SPECOPT: CONTACT-DIFF REASONING FOR AGENTIC MOLECULE OPTIMIZATION TOWARD BINDING SPECIFICITY

Anonymous authors
Paper under double-blind review

ABSTRACT

Off-target protein binding is a major source of adverse effects for small-molecule drugs, yet most structure-based molecular design methods focus on generating selective compounds *de novo* rather than improving the selectivity of existing, well-characterized drugs. We introduce *specificity optimization* (SpecOpt), a molecular design task that seeks constrained structural modifications to an existing compound that increase its binding preference for an intended target over known off-targets while preserving its structural identity and drug-like properties. To enable systematic evaluation of this task, we construct a ChEMBL-derived benchmark from compound–target interaction data, identifying intended targets through curated drug-mechanism annotations and off-targets through measured activities. We then develop an agentic framework for SpecOpt that docks each compound against its intended target and off-targets, compares the resulting poses through residue-aware atom–protein contacts, and provides these differential interactions to a large language model to propose targeted structural modifications. Candidates are retained only if they satisfy molecular similarity, ADMET, and target–off-target docking selectivity criteria. On 915 compounds, the agent improves the target–off-target binding gap for 84.8% of compounds, shifting the mean gap from -0.72 to $+0.47$ kcal/mol while maintaining a mean Tanimoto similarity of 0.72 to the starting compounds. Ablation studies identify residue-specific contact information as the critical optimization signal: replacing residue identities with binary contact indicators eliminates improvement on all 29/29 ablation compounds. These results establish SpecOpt as a distinct molecular design problem and demonstrate residue-aware differential interactions as an effective signal for improving the specificity of existing compounds.

1 INTRODUCTION

Small-molecule drugs rarely interact exclusively with their intended targets. Polypharmacology is widespread (Keiser et al., 2007), and unintended interactions with off-target proteins can contribute to adverse effects, dose-limiting toxicity, and clinical failure. Improving molecular selectivity is therefore a fundamental objective in drug development. Recent computational approaches have increasingly addressed selectivity through structure-based generative design, seeking to generate new molecules that bind strongly to a desired protein while avoiding alternative targets (Chandra et al., 2024; Southiratn et al., 2025; Park et al., 2026). These methods address an important design problem, but they largely formulate selectivity as a property to optimize during *de novo* molecular generation.

A different problem arises once a promising drug or drug candidate already exists. Such a compound may have favorable potency, pharmacokinetic properties, synthetic accessibility, extensive experimental characterization, or even established clinical use, while retaining undesirable activity against one or more off-targets (Lounkine et al., 2012). In this setting, replacing the molecule with a newly generated compound is not necessarily the desired solution. Instead, our objective is to modify the existing molecule as little as possible while selectively weakening its off-target interactions and preserving the properties that made it a viable drug candidate. This goal builds on evidence that structural modifications can alter off-target activity (Wassermann et al., 2010) and on medicinal chemistry principles of addressing liabilities while maintaining on-target potency and favorable

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.

physicochemical properties (Quancard et al., 2025). We refer to this task as *specificity optimization (SpecOpt)*: given a starting compound, an intended target, and one or more known off-targets, identify constrained structural modifications that increase the binding preference for the intended target while preserving the identity and drug-like characteristics of the starting molecule.

To our knowledge, specificity optimization of an existing drug has not previously been formulated and systematically studied as a standalone computational molecular design task. This formulation differs fundamentally from *de novo* selective generation (Chandra et al., 2024). The starting molecule is fixed rather than generated, the permissible structural deviation is constrained, and the optimization objective is inherently comparative: success requires increasing the *specificity gap* between binding to the intended target and binding to off-target proteins. Consequently, a successful method must determine not only which molecular structures are compatible with the target, but which modifications to an existing structure can preferentially alter its interactions with one protein over another.

The fixed starting molecule also makes available a direct structural signal for solving this problem. The same compound can be docked against its intended target and off-targets, allowing its binding modes to be compared atom by atom. Rather than asking only whether a molecular feature contributes to binding, we ask a differential question: *how does each part of the molecule interact with the intended target differently from the off-targets?* This comparison can identify regions that preferentially support target or off-target binding and, importantly, regions that occupy similar positions across binding pockets but interact with different residues. The latter distinction is especially relevant for closely related proteins, where a ligand may contact nearly the same regions of homologous pockets even though the identities and chemical environments of those contacts differ.

We introduce an agentic framework for this specificity-optimization task (Figure 1). Given a compound, its intended target, and known off-targets, the framework docks the compound against each protein and constructs a residue-aware, per-atom representation of the differences between the resulting binding poses. These signals are aggregated over molecular fragments and provided to a large language model (LLM) agent, which reasons over the differential interaction environment to propose targeted structural edits. The agent can remove fragments that preferentially support off-target interactions, modify regions exposed to different residue environments, or extend the molecule toward target-specific interaction opportunities. Proposed molecules are then evaluated using structural similarity, ADMET, and re-docking criteria, ensuring that optimization remains anchored to the original compound while improving its predicted specificity.

Our main contributions are:

- **We formulate specificity optimization as a computational molecular design task.** To our knowledge, this is the first framework to explicitly address the problem of modifying an existing drug or drug candidate to increase its target–off-target binding preference while constraining structural and drug-like-property changes. This task complements, rather than replaces, existing work on selective *de novo* molecular generation.
- **We introduce a benchmark and agentic framework for the task.** We construct a ChEMBL-derived dataset of 1,848 compounds annotated with intended targets and experimentally measured off-targets and develop an iterative agent combining differential structural analysis, LLM-guided molecular editing, and property-based filtering. On 915 evaluated compounds, the framework improves the docking-predicted target–off-target specificity gap for 84.8% of compounds, while maintaining a mean Tanimoto similarity of 0.72 to the starting molecules.
- **We identify residue-aware differential contacts as the key optimization signal.** Successful optimization depends not merely on whether ligand atoms contact the target or off-target, but on *which residues* they contact. Replacing residue-aware contact differences with a simpler target/off-target/shared classification eliminates improvement on all 29/29 compounds in our ablation study, highlighting the importance of residue identity for distinguishing interactions across similar binding pockets.
- **We systematically evaluate additional agentic mechanisms.** We investigate cross-compound memory, pose-consistency caution, and block-synergy mechanisms designed to provide additional experience and structural context. Although these mechanisms produce their intended intermediate behaviors, they do not measurably improve final optimization outcomes, helping isolate which information is consequential for this task.

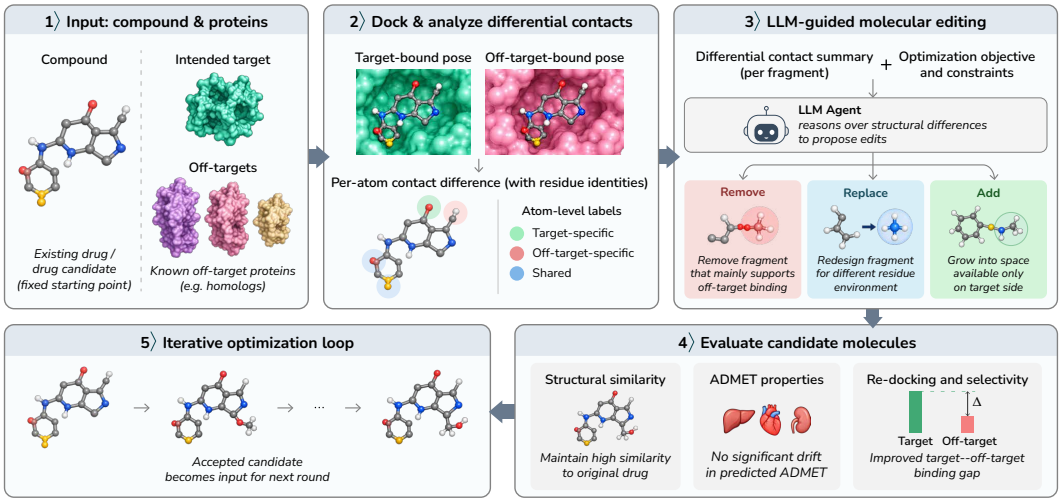


Figure 1: **Overview of SpecOpt.** Given a compound, its intended target, and known off-targets, SpecOpt docks the compound against each protein, computes per-atom contact differences with residue identities, and uses an LLM agent to propose targeted molecular edits. Candidate molecules are evaluated using structural similarity, ADMET, and re-docking criteria, and accepted candidates are iteratively refined in subsequent rounds.

2 Related Work

Selectivity-aware molecular generation. Recent structure-based generative methods have incorporated selectivity by optimizing binding to an intended target while penalizing activity against off-targets. Chandra et al. (2024) use Bayesian optimization over the latent space of a junction-tree VAE (Jin et al., 2018), rewarding docking to FLT3 while penalizing docking to homologous kinases. Park et al. (2026) steer molecular generation with dual affinity guidance, while Southiratn et al. (2025) use Pareto Monte Carlo tree search over fragment space for the related but inverse objective of dual-target binding. These approaches construct molecules *de novo*, whereas SpecOpt begins with an existing compound and seeks constrained structural modifications that improve its binding preference for the intended target while preserving similarity to the original molecule. This setting also enables direct comparison of how the same molecule interacts with its target and off-targets. We compare with Chandra et al. (2024) on FLT3 in Section 4.5, while noting that absolute selectivity is not directly comparable because the two methods address different molecular design settings.

Structure-based lead optimization. Several structure-based methods modify existing ligands using information from the protein binding pocket. DeepFrag (Green et al., 2021) predicts fragment additions from a receptor–ligand complex, while STRIFE (Hadfield et al., 2022) uses target-derived pharmacophoric information to guide fragment elaboration. AutoGrow4 (Spiegel & Durrant, 2020) applies docking-guided genetic search to evolve existing ligands toward improved predicted binding, and FRAME (Powers et al., 2023) uses geometric deep learning to iteratively grow fragments from a bound ligand, reporting improvements in both predicted affinity and selectivity. These methods demonstrate that protein structural context can effectively guide modification of an existing molecular starting point. SpecOpt differs in explicitly reasoning over the *difference* between how the same molecule interacts with its intended target and known off-targets, and using these differential interactions to guide constrained molecular edits. To our knowledge, SpecOpt is the first framework to formulate specificity optimization of an existing compound against known off-targets as a standalone molecular design task.

LLM agents for molecular design. Tool-augmented LLMs have been applied broadly to chemistry (M. Bran et al., 2024) and more recently to lead optimization. Li et al. (2026) formulate tool selection as sequential decision-making over molecular edit trajectories, while Wang et al. (2026) incorporate memory into an agentic reinforcement learning optimizer. MolLingo (Nguyen & Ji, 2026) further demonstrates that LLMs can reason over chemically meaningful molecular fragments together with docking-derived residue and binding-pocket context to propose molecular modifications

that improve predicted target binding. Building on this capability, SpecOpt focuses specifically on target-versus-off-target specificity, providing the LLM with differential structural information between target- and off-target-bound poses to guide molecular editing.

3 SPECOPT: CONTACT-DIFF-GUIDED SPECIFICITY OPTIMIZATION

3.1 TASK FORMULATION

Given a compound m, an intended target T, and a set of off-targets $\mathcal{O}$, let $s(m, P)$ be the AutoDock Vina (Trott & Olson, 2010) score of m against protein P (more negative = tighter binding). We define the **specificity gap**

$$g(m) \;=\; \min_{O \in \mathcal{O}} s(m, O) \;-\; s(m, T), \tag{1}$$

the margin between the target and the *tightest-binding* off-target. A positive gap means the molecule prefers its target over every off-target; the optimization objective is to increase g relative to the original compound m_0, subject to staying chemically close to m_0.

3.2 BENCHMARK CONSTRUCTION

We derive a specificity benchmark from ChEMBL (Mendez et al., 2019)'s compound–target interaction data. A compound's *intended* targets are those with a curated drug-mechanism designation; its *off-targets* are proteins with measured activity but no such designation. Constructing this cleanly required correcting three systematic over-counting problems: mutation panels, where different point mutations of the same protein are recorded as separate targets but are collapsed here because we focus on cross-protein specificity rather than selectivity among protein variants; protein complexes and subunits, where a subunit is counted separately from its parent complex; and generic-organism duplicates, where the same protein is registered under both a specific and an unspecified organism. The resulting dataset contains **1,848 compounds**, with a median of 3 and a mean of 7.5 off-targets each.

For docking-based experiments we filter to compounds where both the intended target and at least one off-target are single proteins with a resolvable experimental or predicted structure, and additionally exclude conformationally extreme molecules (> 45 heavy atoms or > 10 rotatable bonds), since docking cost scales with ligand flexibility. This yields a working pool of **916 compounds**. To make repeated experimentation tractable we precompute baseline docking for all 15,195 dockable (compound, target) pairs in the dataset, of which **13,636** (89.7%) resolve successfully.

3.3 DIFFERENTIAL INTERACTION ANALYSIS

The central signal in SpecOpt comes from comparing how the same ligand interacts with its intended target and off-targets. For each ligand atom i, we identify protein residues within 7 Å in the target-bound pose and each off-target-bound pose. Based on these contacts, atoms are classified as TARGET-ONLY, OFFTARGET-ONLY, SHARED-SAME-RESIDUES, or SHARED-DIFFERENT-RESIDUES. The last two categories distinguish atoms that contact the same residue identities across proteins from those that occupy similar regions of the binding pockets but interact with different residues. Independently, an atom is marked as a GROWTH-OPPORTUNITY when there is measurably more unoccupied pocket volume near that atom in the target than in the off-target.

We aggregate these atom-level signals over BRICS (Degen et al., 2008) fragments and assign each fragment an editing suggestion: REMOVE, KEEP, REDESIGN, GROW, or NEUTRAL. These suggestions determine how each fragment and its local interaction environment are presented to GPT-5.4 (OpenAI, 2026), which serves as our LLM agent. Importantly, contact with both proteins does not imply that an atom is uninformative: when the contacted residues differ between the target and off-target, the position provides an opportunity for a chemical modification to preferentially affect one interaction environment. As shown in Section 4.3, preserving this residue-level distinction is critical to the effectiveness of SpecOpt.

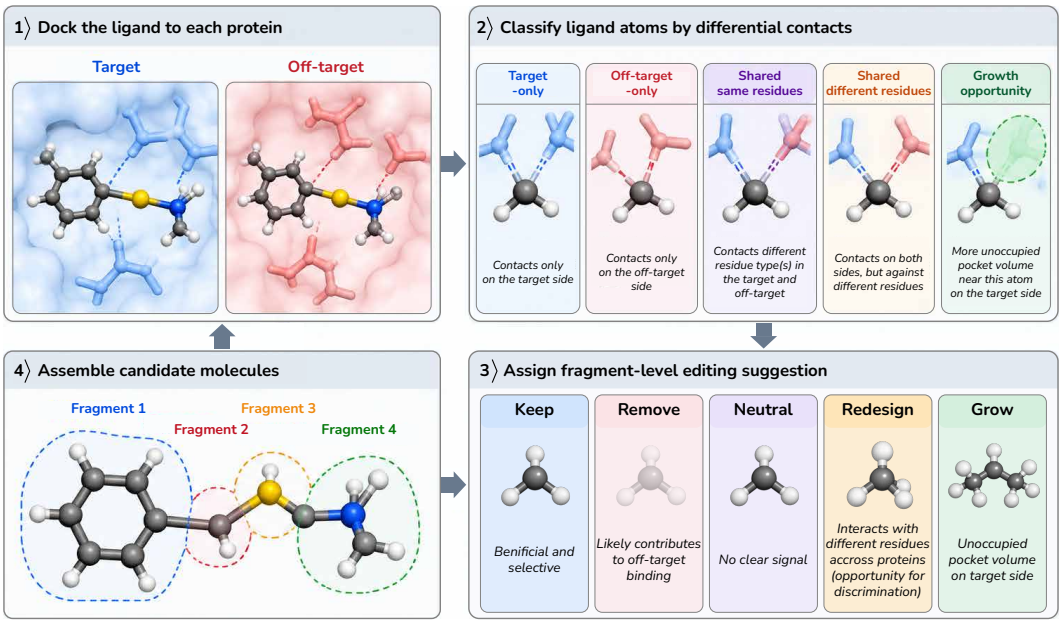


Figure 2: **Differential interaction analysis and molecular editing in SpecOpt.** Target- and off-target-bound poses provide atom-level contact differences and growth opportunities, which are aggregated into fragment-level editing suggestions to guide candidate molecule construction.

3.4 Iterative optimization and candidate filtering

Starting from the original compound, SpecOpt performs iterative optimization by alternating between LLM-guided molecular editing and candidate evaluation. At each round, the fragment-level interaction summary, current docking scores, and previously attempted edits are provided to the LLM, which proposes fragment-level modifications. These proposals are applied to generate candidate molecules, with one fragment modified at a time. Each candidate must pass three criteria before acceptance: Tanimoto similarity to the *original* molecule above 0.4, ADMET changes within tolerance for direction-unambiguous toxicity endpoints (AMES, DILI, hERG, and P-gp), and an improved target–off-target specificity gap after re-docking. We evaluate ADMET properties using the prediction model from mCLM (Edwards et al., 2026). Accepted candidates become the starting molecules for the next round, allowing structural modifications to accumulate over successive iterations. We additionally evaluate two extensions in Section 4.3: a *combo* mechanism that combines independently beneficial edits, and a two-level memory mechanism consisting of within-compound edit history and cross-compound memory retrieval.

Because both docking and LLM generation introduce stochasticity, we fix the RDKit (Landrum, n.d.) conformer seed and Vina search seed throughout the experiments. Fixing both is necessary for reproducible docking scores; otherwise, identical inputs can produce slightly different scores (e.g., -7.1 vs. -7.2 kcal/mol), which can in turn change which edits are selected. LLM proposals retain residual stochasticity, which defines the noise floor against which the ablation results are interpreted.

4 Experiments

4.1 Main result

We run the full system over the 916-compound pool with a budget of 8 candidate proposals per round and 4 rounds per compound. Failed runs are retried up to two additional times, for a maximum of three attempts per compound; a compound is excluded if all three attempts fail. Across 945 attempts, including retries, 915 distinct compounds complete successfully, while one compound exhausts the retry limit. The 30 failed attempts (3.2% of all attempts) are almost all due to baseline-docking failures

Table 1: Main result over the full compound pool. Gap is defined in Eq. 1; positive means target-preferring. The paired test compares each compound's own baseline against its own optimized result.

	Baseline (original drug)	**After optimization**
Mean specificity gap (kcal/mol)	-0.718	$\mathbf{+0.472}$
Median specificity gap (kcal/mol)	-0.600	$+0.300$
Compounds improved	—	$776/915$ (84.8%)
Mean Tanimoto to original	1.000	0.720

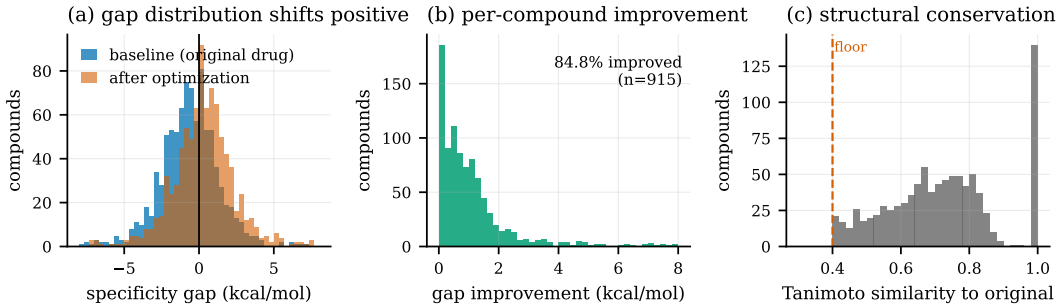


Figure 3: **Main result across 915 compounds.** (a) The specificity-gap distribution shifts from predominantly negative (off-target-preferring) to predominantly positive. (b) Per-compound improvement. (c) Structural conservation; the spike at 1.0 is compounds returned unchanged because no candidate passed the admission gates.

on large, flexible molecules. All subsequent main results are computed over the 915 successfully completed compounds.

Table 1 and Figure 3 summarize the overall results. On average, compounds shift from preferential binding to off-targets at baseline to preferential binding to their intended targets after optimization. Among the 575 compounds with a negative baseline specificity gap, 190 (33%) achieve a positive gap after optimization. The similarity distribution (Figure 3c) shows that these improvements are achieved while largely preserving the original molecular structures.

4.2 Optimization examples and case studies

Because SpecOpt records the contact-difference signal, LLM proposals, and accepted edits at each round, individual optimization trajectories are fully traceable. Figure 4 illustrates this process for edoxaban, a factor Xa inhibitor with prothrombin as an off-target. The contact-difference analysis identifies the cyclohexane linker as a GROW region, with three atoms having greater available pocket volume in factor Xa, while two other fragments are marked for REDESIGN. Consistent with this signal, the LLM proposes extending the cyclohexane ring. A methyl and fluorine modification are accepted in the first round, followed by addition of a trifluoromethyl group at the same position in the second round. Across these edits, the specificity gap improves from -5.6 to $+1.5$ kcal/mol while retaining a Tanimoto similarity of 0.76 to the original compound.

Table 2 presents additional examples that begin with negative specificity gaps, achieve positive gaps after optimization, and retain Tanimoto similarity ≥ 0.6 to the original compound. The examples span multiple target classes, including proteases, GPCRs, kinases, and metabolic enzymes. The resulting modifications are generally small, localized substitutions, including fluorine and CF_3 additions, nitrile addition, substituent repositioning, and methylation, consistent with the constrained lead-optimization setting of SpecOpt.

Contact-difference summary supplied to the LLM (edoxaban, factor Xa vs. prothrombin):

- `[1*][C@@H]1C[C@@H](C(=O)N(C)C)CC[C@@H]1[2*]`: contacts are shared with the off-target, but **3 atoms have substantially more available space in the target pocket**; candidate to GROW.
- `[2*]c1nc2c(s1)CN(C)CC2`: contacts occur in both poses at 7 atoms, but **against different residues**; candidate to REDESIGN.
- `[1*]C(=O)N[2*]`: contacts occur in both poses at 3 atoms, against different residues; candidate to REDESIGN.
- `[1*]Nc1ccc(Cl)cn1`, `[1*]NC(=O)C([2*])=O`: no docking contact data available.

Accepted edits:

- Round 1: extend cyclohexane with methyl and fluorine modifications (gap $-5.6 \rightarrow -1.0$).
- Round 2: add CF_3 at the same position (gap $-1.0 \rightarrow +1.5$).

Final: gap $-5.6 \rightarrow +1.5$ kcal/mol; Tanimoto similarity 0.76.

Figure 4: Example optimization trajectory for edoxaban. The contact-difference analysis identifies candidate regions for modification, and the LLM proposes corresponding structural edits that are evaluated and accumulated across optimization rounds.

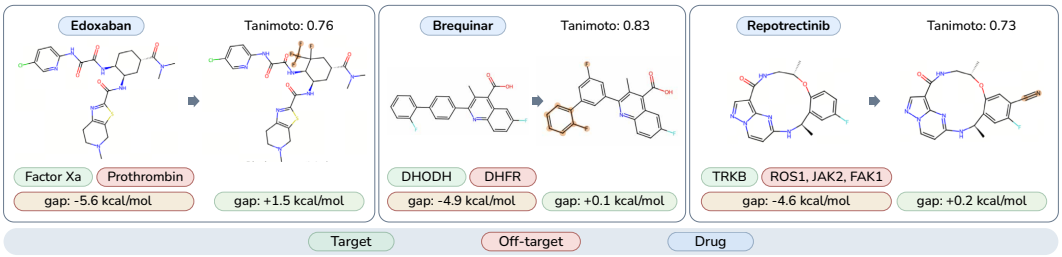


Figure 5: Selected SpecOpt case studies showing the original (left) and optimized (right) structures. Highlighted atoms indicate structural changes relative to the maximum common substructure of the original compound. Specificity gaps are reported in kcal/mol.

4.3 Ablation Study: Identifying the Key Signal

To isolate the contribution of each component, we fix a 30-compound subset stratified by baseline specificity gap and evaluate the full system against five ablated variants. Each ablation is compared with the full-system result on the same compound (Figure 6a).

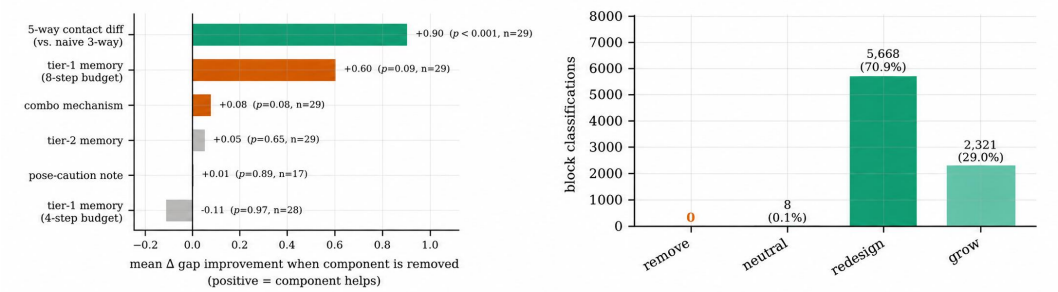
Residue identity provides the critical optimization signal. Replacing the residue-aware classification with a naive three-way classification (target-only, off-target-only, or shared contact) reduces the mean specificity-gap improvement from $+0.903$ to 0.000 **across all 29/29 compounds**: without residue identity, the system detects no differentiating interaction signal and terminates without modification. Figure 6b explains this behavior. Across the full 915-compound evaluation, the dominant signals are SHARED-DIFFERENT-RESIDUES (5,668 fragment classifications) and GROWTH-OPPORTUNITY (2,321). Thus, simply identifying whether a ligand region contacts the target or off-target is insufficient; the useful signal comes primarily from differences in *which residues* are contacted.

This observation also explains why the deterministic REMOVE operation contributes no edits. The operation is designed to delete fragments that interact exclusively with an off-target, but its trigger condition does not occur in our evaluation. We therefore retain this operation in the framework but report its lack of activation rather than relaxing its definition post hoc.

Mechanisms with no measurable performance benefit. We evaluate three additional mechanisms that operate as intended but do not significantly improve optimization performance. (i) *Cross-compound memory* retrieves an average of 2.5 relevant records from previous compounds, but removing it has no significant effect on performance ($+0.052$, $p = 0.65$). To test whether this is caused by limited relevance of the retrieved information, we repeat the experiment on 24 drugs with identical target and off-target sets, where cross-compound information should be maximally

Table 2: Selected case studies from the main evaluation. "Edit" summarizes the net structural change relative to the original compound.

Compound	Target / off-target	Base	Opt.	Sim.	Edit
Edoxaban	factor Xa / prothrombin	-5.6	$+1.5$	0.76	F + CF_3 on cyclohexane
Ibipinabant	CB_1 / CB_2	-1.5	$+5.1$	0.73	F on arylsulfonyl ring
Brequinar	DHODH / DHFR	-4.9	$+0.1$	0.83	F repositioned on biaryl
Repotrectinib	TRKB / ROS1, JAK2, FAK1	-4.6	$+0.2$	0.73	nitrile on fluoroaryl
Berotralstat	kallikrein / plasminogen	-3.5	$+1.9$	0.78	NH $\rightarrow$ N-methyl



(a) Paired ablation effects.

(b) Distribution of interaction signals.

Figure 6: Ablation and signal analysis. (a) Mean change in specificity-gap improvement when each component is removed; positive values indicate that the component improves performance. Green indicates $p < 0.05$, orange $p < 0.1$, and grey no statistical significance. (b) Fragment-level interaction signals aggregated across the 915-compound evaluation. Exclusive target- or off-target-contact signals do not occur, while most informative signals arise from different residue identities across shared contacts and target-specific growth opportunities.

transferable, but again observe no benefit (-0.067, $p = 0.33$). (ii) The *pose-consistency* mechanism identifies cases where a ligand adopts substantially different conformations in the target and off-target pockets and warns the LLM accordingly. We evaluate it on flexible molecules, where 10/17 compounds receive medium- or high-risk warnings, but observe no improvement (-0.07). (iii) The *combo* mechanism attempts to combine individually beneficial edits and activates for approximately 60% of eligible compounds. Most combined edits (6/8) fail the ADMET acceptance criteria, and the overall improvement is small and not statistically significant ($+0.079$, $p = 0.08$).

A budget-dependent effect. Within-compound memory, which prevents the agent from repeating previously attempted edits, shows no measurable benefit under a 4-round optimization budget (-0.111, $p = 0.97$). However, when the budget is increased to 8 rounds, memory produces a larger improvement ($+0.603$, $p = 0.091$). The memory-enabled condition also evaluates more candidates on average (27.3 vs. 23.6), suggesting that avoiding repeated proposals allows the optimization to continue exploring rather than triggering the "no improvement" early-stopping criterion. Because the main evaluation uses the 4-round budget, it may underestimate the benefit of within-compound memory for longer optimization runs.

4.4 Multiple off-targets

For compounds with multiple off-targets, we retain up to three off-targets based first on structural availability and then on binding affinity. The contact-difference signals from these off-targets are pooled to construct a single fragment-level summary for the LLM, allowing SpecOpt to optimize against multiple competing proteins simultaneously. To examine the effect on a broader set of off-targets, we additionally re-dock the FLT3 benchmark compounds against five common off-targets (Appendix A.1, Figure 7). On average, optimization strengthens binding to the intended target while weakening binding to most off-targets. Only $5/20$ compounds weaken all five off-targets simultaneously, while others show mixed effects. This behavior is expected because the current

Table 3: Performance grouped by the relationship between the intended target and the off-target defining the specificity gap. Performance is consistent across relationship types.

Relationship	n	**Base**	**Opt.**	Δ	**% improved**	**Wilcoxon** p
Isoform / subtype	149	-0.72	$+0.48$	$+1.20$	85.9%	9×10^{-23}
Homologous (same class)	335	-0.43	$+0.94$	$+1.38$	83.3%	2×10^{-47}
Different (no shared class)	431	-0.94	$+0.10$	$+1.04$	85.6%	3×10^{-62}

implementation optimizes against at most three off-targets; proteins outside this selected set are not explicitly considered during optimization. Extending SpecOpt to jointly optimize against a larger off-target panel is a natural direction for future work.

4.5 External comparison: FLT3

To compare SpecOpt with prior selectivity-aware molecular design, we adopt the setting of Chandra et al. (2024), using FLT3 as the intended target and PDGFRA, KIT, VEGFR2, MK2, and JAK2 as off-targets. Rather than optimizing newly generated molecules, we apply SpecOpt to 24 real FLT3-targeting drugs in our benchmark, including sunitinib, midostaurin, gilteritinib, and crenolanib. Among 23 successful runs, SpecOpt improves the specificity gap for 20 compounds, shifting the mean gap from -2.03 to -0.91 kcal/mol ($p = 8.8 \times 10^{-5}$).

Absolute selectivity values are not directly comparable between the two methods because they address different tasks and use different docking pipelines. Chandra et al. (2024) generate molecules *de novo*, whereas SpecOpt modifies existing drugs under structural constraints.

5 Discussion and Limitations

Our results highlight the importance of residue-aware differential contacts rather than contact presence alone for constrained molecular optimization. The LLM translates these signals into specific chemical edits, and richer descriptions of the local residue environment may further improve this process. Additional agentic mechanisms do not consistently improve performance, emphasizing the need to validate their contributions rather than assume that greater complexity is beneficial. Several limitations remain. First, SpecOpt optimizes against at most three selected off-targets. Evaluation against five off-targets shows that improvements can extend beyond this set, but some compounds also strengthen binding to individual off-targets; improved specificity against the selected proteins therefore does not guarantee broad selectivity. Second, all reported binding improvements are based on Vina docking scores. Fixed docking seeds support reproducible evaluation but do not establish experimental affinity or selectivity, which require experimental validation or assessment with higher-accuracy free-energy calculations.

6 Conclusion

We introduced *specificity optimization* (SpecOpt), the task of modifying an existing drug to improve its binding preference for an intended target over known off-targets while preserving the original molecule. We constructed a benchmark for this task and developed an agentic framework combining differential structural analysis with LLM-guided molecular editing. Our results identify residue-aware differential contacts as the key optimization signal and demonstrate that existing drugs can be systematically modified toward improved predicted specificity while retaining substantial structural similarity. SpecOpt therefore provides a complementary direction to selective *de novo* generation and a foundation for further development of specificity-aware lead optimization.

Reproducibility Statement

The benchmark construction, filtering criteria, agent implementation, ablation toggles, and the exact per-compound outputs behind every number in this paper are contained in the accompanying code repository. All docking uses fixed seeds; the precomputed docking cache is released to make the main run reproducible without repeating $\sim$15k docking calls. Code is available at `https://anonymous.4open.science/r/specopt-611E`.

REFERENCES

Helen M. Berman, John Westbrook, Zukang Feng, Gary Gilliland, T. N. Bhat, Helge Weissig, Ilya N. Shindyalov, and Philip E. Bourne. The protein data bank. *Nucleic Acids Research*, 28(1):235–242, 2000. doi: 10.1093/nar/28.1.235.

G. Richard Bickerton, Gaia V. Paolini, Jérémy Besnard, Sorel Muresan, and Andrew L. Hopkins. Quantifying the chemical beauty of drugs. *Nature Chemistry*, 4:90–98, 2012. doi: 10.1038/nchem.1243.

Raghav Chandra, Robert I. Horne, and Michele Vendruscolo. Bayesian optimization in the latent space of a variational autoencoder for the generation of selective FLT3 inhibitors. *Journal of Chemical Theory and Computation*, 20(1):469–476, 2024. doi: 10.1021/acs.jctc.3c01224.

Jörg Degen, Christof Wegscheid-Gerlach, Andrea Zaliani, and Matthias Rarey. On the art of compiling and using 'drug-like' chemical fragment spaces. *ChemMedChem*, 3(10):1503–1507, 2008. doi: 10.1002/cmdc.200800178.

Carl Edwards, Chi Han, Gawon Lee, Thao Nguyen, Sara Szymkuć, Chetan Kumar Prasad, Bowen Jin, Jiawei Han, Ying Diao, Ge Liu, Hao Peng, Bartosz A. Grzybowski, Martin D. Burke, and Heng Ji. mCLM: A modular chemical language model that generates functional and makeable molecules. In *International Conference on Learning Representations (ICLR)*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/76f2614f3cabc22b8a1e0d9a6aa35b61-Abstract-Conference.html.

Harrison Green, David R. Koes, and Jacob D. Durrant. Deepfrag: A deep convolutional neural network for fragment-based lead optimization. *Chemical Science*, 12(23):8036–8047, 2021. doi: 10.1039/D1SC00163A.

Thomas E. Hadfield, Fergus Imrie, Andy Merritt, Kristian Birchall, and Charlotte M. Deane. Incorporating target-specific pharmacophoric information into deep generative models for fragment elaboration. *Journal of Chemical Information and Modeling*, 62(10):2280–2292, 2022. doi: 10.1021/acs.jcim.1c01311.

Thomas A. Halgren. Merck molecular force field. I. basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry*, 17(5–6):490–519, 1996. doi: 10.1002/(SICI)1096-987X(199604)17:5/6<490::AID-JCC1>3.0.CO;2-P.

Wengong Jin, Regina Barzilay, and Tommi Jaakkola. Junction tree variational autoencoder for molecular graph generation. *Proceedings of the 35th International Conference on Machine Learning (ICML)*, 2018.

John Jumper, Richard Evans, Alexander Pritzel, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596:583–589, 2021. doi: 10.1038/s41586-021-03819-2.

Michael J. Keiser, Bryan L. Roth, Blaine N. Armbruster, Paul Ernsberger, John J. Irwin, and Brian K. Shoichet. Relating protein pharmacology by ligand chemistry. *Nature Biotechnology*, 25(2):197–206, 2007. doi: 10.1038/nbt1284.

Greg Landrum. RDKit: Open-source cheminformatics. https://www.rdkit.org, n.d. Accessed 2026.

Vincent Le Guilloux, Peter Schmidtke, and Pierre Tuffery. Fpocket: An open source platform for ligand pocket detection. *BMC Bioinformatics*, 10:168, 2009. doi: 10.1186/1471-2105-10-168.

Lingxiao Li, Haobo Zhang, Ruohao Fan, Bin Chen, and Jiayu Zhou. Molecular lead optimization via agentic tool planning. *arXiv preprint arXiv:2605.28862*, 2026.

Eugen Lounkine, Michael J. Keiser, Steven Whitebread, Dmitri Mikhailov, Jacques Hamon, Jeremy L. Jenkins, Paul Lavan, Eckhard Weber, Allison K. Doak, Serge Côté, Brian K. Shoichet, and Laszlo Urban. Large-scale prediction and testing of drug activity on side-effect targets. *Nature*, 486(7403):361–367, 2012. doi: 10.1038/nature11159.

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*, 6:525–535, 2024. doi: 10.1038/s42256-024-00832-8.

David Mendez, Anna Gaulton, A. Patrícia Bento, Jon Chambers, Marleen De Veij, Eloy Félix, María Paula Magariños, Juan F. Mosquera, Prudence Mutowo, Michał Nowotka, et al. ChEMBL: towards direct deposition of bioassay data. *Nucleic Acids Research*, 47(D1):D930–D940, 2019. doi: 10.1093/nar/gky1075.

Garrett M. Morris, Ruth Huey, William Lindstrom, Michel F. Sanner, Richard K. Belew, David S. Goodsell, and Arthur J. Olson. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, 30(16):2785–2791, 2009. doi: 10.1002/jcc.21256.

Thao Nguyen and Heng Ji. MolLingo: Molecule-native representations for LLM-powered scientific agents. *arXiv preprint arXiv:2605.27853*, 2026.

OpenAI. Introducing GPT-5.4. `https://openai.com/index/introducing-gpt-5-4/`, March 2026.

Hyoungjoon Park, Hwanhee Kim, Seungyeon Choi, Seungyong Lee, Yoonju Kim, and Sanghyun Park. TheSelective: Dual affinity-guided diffusion for selective molecular generation. In *Advances in Knowledge Discovery and Data Mining (PAKDD 2026)*, volume 16598 of *Lecture Notes in Computer Science*, pp. 16–27. Springer Nature, 2026. doi: 10.1007/978-981-92-1462-4_2.

Alexander S. Powers, Helen H. Yu, Patricia Suriana, Rohan V. Koodli, Tianyu Lu, Joseph M. Paggi, and Ron O. Dror. Geometric deep learning for structure-based ligand design. *ACS Central Science*, 9(12):2257–2267, 2023. doi: 10.1021/acscentsci.3c00572.

Jean Quancard, Anders Bach, Chiara Borsari, Russell Craft, Christian Gnamm, Stéphanie M. Guéret, Ingo V. Hartung, Hannes F. Koolman, Stefan Laufer, Susan Lepri, Josef Messinger, Kurt Ritter, Gianluca Sbardella, Andrea Unzue Lopez, Marina K. Willwacher, Brian Cox, and Robert J. Young. The european federation for medicinal chemistry and chemical biology (EFMC) best practice initiative: Hit to lead. *ChemMedChem*, 20(8):e202400931, 2025. doi: 10.1002/cmdc.202400931.

Peter Schmidtke and Xavier Barril. Understanding and predicting druggability. a high-throughput method for detection of drug binding sites. *Journal of Medicinal Chemistry*, 53(15):5858–5867, 2010. doi: 10.1021/jm100574m.

Thibaud Southiratn, Bonil Koo, Yijingxiu Lu, and Sun Kim. CombiMOTS: Combinatorial multi-objective tree search for dual-target molecule generation. In *Proceedings of the 42nd International Conference on Machine Learning (ICML)*, volume 267 of *Proceedings of Machine Learning Research*, pp. 56650–56691, 2025.

Jacob O. Spiegel and Jacob D. Durrant. Autogrow4: An open-source genetic algorithm for de novo drug design and lead optimization. *Journal of Cheminformatics*, 12:25, 2020. doi: 10.1186/s13321-020-00429-4.

The UniProt Consortium. UniProt: the universal protein knowledgebase in 2025. *Nucleic Acids Research*, 53(D1):D609–D617, 2025. doi: 10.1093/nar/gkae1010.

Oleg Trott and Arthur J. Olson. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31(2):455–461, 2010. doi: 10.1002/jcc.21334.

Shuzhe Wang, Jagna Witek, Gregory A. Landrum, and Sereina Riniker. Improving conformer generation for small rings and macrocycles based on distance geometry and experimental torsional-angle preferences. *Journal of Chemical Information and Modeling*, 60:2044–2058, 2020. doi: 10.1021/acs.jcim.0c00025.

Ziqing Wang, Yibo Wen, Abhishek Pandey, Han Liu, and Kaize Ding. MolMem: Memory-augmented agentic reinforcement learning for sample-efficient molecular optimization. In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 43694–43712, San Diego, California, United States, July 2026. Association for Computational Linguistics. doi: 10.18653/v1/2026.acl-long.2024. URL https://aclanthology.org/2026.acl-long.2024/.

Anne Mai Wassermann, Lisa Peltason, and Jürgen Bajorath. Computational analysis of multi-target structure–activity relationships to derive preference orders for chemical modifications toward target selectivity. *ChemMedChem*, 5(6):847–858, 2010. doi: 10.1002/cmdc.201000064.

A Appendix

A.1 Generalization across a broader off-target panel

Figure 7 examines whether optimization generalizes across a broader off-target panel. Most compounds show weaker predicted binding to multiple off-targets, and five compounds weaken all five simultaneously, indicating that improvements can generalize beyond individual proteins. However, several compounds strengthen binding to at least one off-target. This mixed behavior is expected because SpecOpt currently considers at most three off-targets during optimization, whereas the figure evaluates five; off-targets outside the selected optimization set are not explicitly constrained.

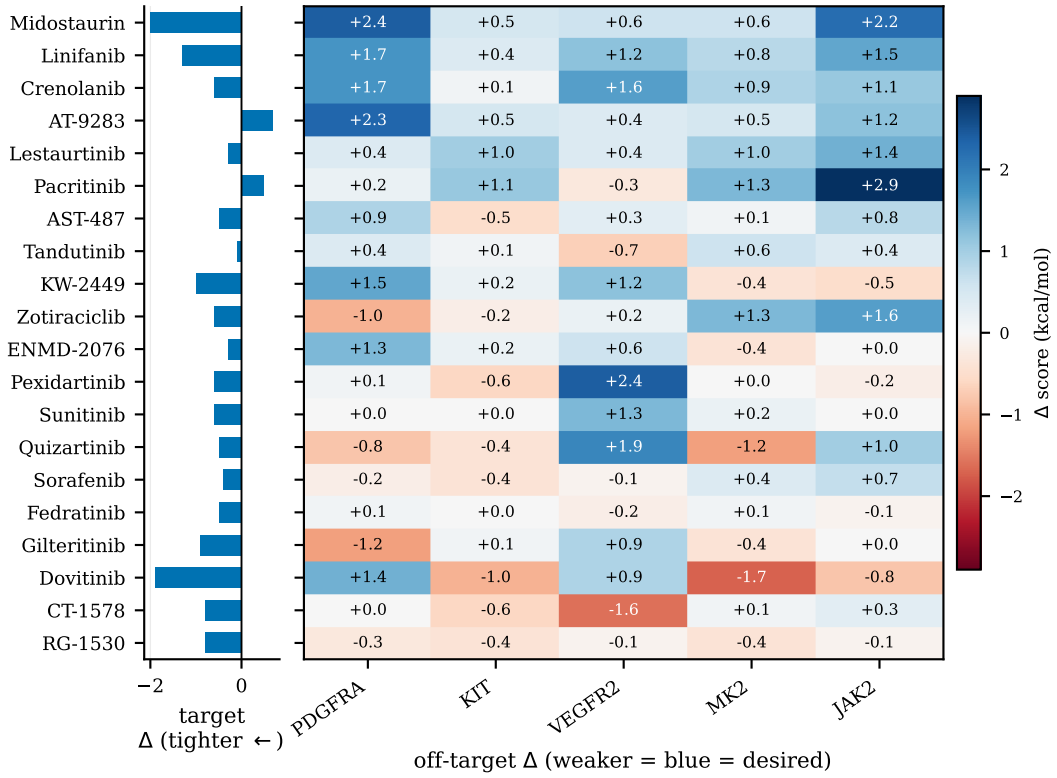


Figure 7: Binding-score changes after optimization for FLT3-targeting compounds. Left: change in target binding, where negative values indicate stronger binding. Right: changes across five off-targets, where positive values (blue) indicate weaker binding. Five compounds weaken all five off-targets simultaneously, while the remaining compounds show mixed effects across off-targets.

A.2 Benchmark characterization and docking validation

We characterize the benchmark composition and assess the reliability and limitations of its docking scores.

Composition. The benchmark contains 1,848 compounds spanning 994 distinct protein targets (Figure 8). The number of annotated off-targets per compound follows a heavy-tailed distribution, with a median of 3, a mean of 7.5, a 90th percentile of 18, and a maximum of 148. This variation is relevant to optimization because the specificity gap is defined by the tightest-binding off-target. Compounds with larger off-target sets may present more competing interactions, increasing the possibility that reducing binding to one off-target leaves another as the principal determinant of the specificity gap (Section 4.4).

The benchmark consists of drugs and drug candidates rather than generated molecules: 1,214 of 1,848 compounds (66%) are approved (ChEMBL `max_phase` 4), with the remainder in Phase 1–3.

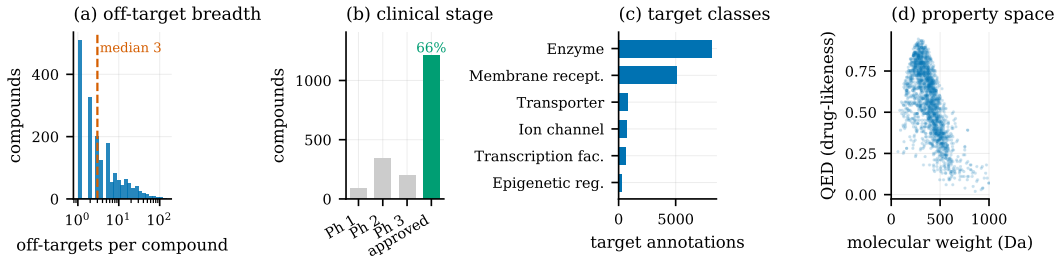


Figure 8: Benchmark composition. (a) Number of off-targets per compound on a logarithmic axis, with a median of 3. (b) Clinical stage; approximately two thirds of compounds are approved drugs. (c) Target-class annotations. (d) Molecular weight versus QED for the starting compounds.

The median molecular weight is 383 Da, and the median quantitative estimate of drug-likeness (QED) (Bickerton et al., 2012) is 0.57. These baseline properties motivate the use of similarity and ADMET constraints to preserve characteristics of the starting compounds. Target annotations are dominated by enzymes (8,163) and membrane receptors (5,110), followed by transporters (785), ion channels (744), transcription factors, and epigenetic regulators. The benchmark therefore emphasizes kinases and GPCRs, and generalization to other target classes, such as protein–protein interaction targets, requires further evaluation.

Docking data. We attempt to dock each compound against every protein annotated as an intended target or off-target. Receptors are resolved per UniProt (The UniProt Consortium, 2025) accession, preferring an experimental PDB structure (Berman et al., 2000) over an AlphaFold model (Jumper et al., 2021) and, among PDB candidates, the highest-resolution holo structure (one with a bound non-polymer ligand) at ≤ 3.5 Å; ligands are embedded with RDKit ETKDGv3 (Wang et al., 2020) and docked with AutoDock Vina. Both the conformer seed and the Vina seed are fixed to support reproducibility. We attempted 15,195 compound–protein pairs. 13,636 returned a score; of those, 143 (1.05%) were physically implausible (> -1 kcal/mol, 117 of them positive, maximum $+43.9$) and were discarded as failed poses, leaving **13,493 usable docking scores**: 1,670 against intended targets and 11,823 against off-targets. The remaining 1,559 pairs did not return a score, primarily because no protein structure could be resolved. The full cache is released with the benchmark.

Agreement with measured activity. pChEMBL values can represent several endpoints (K_i, K_d, IC50, AC50, EC50, and Potency), so direct correlation with docking scores would combine binding constants with assay-dependent measures of activity. To reduce this heterogeneity, we queried ChEMBL at the *activity* level for the 20 most frequent proteins in the benchmark and retained only binding assays (`assay_type=B`), values reported in nM, and uncensored measurements (`standard_relation` '='). For each protein, we selected the endpoint type with the greatest compound coverage, recomputed p-activity from the raw values, and summarized replicate measurements by their median. Coverage ranged from 89–99% of compounds associated with each protein.

Table 4 reports the five proteins with the strongest correlations. Muscarinic M2 achieves $\rho = -0.574$ across 132 compounds, and all five correlations are statistically significant and have the expected negative sign. These results indicate that docking scores provide information about relative activity for some well-characterized targets. However, these five proteins represent the strongest results among 19 evaluated proteins: the overall median is $\rho = -0.075$, and only 6/19 correlations reach $p < 0.05$. Agreement is therefore target-dependent. Restricting the analysis to K_i measurements yields negative correlations for 15/18 proteins, of which 7 are statistically significant, with a median correlation of -0.092.

Intended-target and off-target score distributions. Figure 9 groups all usable docking scores according to whether the protein is an intended target or an off-target of the compound. Both distributions are unimodal, centered near -8.2 kcal/mol, and span approximately -14 to -2 kcal/mol. The two distributions are closely aligned.

Table 4: The five proteins with the strongest docking–affinity agreement, after filtering to a single assay type and unit. Negative ρ is the expected direction (a more negative Vina score should accompany a higher p-activity). These are the strongest correlations among the 19 proteins tested; the median over all 19 is -0.075.

Protein	Endpoint	n	Spearman ρ	p
Muscarinic acetylcholine receptor M2	AC50	132	-0.574	6.3×10^{-13}
Mu-type opioid receptor	AC50	119	-0.316	4.7×10^{-4}
5-HT receptor 2B	Ki	99	-0.258	0.010
KCNH2 (hERG)	IC50	124	-0.237	0.008
5-HT receptor 2A	AC50	117	-0.229	0.013

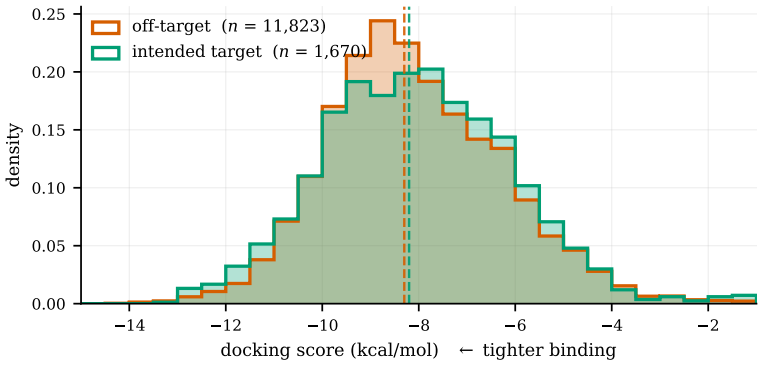


Figure 9: Docking scores pooled into two sets over the whole benchmark: compound–protein pairs where the protein is the compound's intended target ($n = 1{,}670$) and bindings where it is an off-target ($n = 11{,}823$). Density-normalized; dashed lines mark medians. The comparison shows no statistically significant difference and a small effect size: medians -8.20 vs. -8.30 kcal/mol, Mann–Whitney $p = 0.21$, Cohen's $d = 0.02$, with quantiles agreeing to within 0.1–0.4 kcal/mol at the 5th, 25th, 75th and 95th percentiles.

Binding-strength categories. We interpret Vina scores as approximate binding free energies to define nominal dissociation-constant thresholds using $\Delta G = RT \ln K_d$ ($RT = 0.593$ kcal/mol at 298 K). Table 5 groups all usable docking scores by these thresholds and categorizes experimental measurements using the corresponding pChEMBL cutoffs.

Table 5: Binding-strength categories. Left: our docking scores. Right: the experimental affinities binned at the pChEMBL equivalents of the same cut points ($-9.0 \to 6.59$, $-7.0 \to 5.13$, $-5.0 \to 3.66$). Docking places intended targets and off-targets in nearly identical proportions; the experimental labels differ by approximately 54 percentage points in the strong category. The experimental "negligible" row is empty because off-targets enter the benchmark only at pChEMBL ≥ 5.0.

		Docking score		Experimental (pChEMBL)	
Category	Vina / K_d	target	off-target	target	off-target
Strong	≤ -9.0 ($\lesssim 260$ nM)	34.7%	34.5%	81.5%	27.7%
Moderate	-9.0 to -7.0 (260 nM–7.5 µM)	37.6%	40.2%	15.4%	64.1%
Weak	-7.0 to -5.0 (7.5–220 µM)	22.6%	20.4%	3.1%	8.2%
Negligible	> -5.0 ($\gtrsim 220$ µM)	5.1%	4.9%	0.0%	0.0%
Strong or moderate		72.3%	74.7%	96.9%	91.8%
n		1,670	11,823	1,911	13,813

These results illustrate both the utility and the limitations of the docking scores. Under the thresholds in Table 5, 74.6% of compound–protein pairs fall within the strong or moderate categories, while 4.9% fall within the negligible category. Although this distribution does not independently establish

predictive accuracy, the target-specific correlations in Table 4 support the use of docking scores for relative comparisons in some settings.

However, the scores provide limited separation between intended targets and off-targets. Docking assigns similar proportions to the strong or moderate categories (72.3% and 74.7%, respectively), with differences of at most 2.6 percentage points in any category. In contrast, experimental measurements place 81.5% of intended targets and 27.7% of off-targets in the strong category, a difference of approximately 54 percentage points; the corresponding docking difference is only 0.2 percentage points. Absolute Vina scores therefore do not reliably distinguish intended targets from off-targets in this benchmark. Our optimization evaluation instead uses pre- and post-modification *differences* computed with fixed receptors and seeds. This comparison cancels systematic offsets to the extent that they are shared between the two evaluations, but does not eliminate all scoring errors.

A.3 Docking setup

All docking uses AutoDock Vina with receptors and ligands prepared using MGLTools/AutoDockTools (Morris et al., 2009). We report the configuration and implementation safeguards used to address failures observed during large-scale docking.

Receptor selection. ChEMBL target identifiers are mapped to UniProt accessions and then to structures through RCSB PDB (Berman et al., 2000). We query X-ray or cryo-EM entries with resolution ≤ 3.5 Å and prefer a holo structure (one containing at least one bound non-polymer entity) over an apo structure with marginally higher resolution. If no holo candidate is available, we select the highest-resolution structure. For example, ranking FLT3 structures solely by resolution selects the apo entry 1RJB (2.10 Å) over the holo entry 6JQR, despite a resolution difference of only 0.1 Å. Prioritizing holo structures accounts for differences in binding-pocket geometry between ligand-bound and unbound conformations.

Search box. The search box is defined using the following procedures, in order of priority.

1. **Co-crystal ligand present.** The box is centered on the bound ligand and sized to its coordinate span plus 12 Å of padding. Coordinates are taken from the *first occurrence* of the ligand rather than pooled across the structure. In homo-oligomeric structures, pooling ligand copies from multiple subunits can expand a pocket-centered box to encompass the full assembly. For ALDH2/1O04, this produced a $53 \times 161 \times 131$ Å box, exceeding the volume of a typical $\sim$20 Å pocket box by more than 100-fold, and increased Vina runtime to 475 s compared with a typical 10–20 s.
2. **No co-crystal ligand.** We apply `fpocket` (Le Guilloux et al., 2009) to the protein-only receptor and define the box around the highest-ranked pocket with 12 Å of padding. Pockets are ranked by fpocket's *druggability* score (Schmidtke & Barril, 2010) rather than its default geometric score. These rankings can differ: for 2ATX, the default first-ranked pocket had a druggability score of 0.379, whereas the third-ranked pocket scored 0.853. If no suitable pocket is identified, the fallback is a whole-protein search box with 10 Å of padding.

Vina parameters. `--exhaustiveness 8`, `--cpu 1`, `--seed 42`; the reported score is the top-ranked mode from the Vina log. Ligands are embedded with RDKit ETKDGv3 (Wang et al., 2020) and optimized with MMFF (Halgren, 1996) before conversion to PDBQT.

Determinism. Our reproducibility configuration fixes three parameters: the Vina seed, the Vina thread count (`--cpu 1`), and the RDKit conformer seed. The single-thread setting avoids variation associated with thread scheduling. RDKit's `EmbedMolecule` uses a default `randomSeed` of -1, so failing to specify this seed can produce different starting conformers even when the Vina seed is fixed. Before fixing these parameters, repeated docking of granisetron against the same target produced different binding modes and scores. With all three parameters fixed, repeated runs produced byte-identical cache contents.

Concurrency. By default, Vina detects all available CPUs (128 on our system). Concurrent docking calls can therefore oversubscribe computational resources. With six simultaneous processes, we observed SIGSEGV failures and nonzero exit codes that did not occur when the same commands were

executed individually. We consequently restrict each Vina invocation to one thread and parallelize across independent docking processes. Each call uses a separate temporary directory to prevent concurrent processes from overwriting intermediate files and corrupting results.

Timeouts. Each subprocess call has an explicit timeout: 300 s for Vina and 120 s for MGLTools. The Python 2 script `prepare_ligand4.py` can fail to terminate for certain ligand topologies; in two observed cases, calls continued at 99.9% CPU utilization for 70 minutes without producing output. Because `subprocess.run` has no default timeout, explicit limits are necessary to prevent individual ligand-processing failures from indefinitely occupying workers and progressively reducing available processing capacity.

Contacts. For contact-difference analysis, we extract per-ligand-atom contacts using a 7.0 Å threshold and set `max_hotspots` to 150. The underlying routine was designed for fragment growth and prioritizes a small number of high-volume sites. We increase the default limit of 5 to obtain broader coverage of the ligand–protein interface.

Caching. Scores are cached for each (compound, protein) pair in a JSON file shared across the docking campaign. Writes use a temporary file followed by atomic replacement with `os.replace`. Readers retry after incomplete reads and retain the most recent valid copy. These safeguards address `JSONDecodeError` failures observed when concurrent processes accessed partially written cache files. The cache is released with the benchmark to enable reproduction of the main analysis without repeating approximately 15,000 docking calls.

A.4 Ablation detail

Table 6 reports the numerical results corresponding to Figure 6a. Each row compares the full system with one ablated variant using paired observations from the same compounds.

Table 6: Paired ablation results. Δ is mean(full $-$ ablated); positive means the component contributes.

Component removed	n	**full**	**ablated**	Δ	**Wilcoxon** p
Five-way contact diff (naive 3-way instead)	29	$+0.903$	$+0.000$	$\mathbf{+0.903}$	< 0.001
Within-compound memory (8-round budget)	29	$+1.390$	$+0.786$	$+0.603$	0.091
Combo mechanism	29	$+0.903$	$+0.824$	$+0.079$	0.082
Cross-compound memory	29	$+0.903$	$+0.852$	$+0.052$	0.648
Pose-consistency caution	17	$+1.147$	$+1.141$	$+0.006$	0.893
Within-compound memory (4-round budget)	28	$+0.936$	$+1.046$	-0.111	0.972

A.5 ADMET comparison on the FLT3 setting

Using the same six-endpoint ADMET oracle, we compare the optimized molecules with both their starting compounds and the reference method's generated candidates (Table 7). The optimized molecules have lower predicted AMES, P-gp, and DILI probabilities and higher predicted HIA than both comparison groups. In particular, the mean DILI probability is 0.745, compared with 0.753 for the starting compounds and 0.787 for the generated candidates. The improvements are not uniform across endpoints: predicted hERG probability increases slightly from 0.696 to 0.702, while BBBP decreases from 0.680 to 0.619 but remains above the generated candidates' 0.387. These results indicate favorable changes in several predicted properties alongside endpoint-specific trade-offs. Because the methods start from different molecular populations and these scores are model predictions rather than experimental measurements, the comparison does not establish overall pharmacological superiority.

A.6 Implementation details

The complete docking configuration is provided in Section A.3. Molecule handling uses RDKit (Landrum, n.d.), and fragmentation follows BRICS (Degen et al., 2008). The minimum Tanimoto

Table 7: ADMET comparison (mean predicted probability). Arrows indicate the preferred direction: ↓ lower is better; ↑ higher is better.

Endpoint	Generated (top-20)	Real drugs (baseline)	Ours (optimized)
AMES (mutagenicity) ↓	0.539	0.288	**0.178**
hERG (cardiotoxicity) ↓	**0.691**	0.696	0.702
P-gp (efflux) ↓	0.349	0.315	**0.290**
DILI (liver injury) ↓	0.787	0.753	**0.745**
HIA (absorption) ↑	0.442	0.857	**0.876**
BBBP (BBB penetration)	0.387	0.680	0.619

similarity is 0.4 relative to the original molecule, rather than the molecule from the preceding round, to constrain cumulative structural deviation. The ADMET tolerance is 0.15 for AMES, DILI, hERG, and P-gp. Unless otherwise specified, runs use 8 proposals per round and 4 rounds.