: 10.1016/j.inffus.2025.103907.

Yuanxin Zhuang, Dazhong Shen, and Ying Sun. Moleditrl: Structure-preserving molecular editing via discrete diffusion and reinforcement learning. In *International Conference on Learning Representations*, volume 2026, pp. 125947–125983, 2026.

APPENDIX TABLE OF CONTENTS

A Notation and Condition Definitions **15**

A.1 Notation and Indexing 15

A.2 Global and Structural Conditions 15

A.3 Semantic Axes, Guidance, and Windows 15

B Pseudocode for Module Algorithms **18**

B.1 Overall Architecture Algorithm 18

B.2 Semantic Field Algorithms 18

B.3 Hierarchical Diffusion Module Algorithms 22

C Multi-Agent Coordination and Evaluation **24**

C.1 Evidence Preparation and Task Initialization 24

C.2 Memory Objects and Access 24

C.3 Property Evaluation 25

C.3.1 Reference normalization and utility functions 25

C.3.2 Eligibility and structural satisfaction 26

C.4 Batch-level Feedback 26

C.5 Stage and Semantic-loop Coordination 26

C.6 Candidate and Memory Updates 26

C.7 Termination and Generation Budget 27

D Molecular Semantic Field Construction and Agent Learning **27**

D.1 Semantic-field database 27

D.2 Data splitting and paired-window retrieval 27

D.3 Core-centered semantic windows 28

D.4 Phase I: contextual property-response learning 28

D.5 Phase II-A: executable semantic-window decisions 29

D.6 Semantic reward construction 29

D.7 Phase II-B: group-relative policy optimization 30

E HDiff Training and Reference-Guided Sampling **31**

E.1 Conditional representation and generative architecture 31

E.2 Semantic-window encoding 32

E.3 Residual semantic conditioning 32

E.4 Training stages 32

E.4.1 T1: conditional reference generation and length modeling 32

E.4.2 T2: semantic-window conditioning 33

E.4.3 T3: Reference Value Network 33

E.5 Absorbing-mask reverse diffusion 33
E.6 Reference–semantic mixture and importance correction 34
E.7 Reference feasibility bound 34
E.8 Minimum-KL reference-bounded sampling 34
E.9 Reference-feasibility relationship 35
E.10 Deployment 36

F Additional Information on Baseline Models **36**

G Structural Reliability Metrics and Additional Results **36**
G.1 Structural Reliability Metrics Calculation Methods 36
G.2 Additional Structural Reliability Results 37

H Molecular Property Metrics Methods **37**

I Ablation Studies **38**

J Structural and Physicochemical Profile Analysis of Generated Molecules **38**

K Limitations and Future Directions **40**

A NOTATION AND CONDITION DEFINITIONS

We distinguish run-level global conditions, stage-level structural conditions, and loop-level semantic guidance. This appendix specifies their notation, domains, and satisfaction rules, following Section 3. We write $1[\cdot]$ for an indicator function.

A.1 NOTATION AND INDEXING

Stage and loop indices s, r start at zero. The global conditions C remain fixed throughout a run, and F_s remains fixed within stage s. For a single generation request or training example, F denotes its structural conditions; at runtime in stage s, $F = F_s$. The accepted guidance and selected windows remain fixed during the corresponding batch. For a single generation request, F_s and $W_{T,s,r}$ denote the structural conditions and selected target-window collection for the current stage and semantic loop. The T in $W_{T,s,r}$ and W_T labels target windows and is not a diffusion time index.

Indices are local to their expressions: H_i is an attempt-level reward gate, whereas H_l is a hidden state. The generation-attempt record $R_{s,r}$ and reward $R(a; z)$ are distinct objects; ϕ_j denotes a property component, whereas ϕ parameterizes the length model. All learned parameters remain fixed during deployment.

A.2 GLOBAL AND STRUCTURAL CONDITIONS

Global conditions. For a complete molecule m, let $c_1(m)$ be its element set, including implicitly represented hydrogen, and let $c_2(m)$ be its molecular weight in g/mol. C_1 is a nonempty set of allowed element identifiers. $C_2 = [M_{\min}, M_{\max}]$ is a closed interval with $0 \le M_{\min} \le M_{\max}$ and $M_{\max} > 0$. Under the fixed molecular-processing protocol, their joint satisfaction is

$$C_{\text{pass}}(m; C) = \mathbf{1}[c_1(m) \subseteq C_1]\ \mathbf{1}[M_{\min} \le c_2(m) \le M_{\max}]\,. \tag{A7}$$

Structural conditions. F_1, F_2, and F_3 denote scaffold archetype, ring-system profile, and motif topology. Let I_F index the eleven fields in Table A2. Field k has a versioned category vocabulary V_k, an extracted category $f_k(m)$, and a nonempty allowed set $A_{(s,k)} \subseteq V_k$. Its `allowed_values` array contains distinct category labels; enumeration order does not affect meaning. The complete structural condition is

$$F_{\text{pass}}(m; F_s) = \prod_{k \in I_F} \mathbf{1}\big[f_k(m) \in A_{(s,k)}\big]\,. \tag{A8}$$

Thus, alternatives within a field are combined by OR, and requirements across fields by AND. Each value of `key_motif_set` is one registered category for a combination of motifs. Feature extraction, condition encoding, and terminal checks share the same field definitions and category vocabularies, fixed throughout a run.

Terminal qualification. Set $H_{\text{qual}}(x_0; C, F_s) = 0$ if decoding x_0 fails to yield a nonempty, chemically valid, single-component molecule under the fixed run protocol. Otherwise, for the decoded molecule m,

$$H_{\text{qual}}(x_0; C, F_s) = C_{\text{pass}}(m; C)\, F_{\text{pass}}(m; F_s). \tag{A9}$$

Target windows guide generation and are not an additional predicate in H_{qual}. This terminal event is distinct from global candidate eligibility E_{run} in Section 3.1, which uses validity, C, completed evaluation, and task admission criteria.

A.3 SEMANTIC AXES, GUIDANCE, AND WINDOWS

Semantic axes and response labels. The components of $\Phi(m)$ follows the axis order in Table A3. For a directed pair,

$$\Delta\phi_j = \phi_j(m_{\text{tgt}}) - \phi_j(m_{\text{src}}),$$

consistently with Equation 3. Let J contain axes with finite values at both endpoints under compatible units and calculation protocols. For $j \in J$, the response direction is

$$\begin{aligned} d_j = \text{increase} &\iff \Delta\phi_j > \tau_j, \\ d_j = \text{decrease} &\iff \Delta\phi_j < -\tau_j, \\ d_j = \text{maintain} &\iff |\Delta\phi_j| \le \tau_j. \end{aligned} \tag{A10}$$

Table A1: Principal symbols and their scopes

Symbols	Meaning
$m, m_{\mathrm{src}}, m_{\mathrm{tgt}}$	Complete molecule, source molecule, and target molecule of a directed pair.
$s, r; R_{s,r}, b$	Stage and semantic-loop indices; generation-attempt record for the current stage and loop, and batch size.
$t, T, x_t; L, l$	Diffusion time, denoising horizon, and sequence state; SELFIES sequence length and network-layer index. Denoising proceeds from T to 0.
$C = (C_1, C_2); F_s$	Global conditions; structural conditions $(F_{1,s}, F_{2,s}, F_{3,s})$ at stage s.
$S(m), W_{\mathrm{pair}}, W_T, W_T^{\mathrm{sel}}, W_{T,s,r}$	Chemical sentence, complete paired-window group, individual target window, target-window collection for one request or training example, and selected collection at stage s and loop r.
$SFG_{s,r}, P_{s,r}$	Semantic Field Guidance and proposed target windows for the current loop.
$\Phi(m), \phi_j(m), \Delta\Phi$	Ten-axis molecular state, its jth component, and directed target-minus-source response.
$G_1, G_2, G_3, G_{4,s}$; SysLog	Literature evidence, reference molecules, user-provided examples, and stage-specific generated-molecule archive; linked execution and feedback records.
$\Theta = (\theta_{ref}, \phi, \eta, \psi); \vartheta$	Complete HDiff parameter collection: reference backbone, length model, semantic-conditioning modules, and Reference Value Network; Semantic Field Agent LoRA parameters.
$\alpha; \pi_t^{\mathrm{OFF}}, \pi_t^{\mathrm{ON}}$	Adapter switch; OFF and ON next-state kernels. OFF uses $\alpha = 0$ and ON uses $\alpha = 1$.
$H_l, H_l', \Delta H_l; H_W, g_l(t)$	Layer state, pre-adapter output, and semantic residual; joint window memory and learned time gate.
$H_{\mathrm{qual}}; V_\tau^{\mathrm{OFF}}, V_\psi$	Terminal C/F_s qualification indicator; expected qualification under OFF continuation from time τ, and its learned approximation.
$a, z, E_i; R(a; z)$	Complete proposal–selection decision, shared training scenario, attempt evaluation, and batch semantic reward.
$H_i, M_i, K_i; J_{SFG}, \omega_j, u_j$	Reward gates for joint qualification, finite required measurements, and hard maintenance requirements, respectively; priority-axis set, normalized utility weights, and utilities.
$Y_i, n_{\mathrm{OFF}}, n_{\mathrm{ON}}, N; q_t, \rho_i, p_i$	Candidate occurrences, branch counts, and pool size; mixed proposal, importance ratio, and normalized semantic weight.
$v_i, w^*; b_0, \beta_t, \epsilon_t$	Candidate qualification value and adjusted sampling-weight vector; direct-OFF mean, required mean-value bound, and absolute probability tolerance.
T_{top}	Persistent global top-ten candidate set (SysLog.top10)

The axis-specific tolerance $\tau_j \geq 0$ is expressed in the original units and fixed using training data and calculation precision. It remains unchanged for validation and testing. Direction-label tolerances are distinct from runtime maintenance intervals, and each complete paired-window group receives one whole-molecule response.

Descriptor definitions, ionization conventions, and calculation provenance are retained with the data and shared across the corresponding learning and evaluation views.

Semantic Field Guidance. $SFG_{s,r}$ contains intent for the overall property aim, an ordered list of priorities, and trade-offs for joint improvement and maintenance considerations. Each guidance instance contains exactly three distinct priority axes. Each entry supplies an axis key, a direction from increase/decrease/maintain, and its purpose; explicit target or maintenance intervals accompany the guidance when specified. The reference consists of distinct valid C-satisfying molecules from G_3

Table A2: Structural fields and their corresponding categories

Level	Field	Structural category
F_1	`scaffold_family`	Scaffold family
F_1	`core_atom_class`	Core atomic composition
F_1	`core_size_class`	Core size
F_2	`ring_system_count_class`	Number of ring systems
F_2	`aromatic_ring_count_class`	Number of aromatic rings
F_2	`heteroaromatic_state`	Heteroaromatic composition
F_2	`ring_connectivity_state`	Ring-system connectivity
F_3	`layout_type`	Layout of key motifs
F_3	`key_motif_set`	Registered combination of key motifs
F_3	`attachment_pattern`	Distribution of motif attachments to the main structure
F_3	`symmetry`	Symmetry of the motif layout

Table A3: Semantic axes in their fixed order (counts and logarithmic values are dimensionless; net charge is expressed in units of elementary charge e)

Axis key	Meaning	Unit or scale
`clogp`	Calculated log partition coefficient (cLogP)	Logarithmic
`tpsa`	Topological polar surface area (TPSA)	Å^2
`hbd`	Hydrogen-bond donor count (HBD)	Count
`hba`	Hydrogen-bond acceptor count (HBA)	Count
`rotatable_bonds`	Rotatable-bond count	Count
`fraction_csp3`	Fraction of carbon atoms with sp^3 hybridization	Fraction
`primary_amine_count`	Primary-amine count	Count
`clogd_ph7_4`	Calculated log distribution coefficient at pH 7.4	Logarithmic
`most_basic_pka`	Most basic pK_a under the fixed site aggregation rule	Logarithmic
`net_charge_ph7_4`	Net charge at pH 7.4	e

for the run's first loop, and from the latest completed batch thereafter, including across stage transitions. Reference statistics are fixed before comparing window decisions. The Semantic Field Agent receives the task, C, SFG, relevant results, and window history through its dedicated view.

Paired and target windows. A chemical sentence $S(m) = (\xi_0, \ldots, \xi_{n-1})$ is an ordered sequence over a fixed chemical vocabulary V_{chem}. A paired window stores ordered source/target word pairs and their signed positions relative to a shared core. The core is an unchanged word occurrence adjacent to a changed block and has relative position zero. Gaps use a marker $X \notin V_{\text{chem}}$. The complete group W_{pair} retains distinct core occurrences, overlapping windows, and repeated words.

A target window is a complete word-and-core object,

$$W_T = ((\xi_0, \ldots, \xi_{n_W-1}), c_W), \qquad n_W \in \mathbb{N}_{\geq 1}, \qquad \xi_i \in V_{\text{chem}}, \qquad c_W \in \{0, \ldots, n_W - 1\}. \tag{A11}$$

Here, c_W is the zero-based `core_index`. Target-side projection removes target gaps within each paired window and recomputes this index while preserving word order, repetitions, and window boundaries. The executed collection $W_{T,s,r}$ contains one to four complete target windows. The collection retains window occurrences; window order within $W_{T,s,r}$ does not specify a chemical adjacency relation. The core identifies a word occurrence within its window, without assigning a SELFIES position in the generated molecule.

B PSEUDOCODE FOR MODULE ALGORITHMS

B.1 OVERALL ARCHITECTURE ALGORITHM

Algorithm 1 OVERALL ARCHITECTURE

Given: D_{user}: user-provided molecular examples.
function OVERALL(Task, D_{user})
  $(G_1, G_2, G_3, C) \leftarrow$ LIBRARIAN(Task, D_{user})
  $(\text{TaskSpec}, G_3) \leftarrow$ ESTABLISHBASELINE(Task, C, G_3)
  $M \leftarrow$ INITMEMORY(G_1, G_2, G_3, TaskSpec)
  SysLog.top10 $\leftarrow$ RANKRUN($G_3, 10$)
  $N_{\text{started}} \leftarrow 0$; $s \leftarrow -1$ ▷ N_{started}: cumulative attempts
  while $N_{\text{started}} + b \leq B_{\max}$ **do**
    $(F, SFG) \leftarrow$ DISCOVERYSTART(READCONTEXT$_D(M)$)
    $s \leftarrow s + 1$; $F_s \leftarrow F$; $SFG_{s,0} \leftarrow SFG$
    $G_{4,s} \leftarrow \emptyset$; $W_{\text{prev}} \leftarrow \emptyset$
    for $r = 0, \ldots, n_{\text{loop}} - 1$ **do**
      if $r > 0$ **then**
        $SFG_{s,r} \leftarrow$ DISCOVERYCONTINUE(F_s, READCONTEXT$_D(M)$)
      end if
      context $\leftarrow$ READCONTEXT$_{SF}(M, s)$ ▷ task, feedback, window history
      $P_{s,r} \leftarrow$ SEMANTICPROPOSE($SFG_{s,r}$, context)
      $W_{T,s,r} \leftarrow$ SEMANTICSELECT($P_{s,r} \cup W_{\text{prev}}, SFG_{s,r}$, context)
      $R \leftarrow$ HDIFFBATCH($C, F_s, W_{T,s,r}; b$)
      $E \leftarrow$ EVALUATE(R, F_s; TaskSpec)
      $A \leftarrow \{m \in E : E_{\text{run}}(m)\}$
      $G_{4,s} \leftarrow$ RANKRUN($G_{4,s} \cup A, 1000$)
      SysLog.top10 $\leftarrow$ RANKRUN(SysLog.top10 $\cup A, 10$)
      $N_{\text{started}} \leftarrow N_{\text{started}} + b$ ▷ all generation attempts
      Commit conditions, proposals and full-batch feedback to M
      $W_{\text{prev}} \leftarrow W_{T,s,r}$
      if $N_{\text{started}} + b > B_{\max}$ **then**
        return FINALIZE(M, budget exhausted)
      end if
    end for
  end while
  return FINALIZE(M, budget exhausted)
end function

B.2 SEMANTIC FIELD ALGORITHMS

Algorithm 2 SEMANTIC FIELD DATABASE PREPARATION

Given: Fixed chemical vocabulary V; gap symbol $X \notin V$.

function BUILDDATA(molecules)
  molecules $\leftarrow$ CLEAN(molecules) ▷ standardize; deduplicate
  Form chemical sentences S with V; compute $c(m) = \big(c_1(m), c_2(m)\big)$, $f(m) = \big(f_k(m)\big)_{k \in I_F}$, and $\Phi(m)$ for each m
  $b(m) \leftarrow (\mathrm{SMILES}(m), S(m), c(m), f(m), \Phi(m))$ for each m
  $D \leftarrow \emptyset$
  for each split $\in$ GROUPSPLIT(molecules) **do** ▷ disjoint groups
    for each directed $(m_{\mathrm{src}}, m_{\mathrm{tgt}}) \in$ SIMILARPAIRS(split) **do**
      $W_{\mathrm{pair}} \leftarrow$ PAIREDWINDOWS$(S_{\mathrm{src}}, S_{\mathrm{tgt}})$
      if $1 \le |W_{\mathrm{pair}}| \le 4$ **then**
        $\Delta\Phi \leftarrow \Phi_{\mathrm{tgt}} - \Phi_{\mathrm{src}}$ ▷ whole-molecule response
        APPEND$(D[\mathrm{split}], (b_{\mathrm{src}}, b_{\mathrm{tgt}}, W_{\mathrm{pair}}, \Delta\Phi))$
      end if
    end for
  end for
  return D
end function

function SCORE(A, B) ▷ A: aligned source; B: aligned target
  ▷ $[\ell, r]$: inclusive start/end column indices of a run
  $C \leftarrow$ CHANGEBLOCKS(A, B) ▷ maximal changed runs
  $M \leftarrow$ MATCHBLOCKS(A, B) ▷ maximal matching runs
  $D \leftarrow \sum_k \mathbf{1}[A_k \neq B_k]$ ▷ changed-column count
  $Q \leftarrow \sum_{[\ell,r] \in M} (r - \ell + 1)^2$ ▷ sum of squared matching-run lengths
  if $D > 3$ **then**
    return $(+\infty, +\infty)$
  end if
  if $\exists [\ell, r] \in C : r - \ell + 1 > 2$ **then**
    return $(+\infty, +\infty)$
  end if
  return $(D, -Q)$ ▷ minimize D, then maximize Q
end function

function PAIREDWINDOWS$(S_{\mathrm{src}}, S_{\mathrm{tgt}})$
  $(A, B) \leftarrow$ GLOBALALIGN$(S_{\mathrm{src}}, S_{\mathrm{tgt}}; \min_{\mathrm{lex}}$ SCORE$)$
  if no feasible alignment exists **then**
    return ()
  end if
  $N \leftarrow |A|$ ▷ common length of aligned sentences
  $C \leftarrow$ CHANGEBLOCKS(A, B)
  $P \leftarrow \emptyset$ ▷ distinct core positions
  for each $[\ell, r] \in C$ **do**
    for each $p \in \{\ell - 1, r + 1\} \cap \{0, \ldots, N - 1\}$ **do**
      if $A_p = B_p \neq X$ **then**
        $P \leftarrow P \cup \{p\}$ ▷ stable core
      end if
    end for
  end for
  $W_{\mathrm{pair}} \leftarrow ()$
  for each $p \in$ SORT(P) **do**
    $h \leftarrow \max\{r - \ell + 1 : [\ell, r] \in C,\ p \in \{\ell - 1, r + 1\}\}$ ▷ radius
    APPEND$(W_{\mathrm{pair}}, ((A_{p+i}, B_{p+i}, i))_{i=-h}^{h})$ ▷ $i = 0$: core; pad with X
  end for
  return W_{pair}
end function

Algorithm 3 SEMANTIC FIELD AGENT: TRAINING

Given: Paired data D, scenes Z; frozen base θ_0, HDiff and evaluator.
CE_ϑ: mean answer-token loss; b: generation attempts per decision.

function TRAINSF(D, Z)
  $\vartheta \leftarrow \text{INITLORA}(\theta_0)$ ▷ one LoRA throughout
  $(\{\tau_j\}, \{s_j\}) \leftarrow \text{CALIBRATE}(D_{\text{train}})$ ▷ direction tolerances and reward scales
  for $k = 1, \ldots, N_I$ **do** ▷ I: paired response learning
    $n \sim \text{PAIRSAMPLE}(D_{\text{train}})$
    $x \leftarrow (\text{SMILES}_{\text{src}}, S_{\text{src}}, \Phi_{\text{src}}, W_{\text{pair}})_n$ ▷ source and full window group
    $(y, J) \leftarrow \text{AXISLABELS}(\Delta\Phi_n; \{\tau_j\})$ ▷ direction/delta labels; available axes J
    $c \leftarrow \mathbf{1}[|J| = 10]$; $\ell_{\text{joint}} \leftarrow CE_\vartheta(y \mid x)$ if $c = 1$, else 0
    $\ell_{\text{axis}} \leftarrow |J|^{-1} \sum_{j \in J} CE_\vartheta(y_j \mid x, j)$
    $\vartheta \leftarrow \text{UPDATE}\left(\vartheta, \frac{c\ell_{\text{joint}} + \ell_{\text{axis}}}{c+1}\right)$
  end for
  $\vartheta_I \leftarrow \text{PICKVALID}(\vartheta, D_{\text{val}})$; $\vartheta \leftarrow \vartheta_I$
  $D_A \leftarrow ()$ ▷ II-A: actual-trajectory warm start
  for each $z \in Z_{\text{train}}$ **do**
    $z \leftarrow \text{FIXSCENE}(z; \{s_j\})$; $(a, E, R) \leftarrow \text{ROLLOUT}(\pi_{\vartheta_I}, z)$
    if $\text{QUALIFIED}(E, R)$ **then**
      $\text{APPEND}(D_A, (z, a))$
    end if
  end for
  for $k = 1, \ldots, N_A$ **do**
    $(z, (Y_P, Y_W)) \sim D_A$; $(o, W_{\text{prev}}) \leftarrow \text{READCONTEXT}_{SF}(z)$
    $h \leftarrow \text{SELECTCONTEXT}(o, Y_P, W_{\text{prev}})$ ▷ this trajectory's proposal
    $\ell_A \leftarrow \frac{1}{2} CE_\vartheta(Y_P \mid o, \text{propose}) + \frac{1}{2} CE_\vartheta(Y_W \mid h, \text{select})$
    $\vartheta \leftarrow \text{UPDATE}(\vartheta, \ell_A)$
  end for
  $\vartheta_A \leftarrow \vartheta$; $\pi_{\text{ref}} \leftarrow \text{FROZENCOPY}(\pi_{\vartheta_A})$
  for $k = 1, \ldots, N_B$ **do** ▷ II-B: GRPO
    $z \sim Z_{\text{train}}$; $z \leftarrow \text{FIXSCENE}(z; \{s_j\})$ ▷ shared reference, targets and budget
    $\pi_{\text{old}} \leftarrow \text{FROZENCOPY}(\pi_\vartheta)$
    for $g = 1, \ldots, G$ **do**
      $(a_g, E_g, R_g) \leftarrow \text{ROLLOUT}(\pi_{\text{old}}, z)$ ▷ independent decisions
    end for
    if $\text{STD}(R_{1:G}) \le \tau_R$ **then**
      continue
    end if
    $A_g \leftarrow \dfrac{R_g - \text{MEAN}(R_{1:G})}{\text{STD}(R_{1:G}) + \epsilon_{\text{num}}}$ for each g
    $(J_{\text{clip}}, K) \leftarrow \text{GRPO}(z, \{(a_g, A_g)\}_{g=1}^G; \pi_\vartheta, \pi_{\text{old}}, \pi_{\text{ref}})$
    $\vartheta \leftarrow \text{UPDATE}(\vartheta, -J_{\text{clip}} + \beta K)$ ▷ joint token mean; K: KL
  end for
  $\vartheta^* \leftarrow \text{PICKVALID}(\vartheta, Z_{\text{val}})$;
  return $\text{FREEZE}(\vartheta^*)$
end function

Algorithm 3 SEMANTIC FIELD AGENT: TRAINING (CONTINUED)

function ROLLOUT(π, z)
  $(o, W_{\text{prev}}) \leftarrow \text{READCONTEXT}_{SF}(z)$
  $(a, W_T^{\text{sel}}) \leftarrow \text{DECIDE}(\pi, o, W_{\text{prev}})$
  if a is not executable **then**
    return $(a, (), -1)$
  end if
  $(C, F_s) \leftarrow \text{CONDITIONS}(z)$ ▷ Controller inputs
  $E \leftarrow \text{EVALUATE}(\text{HDIFFBATCH}(C, F_s, W_T^{\text{sel}}; b))$
  $J_{SFG} \leftarrow \text{PRIORITYAXES}(z)$ ▷ three SFG axes
  $(H_i, M_i, K_i) \leftarrow \text{GATES}(E_i, z)$ for $i = 1, \ldots, b$
▷ H_i: joint qualification; M_i: finite required measurements; K_i: hard maintenance requirements
  $q_i \leftarrow \sum_{j \in J_{SFG}} \omega_j u_j(E_i; z)$ if $H_i M_i K_i = 1$, else 0
  $R \leftarrow b^{-1} \sum_{i=1}^{b} q_i$ ▷ all attempts, including repeats
  return (a, E, R) ▷ u_j: fixed SFG utility
end function

Algorithm 4 SEMANTIC FIELD AGENT: INFERENCE

Given: Frozen policy π_{ϑ^*} and chemical vocabulary V; b attempts per batch.

function SFSTEP(C, F_s, SFG, M, s, r) ▷ semantic step in Overall
  reference $\leftarrow G_3$ if $(s, r) = (0, 0)$, else $\text{LATESTBATCH}(M)$
  reference $\leftarrow \text{UNIQUE}(\{m \in \text{reference} : m \text{ valid}, m \models C\})$
  $W_{\text{prev}} \leftarrow \emptyset$ if $r = 0$, else $\text{LASTWINDOWS}(M, s)$ ▷ stage-local active windows
  $o \leftarrow \text{READCONTEXT}_{SF}(M, C, SFG, \text{reference}, V)$
    ▷ o: task and property context; F_s and structural labels stay with Controller
  $(a, W_T^{\text{sel}}) \leftarrow \text{DECIDE}(\pi_{\vartheta^*}, o, W_{\text{prev}})$
  $E \leftarrow \text{EVALUATE}(\text{HDIFFBATCH}(C, F_s, W_T^{\text{sel}}; b))$
  $\text{COMMIT}(M, s, r, SFG, \text{reference}, a, W_T^{\text{sel}}, E)$ ▷ whole-batch feedback
  return (W_T^{sel}, E)
end function

function DECIDE(π, o, W_{prev}) ▷ shared by training and inference
  $Y_P \sim \pi(\cdot \mid o, \text{propose})$
  $h \leftarrow \text{SELECTCONTEXT}(o, Y_P, W_{\text{prev}})$
  $Y_W \sim \pi(\cdot \mid h, \text{select})$
  $W_T^{\text{sel}} \leftarrow \text{PARSE}(Y_W)$ ▷ 1–4 complete windows
  return $((Y_P, Y_W), W_T^{\text{sel}})$ ▷ actual replies and selected windows
end function

function SELECTCONTEXT(o, Y_P, W_{prev})
  $P \leftarrow \text{PARSE}(Y_P)$ ▷ proposed target windows
  return $(o, Y_P, P \cup W_{\text{prev}})$
end function

B.3 HIERARCHICAL DIFFUSION MODULE ALGORITHMS

Algorithm 5 HIERARCHICAL DIFFUSION MODULE (HDIFF): INFERENCE

Given: Frozen $\Theta = (\theta_{ref}, \phi, \eta, \psi)$; $n_{\mathrm{OFF}}, n_{\mathrm{ON}}$; schedule $\{dt\}$; tolerances $\{\epsilon_t\}$.

function HDIFF($C, F, W_T^{\mathrm{sel}}; \Theta$)
  $L \sim p_\phi(\cdot \mid C, F)$; $x_T \leftarrow \mathrm{MASK}^L$
  for $t = T, \ldots, 1$ **do**
    $\pi_t^{\mathrm{OFF}} \leftarrow$ REVERSEKERNEL($x_t, t, C, F; \theta_{ref}$) ▷ adapters off
    $\pi_t^{\mathrm{ON}} \leftarrow$ REVERSEKERNEL($x_t, t, C, F, W_T^{\mathrm{sel}}; \theta_{ref}, \eta$) ▷ adapters on
    $x_{t-1} \leftarrow$ COORDINATEDSTEP($\pi_t^{\mathrm{OFF}}, \pi_t^{\mathrm{ON}}, t, C, F$)
  end for
  return (DECODESELFIES(x_0), $H_{qual}(x_0; C, F)$)
end function

function COORDINATEDSTEP($\pi_t^{\mathrm{OFF}}, \pi_t^{\mathrm{ON}}, t, C, F$)
  $N \leftarrow n_{\mathrm{OFF}} + n_{\mathrm{ON}}$; $\gamma \leftarrow n_{\mathrm{OFF}}/N$
  $Y_{\mathrm{OFF}} \leftarrow$ SAMPLE n_{OFF} next states from π_t^{OFF}
  $Y_{\mathrm{ON}} \leftarrow$ SAMPLE n_{ON} next states from π_t^{ON}
  $(Y_1, \ldots, Y_N) \leftarrow$ CONCAT($Y_{\mathrm{OFF}}, Y_{\mathrm{ON}}$) ▷ retain repeats
  for $i = 1, \ldots, N$ **do**
    $q_t(Y_i) \leftarrow \gamma \pi_t^{\mathrm{OFF}}(Y_i) + (1-\gamma)\pi_t^{\mathrm{ON}}(Y_i)$
    $\rho_i \leftarrow \dfrac{\pi_t^{\mathrm{ON}}(Y_i)}{q_t(Y_i)}$
    if $t = 1$ **then**
      $v_i \leftarrow H_{qual}(Y_i; C, F)$
    else
      $v_i \leftarrow \mathrm{V}_\psi(Y_i, t-1; C, F)$
    end if
  end for
  $p \leftarrow$ NORMALIZE(ρ)
  $b_0 \leftarrow n_{\mathrm{OFF}}^{-1} \sum_{i=1}^{n_{\mathrm{OFF}}} v_i$
  $\beta_t \leftarrow \max(0, b_0 - \epsilon_t)$
  $w^* \leftarrow \underset{w \in \Delta_N,\ w^\top v \geq \beta_t}{\arg\min} D_{\mathrm{KL}}(w \| p)$
  $I \sim$ CAT(w^*)
  return Y_I
end function

Algorithm 6 HIERARCHICAL DIFFUSION MODULE (HDIFF): TRAINING

Given: Molecule data D_0, paired records D_W, and requests R.

function TRAINHDIFF(D_0, D_W, R)
  Initialize θ_{ref}, ϕ and independent ψ
  for $k = 1, \ldots, N_0$ **do** ▷ T1: reference learning
    $m \sim D_0$; $(x_0, x_t, t, C, F) \leftarrow$ NOISYEXAMPLE(m)
    $P^{\mathrm{OFF}} \leftarrow f_{\theta_{ref}}(x_t, t; C, F)$
    $\ell \leftarrow$ MASKEDCE($P^{\mathrm{OFF}}, x_0, x_t$) $- \lambda_{\mathrm{len}} \log p_\phi(|x_0| \mid C, F)$
    $(\theta_{ref}, \phi) \leftarrow$ UPDATE($(\theta_{ref}, \phi), \ell$)
  end for
  FREEZE(θ_{ref}, ϕ); $\eta \leftarrow$ INITSEMANTICMODULES
  for $k = 1, \ldots, N_W$ **do** ▷ T2: semantic adaptation
    $(m_{\mathrm{tar}}, W_T^{\mathrm{sel}}) \leftarrow$ TARGETRECORD(D_W)
    $(x_0, x_t, t, C, F) \leftarrow$ NOISYEXAMPLE(m_{tar})
    $H_W \leftarrow$ JOINTWINDOWENCODE$_\eta$(W_T^{sel})
    $P^{\mathrm{ON}} \leftarrow f_{\theta_{ref},\eta}(x_t, t; C, F, H_W)$
    $\eta \leftarrow$ UPDATE $\big(\eta,$ MASKEDCE$(P^{\mathrm{ON}}, x_0, x_t)\big)$
  end for
  FREEZE(η)
  $S \leftarrow$ COLLECTSTATES(OFF/ON, R) ▷ T3: value learning
  $D_V \leftarrow$ REFERENCERETURNS($S; \theta_{ref}$)
  $\psi \leftarrow$ TRAINBCE(ψ, D_V)
  $S \leftarrow$ COLLECTSTATES(HDIFF, R) ▷ offline augmentation
  $D_V \leftarrow D_V \,\|\,$ REFERENCERETURNS($S; \theta_{ref}$)
  $\psi \leftarrow$ TRAINBCE(ψ, D_V)
  FREEZE(ψ)
  return $(\theta_{ref}, \phi, \eta, \psi)$
end function

function REFERENCERETURNS($S; \theta_{ref}$)
  $D_V \leftarrow ()$
  for each $s = (x_\tau, \tau, C, F) \in S$ **do**
    for $k = 1, \ldots, M$ **do**
      $X_0^{(k)} \sim$ OFFCONTINUE($x_\tau, \tau, C, F; \theta_{ref}$)
    end for
    $y \leftarrow M^{-1} \sum_{k=1}^{M} H_{qual}(X_0^{(k)}; C, F)$
    APPEND($D_V, (s, y)$)
  end for
  return D_V
end function

C Multi-Agent Coordination and Evaluation

This section specifies the coordination protocol linking task evidence, agent decisions, molecular evaluations, and memory updates. The framework operates at multiple hierarchical timescales: global conditions are established at the task level, structural conditions are fixed within each structural stage, and semantic guidance is updated within each stage. All agent decisions and generation outcomes are linked through explicit stage, loop, batch, and molecule identifiers.

C.1 Evidence Preparation and Task Initialization

The Librarian Agent first identifies the target bacterium, relevant strain information, design requirements, and user-specified constraints. It retrieves literature evidence, active molecules from ChEMBL, and user-provided molecular examples. These information sources are represented as G_1, G_2, and G_3, respectively.

Retrieved molecules are standardized to canonical isomeric SMILES before downstream evaluation. Original and standardized representations are linked, exact duplicate structures are merged, and distinct experimental measurements for the same standardized molecule are retained. Overlap between public and user-provided molecules is represented using a shared standardized molecular identity rather than duplicated molecular records. All computational properties and model-derived activity predictions are computed using the same protocols used during candidate evaluation.

The Librarian Agent publishes the global molecular conditions

$$C = (C_1, C_2), \qquad C_2 = [M_{\min}, M_{\max}],$$

where C_1 is the allowed element set and C_2 is the molecular-weight interval. The evidence manifest and associated provenance accompany these conditions, which remain fixed during a run unless a new task is explicitly initialized.

Following evidence preparation, the Controller constructs TaskSpec containing the target, design requirements, C, representation and admission criteria, identity and evaluation protocols, candidate-score parameters, batch size, semantic-loop schedule, and generation budget.

C.2 Memory Objects and Access

The system maintains memory at distinct levels of the hierarchy:

- G_1: literature evidence and extracted task-relevant information;
- G_2: retrieved ChEMBL molecules and associated activity information;
- G_3: user-provided molecular examples;
- $G_{4,s}$: generated molecules retained for structural stage s;
- SysLog: structured records of conditions, agent decisions, generation batches, evaluations, and outcomes;
- $\mathrm{BatchArchive}$: complete records of generation attempts, including invalid and duplicate molecules;
- T_{top}: persistent global top-ten candidate set stored in SysLog.top10.

For each structural stage, $G_{4,s}$ contains at most 1,000 distinct eligible generated molecules. In contrast, T_{top} persists across structural stages and contains at most ten distinct generated molecules. Complete batch archives are not restricted to eligible molecules and therefore retain failed and repeated attempts.

The Discovery Agent reads the initial evidence, previous structural plans, completed structural-stage selections, and feedback from completed batches. The Semantic Field Agent receives the task specification, C, current $SFG_{s,r}$, the relevant reference population, evaluated generation results, and semantic-window history. The raw structural plan F_s and explicit structural labels are not exposed to the Semantic Field Agent during executable window selection. The Controller supplies the full C, F_s, and selected $W_{T,s,r}$ to HDiff.

Each batch record links the selected window collection, semantic guidance, structural conditions, reference population, model configuration, sample size, and generation outcomes. These identifiers allow feedback from a completed batch to be associated with the precise conditions under which it was generated.

C.3 Property Evaluation

The Property Evaluation Module applies the same molecular evaluation protocol to literature-derived, ChEMBL, user-provided, and generated molecules. A molecule is considered valid if its representation is parseable, chemically valid, nonempty, and satisfies the required single-component representation.

For each valid standardized molecule, the system records the realized global and structural conditions, validity flags, ten semantic-field properties, and four ranking properties:

- quantitative estimate of drug-likeness (QED);
- synthetic accessibility (SAS);
- target-specific predicted activity ($pMIC$);
- general toxicity (ClinTox).

The target-specific activity is represented as

$$pMIC = -\log_{10}\left(\frac{MIC}{\mathrm{mol/L}}\right). \tag{A12}$$

Higher QED and $pMIC$, and lower SAS and ClinTox, are considered favorable for candidate ranking.

C.3.1 Reference normalization and utility functions

Let B_P denote the set of distinct standardized valid molecules in G_3 with finite $pMIC$. The reference activity location is defined by

$$p_0 = \mathrm{median}_{m \in B_P}\, pMIC(m), \tag{A13}$$

with scale parameter

$$\kappa = 1. \tag{A14}$$

The four normalized utility functions are

$$u_Q(m) = QED(m), \tag{A15}$$

$$u_{SA}(m) = \frac{10 - SAS(m)}{9}, \tag{A16}$$

$$u_P(m) = \frac{1}{1 + \exp[-(pMIC(m) - p_0)/\kappa]}, \tag{A17}$$

$$u_T(m) = 1 - \mathrm{ClinTox}(m). \tag{A18}$$

The fixed candidate score is

$$\mathrm{Score}(m) = 100\left[0.2u_Q(m) + 0.2u_{SA}(m) + 0.4u_P(m) + 0.2u_T(m)\right]. \tag{A19}$$

The target-specific activity receives the largest contribution to the composite score, while drug-likeness, synthetic accessibility, and toxicity each contribute equally to the remaining weight. The independent composite replaces only the oracle pMIC in the Ranking Score with the external predictor's pMIC; all other property values, normalization parameters, and weights remain unchanged.

C.3.2 Eligibility and structural satisfaction

Let $C_{\text{pass}}(m; C)$ indicate satisfaction of the global conditions and let $\text{EvalComplete}(m)$ and $\text{Admit}_{\text{task}}(m)$ indicate completed evaluation and task admission, respectively. Run-level eligibility is

$$E_{\text{run}}(m) = C_{\text{pass}}(m; C)\, \mathbb{1}\left[\text{EvalComplete}(m)\right]\, \mathbb{1}\left[\text{Admit}_{\text{task}}(m)\right]. \tag{A20}$$

Structural satisfaction is recorded separately:

$$F_{\text{pass}}(m; F_s) = \mathbb{1}\left[m \text{ satisfies } F_s\right]. \tag{A21}$$

It is retained as a structural execution diagnostic and does not independently define global candidate eligibility.

A standardized molecular identity is evaluated once for ranking and eligibility, while every occurrence remains represented in the batch archive with its generation provenance.

C.4 Batch-level Feedback

For a batch of size b, all initiated generation chains contribute to the denominator, including invalid and duplicate attempts. Batch feedback contains counts and distributions describing molecular validity, global-condition satisfaction, structural-condition satisfaction, evaluation completion, task admission, candidate eligibility, and duplicate frequency.

For a condition set (C, F_s), the empirical qualification rate is

$$\hat{p}_{CF}^{s,r} = \frac{1}{b} \sum_{i=1}^{b} H_{\text{qual}}\left(x_0^{(i)}; C, F_s\right) \tag{A22}$$

Here, $x_0^{(i)}$ is the terminal SELFIES state of attempt i, and H_{qual} is defined in Equation A9. Batch feedback additionally contains property distributions, candidate scores, generation consumption, and semantic-window outcomes.

C.5 Stage and Semantic-loop Coordination

At the beginning of structural stage s, the Discovery Agent integrates initial evidence, previous plans, structural-stage selections, and completed-batch feedback to produce

$$(F_s, SFG_{s,0}). \tag{A23}$$

The structural plan F_s is held fixed for all semantic loops within the stage. The Semantic Field Guidance is updated as

$$SFG_{s,r} \leftarrow \text{DiscoveryUpdate}\left(SFG_{s,r-1}, R_{s,r-1}, \text{SysLog}, G_{4,s}\right), \tag{A24}$$

while F_s remains unchanged.

The first semantic loop uses G_3 as the reference population. Subsequent loops use the latest completed batch as the reference population. The Semantic Field Agent proposes candidate windows from the current guidance and may retain previously active windows within the same structural stage. Windows that become inactive or are carried across stages must be re-proposed under the new stage context.

Each generation attempt begins from an independently fully masked SELFIES sequence. The global conditions, structural conditions, semantic-window collection, model parameters, generation configuration, and evaluation protocol remain fixed for the duration of the batch.

C.6 Candidate and Memory Updates

Let $U_{s,r}$ denote the set of distinct eligible molecules generated in semantic loop r of structural stage s. The stage-specific memory is updated as

$$G_{4,s}^{+} = \text{Top}_{1000}\left(G_{4,s}^{-} \cup U_{s,r}; G_{4,s}^{-}\right), \tag{A25}$$

where ranking uses the fixed candidate score and existing incumbents receive tie priority.

The persistent global candidate set is updated as

$$T_{top}^{+} = \mathrm{Top}_{10}\left(T_{top}^{-} \cup U_{s,r}; T_{top}^{-}\right). \tag{A26}$$

Duplicate standardized molecular identities are removed before candidate-set insertion.

The stage-specific memory $G_{4,s}$ is reset when a new structural stage begins, whereas T_{top} persists across all stages.

C.7 Termination and Generation Budget

Each run uses 10 structural stages and 5 semantic loops per stage, with $b = 20$ and $B_{max} = 1,000$ for Table 1 and A4, and $b = 200$ and $B_{max} = 10,000$ for Table 2.

A complete batch is initiated only if

$$B_{\max} - N_{\text{started}} \geq b. \tag{A27}$$

Thus, the system does not initiate an incomplete batch when the remaining generation budget is smaller than the configured batch size. Every initiated generation chain contributes to N_{started}, regardless of whether the resulting molecule is valid, eligible, or duplicated.

The final system output contains the ranked candidate set, molecular properties, structural-stage assignments, semantic guidance, selected windows, generation identifiers, and associated provenance.

D Molecular Semantic Field Construction and Agent Learning

D.1 Semantic-field database

The molecular semantic-field database contains 228,539 molecular-pair records after molecular standardization and deduplication. Each record is represented as a directed transformation

$$(m_{\text{src}}, m_{\text{tgt}}), \tag{A28}$$

where m_{src} is the source molecule and m_{tgt} is the transformed target molecule.

The molecular structures are represented using Functional Group Tokenization (FGT). Let

$$S(m) = (A_1, \ldots, A_K) \tag{A29}$$

denote the resulting chemical-word sequence for molecule m. Each word retains its structural identity and position within the complete molecular representation.

For each molecule, the ten-dimensional property vector is

$$\Phi(m) = (\phi_1(m), \ldots, \phi_{10}(m)), \tag{A30}$$

and the directed property response is

$$\Delta\Phi = \Phi(m_{\text{tgt}}) - \Phi(m_{\text{src}}). \tag{A31}$$

The ten response axes represent key physicochemical and structural properties: calculated lipophilicity (cLogP), topological polar surface area (TPSA), hydrogen-bond donors (HBD), hydrogen-bond acceptors (HBA), rotatable bond count, fraction of sp³-hybridized carbons (FractionCSP3), primary amine count, distribution coefficient at pH 7.4 ($cLogD_{7.4}$), most basic pKa, and net charge at pH 7.4. Numerical responses are retained in their original units, together with direction labels using the axis-specific tolerances described in Appendix A.

D.2 Data splitting and paired-window retrieval

Dataset splitting is performed before paired-window retrieval to prevent transformed molecules from crossing training and held-out partitions. Each paired transformation is then processed independently within its assigned split. Reverse transformations are generated only within the corresponding split.

For each directed pair, source and target FGT sequences are aligned using constrained dynamic programming. Let

$$A = (A_1, \ldots, A_L), \qquad B = (B_1, \ldots, B_L) \tag{A32}$$

denote the aligned source and target sequences after gap insertion. A gap is represented by the symbol X.

The alignment is constrained such that each aligned column contains a source/target word pair and the changed-column indicator is

$$\chi_k = \mathbb{1}[A_k \neq B_k]. \tag{A33}$$

The total number of changed positions is

$$D_{\text{edit}} = \sum_k \chi_k, \qquad D_{\text{edit}} \leq 3, \tag{A34}$$

and the maximum contiguous changed interval is restricted to a length of at most two. Dynamic-programming ties are resolved using the predefined ordering of alignment cost, number of changes, interval length, and positional ordering.

D.3 Core-centered semantic windows

For each localized changed interval, a candidate core position p is selected adjacent to the changed interval. A local window with radius

$$h_p \in \{1, 2\} \tag{A35}$$

is defined as

$$W(p, h_p) = \{(A_{p+i}, B_{p+i}, i) : -h_p \leq i \leq h_p\}. \tag{A36}$$

Positions extending beyond the aligned sequence are padded with X. The core index is retained explicitly, and relative positions are measured with respect to the core. Distinct eligible cores are enumerated in ascending positional order. Overlapping windows and repeated windows are retained because they may correspond to distinct contextual representations.

Only transformations yielding a complete group of one to four paired windows are retained. The complete group is associated with the whole-molecule response $\Delta\Phi$.

For use by HDiff, a paired source/target window is projected to its target-side representation:

$$W_T = \Pi_{\text{tgt}}(W), \tag{A37}$$

where target-side gap symbols are omitted when required by the target representation. The target-side core index is retained as c_W.

The complete semantic representation therefore consists of the source molecular context, the paired structural windows, the whole-molecule response $\Delta\Phi$, and the target-side window representation.

D.4 Phase I: contextual property-response learning

Phase I trains a LoRA adapter ϑ on a frozen base language model. The training sequence is initialized from a frozen base checkpoint:

$$\vartheta_I \to \vartheta_A \to \vartheta^*, \tag{A38}$$

where ϑ_I denotes the initial adapter, ϑ_A denotes the warm-started adapter used for Phase II, and ϑ^* denotes the final selected policy adapter.

For each training example, the model receives:

- source molecular representation m_{src};
- source FGT sequence $S(m_{\text{src}})$;
- source property vector $\Phi(m_{\text{src}})$;
- paired semantic windows.

Let J_n denote the available property-response axes for example n. The model predicts both signed numerical responses and their corresponding directions. Both are learned using mean answer-token cross-entropy.

For a response axis j, the directional label is determined using the predefined tolerance for that property. Let c_n indicate that all ten response axes are available:

$$c_n = \mathbf{1}\left[|J_n| = 10\right]. \tag{A39}$$

The joint objective is

$$L_I = \mathbb{E}_n\left[\frac{c_n \ell_{\text{joint},n} + \ell_{\text{axis},n}}{c_n + 1}\right].$$

where $\ell_{\text{joint},n}$ is the mean answer-token loss for the joint response and $\ell_{\text{axis},n}$ is the average mean answer-token loss over the available axes.

Held-out model selection considers both numerical response error and macro-averaged directional F1. The resulting adapter is used to initialize Phase II.

D.5 Phase II-A: executable semantic-window decisions

Phase II adapts contextual semantic representations to executable window selection. Training scenarios include initial loops, later semantic loops, and structural-stage transitions.

The Semantic Field Agent receives an observation

$$o = \{\text{task}, C, SFG_{s,r}, \text{reference}, \text{results}, \text{window history}, \text{vocabulary}\}. \tag{A40}$$

The backend execution state additionally contains

$$z = \{o, F_s, \text{frozen environment}\}. \tag{A41}$$

The structural plan is available to the execution backend but is not directly exposed as a raw structural label in the agent observation used for executable window selection.

The policy factorizes into a proposal and selection process:

$$\pi_\vartheta = \pi_\vartheta^P \, \pi_\vartheta^W, \tag{A42}$$

where π_ϑ^P proposes a candidate window set Y_P and π_ϑ^W selects the executable target-window collection Y_W.

The Phase II-A warm-start objective is

$$L_A = \frac{1}{2}\mathbb{E}\left[\text{CE}_\vartheta\left(Y_P \mid o, \text{propose}\right) + \text{CE}_\vartheta\left(Y_W \mid Y_P, o, \text{select}\right)\right].$$

Warm-start demonstrations are selected using condition qualification, maintenance of required properties, and downstream reward. The resulting adapter ϑ_A initializes Phase II-B. A frozen reference policy π_{ref} is retained for the KL regularization used during policy optimization.

D.6 Semantic reward construction

For each semantic decision a under scenario z, the generation environment is fixed with respect to guidance, reference population, target windows, property scales, weights, structural conditions, and generation budget.

For each property axis j, b_j is the reference-population median, and s_j is a fixed reward scale calibrated from the training data.

$$u_j(v) = \frac{1}{2} + \frac{1}{2}\,\text{clip}\left(\frac{\sigma_j(v - b_j)}{s_j}, -1, 1\right), \tag{A43}$$

where σ_j specifies the desired direction of improvement.

For interval-valued objectives, let I_j denote the target interval. The corresponding utility is

$$u_j(v) = \max\left(0, 1 - \frac{\mathrm{dist}(v, I_j)}{s_j}\right). \tag{A44}$$

Maintenance intervals are constructed from reference-population quantiles using the predefined tolerance parameter δ_{maint}.

The Semantic Field Guidance assigns normalized weights

$$\boldsymbol{\omega} = (0.5, 0.3, 0.2), \qquad \sum_j \omega_j = 1. \tag{A45}$$

For generated molecule m_i, let H_i denote the joint validity and qualification gate, M_i denote the finite required-measurement gate, and

$$K_i = \prod_q \mathbb{1}\left[\mathrm{Maintain}_q(m_i)\right] \tag{A46}$$

denote the product of maintenance indicators. The per-molecule semantic contribution is therefore

$$H_i M_i K_i \sum_{j \in J_{SFG}} \omega_j u_j\left(\phi_j(m_i)\right). \tag{A47}$$

The batch reward is

$$R(a; z) = \frac{1}{b} \sum_{i=1}^{b} H_i M_i K_i \sum_{j \in J_{SFG}} \omega_j u_j\left(\phi_j(m_i)\right), \qquad R(a; z) \in [0, 1]. \tag{A48}$$

Specifically, H_i requires a valid single-component molecule satisfying the relevant C and F_s conditions, M_i requires all measurements needed for reward evaluation to be finite, and K_i requires all hard maintenance requirements to be satisfied. Failed gates contribute zero. Failed and duplicate attempts remain in the batch denominator, ensuring that the reward captures generation efficiency as well as molecular quality.

Non-executable semantic decisions receive

$$R(a; z) = -1. \tag{A49}$$

The semantic reward is distinct from the global candidate-ranking score in Equation A19; it is used only to optimize semantic-window selection.

D.7 Phase II-B: group-relative policy optimization

Policy optimization groups multiple independent semantic decisions under the same scenario. For group g, let the rewards be

$$\{R_{g,1}, \ldots, R_{g,K_g}\}. \tag{A50}$$

The group mean and standard deviation are

$$\mu_g = \frac{1}{K_g} \sum_{k=1}^{K_g} R_{g,k}, \tag{A51}$$

and

$$s_g = \mathrm{SD}\left(R_{g,1}, \ldots, R_{g,K_g}\right). \tag{A52}$$

The group-relative advantage is

$$A_{g,k} = \frac{R_{g,k} - \mu_g}{s_g + \epsilon}. \tag{A53}$$

When

$$s_g \le \tau_R, \tag{A54}$$

all advantages in the group are set to zero and the policy update for that group is skipped. This prevents numerical amplification of negligible within-group reward differences.

For a sampled token or decision component, the probability ratio is

$$\rho_{g,k} = \frac{\pi_\vartheta(a_{g,k}|o_{g,k})}{\pi_{\text{old}}(a_{g,k}|o_{g,k})}. \tag{A55}$$

The clipped group-relative objective is

$$J_{\text{clip}} = \frac{1}{G}\sum_{g=1}^{G}\frac{1}{K_g}\sum_{k=1}^{K_g}\min\left[\rho_{g,k}A_{g,k}, \text{clip}\left(\rho_{g,k}, 1-\epsilon_{\text{clip}}, 1+\epsilon_{\text{clip}}\right)A_{g,k}\right]. \tag{A56}$$

The policy is regularized toward the frozen reference policy using conditional KL divergence:

$$D_g = D_{\text{KL}}\left(\pi_\vartheta(\cdot|o_g)\|\pi_{\text{ref}}(\cdot|o_g)\right). \tag{A57}$$

The Phase II-B objective is

$$L_B = -J_{\text{clip}} + \frac{\lambda_{\text{KL}}}{G}\sum_{g=1}^{G} D_g. \tag{A58}$$

Only LoRA parameters are updated during Phase II-B. Final model selection is performed on held-out scenarios using reward, executable decision rate, and policy stability. The selected adapter ϑ^* is frozen for deployment.

E HDiff Training and Reference-Guided Sampling

This section specifies the HDiff architecture, training objectives, reference value model, reverse diffusion process, reference-bounded sampling procedure, and feasibility relationship underlying the main-text description.

E.1 Conditional representation and generative architecture

HDiff uses a reference backbone with parameters θ_{ref}, a conditional length model with parameters ϕ, semantic-conditioning modules (window encoders, residual adapters, and time gates) with parameters η, and a Reference Value Network with parameters ψ.

HDiff encodes the allowed element set C_1 as a multi-hot vector and the molecular-weight interval $C_2 = [M_{\min}, M_{\max}]$ by its standardized endpoints. These encodings are concatenated and projected to obtain the global condition embedding:

$$\mathbf{e}_C = \text{MLP}\left([\text{MultiHot}(C_1); \text{Standardize}(M_{\min}, M_{\max})]\right).$$

Structural conditions F_s contain eleven structural fields. Each field is represented using a multi-hot categorical encoding followed by a learned projection. A field-identity embedding is added to distinguish otherwise identical categorical values occurring in different structural fields. The resulting field representations are

$$E_F = \{e_{F_1}, \ldots, e_{F_{11}}\}. \tag{A59}$$

For a SELFIES sequence

$$x = (x_1, \ldots, x_L), \tag{A60}$$

token and positional embeddings are combined with a diffusion-time representation. Let

$$e_t = \text{MLP}_t\left(\text{Sinusoidal}(t)\right). \tag{A61}$$

The backbone uses adaptive layer normalization conditioned on e_C and e_t, structural cross-attention to E_F, and a feed-forward transformation.

E.2 SEMANTIC-WINDOW ENCODING

For target semantic window w, let the sequence of target chemical words be indexed by k and let c_w denote the core position. Each local token representation is

$$u_{w,k} = E_{\text{word}}\left(\text{word}_{w,k}\right) + E_{\text{core}}\left(\mathbb{1}[k = c_w]\right) + E_{\text{rel}}(k - c_w). \tag{A62}$$

A local encoder produces a window-level representation from the complete contextual sequence:

$$z_w = \text{LocalEncoder}\left(\{u_{w,k}\}_k\right)_{c_w}. \tag{A63}$$

The selected windows are aggregated using a permutation-invariant set encoder:

$$H_W = \text{SetEncoder}\left(\{z_w + e_{\text{count}}(K)\}_{w=1}^{K}\right), \tag{A64}$$

where $K = |W_{T,s,r}|$.

The set representation is permutation invariant with respect to the ordering of windows, while within-window order, core location, repeated windows, and the number of selected windows remain encoded.

E.3 RESIDUAL SEMANTIC CONDITIONING

At diffusion layer l, let H'_l denote the backbone representation after global and structural conditioning. A query representation is constructed as

$$Q_l = \text{LayerNorm}(H'_l)W_l^q + \mathbf{1}_L(s_{CF}W_l^c), \tag{A65}$$

where s_{CF} is a compact representation of the global and structural conditions.

The semantic residual is computed through cross-attention:

$$\Delta H_l = \text{MHA}\left(Q_l, H_W W_l^k, H_W W_l^v\right) W_l^u + b_l. \tag{A66}$$

A diffusion-time gate is defined by

$$g_l(t) = \sigma\left(w_t^\top e_t + b_t\right). \tag{A67}$$

The conditioned layer update is

$$H_{l+1} = H'_l + \alpha g_l(t)\Delta H_l, \qquad \alpha \in \{0, 1\}. \tag{A68}$$

The setting $\alpha = 0$ defines the reference or semantic-window-off model, whereas $\alpha = 1$ activates semantic-window conditioning. All backbone weights, diffusion schedules, length models, and structural support are otherwise shared.

E.4 TRAINING STAGES

HDiff is trained in three stages.

E.4.1 T1: CONDITIONAL REFERENCE GENERATION AND LENGTH MODELING

The length model samples

$$L \sim p_\phi(\cdot \mid C, F_s). \tag{A69}$$

For a sampled length L, the initial diffusion state is

$$x_T = [\text{MASK}]^L. \tag{A70}$$

Let d_t denote the masking probability at diffusion step t. The denoising objective for model variant ν is

$$L_{\text{den}}^\nu = \mathbb{E}_{t,x_0,x_t}\left[\frac{1}{|M_t|}\sum_{j\in M_t} -\log p_{\theta^\nu}(x_{0,j} \mid x_t, C, F_s)\right], \tag{A71}$$

where M_t is the set of masked positions.

The T1 objective is

$$L_{T1} = L_{\text{den}}^{\text{OFF}} + \lambda_{\text{len}} \mathbb{E}\left[-\log p_\phi(L \mid C, F_s)\right]. \tag{A72}$$

T1 updates the reference backbone θ_{ref} and length model ϕ.

E.4.2 T2: SEMANTIC-WINDOW CONDITIONING

T2 initializes the semantic-conditioning layers from the T1 backbone and trains the window encoders, residual adapters, and diffusion-time gates. The backbone and reference-generation pathway remain fixed.

The T2 objective is

$$L_{T2} = L_{\text{den}}^{\text{ON}}, \tag{A73}$$

where the semantic windows are supplied to the ON model through H_W and the residual adapters.

E.4.3 T3: REFERENCE VALUE NETWORK

The Reference Value Network estimates the probability of eventual qualification under OFF continuation. For an intermediate state x_τ,

$$V_\tau^{\text{OFF}}(x_\tau; C, F_s) = \mathbb{E}_{X_0 \sim \text{OFF}(\cdot \mid x_\tau, \tau, C, F_s)}\left[H_{\text{qual}}(X_0; C, F_s)\right]. \tag{A74}$$

The value network receives the partial sequence, diffusion time, and global/structural conditions but no semantic-window representation.

Training states are collected from reference-model trajectories. For state ζ, let M independent OFF continuations be sampled. The empirical soft target is

$$y(\zeta) = \frac{1}{M} \sum_{m=1}^{M} H_{\text{qual}}\left(X_0^{(m)}; C, F_s\right). \tag{A75}$$

The value model is trained using binary cross-entropy:

$$L_{T3} = -\mathbb{E}_\zeta\left[y(\zeta) \log V_\psi(\zeta) + (1 - y(\zeta)) \log(1 - V_\psi(\zeta))\right]. \tag{A76}$$

ON states are labelled using independent OFF continuations from the same partial states, ensuring that the value model measures reference-defined feasibility rather than semantic-conditioned success.

T3 is trained offline and independently selected using calibration and predictive performance, including Brier score, log loss, and calibration curves.

E.5 ABSORBING-MASK REVERSE DIFFUSION

Let d_t denote the probability that a token remains masked at diffusion step t. Define the probability of remaining masked during a reverse step

$$r_t = \frac{d_{t-1}}{d_t}. \tag{A77}$$

For a candidate clean token j, the reverse transition kernel for model variant ν is

$$K_{t,j}^\nu = \begin{cases} \mathbf{1}\left[Y_j = x_{t,j}\right], & x_{t,j} \neq [\text{MASK}], \\ r_t, & x_{t,j} = [\text{MASK}], \quad Y_j = [\text{MASK}], \\ (1 - r_t) P_{\theta^\nu}(Y_j \mid x_t, C, F_s, W), & x_{t,j} = [\text{MASK}], \quad Y_j \neq [\text{MASK}]. \end{cases} \tag{A78}$$

The complete transition probability is

$$\pi_t^\nu(Y \mid x_t) = \prod_{j=1}^{L} K_{t,j}^\nu. \tag{A79}$$

The OFF model uses $\alpha = 0$, while the ON model uses $\alpha = 1$. Both models use the same diffusion schedule, token support, revealed-token constraints, and molecular length.

E.6 Reference–semantic mixture and importance correction

At each reverse step, candidate transitions are generated from both the OFF and ON models. Let

$$n_{\mathrm{OFF}} \quad \text{and} \quad n_{\mathrm{ON}} \tag{A80}$$

denote the numbers of candidate transitions contributed by the two sources, and let

$$N = n_{\mathrm{OFF}} + n_{\mathrm{ON}}. \tag{A81}$$

For mixture coefficient γ, the proposal distribution for candidate Y_i is

$$q_t(Y_i) = \gamma \pi_t^{\mathrm{OFF}}(Y_i) + (1-\gamma)\pi_t^{\mathrm{ON}}(Y_i). \tag{A82}$$

The semantic distribution is corrected for the finite-pool mixture using

$$\rho_i = \frac{\pi_t^{\mathrm{ON}}(Y_i)}{q_t(Y_i)}, \qquad p_i = \frac{\rho_i}{\sum_j \rho_j}. \tag{A83}$$

This correction prevents the final semantic proposal from being biased by the relative number of OFF and ON candidates in the finite candidate pool.

E.7 Reference feasibility bound

For each candidate Y_i, define its reference feasibility as

$$v_i = \begin{cases} V_\psi(Y_i, t-1; C, F_s), & t > 1, \\ H_{\mathrm{qual}}(Y_i; C, F_s), & t = 1. \end{cases} \tag{A84}$$

The direct OFF candidates provide the step-specific reference baseline

$$b_0 = \frac{1}{n_{\mathrm{OFF}}} \sum_{i \in I_{\mathrm{OFF}}} v_i. \tag{A85}$$

A tolerance ϵ_t is applied to this baseline:

$$\beta_t = \max\left(0, b_0 - \epsilon_t\right). \tag{A86}$$

The bound therefore requires the final candidate distribution to retain, up to the permitted tolerance, the mean reference-defined feasibility observed among direct OFF candidates.

E.8 Minimum-KL reference-bounded sampling

Let

$$\Delta_N = \left\{ w \in \mathbb{R}^N : w_i \geq 0,\ \sum_i w_i = 1 \right\} \tag{A87}$$

denote the probability simplex.

The final sampling distribution is obtained by minimizing the KL divergence from the importance-corrected semantic distribution:

$$w^* = \arg\min_{w \in \Delta_N} D_{\mathrm{KL}}(w \| p) \tag{A88}$$

subject to

$$w^\top v \geq \beta_t. \tag{A89}$$

Equivalently,

$$w^* = \arg\min_{w \in \Delta_N} \sum_i w_i \log \frac{w_i}{p_i} \quad \text{s.t.} \quad \sum_i w_i v_i \geq \beta_t. \tag{A90}$$

If

$$p^\top v \geq \beta_t, \tag{A91}$$

the constraint is already satisfied and

$$w^* = p. \tag{A92}$$

Otherwise, the active solution has the exponential-tilting form

$$w_i(\lambda) = \frac{p_i \exp(\lambda v_i)}{\sum_j p_j \exp(\lambda v_j)}, \qquad \lambda \geq 0. \tag{A93}$$

The corresponding expected feasibility is

$$\mu(\lambda) = \sum_i w_i(\lambda) v_i. \tag{A94}$$

Its derivative satisfies

$$\frac{d\mu(\lambda)}{d\lambda} = \mathrm{Var}_{w(\lambda)}(v) \geq 0. \tag{A95}$$

Therefore, when

$$\beta_t \in \left(p^\top v, \max_i v_i\right), \tag{A96}$$

the unique nonnegative λ satisfying

$$\mu(\lambda) = \beta_t \tag{A97}$$

can be obtained by one-dimensional root finding.

If

$$\beta_t = \max_i v_i, \tag{A98}$$

the limiting distribution is proportional to p_i among candidates attaining the maximum feasibility. If the constraint is inactive, no tilting occurs.

After solving the constrained optimization, the next diffusion state is sampled as

$$I \sim \mathrm{Categorical}(w^*), \qquad x_{t-1} = Y_I. \tag{A99}$$

The reverse kernels, candidate feasibility values, mixture distribution, and KL-constrained distribution are recomputed independently at every diffusion step.

E.9 Reference-feasibility relationship

Under exact reference values and exact OFF continuation, the OFF value function satisfies the recursion

$$\mathbb{E}\left[b_0 \mid X_t = x\right] = V_t^{\mathrm{OFF}}(x; C, F_s). \tag{A100}$$

The reference-bounded distribution satisfies

$$\mathbb{E}_{w^*}\left[V_{t-1}^{\mathrm{OFF}}(Y; C, F_s)\right] \geq b_0 - \epsilon_t. \tag{A101}$$

Combining these relations across reverse-diffusion steps gives the approximate recursive guarantee

$$V_t^{\mathrm{HDiff}} \geq V_t^{\mathrm{OFF}} - \sum_{k=1}^{t} \epsilon_k.$$

This relationship holds under the idealized assumptions of exact OFF transition probabilities, exact value estimates, consistent candidate support, and exact solution of the KL-constrained optimization. In practice, the effective bound depends on the calibration and approximation error of V_ψ, finite candidate-pool effects, and numerical optimization error.

E.10 DEPLOYMENT

After T1–T3 training and Semantic Field Agent optimization, all deployed model parameters are frozen. During molecular generation, the Controller supplies C, F_s, and $W_{T,s,r}$, while HDiff performs reference-bounded sampling without updating its parameters.

The OFF pathway provides the reference generation distribution and reference-defined feasibility estimate. The ON pathway introduces semantic-window information through the residual adapters. The minimum-KL procedure modifies the semantic proposal only when necessary to satisfy the reference-derived feasibility constraint. Thus, semantic conditioning can alter local molecular configurations while remaining bounded relative to the feasibility established by the reference model.

F ADDITIONAL INFORMATION ON BASELINE MODELS

We compared MSF-HDiff with five publicly available molecular generation methods, including three diffusion models and two molecular language model approaches. LDMol generates molecules from natural-language descriptions using diffusion in a structurally informative latent space (Chang & Ye, 2025). DrugDiff combines latent diffusion over SELFIES-based variational autoencoder representations with property-predictor guidance during sampling (Oestreich et al., 2025). DemoDiff conditions a graph diffusion Transformer on molecule–score demonstrations for in-context molecular design (Liu et al., 2025). MegaMolBART is a BART-based SMILES denoising autoencoder that generates molecular analogues through latent-space perturbation and decoding (NVIDIA Corporation, 2024). AutoMolDesigner combines recurrent chemical language modeling for molecular generation with automated machine learning for antibacterial activity and property prediction (Shen et al., 2024).

G STRUCTURAL RELIABILITY METRICS AND ADDITIONAL RESULTS

G.1 STRUCTURAL RELIABILITY METRICS CALCULATION METHODS

Validity (%). Validity was calculated as the proportion of all generated molecules that could be successfully parsed, sanitized, and canonicalized by RDKit ($N_{\text{valid}}/N_{\text{generated}} \times 100$).

Uniqueness (%). Uniqueness was calculated as the proportion of valid molecules corresponding to distinct canonical SMILES ($N_{\text{unique valid}}/N_{\text{valid}} \times 100$).

Novelty (%). Novelty was calculated as the proportion of unique valid molecules whose canonical SMILES were absent from the ChEMBL database ($N_{\text{novel unique valid}}/N_{\text{unique valid}} \times 100$).

Valid-Unique-Novel (VUN) Rate (%). The VUN rate was calculated as the proportion of all generated molecules that were simultaneously valid, unique, and novel ($N_{\text{novel unique valid}}/N_{\text{generated}} \times 100$).

3D Construction Success (%). 3D construction success was calculated as the proportion of molecules submitted for 3D construction for which a complete 3D molecular structure was successfully generated.

Internal-Energy Validity (%). Internal-energy validity was calculated as the proportion of successfully constructed 3D molecules that passed the defined internal-energy criterion.

End-to-End PoseBusters (PB) Validity (%). End-to-end PB validity was calculated as the proportion of all generation attempts that yielded a successfully constructed 3D structure passing the applicable PoseBusters checks (Buttenschoen et al., 2024).

Bond-Angle Validity (%). Bond-angle validity was calculated as the proportion of evaluated 3D structures that passed the bond-angle plausibility criterion.

Steric-Clash-Free Percentage (%). The steric-clash-free percentage was calculated as the proportion of evaluated 3D structures containing no detected steric clashes.

G.2 Additional Structural Reliability Results

In addition to the structural reliability metrics reported in Table 1, we also measured bond-angle validity and the percentage of steric-clash-free generated compounds for both *E. coli* and *S. aureus*, as reported in Table A4.

Table A4: Supplementary structural validity metrics for small-molecule compounds generated for different bacterial species. Metrics are reported as percentages (%), with values shown as mean ± standard deviation across three random seeds; higher values indicate better performance.

Model	Valid bond angles	Steric clash-free
Escherichia coli (E. coli)		
MSF-HDiff	100.00 ± 0.00	100.00 ± 0.00
MSF-Guided Diff	97.67 ± 0.20	98.34 ± 0.19
MSF-Guided GPT	100.00 ± 0.00	100.00 ± 0.00
GPT	100.00 ± 0.00	100.00 ± 0.00
LDMol	100.00 ± 0.00	98.61 ± 0.20
DrugDiff	99.69 ± 0.10	98.15 ± 0.20
DemoDiff	100.00 ± 0.00	99.89 ± 0.11
MegaMolBART	99.90 ± 0.10	98.95 ± 0.21
AutoMolDesigner	100.00 ± 0.00	100.00 ± 0.00
Staphylococcus aureus (S. aureus)		
MSF-HDiff	100.00 ± 0.00	100.00 ± 0.00
MSF-Guided Diff	100.00 ± 0.00	99.25 ± 0.10
MSF-Guided GPT	100.00 ± 0.00	99.90 ± 0.10
GPT	100.00 ± 0.00	100.00 ± 0.00
LDMol	100.00 ± 0.00	99.23 ± 0.11
DrugDiff	99.28 ± 0.10	98.14 ± 0.20
DemoDiff	100.00 ± 0.00	100.00 ± 0.00
MegaMolBART	100.00 ± 0.00	99.70 ± 0.10
AutoMolDesigner	100.00 ± 0.00	100.00 ± 0.00

H Molecular Property Metrics Methods

Diversity. Diversity was calculated over the selected generated molecules based on pairwise molecular similarity, with molecular similarity represented using Morgan fingerprints. Higher diversity indicates greater structural variation among the generated molecules.

Predicted Antibacterial Activity (pMIC). Antibacterial activity was quantified using predicted pMIC values. During generation and ranking, pMIC was predicted using a Chemprop directed message-passing neural network (D-MPNN), which served as the in-loop activity oracle. Higher predicted pMIC values indicate greater predicted antibacterial activity.

Independent pMIC. To provide an independent assessment of predicted antibacterial activity, pMIC was additionally estimated using a random forest model based on Morgan molecular fingerprints. This model was independent of the Chemprop D-MPNN used during generation and ranking.

Quantitative Estimate of Drug-Likeness (QED). Drug-likeness was assessed using the quantitative estimate of drug-likeness (QED). QED summarizes multiple physicochemical properties into a score ranging from 0 to 1, with higher values indicating greater drug-likeness.

Synthetic Accessibility Score (SAS). Synthetic accessibility was evaluated using the synthetic accessibility score (SAS), which estimates the relative difficulty of synthesizing a molecule based on its structural characteristics. Lower SAS values indicate greater predicted synthetic accessibility.

Toxicity. Toxicity was evaluated using ClinTox predictions from ADMET-AI (Swanson et al., 2024b). Lower values indicate lower predicted toxicity.

Ranking Score. The Ranking Score was calculated by combining the molecular-property objectives used to rank generated candidates, including predicted antibacterial activity and the considered drug-development properties. Higher Ranking Scores indicate better overall performance across the incorporated objectives.

Independent Composite Score. The independent composite score was calculated using the independent Morgan-fingerprint random forest pMIC prediction in place of the in-loop Chemprop D-MPNN prediction, together with the remaining molecular-property objectives. This score provides an evaluation of overall molecular performance that does not rely on the activity model used during generation and ranking. Higher values indicate better overall performance.

I Ablation Studies

To assess the contributions of multi-agent planning and hierarchical conditioning, we performed an ablation study on the *E. coli* task by systematically removing individual agents, global conditions (C_1, C_2), structural conditions (F_1, F_2, F_3), or the complete C and F condition sets. For each configuration, we evaluated diversity, independently predicted and oracle pMIC, QED, SAS, and predicted toxicity (Table A5). Among the component-ablation configurations, the full MSF-HDiff framework achieved the highest independent and oracle pMIC (4.32 and 4.47) and QED (0.71), and the lowest SAS (2.87), while maintaining substantial molecular diversity (0.86) and low predicted toxicity (0.14). Several ablations produced modest increases in diversity, but these gains were generally accompanied by deterioration in activity, drug-likeness, synthesizability, or predicted toxicity. This suggests that the full framework does not simply maximize chemical-space exploration, but instead constrains generation toward a productive region of chemical space while preserving meaningful structural diversity.

Agent ablations showed distinct contributions from the three planning agents. Removing the Librarian Agent reduced independent/oracle pMIC to 3.86/3.92 and QED to 0.49, consistent with its role in integrating literature evidence, reference molecules, and user examples during task initialization. Removing the Discovery Agent produced the largest activity reductions among individual agent ablations, decreasing independent/oracle pMIC to 3.68/3.19 and QED to 0.34. Removing the Semantic Field Agent yielded pMIC scores of 4.14/4.09 and QED of 0.34, together with increased predicted toxicity (0.19), consistent with its role in translating property-level guidance into executable local windows. Removing all agents further reduced independent pMIC to 3.16 and increased SAS to 3.67.

We next assessed the contribution of hierarchical conditions by independently removing C_1, C_2, F_1, F_2, and F_3, followed by removal of the complete C or F blocks. Removing either global condition reduced both pMIC estimates and QED, with QED decreasing from 0.71 for the full model to 0.61 without C_1, 0.58 without C_2, and 0.47 when both were removed. Structural-condition ablations produced distinct effects: removing F_1 increased diversity to 0.92 but reduced independent/oracle pMIC to 4.01/3.94, removing F_2 yielded QED of 0.49, and removing F_3 increased SAS to 3.33 while reducing QED to 0.58. Removing the complete F block reduced independent/oracle pMIC to 3.89/3.71 and QED to 0.37. Finally, removing both C and F produced the lowest oracle pMIC (2.98) and QED (0.29), together with the highest predicted toxicity (0.25). These results demonstrate that the agent and conditioning components contribute at distinct levels, with global constraints, structural planning, and local semantic guidance collectively shaping the resulting molecular-quality profile.

J Structural and Physicochemical Profile Analysis of Generated Molecules

To characterize the chemical space generated by MSF-HDiff, we compared its molecular composition and physicochemical properties with those of the study-specific ablations, MSF-Guided Diff and MSF-Guided GPT, and a reference set of real-world drugs (Figure A1). We quantified Jensen–Shannon (JS) distances between generated and reference distributions for elemental composition, bond composition, ring sizes, and functional groups, with lower values indicating closer agreement (Figure A1)a. We additionally evaluated the proportions of single, double, and aromatic bonds

Table A5: Component ablations of MSF-HDiff on the *E. coli* task. Diversity, independent and oracle pMIC, QED, SAS, and predicted toxicity across agent and condition ablations. "w/o" denotes removal of the indicated agent or condition. Independent denotes the independently trained Morgan-fingerprint random-forest predictor; Oracle denotes the in-loop Chemprop D-MPNN predictor. C_1 and C_2 denote the allowed element set and molecular-weight interval, respectively; F_1, F_2, and F_3 denote scaffold archetype, ring-system profile, and motif topology, respectively.

Configuration	Diversity ↑	pMIC (Independent) ↑	pMIC (Oracle) ↑	QED ↑	SAS ↓	Toxicity ↓
MSF-HDiff (full)	**0.86**	**4.32**	**4.47**	**0.71**	**2.87**	**0.14**
Agent ablations						
w/o Librarian Agent	0.89	3.86	3.92	0.49	3.01	0.17
w/o Discovery Agent	0.88	3.68	3.19	0.34	3.33	0.22
w/o Semantic Field Agent	0.89	4.14	4.09	0.34	3.04	0.19
w/o all agents	0.89	3.16	3.23	0.34	3.67	0.18
Condition ablations						
w/o C_1	0.88	4.23	4.38	0.61	3.31	0.19
w/o C_2	0.86	4.23	4.33	0.58	3.05	0.18
w/o C (C_1, C_2)	0.90	4.06	4.27	0.47	3.12	0.17
w/o F_1	0.92	4.01	3.94	0.60	3.40	0.16
w/o F_2	0.87	4.23	4.21	0.49	3.18	0.16
w/o F_3	0.86	4.22	4.38	0.58	3.33	0.14
w/o F (F_1, F_2, F_3)	0.86	3.89	3.71	0.37	3.16	0.12
w/o C and F	0.86	3.63	2.98	0.29	3.08	0.25

among heavy-atom bonds (Figure A1)b, as well as molecular weight (MW) (Figure A1)c, calculated lipophilicity (cLogP) (Figure A1)d, FractionCSP3 (Figure A1)e, rotatable bond count (Figure A1)f, aromatic-ring count (Figure A1)g, and predicted aqueous solubility (logS; $\log_{10}$(mol/L)) (Figure A1)h.

MSF-HDiff showed close agreement with the drug reference across all four structural descriptors, with JS distances of 0.15 for elemental composition, 0.08 for bond composition, 0.12 for ring sizes, and 0.14 for functional groups. The bond-composition profile further showed a distribution of single and aromatic bonds closely resembling the drug reference, consistent with the low JS distance for this descriptor. Relative to the study-specific ablations, these results indicate that MSF-HDiff more closely reproduces the overall structural composition of real-world drugs across multiple complementary molecular features.

The physicochemical analysis further revealed a focused molecular-property profile for MSF-HDiff. Generated molecules showed a molecular-weight distribution concentrated around 300 Da and moderate cLogP values. The FractionCSP3 distribution substantially overlapped with the drug reference, while aromatic-ring and rotatable-bond distributions were concentrated around two aromatic rings and a relatively small number of rotatable bonds, respectively. The generated molecules therefore exhibited a combination of saturated carbon content, aromatic structural elements, and moderate molecular flexibility. Predicted logS values were shifted toward higher aqueous solubility relative to the drug reference, indicating a distinct solubility profile despite the closer agreement in structural composition.

Compared with MSF-Guided Diff, MSF-HDiff exhibited closer compositional agreement with the drug reference while generating larger molecules with greater saturated-carbon content and a stronger preference for two aromatic rings. Together, these analyses show that MSF-HDiff generates molecules within a focused region of chemical space while maintaining close agreement with real-world drugs across elemental, bonding, ring, and functional-group composition. This structural consistency is accompanied by a coherent physicochemical profile spanning molecular size, lipophilicity, saturation, aromaticity, molecular flexibility, and aqueous solubility.

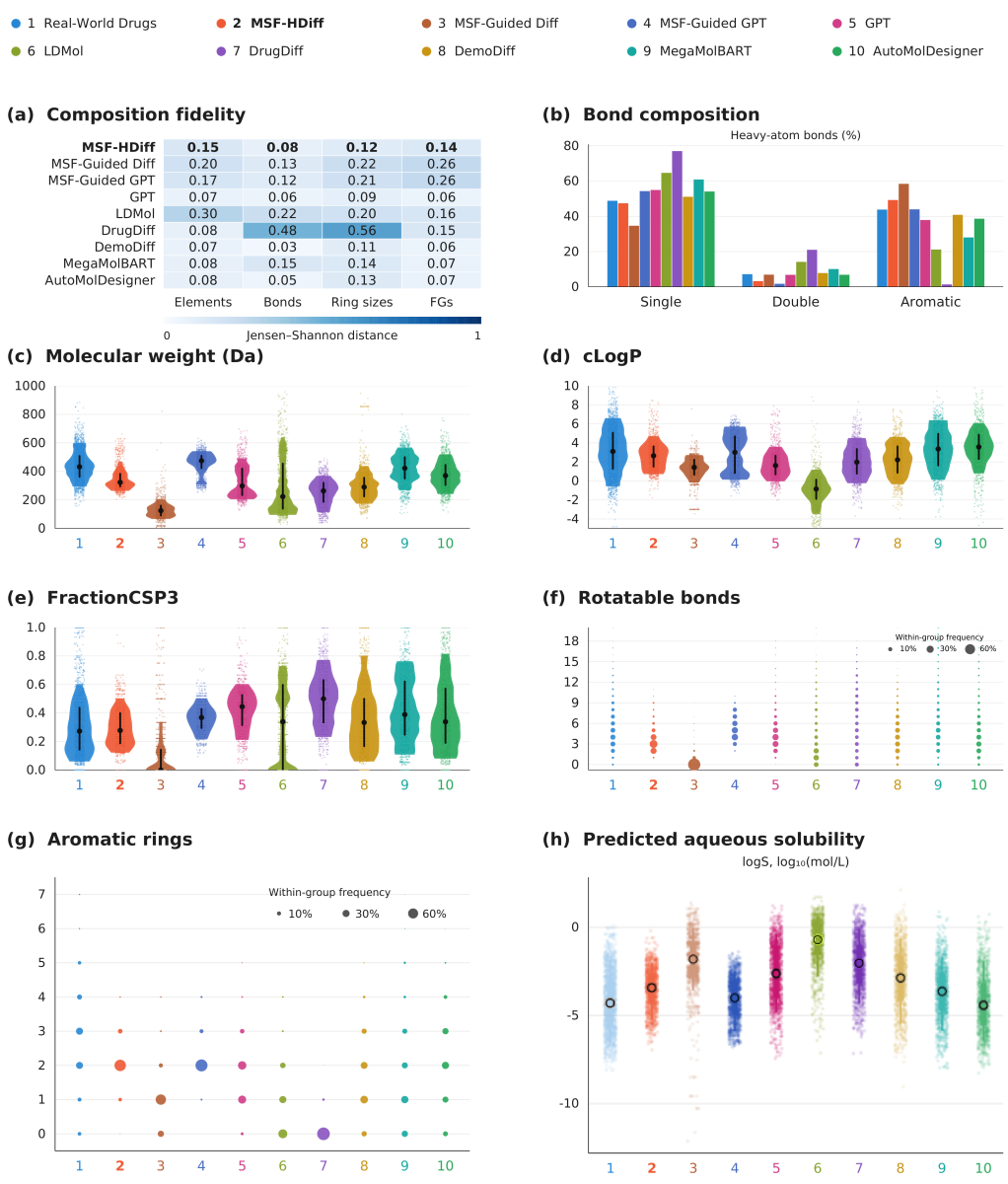

Figure A1: Composition and physicochemical profiles of generated molecules. Comparison of MSF-HDiff, study-specific ablations, and public baselines, with real-world drugs as a reference. (a) Jensen–Shannon distances from the drug reference for elemental composition, bond composition, ring sizes, and functional groups (FGs); smaller values indicate closer agreement. (b) Proportions of single, double, and aromatic bonds among heavy-atom bonds. (c) Molecular weight. (d) Calculated lipophilicity (cLogP). (e) Fraction of sp^3-hybridized carbon atoms (FractionCSP3). Violin plots in (c–e) show distribution shapes with overlaid molecular observations. (f) Rotatable-bond count. (g) Aromatic-ring count. Bubble size in (f, g) encodes within-group frequency. (h) Predicted aqueous solubility ($\log S$, $\log_{10}(\mathrm{mol/L})$), shown as molecule-level distributions. Model IDs and colors follow the common key.

K Limitations and Future Directions

While MSF-HDiff demonstrates strong computational performance, there are limitations that provide directions for future work. First, antibacterial activity, toxicity, and other molecular properties

are primarily evaluated using computational predictors and therefore do not establish experimental efficacy or safety. Future work should synthesize high-ranking candidates and validate them through in vitro antibacterial and cytotoxicity assays. Second, the current study evaluates MSF-HDiff on two bacterial species, *E. coli* and *S. aureus*, representing Gram-negative and Gram-positive bacteria, respectively, thereby assessing its performance across two distinct bacterial classes. However, evaluation was limited to these two species and did not encompass a broader range of bacterial pathogens. Future work should extend evaluation to additional clinically relevant pathogens and resistant strains to further assess the generalizability of the framework across diverse antibacterial design settings.