COMPARING SELECTIVITY-AWARE GENERATIVE AI AND LIBRARY SCREENING IN A VIRTUAL DMT CYCLE

Amit Kadan *
SandboxAQ
amit.kadan@sandboxaq.com

Erika Lloyd *
SandboxAQ
erika.lloyd@sandboxaq.com

Andrew Wildman *
SandboxAQ
andrew.wildman@sandboxaq.com

Leo Zhang
University of Oxford, UK
leo.zhang@stx.ox.ac.uk

Steffen Ridderbusch
University of Oxford, UK
steffen.ridderbusch@eng.ox.ac.uk

ABSTRACT

The rise of generative AI for *de novo* molecular design has fueled optimism regarding our ability to rapidly design compounds that yield favorable experimental outcomes. To align computational predictions with practical drug discovery requirements, generative workflows must expand beyond optimizing solely for primary target affinity. Generative models should optimize selectivity, which ensures efficacy and safety; prioritize synthetic accessibility, which ensures that compounds can be created in the lab; and directly integrate experimental data, which ensures the properties that guide the generative model reflect reality. In this work, we present ALPAQAFlow, a generative framework that optimizes compounds by balancing on-target affinity and off-target selectivity, aiming to widen the therapeutic window. Our framework also enhances predictive accuracy by conditioning on experimental values, grounding heuristic scores with empirical data. Lastly, ALPAQAFlow extends recent advancements in reaction-pathway-based generation to ensure the resulting molecules are explicitly synthesizable. We put our platform to the test in a simulated drug discovery campaign, using Boltz-2 as a proxy for experimental affinity. Starting with a commercial library, we evaluate a selection of molecules with our proxy, then use the yielded IC_{50} values to train surrogate models. These models are used to guide ALPAQAFlow and to select highly potent and selective compounds. We show that generated molecules exhibit improved binding profiles relative to the best commercial compounds. This work demonstrates how generative AI can effectively balance complex design parameters *in silico*, aligning computational outputs more closely with the requirements of downstream experimental validation.

1 INTRODUCTION

Designing novel molecules with desired properties is a key challenge in drug discovery. *In silico* approaches have greatly accelerated pre-clinical timelines, with generative AI for *de novo* design playing a crucial role (Stanley & Segler, 2023; Kanakia et al., 2025). Generative methods directly model the distribution of molecular structures, allowing for efficient exploration of a larger chemical space than is afforded by virtual screening (Walters, 2018; Graff et al., 2021). In particular structure-based drug design (SBDD) has become increasingly popular (Luo et al., 2021; Schneuing et al.,

*Corresponding Author

2024; Shen et al., 2025) due to the availability of high resolution protein complexes, and further enabled by the accuracy of co-folding models (Abramson et al., 2024; Passaro et al., 2025; Lemos et al., 2025).

Despite the successes of generative AI in the early stages of the drug discovery pipeline, there remains a gap when translating generated compounds into successful leads in the wet-lab (Mandal et al., 2009; Gangwal et al., 2024). Most models do not guarantee synthesizability, which is a prerequisite for in-vitro validation (Gao & Coley, 2020). Furthermore, there has been little focus on designing selective inhibitors, with most structure-based models focusing on a single pocket and ignoring structural homologs (Yoshizawa et al., 2022; Zou et al., 2025). Another hurdle is that models are often validated by simplistic heuristic scores, that, despite being computationally efficient, correlate weakly with experiment (Cheng et al., 2009; Guo & Schwaller, 2025).

While some models optimize heurstic proxies for synthetic accessibility (Cremer et al., 2024; Kadan et al., 2025), they do not guarantee synthesizability and introduce a complex multi-objective optimization with competing scores when co-optimized with affinity. Recent generative models sequentially construct compounds from molecular building blocks and reaction templates, guaranteeing synthesizability by construction (Cretu et al., 2025; Gao et al., 2025; Seo et al., 2025).

Some of these works employ Generative Flow Networks (GFlowNets) which are designed to sample diverse objects generated by a sequence of actions, making them particularly suitable for synthesis-based molecular design (Bengio et al., 2021). Notably, Shen et al. (2025) recently proposed CGFlow, a model that interleaves pose prediction with synthesis-based molecular construction, bridging the gap between SBDD and reaction-pathway-based design.

To address selectivity, it is possible to frame generation as a reinforcement learning (RL) problem, where an agent learns to maximize the potency to a target, while jointly minimizing the affinity to off-targets (Liu et al., 2021; Yoshizawa et al., 2022). Outside of RL, some works propose conditioning on a joint embedding of target and off-target sequences (Chenthamarakshan et al., 2020; Vincoff et al., 2025). Within SBDD, Zou et al. (2025) propose a diffusion model capable of simultaneously processing both on-target and selective pharmacophore features. They validate their approach in a wet-lab experiment, successfully designing selective inhibitors for PARP1/2.

In this work, we introduce ALPAQAFlow (**A**ugmented **L**igand **P**ose **A**gainst **Q**uantitative **A**nti-targets G**Flow**Net). ALPAQAFlow extends CGFlow (Shen et al., 2025) while incorporating two new key innovations; structural selectivity – ALPAQAFlow explicitly models the ligand's pose within off-targets, choosing successive actions by integrating structural information from multiple environments; integration of experimental feedback – we use off-policy samples (Jain et al., 2022) to directly integrate experimental observations into the training reward. We evaluate ALPAQAFlow in a simulated campaign, showing that in a single virtual design-make-test (DMT) cycle, it is able to generate molecules with superior potency and selectivity compared to the best molecules from a large commercial library.

2 WORKFLOW

2.1 ALPAQAFLOW

GFlowNets (Bengio et al., 2021; 2023) are family of probabilistic models designed to amortize the sampling of discrete, compositional objects in proportion to some reward model. In the context of synthesizable molecular generation, consider the chemical space $\mathcal{S}$, spanned by some chosen molecular building block library $\mathcal{B}$ and a collection of reaction templates $\mathcal{R}$, where each molecule $\mathbf{x}_N \in \mathcal{S}$ can be represented as a sequence $\mathbf{x}_N = ((\mathbf{C}^{(i)}, \mathbf{S}^{(i)}))_{i=0}^{N}$ of building blocks $\mathbf{C}^{(i)} \in \mathcal{B}$, that are able to be composed according to $\mathcal{R}$, and their associated 3D coordinates $\mathbf{S}^{(i)}$.

ALPAQAFlow constructs a GFlowNet to sample synthesizable molecules in the chemical space $\mathcal{S}$, in proportion to a reward model R, conditional on some chosen target T and off-target[1] OT proteins, rewarding target potency and selectivity. In particular, ALPAQAFlow builds on CGFlow (Shen et al., 2025) by parameterizing a policy network π_θ to sample both building blocks and coordinates. This

[1]The setting of multiple off-targets follows in the natural way.

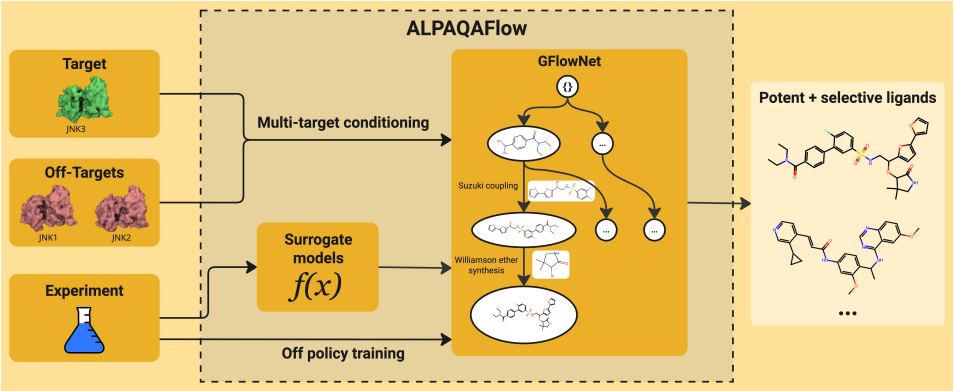

Figure 1: Overview of ALPAQAFlow. ALPAQAFlow sequentially samples synthesis steps within a chemical space composed of realistic reaction templates and building blocks filtered for unwanted moieties. ALPAQAFlow chooses an action by conditioning on the partially built molecule $\mathbf{C}^{(i)}$ and it's associated conformation in both the on-target, $\mathbf{S}^{(i)}$, and off-targets structures, $\mathbf{S}'^{(i)}$. ALPAQAFlow integrates experimental data via surrogate models and off-policy samples, allowing it to create synthesizable, potent, and selective molecules.

is trained with the following *trajectory balance loss* (Malkin et al., 2022):

$$\mathcal{L}_{\text{TB}}(\theta) = \left[\log \frac{Z_\theta(\mathbf{x}_N)\prod_{n=1}^{N}\pi_\theta(\mathbf{C}^{(n)},\mathbf{S}^{(n)},\mathbf{S}'^{(n)}|\mathbf{x}_{n-1},\mathbf{x}'_{n-1},T,OT)}{R(\mathbf{x}_N|T,OT)}\right]^2, \tag{1}$$

where $\mathbf{x}_n = ((\mathbf{C}^{(i)},\mathbf{S}^{(i)}))_{i=0}^{n}$ and $\mathbf{x}'_n = ((\mathbf{C}^{(i)},\mathbf{S}'^{(i)}))_{i=0}^{n}$ represent the same partially built molecules[2] but with conformations adapted to T and OT respectively, and Z_θ denotes the learned normalizing constant. Furthermore, π_θ has the form:

$$\pi_\theta = \pi_\theta^C(\mathbf{C}^{(n)}|\mathbf{x}_{n-1},\mathbf{x}'_{n-1},T)\delta_{\mathbf{S}^{(n)}=F_\phi(\mathbf{C}^{(n)},\mathbf{x}_{n-1},T)}\delta_{\mathbf{S}'^{(n)}=F_\phi(\mathbf{C}^{(n)},\mathbf{x}'_{n-1},OT)}, \tag{2}$$

where π_θ^C is constrained to sample only from compatible building blocks as determined by the reaction templates, δ denotes the Dirac delta distribution, and F_ϕ samples from a pretrained diffusion model[3] with frozen weights ϕ. We see that this allows ALPAQAFlow to integrate structural information from the off-targets to help inform generation. Finally, to implement π_θ^C, we use the same architecture as CGFlow but we now pass in the concatenated embeddings of $\mathbf{x}_{n-1}$ and $\mathbf{x}'_{n-1}$ to the graph transformer backbone used by CGFlow.

2.2 Systems

JNK3: The Jun N-terminal Kinases (JNKs) are part of the MAPK (Mitogen-Activated Protein Kinase) signaling pathway. While inhibiting JNK3 (PDB: 7KSI) is a therapeutic strategy for neurodegenerative diseases like Alzheimer's and Parkinson's (Feng et al., 2020), its highly conserved ATP-binding pocket complicates drug design. High structural homology leads to unintended off-target binding with JNK1 (PDB: 8X5M) and JNK2 (PDB: 8ELC) (Lu et al., 2023).

GLP-1R: Glucagon-like Peptide-1 Receptor, or GLP-1R (PDB: 7LCK), is a G protein-coupled receptor (GPCR) and is widely targeted in metabolic disorders such as type 2 diabetes and obesity (Hinnen, 2017). Recent works has shown the importance of understanding targetable differences in paralogs such as GCGR (PDB :8YW5) and GIPR (PDB: 7RBT) to reduce undesirable side effects and increase target specificity (Sa et al., 2017).

HSP90α: HSP90α (PDB: 2XAB) is an ATP-driven chaperone involved in cell growth and proliferation, and is of current interest as a cancer target due to its unique response to stress (Woodhead et al., 2010). We focus on two other structurally similar chaperones HSP90β (PDB: 1UYM), and GRP94 (PDB: IYT2), as a selectivity challenge for systemic toxicity (Wright et al., 2004).

[2] $\mathbf{x}_0 = \mathbf{x}'_0$ represents the unique initial state.

[3] CGFlow and ALPAQAFlow enforces deterministic sampling from F_ϕ to simplify the form of the policy likelihood.

3 RESULTS

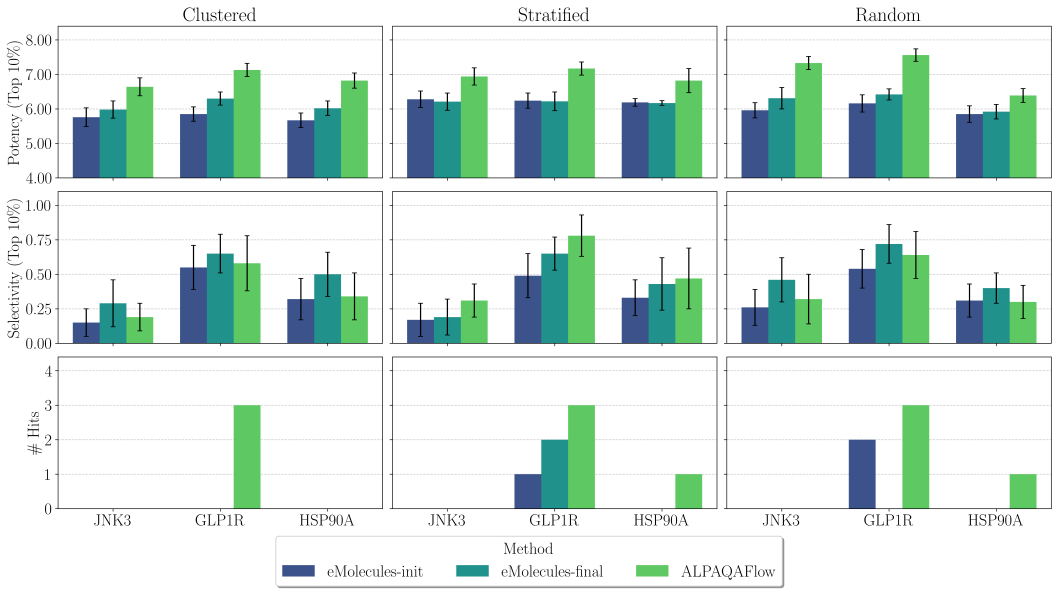

Figure 2: Final evaluation of selected molecules according to Boltz-2 (Passaro et al., 2025). For each initial selection strategy (clustered, stratified, random) and each method ($\text{eMolecules}_{\text{init}}$, $\text{eMolecules}_{\text{final}}$, ALPAQAFlow), we plot the potency and selectivity of top performing molecules (top 10%) in pIC_{50} and the number of hits according to Boltz-2. Selectivity is measured as the difference in pIC_{50} between the on-target and the most potent off-target. Hits are defined as having a potency ≥ 7 and a selectivity ≥ 1.

We assess the utility of ALPAQAFlow by validating its performance within a simulated drug discovery campaign. A typical campaign includes multiple DMT cycles, where each cycle consists of designing, synthesizing, and experimentally testing candidate molecules. The resulting data from each cycle is analyzed in order to guide successive cycles, with the ultimate goal to design viable drug candidates (Ghiandoni et al., 2024). While in practice, this process needs to be repeated many times to achieve success, in this experiment we evaluate the performance of ALPAQAFlow for designing potent and selective molecules in a single virtual DMT cycle.

3.1 EXPERIMENTAL SETUP

Binding Affinity Oracle: We use a computational proxy for experimental binding affinity. Recently, Passaro et al. (2025) proposed Boltz-2 – a co-folding model which can simultaneously predict ligand pose and binding affinity. Boltz-2 was trained on millions of experimental values extracted from ChEMBL (Zdrazil, 2025) and PubChem (Kim et al., 2025). Boltz-2 was shown to be the first AI model to approach the accuracy of free-energy perturbation methods in estimating small-molecule binding affinity while being approximately 1000 times faster.

In-stock Library: Due to the low cost of off-the-shelf molecules, in-stock libraries offer a way to gather large amounts of data in the early stages of a campaign (Lyu et al., 2019). In our simulated campaign, we make repeated use of the eMolecules in-stock library (eMolecules, Inc., 2026). We filter the library to remove unwanted molecules through substructure filters that match moieties in PAINS (Baell & Holloway, 2010), Brenk (Brenk et al., 2008), and other unwanted groups, detailed in Appendix A.5. Additionally, groups that contain more than 1 acidic oxygen and more than 1 basic nitrogen are filtered out. We also filter the library to contain only molecules that are within the chemical space of ALPAQAFlow. This facilitates off-policy training on samples evaluated with the computational oracle. Filtering is achieved by running an exhaustive retrosynthetic tree search (up to 3 synthetic steps) using the reaction templates and building blocks in ALPAQAFlow for each molecule.

Initial Molecule Selection: We begin our simulated campaign by selecting molecules from the in-stock library for evaluation with our oracle, gathering valuable target-specific data for our systems. We use three strategies to select initial molecules from the library – clustered, stratified, and random selection. For clustered selection, we use BitBIRCH (Pérez et al., 2025) to quickly and efficiently cluster the entire library. For stratified selection, we use Vinardo (Quiroga & Villarreal, 2016) to score molecules. Random selection acts as a control for our other strategies. More details are provided in Appendix A.1.

Experimental Feedback: We select 1000 molecules with each selection strategy for evaluation with our computational oracle, evaluating the same molecules for each of our 9 systems. In a typical campaign, molecules undergo an initial single-dose screen to identify significant bioactivity against the intended targets. Promising molecules are subsequently progressed to multi-point dose-response to accurately quantify their affinity. We emulate this process by masking molecules with an IC_{50} greater than 5 μM.

Surrogate Models: For each system, surrogate regression models are trained on the evaluated data. Rather than training only on the unmasked IC_{50} values, we first train random forest binary classifiers on the binder / non-binder data. We use the classifiers to annotate masked samples with approximate IC_{50} values (see Appendix A.3 for details). Afterwards, we train Gaussian process regressors that use a Tanimoto kernel on the combined data.

Chemical Space: We explore a chemical space that is constructed from the eMolecules building block library (eMolecules, Inc., 2026) combined with a set of common and realistic reaction templates. The molecules from the building block catalog were first transformed to synthons (Shen et al., 2025) with each molecule annotated according to 1 or 2 matching substructures from the reaction template library, leading to synthons that can be used as terminal states or linkers, respectively. The unreacted substructures of the synthons were then filtered by the same PAINS, Brenk, and functional group filters described for the in-stock library. This procedure was selected to create molecules that pass these filters by construction, since the reaction templates will not produce functional groups that conflict with the filters, and the synthons themselves are filtered.

Training ALPAQAFlow: In order to guide ALPAQAFlow towards generating compounds which are both potent and selective, we use a reward that promotes simultaneously optimizing both properties. We formulate the *selectivity reward* as,

$$R(\mathbf{x}) = \mathrm{ReLU}\left[\mathrm{pIC}_{50}^{\mathrm{T}}(\mathbf{x}) \cdot \left(\mathrm{pIC}_{50}^{\mathrm{T}}(\mathbf{x}) - \max\left\{\mathrm{pIC}_{50}^{\mathrm{OT}_1}(\mathbf{x}), \mathrm{pIC}_{50}^{\mathrm{OT}_2}(\mathbf{x})\right\}\right)\right], \tag{3}$$

i.e., the on-target affinity multiplied by the minimum selectivity across both off-targets, passed through a ReLU function. For each series, ALPAQAFlow samples 500,000 molecules during training, with a 50/50 split of within-policy and off-policy samples in each batch. More details on off-policy training are included in Appendix A.2. Once trained, we sample 1,000,000 molecules from the model for consideration in final candidate selection.

Final Molecule Selection: We use the trained surrogate regression models to select the top compounds from the in-stock library and from the generated molecules. We first cluster molecules on shared Bemis-Murcko Scaffolds (Bemis & Murcko, 1996), and then iteratively select up to 5 members from the top performing cluster until we reach 1000 molecules. Clusters are ranked by their maximum selectivity reward (Eq. (3)).

3.2 Final Evaluation

We report both predicted statistics according to our surrogate models and true statistics according to our computational oracle. We report potency in terms of pIC_{50}. We report selectivity in terms of difference in pIC_{50} values, i.e., the potency to the on-target minus the maximum potency to all off-targets. We also report the number of hits and chemical series for each target. A hit is defined as a molecule with $IC_{50} \leq 100$ nM ($pIC_{50} \geq 7$), and greater than $10\times$ selectivity against both off-targets. A chemical series is defined as a set of 3 or more molecules that share a common Bemis-Murcko Scaffold (Bemis & Murcko, 1996), contain at least one hit, and all have potency better than 1 μM. We report our findings in Fig. 2. Extended results can be found in Appendix A.6. ALPAQAFlow shines in finding molecules which are simultaneously potent and selective, outperforming selections from the in-stock library in all scenarios. It can be seen that the initial selection strategy has some effect on the performance of ALPAQAFlow – when seeded with a stratified selection of compounds,

it outperforms an exhaustive search of the in-stock library in terms of both potency and selectivity, while it falls short in terms of selectivity using the two other selection strategies.

4 CONCLUSIONS

In this study, we introduce ALPAQAFlow, a generative model capable of designing synthesizable, potent, and selective molecules. Additionally, we've introduced set of reaction templates and filters to produce synthesizable and stable druglike compounds. We demonstrated ALPAQAFlow's performance against exhaustively searching an in-stock library of compounds in a virtual design-make-test cycle. Our results show that ALPAQAFlow is able to consistently identify a greater number of potent and selective compounds compared to exhaustively searching the library. This performance gain comes from being able to produce highly potent molecules while keeping their selectivity moderate. Future work will focus on generating molecules which are both highly potent and highly selective. We'd also like to expand our set of reaction templates to allow for greater expressivity.

REFERENCES

Josh Abramson, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, Olaf Ronneberger, Lindsay Willmore, Andrew J Ballard, Joshua Bambrick, et al. Accurate structure prediction of biomolecular interactions with alphafold 3. *Nature*, 630(8016):493–500, 2024.

Jonathan B Baell and Georgina A Holloway. New substructure filters for removal of pan assay interference compounds (pains) from screening libraries and for their exclusion in bioassays. *Journal of medicinal chemistry*, 53(7):2719–2740, 2010.

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

Emmanuel Bengio, Moksh Jain, Maksym Korablyov, Doina Precup, and Yoshua Bengio. Flow network based generative models for non-iterative diverse candidate generation. In *Advances in neural information processing systems*, volume 34, pp. 27381–27394, 2021.

Yoshua Bengio, Salem Lahlou, Tristan Deleu, Edward J Hu, Mo Tiwari, and Emmanuel Bengio. Gflownet foundations. *Journal of Machine Learning Research*, 24(210):1–55, 2023.

Ruth Brenk, Alessandro Schipani, Daniel James, Agata Krasowski, Ian Hugh Gilbert, Julie Frearson, and Paul Graham Wyatt. Lessons learnt from assembling screening libraries for drug discovery for neglected diseases. *ChemMedChem: Chemistry Enabling Drug Discovery*, 3(3):435–444, 2008.

Tiejun Cheng, Xun Li, Yan Li, Zhihai Liu, and Renxiao Wang. Comparative assessment of scoring functions on a diverse test set. *Journal of chemical information and modeling*, 49(4):1079–1093, 2009.

Vijil Chenthamarakshan, Payel Das, Samuel Hoffman, Hendrik Strobelt, Inkit Padhi, Kar Wai Lim, Benjamin Hoover, Matteo Manica, Jannis Born, Teodoro Laino, and Aleksandra Mojsilovic. Cogmol: Target-specific and selective drug design for covid-19 using deep generative models. In *Advances in Neural Information Processing Systems*, volume 33, pp. 4320–4332, 2020.

Julian Cremer, Tuan Le, Frank Noé, Djork-Arné Clevert, and Kristof T Schütt. Pilot: equivariant diffusion for pocket-conditioned de novo ligand generation with multi-objective guidance via importance sampling. *Chemical Science*, 15(36):14954–14967, 2024.

Miruna Cretu, Charles Harris, Ilia Igashov, Arne Schneuing, Marwin Segler, Bruno Correia, Julien Roy, Emmanuel Bengio, and Pietro Lio. Synflownet: Design of diverse and novel molecules with synthesis constraints. In *The Thirteenth International Conference on Learning Representations*, 2025.

eMolecules, Inc. eMolecules Database, 2026. URL https://www.emolecules.com. Accessed: 2026-01-27.

Yangbo Feng, HaJeung Park, Luke Bauer, Jae Cheon Ryu, and Sung OK Yoon. Thiophene-pyrazolourea derivatives as potent, orally bioavailable, and isoform-selective jnk3 inhibitors. *ACS medicinal chemistry letters*, 12(1):24–29, 2020.

Amit Gangwal, Azim Ansari, Iqrar Ahmad, Abul Kalam Azad, Vinoth Kumarasamy, Vetriselvan Subramaniyan, and Ling Shing Wong. Generative artificial intelligence in drug discovery: basic framework, recent advances, challenges, and opportunities. *Frontiers in pharmacology*, 15: 1331062, 2024.

Wenhao Gao and Connor W Coley. The synthesizability of molecules proposed by generative models. *Journal of chemical information and modeling*, 60(12):5714–5723, 2020.

Wenhao Gao, Shitong Luo, and Connor W Coley. Generative ai for navigating synthesizable chemical space. *Proceedings of the National Academy of Sciences*, 122(41):e2415665122, 2025.

Gian Marco Ghiandoni, Emma Evertsson, David J Riley, Christian Tyrchan, and Prakash Chandra Rathi. Augmenting dmta using predictive ai modelling at astrazeneca. *Drug discovery today*, 29 (4):103945, 2024.

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.

Jeff Guo and Philippe Schwaller. Directly optimizing for synthesizability in generative molecular design using retrosynthesis models. *Chemical Science*, 16(16):6943–6956, 2025.

Deborah Hinnen. Glucagon-like peptide 1 receptor agonists for type 2 diabetes. *Diabetes spectrum*, 30(3):202–210, 2017.

Moksh Jain, Emmanuel Bengio, Alex Hernandez-Garcia, Jarrid Rector-Brooks, Bonaventure FP Dossou, Chanakya Ajit Ekbote, Jie Fu, Tianyu Zhang, Michael Kilgour, Dinghuai Zhang, et al. Biological sequence design with gflownets. In *Thirty-Ninth International Conference on Machine Learning*, pp. 9786–9801, 2022.

Amit Kadan, Kevin Ryczko, Erika Lloyd, Adrian Roitberg, and Takeshi Yamazaki. Guided multi-objective generative ai to enhance structure-based drug design. *Chemical Science*, 15(29):13196–13210, 2025.

Anshul Kanakia, Mark Sale, Liang Zhao, and Zhu Zhou. Ai in action: redefining drug discovery and development. *Clinical and Translational Science*, 18(2):e70149, 2025.

Sunghwan Kim, Jie Chen, Tiejun Cheng, Asta Gindulyte, Jia He, Siqian He, Qingliang Li, Benjamin A Shoemaker, Paul A Thiessen, Bo Yu, et al. Pubchem 2025 update. *Nucleic acids research*, 53(D1):D1516–D1525, 2025.

Pablo Lemos, Zane Beckwith, Sasaank Bandi, Maarten van Damme, Jordan Crivelli-Decker, Benjamin J. Shields, Thomas Merth, Punit K. Jha, Nicola De Mitri, Tiffany J. Callahan, AJ Nish, Paul Abruzzo, Romelia Salomon-Ferrer, and Martin Ganahl. Sair: Enabling deep learning for protein-ligand interactions with a synthetic structural dataset. *bioRxiv*, 2025. doi: 10.1101/2025.06.17.660168.

Xuhan Liu, Kai Ye, Herman WT van Vlijmen, Michael TM Emmerich, Adriaan P IJzerman, and Gerard JP van Westen. Drugex v2: de novo design of drug molecules by pareto-based multi-objective reinforcement learning in polypharmacology. *Journal of cheminformatics*, 13(1):85, 2021.

Kenneth López-Pérez, Taewon D Kim, and Ramón Alain Miranda-Quintana. isim: instant similarity. *Digital Discovery*, 3(6):1160–1171, 2024.

Wenchao Lu, Yao Liu, Yang Gao, Qixiang Geng, Deepak Gurbani, Lianbo Li, Scott B Ficarro, Cynthia J Meyer, Dhiraj Sinha, Inchul You, et al. Development of a covalent inhibitor of c-jun n-terminal protein kinase (jnk) 2/3 with selectivity over jnk1. *Journal of medicinal chemistry*, 66 (5):3356–3371, 2023.

Shitong Luo, Jiaqi Guan, Jianzhu Ma, and Jian Peng. A 3d generative model for structure-based drug design. In *Advances in Neural Information Processing Systems*, volume 34, pp. 6229–6239, 2021.

Jiankun Lyu, Sheng Wang, Trent E Balius, Isha Singh, Anat Levit, Yurii S Moroz, Matthew J O'Meara, Tao Che, Enkhjargal Algaa, Kateryna Tolmachova, et al. Ultra-large library docking for discovering new chemotypes. *Nature*, 566(7743):224–229, 2019.

Nikolay Malkin, Moksh Jain, Emmanuel Bengio, Chen Sun, and Yoshua Bengio. Trajectory balance: Improved credit assignment in gflownets. In *Advances in Neural Information Processing Systems*, volume 35, pp. 5955–5967, 2022.

Soma Mandal, Mee'nal Moudgil, and Sanat K Mandal. Rational drug design. *European journal of pharmacology*, 625(1-3):90–100, 2009.

Andrew T McNutt, Yanjing Li, Rocco Meli, Rishal Aggarwal, and David Ryan Koes. Gnina 1.3: the next increment in molecular docking with deep learning. *Journal of Cheminformatics*, 17(1): 28, 2025.

Saro Passaro, Gabriele Corso, Jeremy Wohlwend, Mateo Reveiz, Stephan Thaler, Vignesh Ram Somnath, Noah Getz, Tally Portnoi, Julien Roy, Hannes Stark, et al. Boltz-2: Towards accurate and efficient binding affinity prediction. *bioRxiv*, 2025. doi: 10.1101/2025.06.14.659707.

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

Kenneth López Pérez, Vicky Jung, Lexin Chen, Kate Huddleston, and Ramón Alain Miranda-Quintana. Bitbirch: efficient clustering of large molecular libraries. *Digital Discovery*, 4(4): 1042–1051, 2025.

Ignacio Pickering, Krisztina Zsigmond, Kenneth López Pérez, Miroslav Lžičař, and Ramón Alain Miranda-Quintana. BitBIRCH-Lean: A memory-efficient implementation of the BitBIRCH clustering algorithm, 2025a. URL https://github.com/mqcomplab/bblean.

Ignacio Pickering, Krisztina Zsigmond, Kenneth López Pérez, Miroslav Lžičař, and Ramón Alain Miranda-Quintana. Bitbirch-lean: chemical space in the palm of your workstation. *bioRxiv*, 2025b. doi: 10.1101/2025.10.22.684015.

Rodrigo Quiroga and Marcos A Villarreal. Vinardo: A scoring function based on autodock vina improves scoring, docking, and virtual screening. *PloS one*, 11(5):e0155183, 2016.

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

Zhining Sa, Jingqi Zhou, Yangyun Zou, Zhixi Su, and Xun Gu. Paralog-divergent features may help reduce off-target effects of drugs: hints from glucagon subfamily analysis. *Genomics, Proteomics & Bioinformatics*, 15(4):246–254, 2017.

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.

Seonghwan Seo, Minsu Kim, Tony Shen, Martin Ester, Jinkyoo Park, Sungsoo Ahn, and Woo Youn Kim. Generative flows on synthetic pathway for drug design. In *The Thirteenth International Conference on Learning Representations*, 2025.

Tony Shen, Seonghwan Seo, Ross Irwin, Kieran Didi, Simon Olsson, Woo Youn Kim, and Martin Ester. Compositional flows for 3d molecule and synthesis pathway co-design. In *Forty-second International Conference on Machine Learning*, 2025.

Megan Stanley and Marwin Segler. Fake it until you make it? generative de novo design and virtual screening of synthesizable molecules. *Current Opinion in Structural Biology*, 82:102658, 2023.

Sophia Vincoff, Oscar Davis, Alexander Tong, Joey Bose, and Pranam Chatterjee. SOAPI: Siamese-guided generation of off-target-avoiding protein interactions. In *Learning Meaningful Representations of Life (LMRL) Workshop at ICLR 2025*, 2025.

W Patrick Walters. Virtual chemical libraries: miniperspective. *Journal of medicinal chemistry*, 62 (3):1116–1124, 2018.

Andrew J Woodhead, Hayley Angove, Maria G Carr, Gianni Chessari, Miles Congreve, Joseph E Coyle, Jose Cosme, Brent Graham, Philip J Day, Robert Downham, et al. Discovery of (2, 4-dihydroxy-5-isopropylphenyl)-[5-(4-methylpiperazin-1-ylmethyl)-1, 3-dihydroisoindol-2-yl] methanone (at13387), a novel inhibitor of the molecular chaperone hsp90 by fragment based drug design. *Journal of medicinal chemistry*, 53(16):5956–5969, 2010.

Lisa Wright, Xavier Barril, Brian Dymock, Louisa Sheridan, Allan Surgenor, Mandy Beswick, Martin Drysdale, Adam Collier, Andy Massey, Nick Davies, et al. Structure-activity relationships in purine-based inhibitor binding to hsp90 isoforms. *Chemistry & biology*, 11(6):775–785, 2004.

Tatsuya Yoshizawa, Shoichi Ishida, Tomohiro Sato, Masateru Ohta, Teruki Honma, and Kei Terayama. Selective inhibitor design for kinase homologs using multiobjective monte carlo tree search. *Journal of Chemical Information and Modeling*, 62(22):5351–5360, 2022.

Barbara Zdrazil. Fifteen years of chembl and its role in cheminformatics and drug discovery. *Journal of Cheminformatics*, 17(1):1–9, 2025.

Yurong Zou, Tao Guo, Zhiyuan Fu, Zhongning Guo, Weichen Bo, Dengjie Yan, Qiantao Wang, Jun Zeng, Dingguo Xu, Taijin Wang, et al. A structure-based framework for selective inhibitor design and optimization. *Communications Biology*, 8(1):422, 2025.

A Appendix

A.1 Selection Strategies

Clustered: We use BitBIRCH-Lean (Pickering et al., 2025b) to cluster the eMolecules in-stock library. We follow the BitBirch best practices (Pickering et al., 2025a) to perfrom clustering. This involves using the instant similarity (iSIM) framework (López-Pérez et al., 2024) to calculate the average and standard deviation of the pairwise similarity of molecules within the library, which yields an optimal threshold for clustering. The threshold is set to 0.336 and the branching factor is set to 1024. This yields 13359 distinct clusters, with 1237 singleton clusters, and 1563 clusters with more than 100 members. From each of the 1000 largest clusters, we select its medoid for inclusion in our initial molecule selection.

Stratified: We evaluated the entire filtered in-stock library on the Vinardo (Quiroga & Villarreal, 2016) score through GNINA (McNutt et al., 2025) on each of the targets and off-targets. For each molecule, plausible ionization states are generated, 3D conformers for each ionization state are generated, and each conformer is docked using GNINA. The lowest Vinardo score across all ionization states, conformers, and docked poses for a given molecule is attributed to that molecule before stratification. The stratification method tries to best represent the joint distribution of these assigned Vinardo scores across each set of target and off-target triplets. For each protein, the distribution of Vinardo scores is separated into deciles, and molecules are assigned into triplet bins based on the decile they fall into for each of the on-target and off-targets. Finally, 1000 samples are drawn in a round-robin manner from each bin, starting from the bins with the lowest average Vinardo score (highest estimated affinity) to the highest. For all joint distributions in this paper, no bin was sampled more than twice.

A.2 Off-policy Training

In ALPAQAFlow, 50% of each training batch comes from a labeled dataset $D = \{(\mathbf{x}_1, y_1), ..., (\mathbf{x}_n, y_n)\}$, where each y_i is an evaluation of our computational oracle on $\mathbf{x_i}$. Trajectory balance (Eq. (1)) allows for off-policy samples because it is *offline*, meaning that predicted flows do not depend on the policy used to sample trajectories (Jain et al., 2022). Specifically, if we

use off-policy samples during training, the optimal policy π_θ^* remains the same under the assumption that sampled trajectories sufficiently cover the space of all possible trajectories (Bengio et al., 2021).

To implement off-policy training in ALPAQAFlow, for each data point $(\mathbf{x}_i, y_i) \in D$, we run exhaustive retrosynthetic tree search on $\mathbf{x}_i$ using the reaction templates $\mathcal{R}$ and building blocks $\mathcal{B}$ that make up the chemical space of ALPAQAFlow, yielding all valid trajectories terminating in $\mathbf{x}_i$. At each training iteration, a random sample is taken from D, and for each sampled data point, a valid trajectory is randomly chosen to be included in the batch.

A.3 Surrogate Models

For all surrogate models, molecules are featurized using the Morgan fingerprint (Rogers & Hahn, 2010) with a bit length of 2048 and a radius of 3.

For each protein target, we fit a random forest classifier on the thresholded IC_{50} data to distinguish between negative and positive samples. We use the random forest implementation from scikit-learn (Pedregosa et al., 2011). We perform a hyperparameter search over the parameters specified in Table 1 using Bayesian optimization for 20 iterations optimizing 5-fold cross-validation.

Hyperparameter	Values
`max_depth`	$\{5, 10, 20, 40, \text{None}\}$
`max_features`	$\{0.3, 0.5, 1.0\}$
`min_samples_split`	$\{2, 3, 4, 8\}$
`n_estimators`	$\{100, 200, 300\}$

Table 1: Hyperparameter search space for random forest classifier.

Once trained, we use the classifier to label negative samples for its respective protein target. Specifically, for a protein target with an associated threshold η, and a classifier $\mathcal{C}$, we label negative samples with a classifier weighted threshold,

$$\tilde{\text{IC}}_{50}(\mathbf{x}) = \max\{P_{\mathcal{C}}(\mathbf{x} \mid \text{IC}_{50}(\mathbf{x}) > \eta) + 0.5, 1\} \cdot \eta,$$

with the rationale that molecules with $P_{\mathcal{C}} \approx 0.5$ should have an IC_{50} value close to η.

Next, we fit a Gaussian process regressor that uses a scaled Tanimoto (Jaccard) distance for the kernel,

$$k(\mathbf{x}_i, \mathbf{x}_j) = \sigma^2 \left(\frac{\mathbf{x}_i \cdot \mathbf{x}_j}{||\mathbf{x}_i||^2 + ||\mathbf{x}_j||^2 - \mathbf{x}_i \cdot \mathbf{x}_j} \right).$$

A hyperparameter search is performed over the variance parameter σ using Bayesian optimization for 20 iterations, optimizing 5-fold cross-validation. This regressor is trained on the pIC_{50} values from positive samples and the approximated $\text{p}\tilde{\text{IC}}_{50}$ values on the negative samples from the classifier weighted threshold in the previous paragraph.

A.4 Training Details

Each individual training run is performed on an NVIDIA A100 GPU with 80 GB of VRAM. We kept most of the same hyperparameters as Shen et al. (2025). The only general hyperparameters relevant to our workflow can be found in `https://github.com/tsa87/cgflow/blob/main/src/gflownet/algo/config.py`. In particular, we set the maximum trajectory length to 3, to only allow for molecules composed of up to 3 synthesis steps. We set the maximum number of nodes allowed in a generated graph to 50, allowing for the generation of molecules with up to 50 heavy atoms. This matches the size of the largest molecule found in the eMolecules in-stock library (eMolecules, Inc., 2026). We also modify the batch size according to our hardware, setting it to 170, and the corresponding number of on-policy and off-policy samples each set to 85. All relevant hyperparameters with their variable name in `gflownet/algo/config.py` can be found in Table 2.

Hyperparameter	Values
`max_len`	3
`max_nodes`	51
`batch_size`	170
`num_from_policy`	85
`num_from_dataset`	85

Table 2: Hyperparameters used when training ALPAQAFlow.

We include a plot of the average sampled potency of molecules during training for both the on-target and two off-targets for each selection strategy and target series in Fig. 3. We plot the exponential moving average of each value (bold lines) with a span of 100, and its volatility (shaded regions) by taking the standard deviation of the rolling average with a window equal to 10.

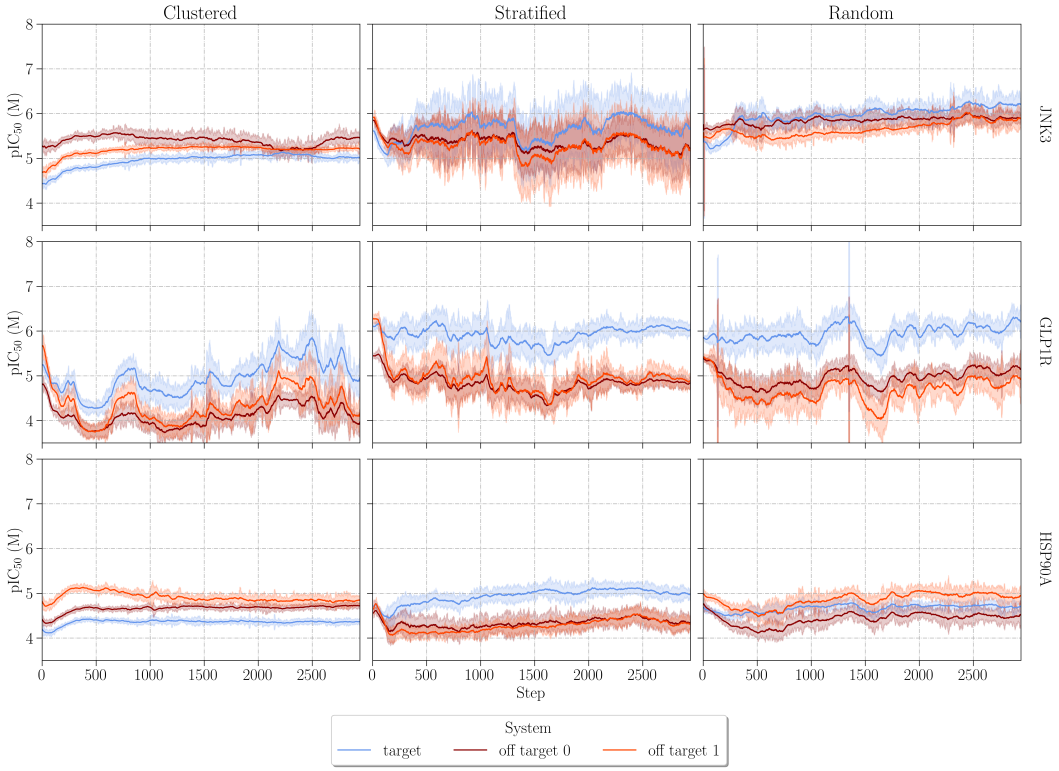

Figure 3: Average potency of molecules to all systems per training batch. On-target potency is shown in blue, while off-targets potency is shown in red. Individual plots for all target series and selection strategies are included.

A.5 Custom filters for unwanted groups

Additional custom filters for unwanted groups applied to both the in-stock library and the molecular building blocks are shown in Table 3.

A.6 Full Results

Table 4 contains the full extended results for predicted statistics according to our surrogate models and true statistics according to Boltz-2 (Passaro et al., 2025) for all selection strategies, methods, and target series.

Filtered groups	SMARTS applied
Br	`[Br]`
I	`[I]`
Aldehyde	`[CX3H1](=O)[#6]`
Azide	`[$([N-][N+]#N),$([N]=[N+]=[N-])]`
Ester	`[#6][CX3](=O)[OX2H0][#6]`
Ketone	`[#6][CX3](=O)[#6]`
Acetal	`[CX4]([OX2H0])[OX2H0]`
Alkyne	`C#C`
Olefin	`[CX3]=[CX3]`
Formamide	`[NX3][CX3H1](=O)`
Hemiacetal	`[CX4]([OX2H1])[OX2H0]`
Alkyl-NO	`[CX4][N][O]`
Thiol	`[SX2H]`
Epoxide	`[#6]1[O][#6]1`
Aminal	`[NX3][CX4][NX3]`
Hemiaminal	`[NX3][CX4][OX2H]`

Table 3: Additional substructure filters to remove potentially unwanted functional groups. Some filters are redundant with substructure filters that match PAINS and Brenk moieties.

	Target	Method	$\text{Potency}_{\text{pred}}$ (↑)		$\text{Selectivity}_{\text{pred}}$ (↑)		Potency (↑)		Selectivity (↑)		# Hits (↑)	# Series (↑)
			Avg	Top 10%	Avg	Top 10%	Avg	Top 10%	Avg	Top 10%		
Clustered	JNK3	$\text{eMolecules}_{\text{init}}$	-	-	-	-	4.77 ± 0.58	5.76 ± 0.27	-0.34 ± 0.29	0.15 ± 0.10	0	0
		$\text{eMolecules}_{\text{final}}$	4.43 ± 0.54	5.29 ± 0.12	0.43 ± 0.20	0.72 ± 0.08	5.08 ± 0.47	5.98 ± 0.25	-0.24 ± 0.29	0.29 ± 0.17	0	0
		ALPAQAFlow	5.52 ± 0.20	5.86 ± 0.08	0.35 ± 0.14	0.53 ± 0.03	5.66 ± 0.54	6.64 ± 0.26	-0.35 ± 0.32	0.19 ± 0.10	0	0
	GLP1R	$\text{eMolecules}_{\text{init}}$	-	-	-	-	5.00 ± 0.48	5.85 ± 0.21	-0.07 ± 0.36	0.55 ± 0.16	0	0
		$\text{eMolecules}_{\text{final}}$	5.34 ± 0.60	6.31 ± 0.23	1.42 ± 0.35	1.91 ± 0.09	5.50 ± 0.44	6.30 ± 0.19	0.07 ± 0.36	0.65 ± 0.14	0	0
		ALPAQAFlow	6.19 ± 0.46	6.86 ± 0.12	1.33 ± 0.28	1.72 ± 0.10	6.22 ± 0.54	7.13 ± 0.19	-0.08 ± 0.39	0.58 ± 0.20	3	1
	HSP90A	$\text{eMolecules}_{\text{init}}$	-	-	-	-	4.85 ± 0.11	5.76 ± 0.08	-0.19 ± 0.11	0.32 ± 0.17	0	0
		$\text{eMolecules}_{\text{final}}$	4.66 ± 0.49	5.50 ± 0.59	0.81 ± 0.44	1.21 ± 0.51	5.26 ± 0.18	6.10 ± 0.14	-0.07 ± 0.13	0.45 ± 0.15	0	0
		ALPAQAFlow	4.70 ± 0.26	5.10 ± 0.07	0.01 ± 0.21	0.29 ± 0.06	5.81 ± 0.48	6.67 ± 0.18	-0.29 ± 0.35	0.26 ± 0.10	0	0
	Summary	$\text{eMolecules}_{\text{init}}$	-	-	-	-	4.77 ± 0.17	5.67 ± 0.17	-0.19 ± 0.09	0.32 ± 0.15	0	0
		$\text{eMolecules}_{\text{final}}$	4.77 ± 0.47	5.62 ± 0.55	0.98 ± 0.48	1.44 ± 0.59	5.26 ± 0.15	6.08 ± 0.13	-0.01 ± 0.16	0.50 ± 0.16	0	0
		ALPAQAFlow	5.47 ± 0.61	5.94 ± 0.72	0.56 ± 0.56	0.85 ± 0.63	5.90 ± 0.24	6.82 ± 0.22	-0.24 ± 0.12	0.34 ± 0.17	3	1
Stratified	JNK3	$\text{eMolecules}_{\text{init}}$	-	-	-	-	5.21 ± 0.60	6.28 ± 0.24	-0.36 ± 0.31	0.17 ± 0.12	0	0
		$\text{eMolecules}_{\text{final}}$	5.36 ± 0.92	6.67 ± 0.16	0.20 ± 0.47	0.86 ± 0.21	5.24 ± 0.53	6.21 ± 0.25	-0.32 ± 0.31	0.19 ± 0.13	0	0
		ALPAQAFlow	6.42 ± 0.54	7.06 ± 0.12	0.65 ± 0.18	0.91 ± 0.06	5.83 ± 0.69	6.94 ± 0.25	-0.24 ± 0.33	0.31 ± 0.12	0	0
	GLP1R	$\text{eMolecules}_{\text{init}}$	-	-	-	-	5.33 ± 0.50	6.24 ± 0.22	-0.11 ± 0.35	0.49 ± 0.16	1	0
		$\text{eMolecules}_{\text{final}}$	5.35 ± 0.70	6.40 ± 0.21	1.07 ± 0.35	1.63 ± 0.15	5.40 ± 0.46	6.22 ± 0.27	0.05 ± 0.34	0.65 ± 0.12	2	0
		ALPAQAFlow	6.66 ± 0.47	7.46 ± 0.05	1.60 ± 0.22	1.92 ± 0.06	6.21 ± 0.57	7.17 ± 0.19	0.18 ± 0.35	0.78 ± 0.15	3	0
	HSP90A	$\text{eMolecules}_{\text{init}}$	-	-	-	-	5.07 ± 0.52	6.04 ± 0.28	-0.12 ± 0.29	0.33 ± 0.10	0	0
		$\text{eMolecules}_{\text{final}}$	4.63 ± 0.47	5.41 ± 0.14	0.80 ± 0.24	1.14 ± 0.08	5.20 ± 0.48	6.07 ± 0.23	-0.02 ± 0.30	0.46 ± 0.13	0	0
		ALPAQAFlow	5.03 ± 0.33	5.70 ± 0.17	0.92 ± 0.16	1.14 ± 0.05	5.52 ± 0.47	6.34 ± 0.18	-0.31 ± 0.40	0.32 ± 0.18	1	0
	Summary	$\text{eMolecules}_{\text{init}}$	-	-	-	-	5.20 ± 0.11	6.19 ± 0.11	-0.20 ± 0.12	0.33 ± 0.13	1	0
		$\text{eMolecules}_{\text{final}}$	5.11 ± 0.34	6.16 ± 0.54	0.69 ± 0.36	1.21 ± 0.32	5.28 ± 0.09	6.17 ± 0.07	-0.10 ± 0.16	0.43 ± 0.19	2	0
		ALPAQAFlow	6.04 ± 0.72	6.74 ± 0.76	1.06 ± 0.40	1.32 ± 0.44	5.85 ± 0.28	6.82 ± 0.35	-0.13 ± 0.22	0.47 ± 0.22	4	0
Random	JNK3	$\text{eMolecules}_{\text{init}}$	-	-	-	-	4.99 ± 0.55	5.96 ± 0.22	-0.32 ± 0.34	0.26 ± 0.13	0	0
		$\text{eMolecules}_{\text{final}}$	4.88 ± 0.72	6.03 ± 0.20	0.63 ± 0.36	1.26 ± 0.30	5.20 ± 0.63	6.31 ± 0.31	-0.16 ± 0.35	0.46 ± 0.16	0	0
		ALPAQAFlow	7.10 ± 0.28	7.50 ± 0.06	0.78 ± 0.16	1.02 ± 0.05	6.40 ± 0.51	7.33 ± 0.19	-0.29 ± 0.34	0.32 ± 0.18	0	0
	GLP1R	$\text{eMolecules}_{\text{init}}$	-	-	-	-	5.26 ± 0.49	6.16 ± 0.25	-0.07 ± 0.36	0.54 ± 0.14	2	0
		$\text{eMolecules}_{\text{final}}$	5.87 ± 0.60	6.89 ± 0.19	1.50 ± 0.29	1.97 ± 0.12	5.68 ± 0.42	6.42 ± 0.16	0.11 ± 0.36	0.72 ± 0.14	0	0
		ALPAQAFlow	7.38 ± 0.33	7.87 ± 0.09	1.81 ± 0.23	2.13 ± 0.05	6.71 ± 0.51	7.56 ± 0.18	-0.02 ± 0.37	0.64 ± 0.17	3	2
	HSP90A	$\text{eMolecules}_{\text{init}}$	-	-	-	-	4.94 ± 0.49	5.85 ± 0.24	-0.13 ± 0.27	0.31 ± 0.12	0	0
		$\text{eMolecules}_{\text{final}}$	4.21 ± 0.45	4.94 ± 0.15	0.77 ± 0.26	1.18 ± 0.10	5.13 ± 0.43	5.92 ± 0.21	-0.05 ± 0.28	0.40 ± 0.11	0	0
		ALPAQAFlow	4.53 ± 0.37	5.05 ± 0.10	0.52 ± 0.23	0.85 ± 0.09	5.49 ± 0.50	6.39 ± 0.20	-0.16 ± 0.32	0.30 ± 0.12	1	0
	Summary	$\text{eMolecules}_{\text{init}}$	-	-	-	-	5.06 ± 0.14	5.99 ± 0.13	-0.17 ± 0.11	0.37 ± 0.12	2	0
		$\text{eMolecules}_{\text{final}}$	4.98 ± 0.68	5.95 ± 0.80	0.97 ± 0.38	1.47 ± 0.36	5.34 ± 0.24	6.22 ± 0.22	-0.03 ± 0.11	0.53 ± 0.14	0	0
		ALPAQAFlow	6.34 ± 1.28	6.81 ± 1.25	1.04 ± 0.56	1.34 ± 0.57	6.20 ± 0.52	7.09 ± 0.50	-0.16 ± 0.11	0.42 ± 0.15	4	2

Table 4: Full results of our simulated campaign. We report the potency and selectivity according to our surrogate models, the potency and selectivity according to Boltz-2 (Passaro et al., 2025), as well as number of hits and number of chemical series for each target series. Additionally, we report aggregate statistics for each selection strategy, reporting the average potency, and selectivity (both according to our oracle and our surrogates) as well as the sum of hits and chemical series found.