DIVERGENT EVOLUTION PATHS FOR MULTI-OBJECTIVE MOLECULAR OPTIMIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Multi-objective molecular optimization is fundamentally constrained not only by the vastness of chemical space, but by the intrinsic conflict among objectives: modifications that improve one property often degrade others. Existing approaches typically optimize molecules within a single evolving search context, which entangles conflicting objectives and discards the path-specific experience needed to reach distinct Pareto regions. We propose **ATOM** (**A**gents on a **T**ree for multi-**O**bjective **M**olecular optimization), a trajectory-centric framework that treats alternative optimization histories, rather than individual molecular candidates alone, as the units of Pareto exploration. Instead of enforcing global consensus, ATOM organizes specialized agents along different branches of a tree-structured search, where each agent evolves molecules toward a particular objective while accumulating trajectory-specific structural experience. To support structured knowledge sharing, ATOM incorporates three complementary coordination mechanisms operating at lateral, hierarchical, and global scales, enabling progressive cross-path optimization. Experiments on molecular optimization benchmarks involving predicted bioactivity, synthetic accessibility, and CYP3A4 non-inhibition show that ATOM attains the best mean aggregate hypervolume and competitive Pareto coverage, with gains more pronounced under severe inter-objective conflict. These results suggest that divergent evolution provides an effective paradigm for complex multi-objective molecular design. Code is available at `https://anonymous.4open.science/r/ATOM-41CE`.

1 INTRODUCTION

Multi-objective molecular optimization is a central task in early-stage drug discovery (De Rycker et al., 2018; Yang et al., 2024; Liu et al., 2025), where the goal is to refine lead compounds to simultaneously satisfy multiple, often conflicting, properties such as bioactivity, drug-likeness, and synthesizability (He et al., 2021). What makes this problem fundamentally difficult is not merely the size of chemical space, but its **path-dependent structure**: scaffold choices made in early optimization steps determine which regions of property space remain reachable downstream—introducing a heteroaromatic core to improve kinase binding, for instance, simultaneously constrains achievable lipophilicity and synthetic complexity. Traditional approaches, including high-throughput screening (Graff et al., 2021) and simulation-based methods (Hsu et al., 2017), are effective but require substantial time and computational resources, limiting scalability.

Driven by advances in machine learning, existing methods for multi-objective molecular design broadly share one structural limitation: they commit to a **single optimization trajectory**. Scalarization-based methods (Maziarka et al., 2020; Ji et al., 2021; Xia et al., 2024) encode a fixed preference ordering over objectives. With a single weighting, they typically emphasize a limited trade-off region; choosing and calibrating weights can be difficult when objectives are strongly conflicting (Xie et al., 2021; Fromer & Coley, 2023). Pareto-based approaches (Yasonik, 2020; Verhellen, 2022) apply post-hoc non-dominated sorting to large candidate pools, but the generative process itself is not trajectory-aware. Bayesian optimization and MCTS-based methods (Yang et al., 2024; Xie et al., 2021; Gao et al., 2022) plan more carefully, yet still maintain a single search frontier over molecular structures rather than over optimization strategies, limiting their ability to explore divergent trade-off regions. The consequence is systematic: **methods with limited trajectory diversity may have difficulty covering the full Pareto front under strong conflict**.

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

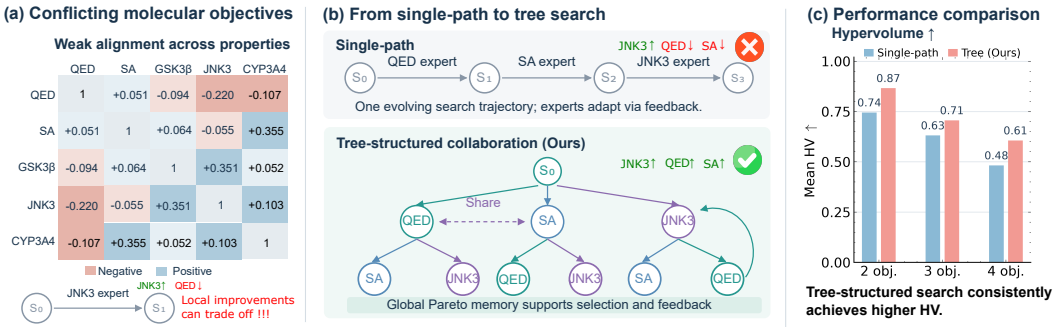

Figure 1: (a) Correlations between representative molecular properties are weak or conflicting, illustrating the intrinsic difficulty of balancing multiple objectives. (b) Conceptual comparison between the single-path baseline, instantiated by EAG in our experiments, and ATOM's tree-structured search: EAG follows one evolving trajectory, whereas ATOM coordinates specialized agents across divergent paths. (c) The resulting tree-structured search attains higher hypervolume as the number of objectives increases.

More recently, the rapid progress of large language models (Achiam et al., 2023; Bai et al., 2023; Dubey et al., 2024) has sparked growing interest in their application to molecular generation (Brahmavar et al., 2024; Wang et al., 2025) and optimization (Yu et al., 2025; Ye et al., 2025). LLMs provide a flexible framework for goal-conditioned generation and reasoning across heterogeneous molecular properties (Nguyen & Grover, 2024). However, existing LLM-based approaches formulate multi-objective optimization as a monolithic generation problem: a single agent, or a flat collection of agents, must simultaneously serve all objectives within one context, without any mechanism to explore how different prioritization strategies lead to structurally different regions of Pareto space.

In this work, we propose a different perspective: **divergent optimization paths can facilitate Pareto convergence**. We propose **ATOM** (**A**gents on a **T**ree for multi-**O**bjective **M**olecular optimization), a multi-agent framework that formulates molecular optimization as a tree-structured search. Each node corresponds to an atomic-level operation on a molecular population and hosts an agent specialized for a particular objective. Agents coordinate *pathwise* along different branches rather than enforcing global consensus, enabling explicit comparison of alternative molecular evolution trajectories. To support long-horizon coordination, ATOM incorporates a global Pareto memory that aggregates non-dominated candidates across all paths, supplying each agent with chemically diverse reference structures while preserving specialization.

In summary, our contributions are as follows:

(i) We introduce a pathwise, tree-structured formulation of multi-objective molecular optimization that explicitly models alternative molecular evolution trajectories, designed to support complementary Pareto exploration through trajectory-specific divergence.
(ii) We propose ATOM, a multi-agent framework in which specialized agents coordinate along different paths rather than collapsing into a single global policy, supported by three complementary knowledge-exchange mechanisms operating at lateral, hierarchical, and global scales.
(iii) We demonstrate empirically that ATOM attains the best mean aggregate hypervolume on the evaluated multi-objective benchmarks, with performance advantages that are larger under higher inter-objective conflict.

2 Related Work

2.1 Molecular Optimization

Molecular optimization is a core problem in drug discovery and materials science, and has gradually shifted from manual experimentation to data-driven computational methods (Gao et al., 2022). Existing approaches can be broadly categorized into two classes. (1) **Combinatorial Optimization.** Traditional approaches treat molecular design as a search problem over a discrete, exponentially large chemical space (Bohacek et al., 1996; Stumpfe & Bajorath, 2012). Common techniques include Monte

Carlo Tree Search (MCTS) (Yang et al., 2023), Genetic Algorithms (GA) (Jensen, 2019; Fu et al., 2022), and Reinforcement Learning (RL) (Bohacek et al., 1996; Stumpfe & Bajorath, 2012). While these methods explore the structural space iteratively, they often struggle with high-dimensional search landscapes and the prohibitive computational cost of evaluating complex objectives. (2) **Generative Models in Molecular Design.** Recent research has moved toward generative modeling (Du et al., 2024), learning the implicit probability distribution of chemical data to propose valid molecular candidates across a variety of architectural paradigms (Gómez-Bombarelli et al., 2018; Jin et al., 2018; Guimaraes et al., 2017; Shi et al., 2020; Hoogeboom et al., 2022; Schneuing et al., 2024).

2.2 LLMs and Agents for Molecular Optimization

Large language models provide a flexible interface for molecular design: they can interpret SMILES strings, follow property-specific instructions, and invoke external evaluators. In our experiments, GPT-4o-mini (Achiam et al., 2023) and Llama-3-8B (Dubey et al., 2024) serve as direct LLM baselines for instruction-conditioned molecular generation. Molleo employs LLM-guided evolutionary operators for molecular optimization (Wang et al., 2024). Agent-based methods extend this setting by assigning optimization and tool-use roles to multiple components. DrugAssist (Ye et al., 2025) formulates molecular optimization as an interactive assistant workflow, while EAG (Gu et al., 2025) coordinates a sequence of specialized agents. We also evaluate MAMO and Biomni as agent-based molecular design systems under the same oracle budget. Across these method families, the search state remains primarily candidate-centric. Multiple objectives may be evaluated jointly, and different agents may sequentially modify the same population, but alternative optimization histories are typically merged into a common context or compressed into a single search frontier. This distinction matters in molecular optimization because the effect of a structural edit depends on the sequence of earlier edits: two populations with similar current scores may admit substantially different future improvements.

ATOM instead adopts a trajectory-centric formulation. It maintains multiple path-conditioned search states, each representing a distinct sequence of objective-specialized transformations, and allocates future evaluations according to their Pareto potential. The tree therefore represents alternative optimization histories rather than merely scheduling agent calls. Knowledge is exchanged across these histories without collapsing them into a single optimization policy.

3 Preliminary

We study the multi-objective molecular optimization problem over the discrete space of chemically valid molecules $\mathcal{X}$, where each molecule $x \in \mathcal{X}$ is represented by a SMILES string and associated with a K-dimensional objective vector $\mathbf{f}(x) = [f_1(x), \ldots, f_K(x)]^\top$. The goal is to find a set of candidate molecules that jointly maximize $\mathbf{f}(x)$:

$$\max_{x \in \mathcal{X}} \mathbf{f}(x). \tag{1}$$

Due to conflicting objectives, we aim to identify a diverse set of trade-off solutions.

A vector Y is said to *dominate* X, denoted $Y \succ X$, if $Y_k \geq X_k$ for all $k \in \{1, \ldots, K\}$ and $Y_{k'} > X_{k'}$ for at least one k'. A solution X is *non-dominated* if no other solution dominates it.

The *first Pareto front* $\mathcal{S}_1$ consists of all non-dominated solutions:

$$\mathcal{S}_1 = \{X \in \mathcal{S} \mid \nexists Y \in \mathcal{S} \text{ s.t. } Y \succ X\}. \tag{2}$$

Subsequent Pareto fronts are defined recursively by removing solutions from the preceding fronts:

$$\mathcal{S}_k = \left\{ X \in \mathcal{S} \setminus \bigcup_{i=1}^{k-1} \mathcal{S}_i \;\middle|\; \nexists Y \in \mathcal{S} \setminus \bigcup_{i=1}^{k-1} \mathcal{S}_i \text{ s.t. } Y \succ X \right\}. \tag{3}$$

3.1 Path-Conditioned Search States

Conventional methods represent search primarily by the current molecular population. ATOM instead defines a node as $N = (P_N, h_N)$, where h_N records the ordered specialists, inherited references, and

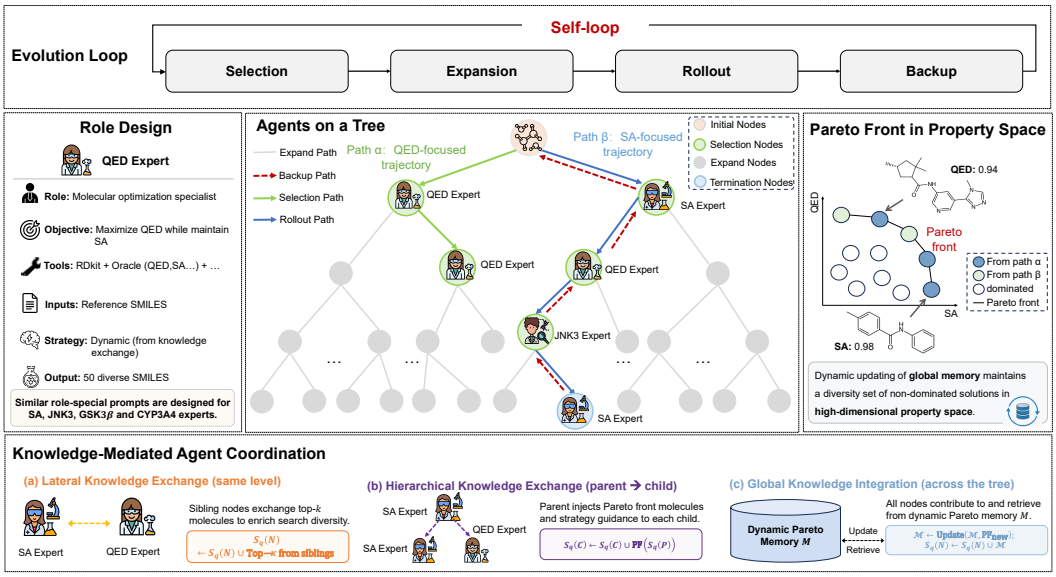

Figure 2: Agents-on-a-Tree framework for multi-objective molecular optimization, combining pathwise MCTS planning with knowledge-mediated coordination among specialized agents to improve Pareto coverage under conflicting objectives.

distilled strategies along its root-to-node path. Specialist q induces the path-conditioned transition $P_{N'} \sim A_q(\cdot \mid P_N, h_N, \mathcal{M})$. Thus, nodes with similar populations but different histories remain distinct because they imply different continuations. Under a fixed oracle budget, MCTS allocates evaluations across these trajectory hypotheses to maximize the hypervolume of their joint Pareto front, separating path-specific search states from globally reusable molecular evidence.

4 METHOD

4.1 ALGORITHMIC FRAMEWORK OF ATOM

Multi-Agent Attribute-Specific Optimization. ATOM instantiates the trajectory-centric formulation as a dynamically constructed search tree. Each root-to-node path represents an optimization hypothesis: an ordered sequence of objective-specialized transformations together with the structural experience accumulated along that sequence. Branching preserves counterfactual futures that would be lost if all specialists continually updated a shared population, while MCTS allocates oracle evaluations among these histories according to their observed Pareto potential. As illustrated in Figure 2, this search is implemented by a collection of expert agents within the proposed ATOM framework, where each agent is explicitly specialized for optimizing a particular molecular property. Each expert is instantiated as an LLM, such as GPT-4o mini (Achiam et al., 2023), and assigned a well-defined domain role corresponding to a specific optimization objective. Concretely, these roles include a QED expert, an SA expert, a CYP3A4 expert and target-specific experts for GSK3β and JNK3, which are closely associated with Alzheimer's disease.

Recent studies have shown that LLM performance on biochemical and molecular reasoning tasks can be substantially enhanced through domain-aware prompt engineering (Luo et al., 2025; Li et al., 2025). Motivated by these findings, ATOM adopts expert-specific prompt templates that explicitly specify the task scope, optimization objective, input representation, and expected output format for each agent. This design enforces functional disentanglement across experts while enabling focused and interpretable decision-making. Each expert agent is additionally equipped with tool-calling capabilities, enabling on-demand access to domain-specific tools such as RDKit and learned oracles for molecular property evaluation. Detailed descriptions of the tool interfaces and invocation protocols are provided in Appendix A.

MCTS Abstractions: States, Actions, and Expansion. We formalize molecular optimization as an MCTS over a dynamically constructed tree. A *node* N encodes a path-conditioned optimization *state*

as the pair (P_N, h_N), where $P_N = \{x_1, \ldots, x_m\} \subset \mathcal{X}$, and each x_i is a valid SMILES string. The history h_N records the ordered specialists, inherited references, and distilled strategies along the root-to-node path. An *action* at N corresponds to invoking a specialized agent A_k, which generates a new candidate population by conditioning on (i) an objective-specific optimization strategy accumulated along the current path, and (ii) reference molecules drawn from the parent population, sibling nodes, and the global Pareto-front memory (Section 4.2):

$$P_{N'} = A_k(P_N,\ \mathrm{Strategy}_k,\ \mathcal{M}_{\mathrm{ref}})\,. \tag{4}$$

Each agent generates exactly 50 SMILES per invocation. Upon *expansion*, node N instantiates K child nodes $\{N'_1, \ldots, N'_K\}$—one per objective—so the branching factor scales naturally with task complexity. This LLM generation step maps directly onto the **expansion** phase (creating K children) and the **rollout** phase (evaluating the generated molecules via oracle calls to update each child's Pareto front).

Local Hypervolume Reward. A fixed weighted sum can obscure multi-objective trade-offs by favoring particular preferences. Instead, we score each newly evaluated node by the normalized hypervolume of its local Pareto front:

$$H(N) = \widehat{\mathrm{HV}}(N) = \frac{\lambda_K\left(\bigcup_{\mathbf{y}\in\mathcal{P}(P_N)}[\mathbf{r}^-,\ \mathbf{y}]\right)}{\prod_{k=1}^{K}(r_k^+ - r_k^-)}, \tag{5}$$

where $\mathcal{P}(P_N)$ is the non-dominated objective set of P_N. Since all objectives are normalized to $[0, 1]$ (Appendix B), we fix $\mathbf{r}^- = \mathbf{0}$ and $\mathbf{r}^+ = \mathbf{1}$, giving

$$H(N) = \lambda_K\left(\bigcup_{\mathbf{y}\in\mathcal{P}(P_N)}[\mathbf{0},\ \mathbf{y}]\right) \in [0, 1]. \tag{6}$$

Thus, $H(N)$ is an immediate, objective-weight-free reward, not the running search estimate used for selection. Empty populations have $H(N) = 0$. Each unvisited child is expanded once before applying UCT to visited children, so no leaf-value prior is required.

UCT Selection and Value Backup. Let $n(N)$ be the visit count and $Q(N)$ the empirical search value of node N. On first expansion, set $n(N) = 1$ and $Q(N) = H(N)$. Thereafter, visited children are selected by

$$\mathrm{UCT}(N) = \underbrace{Q(N)}_{\text{exploitation}} + \underbrace{\sqrt{\frac{2\log n(\mathrm{parent}(N))}{n(N)}}}_{\text{exploration}}. \tag{7}$$

The UCB1 term uses a fixed coefficient because the local rewards, and hence their running averages, lie in $[0, 1]$. After expanding the selected node, let $\mathcal{C}$ be its nonempty set of newly evaluated children and $R_{\mathcal{C}} = |\mathcal{C}|^{-1}\sum_{C\in\mathcal{C}} H(C)$ their mean local-HV reward. For each previously visited node N on the selected path, back up this batch reward using

$$Q(N) \leftarrow \frac{n(N)Q(N) + R_{\mathcal{C}}}{n(N) + 1}, \qquad n(N) \leftarrow n(N) + 1. \tag{8}$$

This separates a node's directly measured local HV, $H(N)$, from the trajectory-level running estimate, $Q(N)$, used by UCT. No smoothing coefficient or scalarization weight is introduced.

4.2 Knowledge-Mediated Agent Coordination

Independent trajectories preserve alternative optimization hypotheses, but complete isolation would waste useful discoveries. ATOM therefore introduces a topology-aware coordination mechanism to enable structured information sharing among specialized expert agents within the MCTS framework. Rather than enforcing a single global consensus, the coordination mechanism operates at three complementary levels—lateral exchange within the same tree depth, hierarchical propagation across parent-child nodes, and global integration via a shared Pareto memory—so that each agent can leverage collective knowledge while retaining its specialized optimization focus. The overall coordination pipeline is illustrated in Figure 2.

Lateral Knowledge Exchange. At the same tree depth, the agent identifies the most deficient property in the current node population and imports top-ranked molecules from sibling nodes to compensate. For node N at depth d, the most deficient non-lead property is

$$k^* = \arg \min_{k \in \{1,\dots,K\} \setminus \{q\}} \bar{f}_k(P_N), \tag{9}$$

where q denotes the agent's primary objective. The effective candidate set for agent A_q is then

$$\widetilde{S}_q(N) = P_N \cup \bigcup_{N' \in \mathcal{N}_d \setminus \{N\}} S_{k^*}(N'), \tag{10}$$

where $S_{k^*}(N')$ denotes the top-κ molecules from sibling node N' ranked by f_{k^*}. This formulation requires no hand-tuned satisfaction thresholds: the weakest property is identified adaptively at each node.

Hierarchical Knowledge Propagation. To bias expansion toward promising regions, parent nodes propagate both high-quality molecules and distilled optimization strategies to their children. Given a parent node N_p and its Pareto front $\mathcal{P}(P_{N_p})$, the top-r non-dominated molecules ranked by the child agent's responsible objective are injected:

$$\Pi(N_p \to N_c) = \text{Top-}r\big(\mathcal{P}(P_{N_p})\big), \qquad P_{N_c} \leftarrow P_{N_c} \cup \Pi(N_p \to N_c). \tag{11}$$

The parent also propagates the optimization strategy accumulated along the current search path, incorporated into the child agents' prompt templates via the `Strategy` field (Appendix A). This combined transfer of structures and strategies enables child nodes to inherit both favorable structural motifs and actionable optimization knowledge.

Global Knowledge Integration. We maintain a dynamic global memory $\mathcal{M}$ that aggregates the most promising molecules discovered across all search trajectories. Unlike conventional top-M selection, $\mathcal{M}$ retains exclusively the Pareto-optimal molecules with respect to all K objectives:

$$\mathcal{M} \leftarrow \text{ParetoFront}(\mathcal{M} \cup \{(m, \mathbf{f}(m))\}). \tag{12}$$

During optimization, each agent uniformly samples up to R molecules from $\mathcal{M}$ and adds them to its reference pool alongside the other reference molecules:

$$\widetilde{S}_q(N) \leftarrow \widetilde{S}_q(N) \cup \text{RandSample}(\mathcal{M}, R). \tag{13}$$

Pareto-front retention preserves diverse trade-offs without favoring any single objective, while random sampling exposes under-explored branches to globally non-dominated structures. Operationally, lateral exchange repairs local population deficiencies with property-targeted examples, hierarchical propagation transfers promising structures and path-specific strategies to warm-start downstream search, and global integration supplies Pareto anchors that calibrate optimization across the tree. Together, hierarchical propagation preserves path-consistent continuation, lateral exchange repairs weak objectives across peer branches, and global Pareto memory reuses non-dominated evidence without merging trajectory identities.

5 Experiments

Implementation Details. We analyze pairwise correlations of five molecular properties (QED, SA, GSK3β, JNK3, CYP3A4) on the ZINC20 (Irwin et al., 2020) dataset; as shown in Figure 1(a), pairwise correlations vary in sign and magnitude, indicating heterogeneous objective alignment and potential trade-offs. To ensure a fair comparison across all methods, we use a maximum oracle evaluation budget of 10,000 calls per task. We evaluate eight settings spanning two to four objectives and increasing conflict, covering drug-likeness, synthesizability, kinase inhibition, and CYP3A4 non-inhibition. Exact task definitions and oracle implementations are provided in Appendix B.

Benchmarks and Evaluation Metrics. All molecular property scores are computed via the oracle functions provided by the Therapeutics Data Commons (TDC) (Huang et al., 2021). We adopt the hypervolume (HV) indicator (Zitzler et al., 2003) as our primary metric, which jointly characterizes both Pareto convergence and distributional diversity. We further report novelty and diversity for holistic assessment.

Baselines. We compare ATOM with representative generative/RL, evolutionary, surrogate-based, LLM, and agent-based methods (Table 1). Appendix C details the implementations and unified evaluation protocol.

Table 1: Performance comparison on multi-objective molecular optimization tasks. Higher values indicate better performance. Values are reported as mean ± standard deviation over five independent random seeds for methods evaluated in this work.

Method	QED +SA	JNK3 +GSK3β	QED+SA +JNK3	QED+SA +GSK3β	QED+JNK3 +GSK3β	QED+SA +CYP3A4	JNK3+GSK3β +CYP3A4	QED+SA+JNK3 +GSK3β	Sum↑
Generative and reinforcement-learning baselines									
SMILES LSTM	0.925±0.004	0.592±0.035	0.512±0.025	0.706±0.037	0.445±0.061	0.864±0.013	0.523±0.011	0.438±0.031	5.005
REINVENT	0.863±0.001	0.783±0.025	0.195±0.045	0.200±0.017	0.226±0.021	0.798±0.022	0.496±0.015	0.167±0.049	3.728
RationaleRL	0.692±0.002	0.728±0.032	0.534±0.019	0.637±0.027	0.570±0.126	0.659±0.013	0.592±0.032	0.547±0.059	4.959
Evolutionary and representation-space search baselines									
SMILES GA	0.885±0.003	0.706±0.014	0.510±0.007	0.655±0.070	0.609±0.021	0.703±0.015	0.594±0.039	0.424±0.034	5.086
GRAPH GA	0.899±0.002	0.725±0.059	0.582±0.072	0.704±0.035	0.610±0.071	0.725±0.024	**0.645**±0.013	0.449±0.096	5.339
STONED	0.876±0.001	0.689±0.023	0.497±0.035	0.623±0.027	0.603±0.037	0.657±0.036	0.544±0.027	0.411±0.041	4.900
Surrogate- and sampling-based baselines									
GP BO	0.843±0.003	0.783±0.043	0.505±0.045	0.629±0.030	0.620±0.033	0.752±0.028	0.633±0.043	0.422±0.101	5.187
qNEHVI	0.840±0.007	0.678±0.025	0.561±0.023	0.692±0.031	0.532±0.026	0.796±0.031	0.617±0.039	0.462±0.059	5.178
MARS	0.919±0.005	0.609±0.022	0.527±0.008	0.665±0.021	0.478±0.013	0.807±0.011	0.504±0.022	0.434±0.017	4.943
LLM baselines									
GPT-4o-mini	0.880±0.001	0.621±0.033	0.583±0.025	0.671±0.024	0.499±0.013	0.854±0.012	0.556±0.027	0.433±0.073	5.097
Llama-3-8B	0.893±0.001	0.611±0.009	0.586±0.031	0.623±0.023	0.449±0.025	0.693±0.017	0.477±0.031	0.411±0.011	4.743
Molleo	0.927±0.004	0.694±0.010	0.602±0.019	0.760±0.026	0.610±0.010	0.925±0.001	0.643±0.031	0.543±0.013	5.704
Agent-based baselines									
Biomni	0.9273±0.003	0.6597±0.041	0.5030±0.029	0.6952±0.026	0.5356±0.018	0.7822±0.031	0.5027±0.019	0.4111±0.057	5.017
MAMO	0.927±0.003	0.677±0.014	0.601±0.022	0.758±0.028	0.597±0.015	**0.936**±0.006	0.615±0.035	0.536±0.018	5.647
DrugAssist	0.874±0.002	0.645±0.011	0.581±0.029	0.692±0.047	0.501±0.021	0.797±0.013	0.618±0.022	0.484±0.061	5.192
EAG	0.895±0.001	0.595±0.021	0.551±0.037	0.711±0.061	0.519±0.049	0.866±0.032	0.507±0.029	0.482±0.071	5.126
Tree-structured agent method									
ATOM	**0.941**±0.001	**0.790**±0.017	**0.608**±0.045	**0.764**±0.051	**0.644**±0.062	0.888±0.031	0.626±0.014	**0.606**±0.041	**5.867**

5.1 RESULTS

Performance Comparison on Multi-Objective Tasks. Table 1 reports HV scores across 8 tasks. ATOM attains the best mean aggregate HV sum (5.867), exceeding Molleo (5.704), the best-performing external baseline by mean aggregate HV. ATOM attains the best mean HV on six of the eight tasks (Tasks 1–5 and 8). MAMO is best on QED+SA+CYP3A4 (Task 6), while GRAPH GA performs best on JNK3+GSK3β+CYP3A4 (Task 7). This comparison shows that ATOM is broadly competitive across the evaluated objective combinations while leaving two ADMET-oriented settings where other methods perform better. On the four-objective QED+SA+JNK3+GSK3β task, REINVENT trails ATOM in HV (0.167 vs. 0.606), a within-task comparison. Similarly, SMILES LSTM and STONED struggle on tasks involving biological targets or ADMET constraints, reflecting their limited capacity to jointly satisfy conflicting objectives. Among LLM-based approaches, single-agent GPT-4o-mini (5.097) and multi-agent EAG (5.126) both have lower aggregate mean HV than ATOM, particularly on the most challenging four-objective task QED+SA+JNK3+GSK3β (0.606 vs. EAG's 0.482). Post-hoc medicinal-chemistry screens on three archived task settings further show high PAINS-free rates, substantial Lipinski compliance, low synthetic-accessibility scores, and appreciable structural diversity (Appendix I). Appendix G provides a detailed analysis of top-ranked molecules, while Appendix H evaluates the diversity and novelty of generated molecules.Property-distribution comparisons are provided in Appendix H.

Multi-Objective Results and Sample Efficiency Analysis. As illustrated in Figure 12, the agent-based methods ATOM and MAMO concentrate more probability mass in jointly favorable regions of the GSK3β, JNK3, QED, and SA distributions than the non-agent baselines. This pattern suggests that agent-based coordination better balances the four competing molecular objectives. Figure 3 compares cumulative HV as a function of oracle calls across three tasks of increasing difficulty. ATOM exhibits characteristic stepwise HV improvements: rather than repeatedly refining a single incumbent, its tree-structured search maintains alternative molecular trajectories and selectively revisits branches with remaining potential. When one branch reaches a local trade-off plateau, exploration of other branches can identify a molecule that improves the global Pareto set, producing a discrete increase in HV. This behavior is particularly evident on the higher-dimensional tasks, where it enables continued progress after flat intervals and keeps ATOM competitive under the same oracle budget. Detailed wall-clock time and external LLM generation-request counts under the same oracle budget are provided in Appendix N.

5.2 Ablation Study

Does tree structure matter beyond flat coordination? Table 2 first separates trajectory structure from agent count and shared memory. The gap between *single-agent GPT-4o-mini* and *w/o All* provides a modest indication that MCTS-based trajectory search helps even without knowledge exchange (+0.013 on JNK3+GSK3β). Flat+memory obtains HVs of 0.679, 0.590, and 0.502, whereas Full ATOM obtains 0.790, 0.644, and 0.606. Thus, matching the agent set and providing shared Pareto memory is insufficient to recover the path-conditioned tree search.

Table 2: Ablation on MCTS-based search, tree structure, and multi-agent coordination.

	Variant	JNK3 +GSK3β	QED+JNK3 +GSK3β	QED+SA+JNK3 +GSK3β
	GPT-4o-mini	0.621±0.033	0.499±0.013	0.433±0.073
	Flat+memory	0.679±0.027	0.590±0.057	0.502±0.046
+MCTS	w/o All	0.634±0.027	0.518±0.037	0.530±0.047
+MCTS	Only G	0.639±0.065	0.571±0.013	0.548±0.052
+MCTS	Only L	0.689±0.037	0.558±0.035	0.564±0.035
+MCTS	Only H	0.679±0.013	0.609±0.045	0.569±0.045
	Full ATOM	**0.790±0.017**	**0.644±0.062**	**0.606±0.041**

The single-scope variants further show that no exchange scope alone matches the complete topology-aware design. Lateral-only is strongest on JNK3+GSK3β, whereas hierarchical-only is strongest on QED+JNK3+GSK3β; their task-dependent ranking is consistent with complementary rather than interchangeable roles. *Full ATOM* achieves the best HV on all three tasks, including an absolute gain of +0.169 over single-agent GPT-4o-mini on JNK3+GSK3β.

Does trajectory identity matter? We ablate path-specific history on the four-objective task. *Path-reset* removes ancestor strategies and parent-derived references, retaining current-population, sibling, and global-memory inputs. *Path-shuffle* exchanges histories among nodes with the same depth and specialist role, whereas *Random-path* samples history from a matched random trajectory while preserving prompt length and reference count. These controls probe path consistency by disrupting different components of trajectory history.

Table 3 reports mean ± standard deviation over five matched random seeds. These controls probe trajectory identity: Path-reset removes continuity, whereas Path-shuffle and Random-path preserve the amount of historical information but break its correspondence with the current branch. Full ATOM improves terminal HV from 0.540–0.561 to 0.606 and 4D coverage from 17.7–20.3 to 27.8 under identical oracle budgets and the same nominal retrieval limits. The gains over the history-shuffling controls support the value of aligning historical information with the current branch, beyond merely supplying historical context. The pairwise union-Pareto analysis additionally reports seed-wise exclusive contributions from Full ATOM and each ablation, showing that path-consistent histories recover complementary trade-off solutions not obtained by the corresponding controls.

Table 3: Path-history ablation on the four-objective task. Higher is better. 4D cov. counts occupied cells in a $5^4 = 625$ grid over normalized QED, SA, JNK3, and GSK3β scores of the final Pareto front (including initial molecules). Full-only PF is Full ATOM's share of pairwise union-Pareto points unique to it.

Method	Final HV↑	4D cov.↑	Full-only PF↑
Full ATOM	**0.606±0.041**	**27.8±3.1**	–
Path-reset	0.540±0.012	17.7±3.2	44.1%±12.9%
Path-shuffle	0.555±0.030	19.0±3.5	27.5%±6.9%
Random-path	0.561±0.002	20.3±3.2	30.8%±5.1%

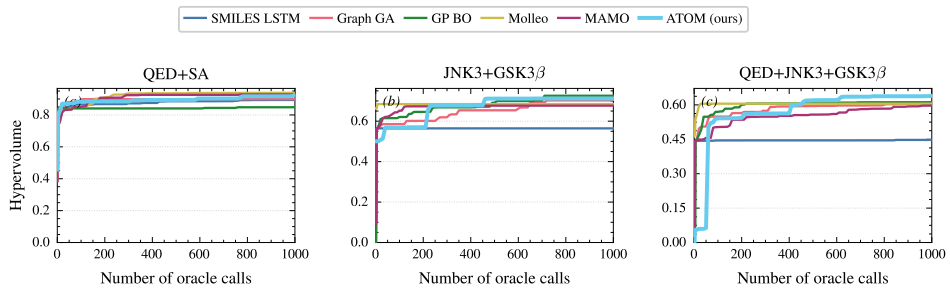

Figure 3: Cumulative HV curves versus oracle calls for three multi-objective tasks.

6 CASE STUDIES

6.1 DUAL-TARGET CANDIDATE CASE STUDY AND OPTIMIZATION PATHS

Figure 4 examines two representative ATOM candidates from the joint GSK3β (PDB: 6Y9S) and JNK3 (PDB: 4WHZ) task. We used CB-Dock (Liu et al., 2020), a cavity-detection-guided blind docking server, to explore their potential binding modes without manually specifying an active site. The resulting poses place both candidates in predicted ATP-binding pockets, with hinge-region contacts and hydrophobic packing characteristic of kinase ligands. For Compound 1, the docking scores are -9.2 and -9.5 kcal/mol for GSK3β and JNK3, respectively. Its poses suggest hinge contacts with VAL135 in GSK3β and MET149 in JNK3, together with hydrophobic contacts near VAL70/LEU132 and ILE70/VAL78, respectively. Compound 2 obtains corresponding scores of -8.8 and -9.0 kcal/mol; its predicted poses retain the hinge-region interaction pattern but differ in peripheral packing. The two molecules also exhibit distinct oracle-score profiles: Compound 1 has QED/SA scores of $0.80/0.89$ and predicted JNK3/GSK3β scores of $0.68/0.78$, whereas Compound 2 has $0.73/0.89$ and $0.65/0.86$, respectively. Thus, the case study captures two related structural solutions with different target-score balances rather than a single uniformly dominant candidate. The docking results offer structural hypotheses for these model-selected molecules, not experimental evidence of dual-target activity.

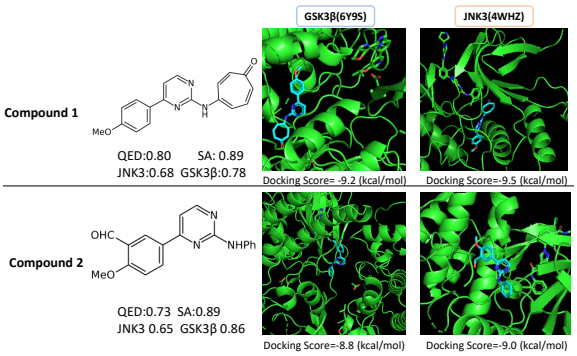

Figure 4: Representative ATOM molecules for the JNK3–GSK3β target pair.

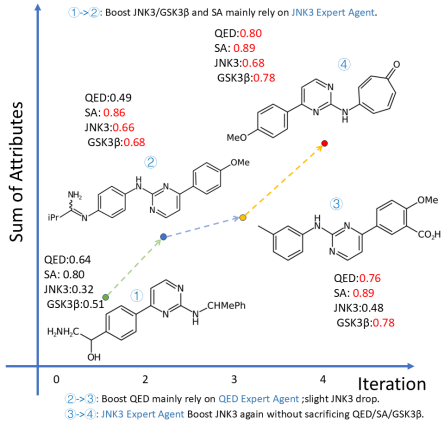

Figure 5: An ATOM optimization trajectory.

Figure 5 traces a structure-aware optimization path from an initial heteroaryl-amine scaffold. Early edits reorganize the polar amino-alcohol side chain into a more constrained amide-like motif and adjust peripheral substituents to change polar-group geometry, coinciding with higher oracle-predicted scores for both targets without a large increase in structural complexity. Subsequent modifications reduce highly basic functionality and introduce a polar acidic group, shifting the balance of polarity, lipophilicity, and chemical feasibility. A final JNK3-focused adjustment to the terminal aromatic group further raises the predicted JNK3 score while retaining favorable QED and SA scores. The sequence illustrates how path-conditioned guidance can preserve a productive scaffold, revisit objective trade-offs after each edit, and improve a selected target through localized functional-group changes rather than a coarse scaffold replacement.

7 CONCLUSION

We introduced ATOM, a trajectory-centric framework that coordinates objective-specialized agents along divergent tree paths while sharing information through global memory. This design preserves path-conditioned states across successive molecular transformations and supports exploration of conflicting trade-offs. Across eight benchmarks, ATOM achieves the best aggregate mean HV and the highest mean HV on six tasks. Matched ablations also show improved Pareto coverage, supporting trajectory-aware coordination for multi-objective molecular optimization.

AI USE STATEMENT

Large language models were used solely to assist with language polishing and improving the clarity of the manuscript. They were not used to generate research ideas, develop the methodology, conduct experiments, analyze results, or make scientific claims. All content was reviewed and revised by the authors, who take full responsibility for this submission.

REFERENCES

Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023.

Jinze Bai, Shuai Bai, Yunfei Chu, Zeyu Cui, Kai Dang, Xiaodong Deng, Yang Fan, Wenbin Ge, Yu Han, Fei Huang, et al. Qwen technical report. *arXiv preprint arXiv:2309.16609*, 2023.

Regine S Bohacek, Colin McMartin, and Wayne C Guida. The art and practice of structure-based drug design: a molecular modeling perspective. *Medicinal research reviews*, 16(1):3–50, 1996.

Shreyas Bhat Brahmavar, Ashwin Srinivasan, Tirtharaj Dash, Sowmya Ramaswamy Krishnan, Lovekesh Vig, Arijit Roy, and Raviprasad Aduri. Generating novel leads for drug discovery using llms with logical feedback. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 21–29, 2024.

Manu De Rycker, Beatriz Baragaña, Suzanne L Duce, and Ian H Gilbert. Challenges and recent progress in drug discovery for tropical diseases. *Nature*, 559(7715):498–506, 2018.

Yuanqi Du, Arian R Jamasb, Jeff Guo, Tianfan Fu, Charles Harris, Yingheng Wang, Chenru Duan, Pietro Liò, Philippe Schwaller, and Tom L Blundell. Machine learning-aided generative molecular design. *Nature Machine Intelligence*, 6(6):589–604, 2024.

Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Amy Yang, Angela Fan, et al. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*, 2024.

Jenna C Fromer and Connor W Coley. Computer-aided multi-objective optimization in small molecule discovery. *Patterns*, 4(2), 2023.

Tianfan Fu, Wenhao Gao, Connor Coley, and Jimeng Sun. Reinforced genetic algorithm for structure-based drug design. *Advances in Neural Information Processing Systems*, 35:12325–12338, 2022.

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

Rafael Gómez-Bombarelli, Jennifer N Wei, David Duvenaud, José Miguel Hernández-Lobato, Benjamín Sánchez-Lengeling, Dennis Sheberla, Jorge Aguilera-Iparraguirre, Timothy D Hirzel, Ryan P Adams, and Alán Aspuru-Guzik. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Central Science*, 4(2):268–276, 2018.

David E Graff, Eugene I Shakhnovich, and Connor W Coley. Accelerating high-throughput virtual screening through molecular pool-based active learning. *Chemical Science*, 12(22):7866–7881, 2021.

Wenyuan Gu, Jiale Han, Haowen Wang, Xiang Li, and Bo Cheng. Explain-analyze-generate: A sequential multi-agent collaboration method for complex reasoning. In *Proceedings of the 31st International Conference on Computational Linguistics*, pp. 7127–7140, 2025.

Gabriel Lima Guimaraes, Benjamin Sanchez-Lengeling, Carlos Outeiral, Pedro Luis Cunha Farias, and Alán Aspuru-Guzik. Objective-reinforced generative adversarial networks (organ) for sequence generation models. *arXiv preprint arXiv:1705.10843*, 2017.

Jiazhen He, Huifang You, Emil Sandström, Eva Nittinger, Esben Jannik Bjerrum, Christian Tyrchan, Werngard Czechtizky, and Ola Engkvist. Molecular optimization by capturing chemist's intuition using deep neural networks. *Journal of cheminformatics*, 13(1):26, 2021.

Emiel Hoogeboom, Vıctor Garcia Satorras, Clément Vignac, and Max Welling. Equivariant diffusion for molecule generation in 3d. In *International conference on machine learning*, pp. 8867–8887. PMLR, 2022.

Hui-Hui Hsu, Yen-Chao Hsu, Li-Jen Chang, and Jinn-Moon Yang. An integrated approach with new strategies for qsar models and lead optimization. *BMC Genomics*, 18:1–9, 2017.

Kexin Huang, Tianfan Fu, Wenhao Gao, Yue Zhao, Yusuf Roohani, Jure Leskovec, Connor W Coley, Cao Xiao, Jimeng Sun, and Marinka Zitnik. Therapeutics data commons: Machine learning datasets and tasks for drug discovery and development. *arXiv preprint arXiv:2102.09548*, 2021.

John J Irwin, Khanh G Tang, Jennifer Young, Chinzorig Dandarchuluun, Benjamin R Wong, Munkhzul Khurelbaatar, Yurii S Moroz, John Mayfield, and Roger A Sayle. Zinc20—a free ultralarge-scale chemical database for ligand discovery. *Journal of chemical information and modeling*, 60(12):6065–6073, 2020.

Jan H Jensen. A graph-based genetic algorithm and generative model/monte carlo tree search for the exploration of chemical space. *Chemical science*, 10(12):3567–3572, 2019.

Chaojie Ji, Yijia Zheng, Ruxin Wang, Yunpeng Cai, and Hongyan Wu. Graph polish: A novel graph generation paradigm for molecular optimization. *IEEE Transactions on Neural Networks and Learning Systems*, 34(5):2323–2337, 2021.

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

Jiatong Li, Wei Liu, Zhihao Ding, Wenqi Fan, Yuqiang Li, and Qing Li. Large language models are in-context molecule learners. *IEEE Transactions on Knowledge and Data Engineering*, 2025.

Yang Liu, Maximilian Grimm, Wen-tao Dai, Mu-chun Hou, Zhi-Xiong Xiao, and Yang Cao. Cb-dock: a web server for cavity detection-guided protein–ligand blind docking. *Acta Pharmacologica Sinica*, 41(1):138–144, 2020.

Yiping Liu, Jiahao Yang, Xuanbai Ren, Zhang Xinyi, Yuansheng Liu, Bosheng Song, Xiangxiang Zeng, and Hisao Ishibuchi. Multi-objective molecular design through learning latent pareto set. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 19006–19014, 2025.

Feifei Luo, Jinglang Zhang, Qilong Wang, and Chunpeng Yang. Leveraging prompt engineering in large language models for accelerating chemical research. *ACS Central Science*, 2025.

Łukasz Maziarka, Agnieszka Pocha, Jan Kaczmarczyk, Krzysztof Rataj, Tomasz Danel, and Michał Warchoł. Mol-cyclegan: a generative model for molecular optimization. *Journal of Cheminformatics*, 12(1):2, 2020.

Tung Nguyen and Aditya Grover. Lico: Large language models for in-context molecular optimization. *arXiv preprint arXiv:2406.18851*, 2024.

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

Chence Shi, Minkai Xu, Zhaocheng Zhu, Weinan Zhang, Ming Zhang, and Jian Tang. Graphaf: a flow-based autoregressive model for molecular graph generation. *arXiv preprint arXiv:2001.09382*, 2020.

Dagmar Stumpfe and Jurgen Bajorath. Exploring activity cliffs in medicinal chemistry: miniperspective. *Journal of medicinal chemistry*, 55(7):2932–2942, 2012.

Jonas Verhellen. Graph-based molecular pareto optimisation. *Chemical Science*, 13(25):7526–7535, 2022.

Haorui Wang, Marta Skreta, Cher-Tian Ser, Wenhao Gao, Lingkai Kong, Felix Strieth-Kalthoff, Chenru Duan, Yuchen Zhuang, Yue Yu, Yanqiao Zhu, et al. Efficient evolutionary search over chemical space with large language models. *arXiv preprint arXiv:2406.16976*, 2024.

Runze Wang, Mingqi Yang, and Yanming Shen. Bridging molecular graphs and large language models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 21234–21242, 2025.

Xin Xia, Yiping Liu, Chunhou Zheng, Xingyi Zhang, Qingwen Wu, Xin Gao, Xiangxiang Zeng, and Yansen Su. Evolutionary multiobjective molecule optimization in an implicit chemical space. *Journal of Chemical Information and Modeling*, 64(13):5161–5174, 2024.

Yutong Xie, Chence Shi, Hao Zhou, Yuwei Yang, Weinan Zhang, Yong Yu, and Lei Li. Mars: Markov molecular sampling for multi-objective drug discovery. *arXiv preprint arXiv:2103.10432*, 2021.

Chengrun Yang, Xuezhi Wang, Yifeng Lu, Hanxiao Liu, Quoc V Le, Denny Zhou, and Xinyun Chen. Large language models as optimizers. In *The Twelfth International Conference on Learning Representations*, 2023.

Yaodong Yang, Guangyong Chen, Jinpeng Li, Junyou Li, Odin Zhang, Xujun Zhang, Lanqing Li, Jianye Hao, Ercheng Wang, and Pheng-Ann Heng. Enabling target-aware molecule generation to follow multi objectives with pareto mcts. *Communications Biology*, 7(1):1074, 2024.

Jacob Yasonik. Multiobjective de novo drug design with recurrent neural networks and nondominated sorting. *Journal of Cheminformatics*, 12(1):14, 2020.

Geyan Ye, Xibao Cai, Houtim Lai, Xing Wang, Junhong Huang, Longyue Wang, Wei Liu, and Xiangxiang Zeng. Drugassist: A large language model for molecule optimization. *Briefings in Bioinformatics*, 26(1):bbae693, 2025.

Jiajun Yu, Yizhen Zheng, Huan Yee Koh, Shirui Pan, Tianyue Wang, and Haishuai Wang. Collaborative expert llms guided multi-objective molecular optimization. *arXiv preprint arXiv:2503.03503*, 2025.

Eckart Zitzler, Lothar Thiele, Marco Laumanns, Carlos M Fonseca, and Viviane Grunert Da Fonseca. Performance assessment of multiobjective optimizers: An analysis and review. *IEEE Transactions on evolutionary computation*, 7(2):117–132, 2003.

A EXPERT PROMPT DESIGN

We provide prompt templates for the five objective-specific experts—GSK3β, JNK3, QED, SA, and CYP3A4—and the Strategy Summarizer in Figures 6–11. Each objective-specific expert proposes 50 molecular candidates from the supplied reference molecules and path-dependent optimization strategies; the output is restricted to numbered, RDKit-parseable SMILES strings. The GSK3β and JNK3 experts seek higher predicted inhibition scores, the QED expert seeks higher drug-likeness, the SA expert seeks a lower raw SA score (equivalently, a higher normalized SA score), and the CYP3A4 expert seeks lower predicted inhibition probability (equivalently, a higher $1 - P_{\mathrm{inh}}$ score). The Strategy Summarizer instead distills recurring structural edits from initial–optimized molecule pairs into concise guidance for later iterations. RDKit validates molecular structures and computes molecular descriptors, while the task oracles evaluate the corresponding objectives.

GSK3β Expert Prompt

Role Description: You are an expert computational chemist and molecular optimization specialist. You have extensive knowledge of medicinal chemistry, SMILES notation, and structure–activity relationships (SAR). Your task is to design novel molecules that show improved GSK3beta inhibitory activity while maintaining good drug-likeness and chemical validity. Your goal is to optimize these molecules to generate 50 new compounds with improved GSK3beta inhibition activity.
Please follow these guidelines carefully:
Input: A list of starting SMILES strings (example below). Objective: Increase the predicted or expected GSK3beta inhibitory activity. Output: Generate exactly 50 valid SMILES strings representing new molecules.

Constraints: The generated molecules should retain the core scaffold or key pharmacophores of the starting compounds. Ensure synthetic accessibility (reasonable SA scores) and good drug-likeness (high QED or Lipinski compliance). Each SMILES must be syntactically valid and parseable by RDKit.

Optimization strategies: {Adjusted according to reference molecular dynamics}

Reference molecular: {Select from parent node, sibling nodes, and dynamic knowledge base}

Output Requirements: Output the 50 optimized SMILES only, one per line, numbered from 1 to 50. Do not include explanations, commentary, or non-SMILES text.

Figure 6: Prompt design template for the GSK3β expert agent.

JNK3 Expert Prompt

Role Description: You are an expert computational chemist and molecular optimization specialist. You have extensive knowledge of medicinal chemistry, SMILES notation, and structure–activity relationships (SAR). Your task is to design novel molecules that show improved JNK3 inhibitory activity while maintaining good drug-likeness and chemical validity. Your goal is to optimize these molecules to generate 50 new compounds with improved JNK3 inhibition activity.
Please follow these guidelines carefully:
Input: A list of starting SMILES strings (example below). Objective: Increase the predicted or expected JNK3 inhibitory activity. Output: Generate exactly 50 valid SMILES strings representing new molecules.

Constraints: The generated molecules should retain the core scaffold or key pharmacophores of the starting compounds. Ensure synthetic accessibility (reasonable SA scores) and good drug-likeness (high QED or Lipinski compliance). Each SMILES must be syntactically valid and parseable by RDKit.

Optimization strategies: {Adjusted according to reference molecular dynamics}

Reference molecular: {Select from parent node, sibling nodes, and dynamic knowledge base}

Output Requirements: Output the 50 optimized SMILES only, one per line, numbered from 1 to 50. Do not include explanations, commentary, or non-SMILES text.

Figure 7: Prompt design template for the JNK3 expert agent.

QED Expert Prompt

Role Description: You are an expert computational chemist and molecular optimization specialist. You have extensive knowledge of medicinal chemistry, SMILES notation, and physicochemical properties relevant to drug-likeness. Your task is to design novel molecules with improved quantitative estimate of drug-likeness (QED) while preserving chemical validity and the key structural features of the starting compounds.
Please follow these guidelines carefully:
Input: A list of starting SMILES strings (example below). Objective: Increase the QED score computed from molecular structure. Output: Generate exactly 50 valid SMILES strings representing new molecules.

Constraints: The generated molecules should retain the core scaffold or key pharmacophores of the starting compounds. Prefer chemically meaningful edits that improve the balance of molecular weight, lipophilicity, polarity, hydrogen-bonding capacity, aromaticity, and flexibility while maintaining reasonable synthetic accessibility and other task-relevant properties. Each SMILES must be syntactically valid and parseable by RDKit.

Optimization strategies: {Adjusted according to reference molecular dynamics}

Reference molecular: {Select from parent node, sibling nodes, and dynamic knowledge base}

Output Requirements: Output the 50 optimized SMILES only, one per line, numbered from 1 to 50. Do not include explanations, commentary, or non-SMILES text.

Figure 8: Prompt design template for the QED expert agent.

SA Expert Prompt

Role Description: You are an expert computational chemist and molecular optimization specialist. You have extensive knowledge of medicinal chemistry, SMILES notation, synthesis planning, and structure–activity relationships (SAR). Your task is to design novel molecules that improve synthetic accessibility (lower SA score) while preserving drug-like properties and other task-relevant scores, including kinase scores where applicable.
Please follow these guidelines carefully:
Input: A list of starting SMILES strings (example below). Objective: Decrease the predicted SA score (make the molecule easier to synthesize). Output: Generate exactly 50 valid SMILES strings representing new molecules.

Constraints: Exactly 50 output SMILES. Each SMILES must be valid and RDKit-parseable. Modifications should be mild and chemically meaningful: simplify overly complex ring linkages, replace exotic fragments with classical synthetic handles, reduce steric congestion or macrocyclic motifs, use robust bioisosteric substitutions, avoid exotic heterocycles or unusual protecting-group-like fragments.

Optimization strategies: {Adjusted according to reference molecular dynamics}

Reference molecular: {Select from parent node, sibling nodes, and dynamic knowledge base}

Output Requirements: Output the 50 optimized SMILES only, one per line, numbered from 1 to 50. Do not include explanations, commentary, or non-SMILES text.

Figure 9: Prompt design template for the SA expert agent.

CYP3A4 Expert Prompt

Role Description: You are an expert computational chemist and molecular optimization specialist. You have extensive knowledge of medicinal chemistry, SMILES notation, and structure–property relationships. Your task is to design novel molecules with a lower predicted probability of CYP3A4 inhibition while preserving chemical validity and the key structural features of the starting compounds.
Please follow these guidelines carefully:
Input: A list of starting SMILES strings (example below). Objective: Decrease the predicted CYP3A4 inhibitor probability, thereby increasing the normalized non-inhibition score $1 - P_{\mathrm{inh}}$.
Output: Generate exactly 50 valid SMILES strings representing new molecules.

Constraints: The generated molecules should retain the core scaffold or key pharmacophores of the starting compounds. Use local, chemically meaningful modifications and seek to maintain reasonable QED, synthetic accessibility, and other task-relevant properties. Treat the CYP3A4 score as a model prediction rather than an experimental measurement. Each SMILES must be syntactically valid and parseable by RDKit.

Optimization strategies: {Adjusted according to reference molecular dynamics}

Reference molecular: {Select from parent node, sibling nodes, and dynamic knowledge base}

Output Requirements: Output the 50 optimized SMILES only, one per line, numbered from 1 to 50. Do not include explanations, commentary, or non-SMILES text.

Figure 10: Prompt design template for the CYP3A4 expert agent.

Strategy Summarizer Expert Prompt

Role Description: You are an expert computational chemist specializing in molecular optimization analysis. Your task is to compare a set of initial molecules and their optimized counterparts, and summarize the key structural modification strategies that were applied during the optimization process.
Please follow these guidelines carefully:
Input:

- **Initial Molecules:** {A list of starting SMILES strings}
- **Optimized Molecules:** {A list of optimized SMILES strings with improved property scores}

Objective: By comparing the initial and optimized molecules pairwise, identify and summarize the recurring structural modification strategies that consistently led to property improvements.

Output Requirements: Output a concise list of no more than 5 optimization strategies, one per line, numbered from 1 to 5. Each strategy should describe a specific and chemically interpretable structural operation (e.g., functional group substitution, ring modification, scaffold extension). Do not include SMILES strings, scores, or any additional commentary in the output.

Figure 11: Prompt design template for the Strategy Summarizer expert agent.

B Properties and Oracle Definitions

We consider five molecular objectives: predicted inhibition of GSK3β and JNK3, drug-likeness (QED), synthetic accessibility (SA), and predicted CYP3A4 non-inhibition.

B.1 Task Configurations

We evaluate eight settings spanning two to four objectives and increasing conflict:

1. **QED+SA**: optimize drug-likeness (QED) and synthetic accessibility (SA), computed with RDKit.
2. **JNK3+GSK3β**: jointly optimize inhibition of the two kinase targets, predicted by random-forest models.
3. **QED+SA+JNK3**: optimize JNK3 inhibition subject to QED and SA.
4. **QED+SA+GSK3β**: optimize GSK3β inhibition subject to QED and SA.
5. **QED+JNK3+GSK3β**: jointly optimize both kinase targets with a QED constraint.
6. **QED+SA+CYP3A4**: optimize drug-likeness, synthesizability, and CYP3A4 non-inhibition.
7. **JNK3+GSK3β+CYP3A4**: jointly optimize kinase inhibition and CYP3A4 non-inhibition.
8. **QED+SA+JNK3+GSK3β**: jointly optimize activity, drug-likeness, and synthesizability across four objectives.

GSK3β. Glycogen synthase kinase 3 beta (GSK3β) is a serine/threonine protein kinase involved in glycogen metabolism, cell proliferation, differentiation, and apoptosis. It has attracted substantial interest in neurodegenerative disease research because it regulates tau phosphorylation, amyloid precursor protein processing, and neuronal survival. Consequently, GSK3β inhibition is considered an important molecular objective in drug discovery targeting neurodegeneration.

JNK3. c-Jun N-terminal kinase 3 (JNK3) is a member of the MAPK family primarily expressed in the central nervous system. JNK3 mediates cellular stress responses including apoptosis and inflammation; inhibiting JNK3 is therefore a relevant objective for mitigating neuronal cell death and inflammatory processes in neurodegenerative disease models.

CYP3A4. Cytochrome P450 3A4 (CYP3A4) is a major drug-metabolizing enzyme predominantly expressed in the liver and small intestine. CYP3A4 catalyzes the oxidation of a wide variety of structurally diverse substrates, including approximately 50% of clinically used drugs. Inhibiting CYP3A4 can lead to increased systemic exposure and potential toxicity of co-administered drugs, making the avoidance of CYP3A4 inhibition a critical criterion in early-stage drug discovery. In

our experiments, we use the CYP3A4_Veith oracle provided by TDC, which outputs the predicted probability of a molecule being a CYP3A4 inhibitor, denoted as $P_{\text{inh}}(x) \in [0, 1]$. Since lower inhibition probability is desirable, we apply a simple normalization:

$$\text{Normalized CYP3A4}(x) = 1 - P_{\text{inh}}(x), \tag{14}$$

so that higher values indicate lower CYP3A4 inhibition risk and thus more favorable metabolic safety profiles.

QED. Quantitative Estimate of Drug-likeness (QED) aggregates several physicochemical properties (molecular weight, lipophilicity/logP, topological polar surface area, counts of H-bond donors/acceptors, aromatic ring count, and rotatable bonds) into a single score on the interval $[0, 1]$. Higher QED values indicate molecules that are more likely to exhibit favorable drug-like properties.

SA and normalized SA. Synthetic Accessibility (SA) is a heuristic score used to estimate the ease of laboratory synthesis; SA typically ranges from 1 (easy) to 10 (difficult). To treat synthetic feasibility as a maximization objective, we follow prior work and convert SA to a normalized score in $[0, 1]$ via

$$\text{Normalized_SA} \;=\; 1 - \frac{\text{SA} - 1}{9}, \tag{15}$$

so that larger values correspond to higher synthetic accessibility.

The Random Forest (RF) oracles for GSK3β, JNK3, and CYP3A4 are provided by the Therapeutics Data Commons (TDC) benchmark. These models are pre-trained on standardized datasets curated by TDC, with train/validation/test splits predefined to ensure reproducibility. Since these oracles are trained on experimental bioactivity data, their predictions may contain measurement noise and model bias. We therefore interpret oracle scores as standardized benchmark signals rather than direct experimental validation.

Reference Point for Hypervolume Computation. Since all objectives are defined on $[0, 1]$, we set the reference point for HV computation to $\mathbf{r} = (0, 0, 0, 0)$, corresponding to the worst optimum in the normalized objective space. This ensures that the HV indicator directly reflects the dominated volume relative to the theoretical upper bound.

C Baseline Details

We group the compared external methods by their primary optimization mechanism.

Unified evaluation protocol. All external baselines were evaluated by us under a unified protocol. With the exception of EAG, which we reimplemented, we used the authors' official codebases for all compared methods. We standardized the oracle functions, evaluation budget, initial molecules, and applicable model settings across methods. The results reported in this paper are computed from these runs under the shared protocol.

Protocol alignment. Across methods, we use the same initial seed population (and count its oracle evaluations in the shared budget), oracle functions, objective directions and normalizations, hypervolume reference point, and candidate aggregation rule. HV is measured from the canonical-SMILES-de-duplicated union of all valid molecules evaluated to date. Invalid outputs do not consume budget, and previously evaluated molecules use cached scores rather than additional oracle calls. We preserve the method-specific settings in official implementations unless changes are required to interface with this common protocol. Because EAG has no official codebase, our reimplementation follows its published algorithmic description and changes only the task interfaces needed for the shared molecular-oracle setting.

Generative and reinforcement-learning baselines. SMILES LSTM, REINVENT, and RationaleRL are learned generative methods; the latter two use reward-guided reinforcement learning.

Evolutionary and representation-space search baselines. SMILES GA and GRAPH GA use genetic search over molecular strings and graphs, respectively. STONED performs local SELFIES-based exploration.

Surrogate- and sampling-based baselines. GP BO and qNEHVI are Bayesian-optimization methods; MARS uses Markov-chain molecular sampling for multi-objective discovery.

LLM baselines. Direct GPT-4o-mini and direct Llama-3-8B generate molecular candidates with a single language model. Molleo is an external LLM-guided evolutionary baseline that uses an LLM as a variation operator in an evolutionary search loop.

Agent-based baselines. DrugAssist and EAG use multi-step or multi-agent workflows to coordinate molecular optimization. MAMO is an external agent-based baseline that we independently ran under the unified evaluation protocol; its reported HV means and standard deviations are computed from our repeated runs. Biomni is an open-source biomedical agent framework. We evaluate it using the official repository (`https://github.com/snap-stanford/biomni`) and report the initial-molecule-inclusive HV protocol.

Implementations. We use the ChemLactica Test Suite for the traditional baselines, BoTorch for qNEHVI, the official DrugAssist repository, and the official Molleo repository (`https://github.com/zoom-wang112358/MOLLEO`). EAG is re-implemented from its published algorithmic description.

D Correlation Analysis

To characterize the relationships among the molecular objectives considered in this work, we conducted a correlation analysis on a random subset of molecules sampled from the ZINC20 database. For each molecule, QED and SA scores were computed using RDKit, while JNK3, GSK3β, and CYP3A4 scores were obtained from the corresponding oracle models. We computed pairwise Pearson correlation coefficients for all five properties as descriptive measures of linear association between objectives.

Figure 1(a) presents the Pearson correlation coefficients among the five properties. QED and SA are nearly uncorrelated ($r = 0.051$); GSK3β and JNK3 are positively correlated ($r = 0.351$), as are SA and CYP3A4 ($r = 0.355$). QED and JNK3 show the strongest negative association ($r = -0.220$), while the remaining pairwise coefficients have magnitudes no greater than 0.107. These task-dependent relationships illustrate that improvement in one objective need not transfer uniformly to the others.

E Hyperparameter Settings

The eight main benchmarks use the configuration in Table 4. In the path-history ablation of Table 3, Full ATOM and all controls retain the same nominal retrieval limits ($\kappa = 10$ and $r = 10$); the controls alter the specified history inputs, not these limits.

Budget-driven MCTS. We use the UCB1 exploration term directly; no exploration coefficient is tuned. Every unvisited child is expanded once before UCT is applied. The eight main benchmarks use a 10,000-oracle-call budget.

Reference retrieval. For the main multi-objective benchmarks, the retrieval counts are $\kappa = 10$ molecules per eligible sibling and deficient objective, $r = 10$ parent Pareto molecules, and $R = 10$ molecules sampled from the global Pareto memory (or all available memory if it contains fewer than R molecules). Candidate sources are concatenated in the order local population, lateral imports, and global samples, then exact-SMILES de-duplicated. The resulting unique reference molecules are supplied to the agent without a separate fixed cap on their total number. The path strategy is a deterministic summary with one status line per active objective plus the current best candidate, i.e., at most $K + 1$ lines for a K-objective task.

LLM agent. All agents are instantiated with `gpt-4o-mini`/`Llama-3-8B`. Each invocation requests 50 candidate SMILES using a shared output-interface contract. When no candidate passes the specialist filter, the same invocation may be retried up to five times.

Table 4: Fixed ATOM configuration used in the reported experiments.

Component	Setting	Value	Role
MCTS	UCB1 exploration term	$\sqrt{2 \log N/n}$	Fixed by normalized rewards
MCTS	Main oracle budget	10,000	Eight-task benchmark
Lateral exchange	κ	10	Eight-task benchmark
Hierarchical transfer	r	10	Eight-task benchmark
Global memory	R	10	Random Pareto references
Prompt	Reference SMILES	Variable	Deduplicated source pool
Prompt	Strategy summary	$\leq K+1$ lines	Path-specific context
LLM agent	Output batch / retries	50 / 5	Generation and empty-output recovery

F Algorithm and Candidate Processing

Algorithm 1: ATOM search and update.
Input: initial population P_0, objective-specific agents $\{A_k\}_{k=1}^{K}$, and oracle budget B.
Output: the global non-dominated set $\mathcal{M}$.

1. Initialize the root node with P_0, compute its Pareto front, and set $\mathcal{M}$ to this front.
2. While the oracle budget remains, traverse the tree by expanding each unvisited child once and otherwise selecting the child with the UCT score in Section 3.
3. For each specialist A_k, form its reference pool from the selected node, eligible sibling imports (κ), and global-memory samples (R); pass the unique references together with the inherited strategy summary to A_k.
4. Parse and filter the proposed SMILES as specified below. Create one child node from the retained candidates, inject the top-r parent Pareto references, evaluate the child population with the task oracles, and update its Pareto front.
5. Update $\mathcal{M} \leftarrow \mathrm{ParetoFront}(\mathcal{M} \cup \mathcal{P}(P_{\mathrm{child}}))$ and back up the mean local-HV reward of newly evaluated children along the selected path.

Query accounting and evaluation. For every method, one oracle-budget unit corresponds to evaluating one previously unevaluated, RDKit-valid candidate molecule on its complete K-dimensional objective vector. Invalid SMILES are discarded before oracle evaluation and do not consume budget. Scores for previously evaluated molecules are retrieved from a cache, so repeated candidates do not trigger additional oracle calls. All methods start from the same fixed seed population; its oracle evaluations count toward the common budget and its molecules are included in the accumulated candidate set. At every evaluation step, HV is computed from the union of all valid molecules evaluated so far, after canonical-SMILES de-duplication and before Pareto-front extraction. Thus, the 10,000-call budget counts new valid molecular evaluations rather than raw attempted SMILES strings or LLM generations.

Candidate parsing, screening, and accounting. Each agent response is requested as 50 numbered SMILES lines. We first extract only lines matching the prescribed SMILES format, then retain molecules that RDKit can parse. A specialist-specific screen is applied before a child is created. Because every run begins from the fixed seed population, the root and descendant thresholds are both determined from the current parent population: QED, normalized SA, GSK3β, JNK3, and CYP3A4 must exceed the parent mean. Exact duplicate strings within an agent response are collapsed by retaining the first occurrence. Repeated exact SMILES across the tree may remain as local references, but are counted only once in the global set of newly generated valid molecules and therefore cannot inflate the target-based stopping count or the final molecular library. Canonicalization and canonical-SMILES de-duplication are applied subsequently for the diversity, novelty, and post-hoc medicinal-chemistry analyses.

Candidate-count definitions. We distinguish raw parsed strings, RDKit-valid strings, and globally new valid strings. Validity is the fraction of RDKit-valid strings among parsed strings; the repeated-

string fraction is computed after RDKit validation as one minus the fraction of globally new exact SMILES. These accounting statistics are separate from Pareto selection: all retained valid candidates are kept in the search record, whereas only non-dominated candidates enter global memory.

G Top-Ranked Candidate Performance

We further evaluate method effectiveness by reporting top-10 and top-50 average scores for each objective, together with overall average score and average rank (Tables 5 and 6). Avg Rank is the rank of the arithmetic mean of the two average property scores. For the biological objective pair (GSK3β + JNK3), ATOM obtains the best Avg Rank (1) and the best mean GSK3β averages, while Molleo remains competitive on JNK3 top-ranked scores. For the non-biological task (QED + SA), MAMO obtains the best mean QED score (0.799) and Molleo the best mean SA score (0.916). These summary metrics complement the Pareto-oriented evaluation used in the main experiments.

Table 5: Comparison of GSK3β and JNK3 performance across different methods. Avg Rank is the rank of the arithmetic mean of the two average property scores.

Method	Top10 GSK3β	Top10 JNK3	Top50 GSK3β	Top50 JNK3	Avg GSK3β	Avg JNK3	Avg Rank
SMILES LSTM	0.840	0.608	0.678	0.428	0.090	0.040	16
SMILES GA	0.906	0.763	0.896	0.729	0.572	0.388	5
GRAPH GA	0.927	0.772	0.902	0.753	0.501	0.354	8
STONED	0.910	0.738	0.897	0.726	0.342	0.207	10
GP BO	0.932	0.787	0.914	0.776	0.764	0.541	2
MARS	0.761	0.621	0.611	0.446	0.077	0.144	13
RationalRL	0.948	0.767	0.927	0.733	0.677	0.562	3
REINVENT	0.923	0.823	0.907	0.797	0.413	0.512	7
qNEHVI	0.897	0.713	0.861	0.655	0.669	0.464	4
Drugassist	0.860	0.688	0.810	0.625	0.330	0.211	11
GPT4o-mini	0.841	0.704	0.834	0.685	0.151	0.085	12
EAG	0.871	0.680	0.847	0.658	0.120	0.080	14
Biomni	0.916	0.733	0.877	0.666	0.0998	0.0663	15
MAMO	0.690	0.490	0.690	0.490	0.542	0.387	6
Molleo	0.940	0.783	0.931	0.763	0.357	0.401	9
ATOM	0.952	0.792	0.943	0.754	0.765	0.543	1

Table 6: Comparison of QED and SA performance across different methods. Avg Rank is the rank of the arithmetic mean of the two average property scores.

Method	Top10 Avg QED	Top10 Avg SA	Top50 Avg QED	Top50 Avg SA	Avg QED	Avg SA	Avg Rank
SMILES LSTM	0.940	0.953	0.925	0.926	0.572	0.746	8
SMILES GA	0.814	0.997	0.762	0.880	0.437	0.369	14
GRAPH GA	0.933	0.958	0.890	0.879	0.346	0.418	16
STONED	0.855	0.910	0.794	0.786	0.379	0.401	15
GP BO	0.924	0.806	0.883	0.724	0.551	0.524	10
MARS	0.924	0.984	0.885	0.954	0.486	0.705	9
RationalRL	0.718	0.907	0.630	0.895	0.275	0.767	11
REINVENT	0.939	0.904	0.921	0.885	0.587	0.454	13
qNEHVI	0.943	0.913	0.914	0.899	0.714	0.781	4
Drugassist	0.941	0.930	0.941	0.901	0.339	0.703	11
GPT4o-mini	0.944	0.937	0.942	0.931	0.675	0.744	6
EAG	0.943	0.938	0.938	0.938	0.683	0.711	7
Biomni	0.943	0.975	0.934	0.954	0.660	0.819	5
MAMO	0.946	0.977	0.943	0.977	0.799	0.889	1
Molleo	0.945	1.00	0.943	1.00	0.763	0.916	2
ATOM	0.947	0.985	0.946	0.977	0.794	0.819	3

H Diversity and Novelty Analysis

In Table 7, we present the diversity (Div) and novelty (Nov) scores of molecular libraries generated by different methods under six multi-objective optimization settings. Novelty is computed as the fraction of generated molecules not present in the ZINC20 database (Irwin et al., 2020). ATOM demonstrates consistently strong and balanced performance across nearly all experimental configurations, achieving exceptionally high novelty scores ($\geq$0.9994) in most settings while maintaining competitive diversity scores in the 0.75–0.85 range.

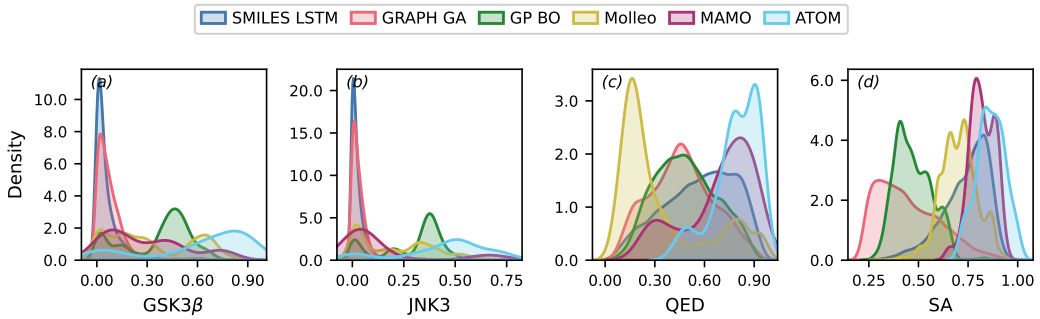

Figure 12: Normalized distributions of generated molecules from the four-objective QED+SA+JNK3+GSK3β task. Agent-based methods (ATOM and MAMO) show more favorable joint distributions across the four properties than the non-agent baselines.

Table 7: Diversity (Div) and Novelty (Nov) of generated molecules under different multi-objective settings. All values are rounded to 4 decimal places.

Method	QED+SA		GSK3β+JNK3		QED+JNK3+GSK3β		GSK3β+QED+SA		JNK3+QED+SA		QED+JNK3+GSK3β+SA	
	Div	Nov	Div	Nov	Div	Nov	Div	Nov	Div	Nov	Div	Nov
SMILES LSTM	0.8833	0.9996	0.8799	0.9994	0.8780	0.9994	0.8780	0.9994	0.8800	0.9994	0.8800	0.9994
SMILES GA	0.8555	0.9994	0.7465	0.9994	0.7444	0.9996	0.8253	0.9997	0.8689	0.9997	0.7776	0.9997
GRAPH GA	0.8690	0.9997	0.7717	0.9995	0.8162	0.9997	0.8518	0.9997	0.8899	0.9997	0.8485	0.9997
STONED	0.7700	0.9996	0.8313	0.9997	0.8389	0.9997	0.7828	0.9997	0.9063	0.9991	0.7878	0.9996
GP BO	0.5944	0.9967	0.6721	0.9967	0.6560	0.9967	0.6253	0.9967	0.6000	0.9967	0.5475	0.9967
MARS	0.6739	0.9987	0.7121	0.9967	0.6352	0.9967	0.6732	0.9989	0.6324	0.9967	0.6174	0.9973
RationaleRL	0.5697	0.9997	0.7921	0.9987	0.7580	0.9995	0.7253	0.9997	0.8023	0.9996	0.6485	0.9950
qNEHVI	0.8697	0.9999	0.8017	0.9996	0.7970	0.9999	0.8213	0.9997	0.8137	0.9999	0.7235	0.9989
GPT4o-mini	0.7914	0.9667	0.8067	0.9778	0.7911	0.9720	0.6890	0.9667	0.8101	0.9745	0.8040	0.9750
Drugassist	0.7544	0.9934	0.7507	0.9934	0.7553	0.9934	0.7501	0.9934	0.7511	0.9934	0.7655	0.9934
EAG	0.8445	0.9766	0.8045	0.9900	0.8244	0.9899	0.8401	0.9912	0.8341	0.9911	0.8130	0.9912
Biomni	0.7695	0.9865	0.7524	0.9910	0.8063	0.9997	0.8051	0.9997	0.7964	0.9919	0.7567	0.9993
MAMO	0.7415	0.8411	0.3483	0.8941	0.6943	0.9918	0.7557	0.9925	0.7382	0.9903	0.7334	0.9261
Molleo	0.7471	0.9017	0.7068	0.9246	0.7283	0.9942	0.7694	0.9939	0.7638	0.9947	0.8563	0.9474
ATOM	0.8394	0.9997	0.7020	0.9999	0.8182	0.9991	0.8278	0.9998	0.8482	0.9998	0.7521	0.9994

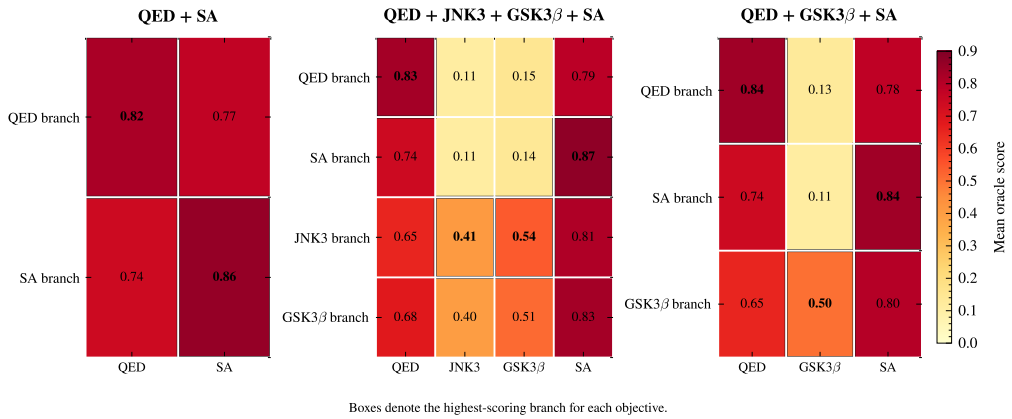

Figure 13: Branch-wise objective specialization in ATOM. Each entry is the mean raw oracle score of molecules generated by one objective-specific branch. Black boxes mark the top-scoring branch for each objective; in all panels, it coincides with the branch specialized for that objective.

I Post-hoc Medicinal-Chemistry Screens

To probe whether strong oracle scores coincide with basic medicinal-chemistry quality, we conduct a post-hoc screen on archived outputs from three primary settings with complete RDKit-valid records. After canonical SMILES de-duplication, we evaluate PAINS alerts, Lipinski rule-of-five (Ro5) compliance, the raw synthetic-accessibility score (SA; lower is easier to synthesize), and internal structural diversity. Internal diversity is defined as one minus the mean pairwise Tanimoto similarity of Morgan fingerprints; it is estimated from a fixed random sample of at most 1,000 unique molecules.

Table 8: Post-hoc chemical-quality screens of unique RDKit-valid ATOM molecules. Higher is better except raw SA.

Task	PAINS-free	Ro5 pass	Raw SA	Int. Div.
JNK3+GSK3β	97.6%	73.2%	2.81$\pm$0.48	0.586
QED+JNK3+GSK3β	99.0%	77.6%	3.36$\pm$0.80	0.799
QED+SA+JNK3+GSK3β	98.5%	81.9%	2.74$\pm$0.73	0.805

The three screens show that the generated libraries are largely PAINS-free, predominantly Ro5-compliant, and synthetically accessible while retaining appreciable structural diversity. These checks do not replace experimental validation, but reduce the possibility that the reported gains are driven solely by readily identifiable structural liabilities.

J Branch-wise Objective Specialization

To assess whether the designated branches exhibit measurable objective specialization rather than merely nominal role assignment, we compute the mean raw oracle score of the molecules produced by each branch. Figure 13 summarizes three representative tasks. In every objective column, the corresponding specialized branch attains the top mean score (black boxes), including the jointly challenging four-objective setting. This pattern shows that branch-specific optimization pressure is reflected in the generated molecular populations, yielding complementary regions of the objective space that can subsequently be integrated through global coordination.

K Multi-Objective Results and Pareto-Front Analysis

To evaluate optimization efficiency and trade-off management, we analyze the Pareto frontiers of the generated molecules. As shown in Figure 14, ATOM yields non-dominated solution sets with competitive convergence and superior trade-off balance, particularly in drug-likeness-oriented tasks. In the challenging multi-objective scenarios of GSK3β+QED and JNK3+QED, ATOM significantly outperforms baselines by maintaining high biological activity scores (> 0.8) even as QED values

approach the high-quality region (~ 0.9). While LLM-based methods such as DrugAssist and EAG experience a sharp performance decay in target affinity when optimizing for molecular properties, ATOM effectively maintains broader coverage near the knee region of the frontier.

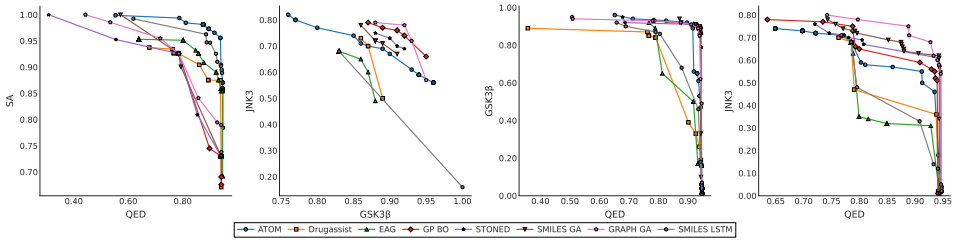

Figure 14: Non-dominated solutions of various methods on GSK3β+JNK3, QED+SA, GSK3β+QED and JNK3+QED objectives.

L Additional Pareto-front Analysis

For completeness, Table 9 reports the matched-budget Pareto-front comparison that was previously included in the main text. Although Flat+memory can retain more taskwise non-dominated molecules, ATOM contributes a larger fraction of the union front on all three tasks, indicating that tree-structured trajectories recover complementary trade-off solutions.

Table 9: Pareto-front comparison under matched budgets. HV and ΔHV are from Table 2.

Task	Method	HV	ΔHV	Unique Pareto	Union (%)
QED+SA+JNK3	**ATOM**	**0.606**	**+0.104**	51	**60.5**
+GSK3β	Flat+memory	0.502	–	72	42.0
JNK3+GSK3β	**ATOM**	**0.790**	**+0.111**	6	**100.0**
	Flat+memory	0.679	–	5	16.7
QED+JNK3	**ATOM**	**0.644**	**+0.054**	26	**83.9**
+GSK3β	Flat+memory	0.590	–	24	19.4

M Backbone Robustness with an Open-weight LLM

To examine whether ATOM depends on a particular proprietary model, we replace GPT-4o-mini with the open-weight Llama-3-8B backbone (Dubey et al., 2024), while keeping the task prompts, tree search, coordination mechanisms, and 10,000-oracle-call budget unchanged. Table 10 compares direct Llama-3-8B generation, ATOM with Llama-3-8B, and the default GPT-4o-mini implementation. ATOM raises the eight-task HV sum from 4.743 to 5.311 with the same smaller backbone. The improvement is also retained on representative dual-, triple-, and four-objective tasks. Although the smaller model remains below the GPT-4o-mini implementation in absolute HV, these results indicate that the gains from tree-structured coordination are not tied to a single frontier-model backbone.

N Computational Cost Analysis

To ensure transparency and reproducibility, we report wall-clock runtime and external LLM generation requests under the common budget of 10,000 oracle evaluations. All experiments use the same hardware configuration (Intel Xeon CPU and Nvidia A800 GPU), and results are averaged over five random seeds.

ATOM completes a run in 2h 23m on average, which is comparable to the other agent-based systems evaluated here. Its 2,691 external LLM generation requests are lower than those of Biomni and EAG, while remaining close to MAMO. The tree-structured search therefore incurs a moderate computational cost relative to direct LLM generation, but it achieves substantially higher Pareto hypervolume under the same 10,000-oracle-call budget. Future work may further reduce runtime through asynchronous expansion or locally hosted LLMs.

Table 10: Backbone robustness under the common 10,000-oracle-call budget. Values are HV (higher is better); four representative tasks are shown; the last column sums all eight benchmark tasks, including four not displayed here.

Method	QED+SA	JNK3+GSK3β	QED+JNK3+GSK3β	QED+SA+JNK3+GSK3β	Sum ↑
Direct Llama-3-8B	0.893±0.001	0.611±0.009	0.449±0.025	0.411±0.011	4.743
ATOM (Llama-3-8B)	0.909±0.001	0.668±0.011	0.496±0.031	0.498±0.027	5.311
ATOM (GPT-4o-mini)	**0.941±0.001**	**0.790±0.017**	**0.644±0.062**	**0.606±0.041**	**5.867**

Table 11: Runtime and external LLM generation requests under 10,000 oracle evaluations.

Method	**Type**	**Runtime**	**Ext. LLM Requests**
SMILES LSTM-HC	Traditional	7m 58s ± 1m 24s	–
SMILES GA	Traditional	11m 42s ± 2m 04s	–
GRAPH GA	Traditional	14m 30s ± 1m 24s	–
STONED	Traditional	13m 53s ± 1m 40s	–
REINVENT	RL-based	10m 53s ± 1m 24s	–
GPBO	Bayesian	19m 37s ± 1m 28s	–
RationaleRL	RL-based	2m 23s ± 44s	–
MARS	MCMC-based	1h 21m 40s ± 13m 23s	–
GPT-4o-mini (single)	LLM-based	1h 07m 20s ± 22m 17s	1,598 ± 150
Molleo	LLM-guided	47m ± 28m 19s	–
DrugAssist	LLM-based	4h 15m 23s ± 22m 18s	–
EAG	LLM-based	3h 05m 05s ± 44m 03s	4,723 ± 423
Biomni	Agent-based	2h 05m ± 1h 20m	2,990 ± 870
MAMO	Agent-based	2h 57m ± 1h 11m	2,328 ± 814
ATOM (ours)	LLM-based	**2h 23m ± 57m 13s**	**2,691 ± 748**

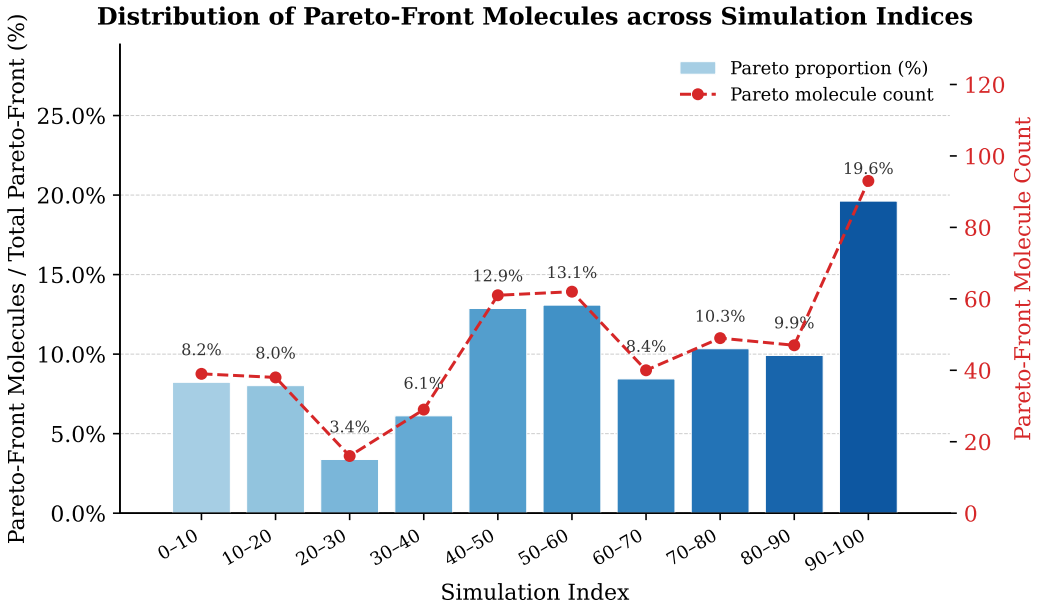

Figure 15: Pareto-front molecule distribution across MCTS simulation indices.

Figure 15 further highlights the role of global memory: a substantial fraction of Pareto-front molecules originates from early MCTS simulations, indicating that candidates discovered along diverse search trajectories are critical for populating the Pareto front. Persistent retention of non-dominated molecules across *all* simulations gives each agent access to a chemically diverse reference pool. This diagnostic observation complements the controlled ablation results for global integration.

O Post-hoc Cross-Model Kinase Re-scoring and Overlap Audit

To examine whether benchmark-oracle trends remain visible under a different predictive pipeline, we conducted a post-hoc re-scoring analysis on the QED+SA+JNK3+GSK3β task. We collected target-specific ChEMBL records and trained separate ECFP4–XGBoost regressors for GSK3β and

JNK3 using Bemis–Murcko scaffold splits. Their test-set Spearman correlations were 0.759 and 0.778, respectively. These predictors were not used during optimization. All results in this section are therefore predictor-based consistency checks, not measurements of biological activity.

For each method, we deduplicated the final Pareto front and selected 50 molecules by farthest-point sampling in the original four-dimensional objective space. Table 12 reports predicted pIC_{50} summaries. ATOM has the highest JNK3 scores, whereas MAMO is marginally higher under the GSK3β predictor. These differences indicate that the methods remain distinguishable under the separately trained predictors; they do not establish experimental activity or prospective target engagement.

Table 12: Post-hoc cross-model kinase re-scoring of equal-sized, deduplicated final Pareto-front samples ($n = 50$ per method) for QED+SA+JNK3+GSK3β. Mean and Top-10 are predicted pIC_{50}; Active is the fraction with predicted $\mathrm{pIC}_{50} \geq 6$.

	GSK3β predictor			JNK3 predictor		
Method	Mean	Top-10	Active	Mean	Top-10	Active
ATOM	6.315	6.847	80%	**6.599**	**7.385**	**76%**
MAMO	**6.347**	**6.887**	**84%**	5.812	6.799	34%
Molleo	6.153	6.725	64%	5.916	7.222	44%
Biomni	5.740	6.187	22%	5.094	5.722	4%

Training-set overlap and scaffold audit. We reconstructed the exact scaffold-based training partitions used by both re-scoring models and compared them with the same 50 sampled candidates per method. Canonical-SMILES overlap with either model's training molecules is zero for every method (Table 13), excluding direct candidate duplication in these re-scoring models. ATOM retains 92% scaffold novelty relative to each training partition. Against the full cleaned ChEMBL collections, which also include held-out records, ATOM has no exact molecular overlap and 62% scaffold novelty for both targets.

After removing candidates whose scaffold occurs in the corresponding training partition, 46 of 50 ATOM candidates remain for each target. On these scaffold-novel subsets, ATOM's mean predicted pIC_{50} and predicted-active fraction are 6.346 and 82.6% for GSK3β, and 6.731 and 82.6% for JNK3. The corresponding JNK3 means for MAMO, Molleo, and Biomni are 5.828, 5.972, and 5.124; thus, the model-based difference persists after this exclusion. This audit addresses reuse of molecules and scaffolds in the post-hoc models. It does not establish that their ChEMBL records are disjoint from the benchmark-oracle training data, whose complete provenance is unavailable. The results remain a predictive-model check rather than biological validation.

Table 13: Candidate overlap with the external predictor training partitions ($n = 50$ per method). Exact is canonical-SMILES overlap; novel scaffold is the fraction with a Bemis–Murcko scaffold absent from the corresponding training partition.

	GSK3β		JNK3	
Method	Exact	Novel scaffold	Exact	Novel scaffold
ATOM	0%	92%	0%	92%
MAMO	0%	98%	0%	98%
Molleo	0%	88%	0%	96%
Biomni	0%	70%	0%	70%

P Limitations and Future Directions

The current evaluation considers a fixed collection of benchmark objectives under predefined oracle-call budgets. Extending it to a broader set of targets, ADMET endpoints, and budget regimes would clarify how the benefits of tree-structured coordination scale with task diversity and search resources. The present implementation also uses manually specified objective prompts and branch assignments. Future work could allocate agents dynamically, prune unproductive branches, and learn task-specific coordination policies from search history. Finally, incorporating higher-fidelity docking, synthesis-

planning feedback, or experimental measurements would further strengthen the practical utility of the framework for downstream lead optimization.