PREFERENTIAL MULTI-OBJECTIVE BAYESIAN OPTIMIZATION FOR DRUG DISCOVERY

Anonymous authors
Paper under double-blind review

ABSTRACT

Despite decades of advancements in automated ligand screening, large-scale docking remains resource-intensive and requires post-processing hit selection, a step where chemists manually select a few promising molecules based on their chemical intuition. This creates a major bottleneck in the virtual screening process for drug discovery, demanding experts to repeatedly balance complex trade-offs among drug properties across a vast pool of candidates. To improve the efficiency and reliability of this process, we propose a novel human-centered framework *CheapVS* that allows chemists to guide the ligand selection process through pairwise preference feedback. Our framework combines preferential multi-objective Bayesian optimization with an efficient diffusion docking model to capture human chemical intuition for improving hit identification. Specifically, on a library of 100K chemical candidates that target EGFR, a cancer-associated protein, *CheapVS* outperforms state-of-the-art docking methods in identifying drugs within a limited computational budget. Notably, our multi-objective algorithm can recover up to 16 out of 37 known drugs while scanning only 6% of the library, showcasing its potential to advance drug discovery[1].

1 INTRODUCTION

Virtual screening (VS) is a key pillar of modern computational drug discovery, acting as a rapid sift through massive molecular libraries-ranging from millions to billions of compounds-to identify a set of "hits" with promising therapeutic potential (Shoichet, 2004; Lyu et al., 2023). At the core of VS lies hit selection: the practical step in which a set of candidate compounds is manually chosen from top-ranked VS results, informed not only by binding affinity scores but also by key factors such as solubility, toxicity, and pharmacokinetic properties—all of which collectively determine a compound's potential. Despite its centrality to drug discovery, VS and hit selection remain both resource-intensive and laborious. Traditional pipelines rely on exhaustive docking of the entire library, which demands substantial time and computational resources (Lyu et al., 2019; Gorgulla et al., 2020). Moreover, human expertise is required in the loop: medicinal chemists must examine the docking results to finalize which hits are worthy of costly experimental validation. In large-scale projects with millions of compounds, this process quickly becomes bottlenecked by both the computational costs and the limited bandwidth of expert reviewers. On top of the issue, vast computational resources are spent on docking a large number of unpromising (later-known low-scored) compounds, even though only a small fraction of top-ranked molecules typically move forward for hit selection and experimental validation. To address this problem, recent methods have combined active learning with molecular docking to query compounds based on predicted binding affinity, substantially reducing computational overhead while maintaining high accuracy (Graff et al., 2021; Zhou et al., 2024; Gentile et al., 2020).

Although binding molecules are good starting points for screening campaigns, it is critical to emphasize that the drug discovery process, in its entirety, is a complex multi-objective optimization problem. Indeed, VS presents a unique challenge due to its operation in a high-dimensional search space where these objectives (e.g., binding affinity, solubility, toxicity, pharmacokinetic properties, etc.) exhibit complex and often poorly understood inter-dependencies (Hann & Keserü, 2012). For instance, adding bulky functional groups can enhance binding affinity but simultaneously lower solubility or

[1]Code and data for these experiments can be found at `https://anonymous.4open.science/r/vs-9A83`

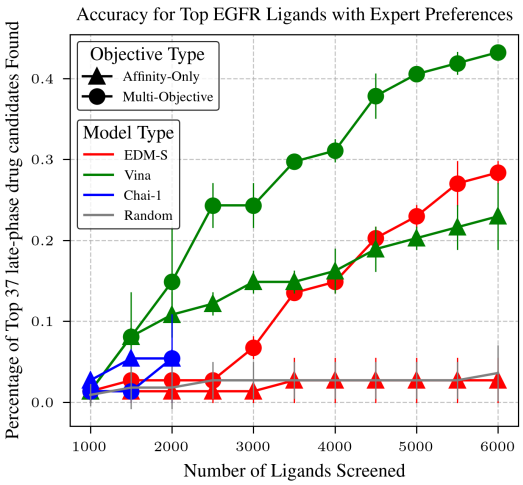


Figure 1: Performance of Chemist-guided Preferential Multi-Objective Bayesian Optimization in identifying EGFR drugs using qEUBO acquisition function. The search is conducted on a 100K ligand library, screened for a maximum of 6% of the library. The plot compares different docking models (Vina, EDM-S, Chai-1) and objective types. The y-axis shows the percentage of the top 37 approved drugs identified, while the x-axis represents the number of ligands screened. Multi-objective optimization (circles) across all docking models outperforms affinity-only selection (triangles) and random screening (gray line), with the best-performing method discovering up to 16 out of 37 drugs. The limited 48-hour docking time constraint explains the incomplete results for Chai-1.

increase off-target toxicity, complicating the search for high-potential candidates. This underscores the multi-objective nature of VS, where optimizing multiple correlated and potentially competing properties is central to finding robust drug leads. The ultra-large scale of ligand libraries and the sparse availability of experimental data make VS particularly challenging compared to conventional multi-objective optimization problems.

To address these limitations, we present a novel framework named *CheapVS* (CHEmist-guided Active Preferential Virtual Screening) that assists chemists in expert-guided VS by leveraging a preferential multi-objective Bayesian optimization (BO) toolbox. By translating expert chemists' nuanced understanding into multi-objective utility functions - incorporating factors such as binding affinity, solubility, or toxicity—our framework ensures that computational optimization captures subtle trade-offs that purely algorithmic methods often overlook. This expert-guided approach refines the VS process, prioritizing candidates based on binding profiles and broader criteria crucial for downstream development. In doing so, we aim to make the hit identification process more efficient and aligned with expert preference and, ultimately, more effective in discovering promising drug leads from vast chemical spaces. Our key contributions are:

- **Expert-Driven Preference Elicitation:** We capture expert chemists' priorities via preference rankings, translating their insights into multi-objective utility functions. This ensures computational optimization mirrors the nuanced decision-making in drug discovery.
- **Multi-Objective Active Virtual Screening:** Leveraging these utility functions, *CheapVS* optimizes multiple criteria (e.g., binding affinity, solubility, toxicity) concurrently rather than focusing on a single objective, leading to more balanced and clinically relevant candidates.
- **Enhanced Diffusion Models for Molecular Docking:** We introduce an improved diffusion-based docking approach that employs extensive data augmentation, training on both a large synthetic and a curated high-quality dataset.

2 METHODS

We tackle the problem of identifying the top k candidate ligands for a given protein, starting from a screening library $\mathcal{L} = \{l_1, \ldots, l_N\}$. The goal is to select ligands with the highest potential to succeed as drugs, taking into account multiple objectives (also referred to as features), such as binding affinity, toxicity, solubility, or synthesizability. To achieve this, we define a descriptor function $\phi : \mathcal{L} \to \mathbb{R}^d$, which maps ligands from the space $\mathcal{L}$ to a d-dimensional feature space. This mapping is often computationally expensive to evaluate, especially for objectives like binding affinity. While binding affinity is often prioritized, other factors are equally important. This task is commonly formulated as a multi-objective optimization problem in VS. Here, balancing objectives is often challenging, as the optimal trade-offs are influenced by expert preferences. A key aspect of this method involves eliciting

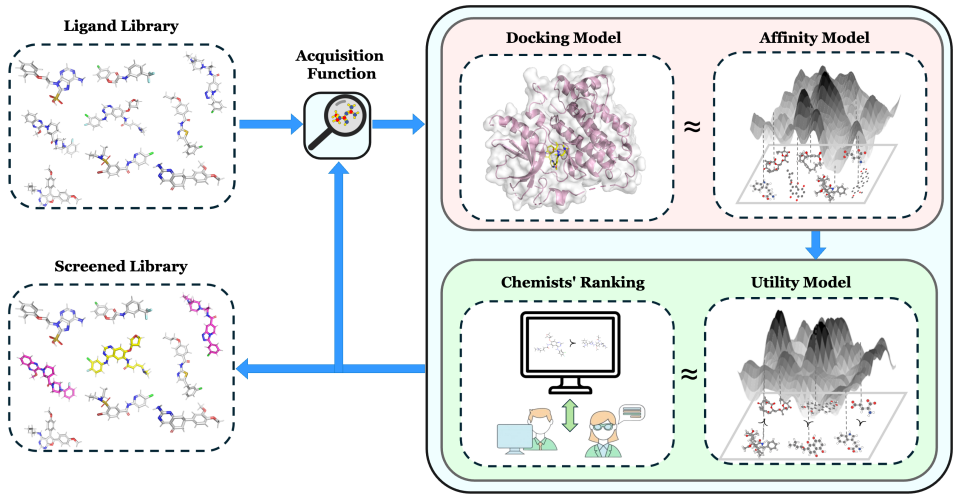


Figure 2: Overview of Human Preferential Bayesian Optimization for multi-objective virtual screening. Ligands from a large library are selected using an acquisition function and evaluated through docking and affinity models. Chemists provide preference rankings, which inform a utility model to refine the selection process. The screened library iteratively improves, prioritizing ligands with desirable properties. Yellow-colored ligands represent found drug compounds, while purple ligands indicate screened compounds.

expert preferences to guide the optimization process. Preferences are gathered through pairwise comparisons of ligands, which are then used to model a latent utility function.

We model two unknown functions, a utility function f and an affinity function g, using Gaussian Processes (GPs). Following Chu & Ghahramani (2005); Brochu et al. (2010), we assume that g is associated with a Gaussian likelihood, while f uses a Bernoulli likelihood to capture preferences. Concretely, the probability of preferring ligand $\mathbf{x}_1$ over $\mathbf{x}_2$ is defined by $p(y = 1 \mid \mathbf{x}_1, \mathbf{x}_2) = \sigma(f(\mathbf{x}_1) - f(\mathbf{x}_2))$, where σ is the sigmoid function. Here, $\mathbf{x}$ represents the ligand features. The function f itself is drawn from a GP prior, $f \sim \mathcal{GP}(\mu(\mathbf{X}), k(\mathbf{X}, \mathbf{X}'))$, with $\mu(\mathbf{X})$ and $k(\mathbf{X}, \mathbf{X}')$ denoting the mean and kernel functions, respectively.

The ligand l_i is represented by its Morgan fingerprint t_i, which is input to the predicted binding affinity function $g(t_i)$. The predicted binding affinity $g(t)$ then serves as an input to f, enabling an online optimization process that balances multiple objectives. Because g influences f, any uncertainty in g's predictions can propagate through the optimization pipeline, ultimately degrading the utility predictions and affecting the performance of multi-objective optimization. To mitigate this, we explicitly model the uncertainty in g. We treat the predicted binding affinity $a = g(t)$ as a random variable and let its observed value a_{obs} follow a distribution $p(a_{\text{obs}} \mid a)$. Then, instead of maximizing the acquisition function α at a single point estimate, we integrate over this uncertainty: $a^* = \arg\max_a \mathbb{E}_{p(a_{\text{obs}}|a)}[\alpha(a_{\text{obs}})]$. We assume $p(a_{\text{obs}} \mid a)$ follows a Gaussian posterior predictive distribution $p(a_{\text{obs}} \mid t, \mathcal{L})$. This expectation can be approximated via Monte Carlo sampling. As a result, its impact on f is reduced, making the VS more robust.

The process starts by randomly sampling a small fraction of the entire ligand library to form the initial training set. These ligands are docked with a model θ, yielding binding affinities a, which train an affinity predictor g. Random ligand pairs drawn from this subset receive preference labels (indicating which ligand is better based on expert criteria) to train a utility predictor f. After this initialization, an iterative optimization phase begins. At each iteration, newly docked affinities update g, whose refined predictions are then used to further improve f. The utility model then predicts utilities for the remaining compounds, and an acquisition function $\alpha(t)$ ranks them by balancing exploration and exploitation. The top-ranked ligands are docked into the target protein P, their affinities a are recorded, and the data is added to retrain g. This iterative process continually improves the search for optimal ligands. The choice of both f and g as GPs is due to their superior performance over Neural Networks and Decision Trees as detailed in Appendix F. To evaluate the optimization procedure, we

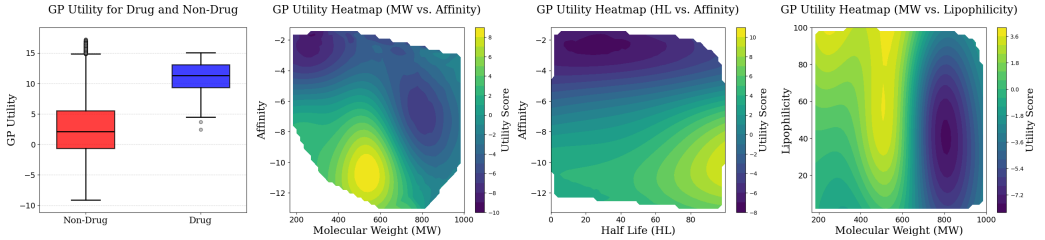


Figure 3: Predictive utility scores after BO on expert preference elicitation. The box plot compares drugs and non-drug compounds, highlighting the distinctive expert intuition. Heatmaps illustrate utility over two objectives while keeping others at their mean. Results align well with established medicinal chemistry ranges, favoring optimal MW (400-600) for absorption, while maximizing binding affinity and half-life (drug retention time). Lipophilicity can be flexibly adjusted for either solubility or permeability. The properties' distribution further demonstrated the algorithm's explainability and its ability to capture domain knowledge in distinguishing drug-like compounds.

employ two key metrics. *Regret* at iteration i is defined as $R_i = U^* - U(i)$, where U^* is the highest possible utility in the library and $U(i)$ is the highest utility found by the model at iteration i, with lower regret indicating more efficient identification of optimal candidates. *Top-k Accuracy* measures the proportion of correctly identified compounds within the top-k set. These metrics comprehensively assess the model's efficiency and accuracy in discovering drug-like ligands.

In addition, we improve the speed of molecular docking by integrating a diffusion model. This type of model refines ligand poses by simulating natural diffusion, improving the exploration of conformational space. The resulting refined poses and energies enhance the accuracy of our binding affinity predictor, $g(t)$. This integration bridges the gap between high-throughput screening and accurate binding prediction, resulting in a more robust and reliable VS process. Traditional force field-based docking methods, e.g., (McNutt et al., 2021; Eberhardt et al., 2021; Koes et al., 2013), are computationally expensive for large-scale VS. Deep learning offers a faster alternative, maintaining accuracy while significantly reducing computational time for the pose sampling process. Our diffusion model is based on EDM (Abramson et al., 2024; Karras et al., 2022). The model iteratively refines ligand coordinates, mimicking binding adjustments while keeping protein coordinates fixed. Training occurs in two phases: pretraining on our 11 million synthetic protein-ligand pairs (generated via pharmacophore alignment) to learn general docking patterns, followed by fine-tuning on a combination set of PDBScan22 ($\sim$180,000 experimental protein-ligand complexes). The rationale behind these dataset choices is detailed in Appendix G.

3 Experiments

We conduct the experiments in two main stages. First, we explore real-world human-labeled preferences for a realistic drug discovery problem targeting the EGFR protein. Next, we compare different docking models to assess how computational cost and accuracy trade-offs shape the final optimization performance. We investigate the `qEUBO` acquisition strategy (Astudillo et al., 2023). A comprehensive analysis of these acquisition functions can be found in Appendix C. We calculate physicochemical properties using (RDKit, online) and incorporate ADMET prediction (absorption, distribution, metabolism, excretion, and toxicity) (Swanson et al., 2024) to better capture realistic multi-objective trade-offs. We aim to address the following Research Questions (RQ):

- **RQ1**: How effectively does *Human-Preferential Multi-Objective Optimization* identify clinically relevant EGFR ligands compared to single-objective (affinity-only) or random baselines, and does incorporating additional drug-like properties beyond affinity lead to better real-world performance?
- **RQ2**: What is the effect of data augmentation on model generalization in molecular docking?

3.1 Human-in-the-loop Experiment

To address *RQ1*, we focus on the Epidermal Growth Factor Receptor (EGFR) protein due to its clinical importance and the availability of multiple FDA-approved drugs (Cohen et al., 2021). We collect

37 FDA-approved or late-stage clinical candidates from the PKIDB and DrugBank (Carles et al., 2018; Knox et al., 2024), treating them as "goal-optimal" molecules. The screening library comprises $\sim 260,000$ molecules from García-Ortegón et al. (2022), in which a random subset of $100,000$ is used to simulate a realistic campaign. Expert chemists provide preference labels, thereby constructing nuanced *multi-objective* utility functions reflecting real-world considerations of physicochemical and ADMET properties, including affinity, molecular weight, lipophilicity, and half-life. In each BO iteration, these experts complete 200 pairwise comparisons, yielding a total of $2,200$ pairs over one full experiment of the study. We also evaluate a *single-objective* baseline focusing solely on binding affinity to see whether multi-property feedback yields more meaningful optimization for VS.

The BO pipeline begins by randomly sampling 1.0% of the $100,000$-compound library, then screening an additional 0.5% per iteration for 10 iterations (covering a total of 6% of the library). All experiments are run with two random seeds, and the overall docking time is constrained to **48 hours**. Chai-1 is too computationally expensive to complete 6000 dockings, EDM-S finishes in 40 hours, and Vina requires no docking as affinities are precomputed from García-Ortegón et al. (2022).

Figure 1 shows how effectively each approach (Multi-Objective, Affinity-Only, and Random) with different dock models (Vina, Chai, EDM-S) identifies the 37 known EGFR ligands. The Vina (Multi-Objective) strategy, guided by expert preferences, attains about **42%** accuracy in retrieving these known drugs, substantially surpassing the **22%** accuracy of the affinity-only approach. EDM-S (Multi-Objective) reaches up to 28%, whereas its Affinity-Only counterpart lags around 2%. Although Chai-1's multi-objective performance improves over time, the strict 48-hour time constraint restricts the number of iterations, limiting its final performance. Random screening performs poorly. These findings *directly address RQ1*: leveraging expert preference leads to more clinically relevant molecules than relying solely on affinity. Moreover, adding additional drug-like properties significantly enhances performance in a realistic drug discovery pipeline. We observe that *all three docking models* (EDM-S, Chai-1, Vina) show improved performance when incorporating multi-objective preferences over single-objective affinity, emphasizing the broad advantage of reflecting real-world trade-offs in the BO process.

3.2 Diffusion Training Result

Binding affinity (measured in kilocalories per mole) was evaluated using EGFR as the target, and results from all VS experiments were collected using the Vinardo scoring function. As shown in Figure 13, EDM-S outperforms Chai in binding affinity predictions, highlighting its advantage in robust sampling. While EDM-S is slightly outperformed by Vina, it achieves comparable binding affinity distributions, showcasing its ability to balance efficiency and precision. These results, combined with its faster runtime, make EDM-S highly suitable for large-scale VS.

Regarding Root Mean Square Deviation (RMSD) performance, EDM-S, and DockScan22 (a DiffDock-S trained on our PDBScan22) employ distinct training strategies. RMSD, measured in Ångströms (Å, where 1Å = 0.1 nm), quantifies structural deviations between predicted and reference ligand poses, with lower values indicating higher accuracy. EDM-S combines pretraining on the PapyrusScan dataset (11M synthetic pairs) with fine-tuning on PDBScan22 (322K validated complexes), while DockScan22 trains solely on PDBScan22 using DiffDock-S as its backbone. Experimental results (6) show DockScan22 achieves 54.1% and 34.1% accuracy (RMSD $<$ 2Å) on PoseBuster V1 and PDBBind, outperforming DiffDock-S and the original DiffDock. EDM-S achieves 91% accuracy (RMSD $<$ 5Å) on PoseBuster V1. Both models are 34 times faster than folding models like AlphaFold, demonstrating their practicality for large-scale applications.

4 Conclusion

We present a framework for accelerating drug discovery with preferential multi-objective BO. *CheapVS* enables a deeper understanding of how incorporating chemical intuition can enhance the practicality of the VS process. By addressing the challenges chemists often face during hit identification, *CheapVS* speeds up the VS process. It requires screening only a small subset of the ligand library and leveraging a few chemists' pairwise preferences to efficiently identify drug-like compounds. This paper opens exciting avenues for future research. *CheapVS* relies on pairwise preference and is well-suited for listwise preference. Here, chemists can select the best ligand from a list, providing richer preference information and further boosting algorithm performance. Future work would benefit from exploring advanced preference modeling, to enable deeper insights into and further accelerating the drug discovery process.

REFERENCES

Josh Abramson, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, Olaf Ronneberger, Lindsay Willmore, Andrew J. Ballard, Joshua Bambrick, Sebastian W. Bodenstein, David A. Evans, Chia-Chun Hung, Michael O'Neill, David Reiman, Kathryn Tunyasuvunakool, Zachary Wu, Akvilė Žemgulytė, Eirini Arvaniti, Charles Beattie, Ottavia Bertolli, Alex Bridgland, Alexey Cherepanov, Miles Congreve, Alexander I. Cowen-Rivers, Andrew Cowie, Michael Figurnov, Fabian B. Fuchs, Hannah Gladman, Rishub Jain, Yousuf A. Khan, Caroline M. R. Low, Kuba Perlin, Anna Potapenko, Pascal Savy, Sukhdeep Singh, Adrian Stecula, Ashok Thillaisundaram, Catherine Tong, Sergei Yakneen, Ellen D. Zhong, Michal Zielinski, Augustin Žídek, Victor Bapst, Pushmeet Kohli, Max Jaderberg, Demis Hassabis, and John M. Jumper. Accurate structure prediction of biomolecular interactions with alphafold 3. *Nature*, 630(8016):493–500, 2024. doi: 10.1038/s41586-024-07487-w. URL https://doi.org/10.1038/s41586-024-07487-w.

A. Acharya, R. Agarwal, M. B. Baker, J. Baudry, D. Bhowmik, S. Boehm, K. G. Byler, S. Y. Chen, L. Coates, C. J. Cooper, O. Demerdash, I. Daidone, J. D. Eblen, S. Ellingson, S. Forli, J. Glaser, J. C. Gumbart, J. Gunnels, O. Hernandez, S. Irle, D. W. Kneller, A. Kovalevsky, J. Larkin, T. J. Lawrence, S. LeGrand, S. H. Liu, J. C. Mitchell, G. Park, J. M. Parks, A. Pavlova, L. Petridis, D. Poole, L. Pouchard, A. Ramanathan, D. M. Rogers, D. Santos-Martins, A. Scheinberg, A. Sedova, Y. Shen, J. C. Smith, M. D. Smith, C. Soto, A. Tsaris, M. Thavappiragasam, A. F. Tillack, J. V. Vermaas, V. Q. Vuong, J. Yin, S. Yoo, M. Zahran, and L. Zanetti-Polzi. Supercomputer-based ensemble docking drug discovery pipeline with application to covid-19. *Journal of Chemical Information and Modeling*, 60(12):5832–5852, 12 2020.

Raul Astudillo, Zhiyuan Jerry Lin, Eytan Bakshy, and Peter Frazier. qeubo: A decision-theoretic acquisition function for preferential bayesian optimization. In *International Conference on Artificial Intelligence and Statistics*, pp. 1093–1114. PMLR, 2023.

Syrine Belakaria, Aryan Deshwal, and Janardhan Rao Doppa. Max-value entropy search for multi-objective bayesian optimization with constraints. *arXiv preprint arXiv:2009.01721*, 2020a.

Syrine Belakaria, Aryan Deshwal, Nitthilan Kannappan Jayakodi, and Janardhan Rao Doppa. Uncertainty-aware search framework for multi-objective bayesian optimization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pp. 10044–10052, 2020b.

Olivier JM Béquignon, Brandon J Bongers, Willem Jespers, Adriaan P IJzerman, B van der Water, and Gerard JP van Westen. Papyrus: a large-scale curated dataset aimed at bioactivity predictions. *Journal of cheminformatics*, 15(1):3, 2023.

CGE Boender. Bayesian approach to global optimization–theory and applications., 1991.

Eric Brochu, Vlad M Cora, and Nando De Freitas. A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv preprint arXiv:1012.2599*, 2010.

Martin Buttenschoen, Garrett M Morris, and Charlotte M Deane. Posebusters: Ai-based docking methods fail to generate physically valid poses or generalise to novel sequences. *Chemical Science*, 2024.

Heng Cai, Chao Shen, Tianye Jian, Xujun Zhang, Tong Chen, Xiaoqi Han, Zhuo Yang, Wei Dang, Chang-Yu Hsieh, Yu Kang, et al. Carsidock: a deep learning paradigm for accurate protein–ligand docking and screening based on large-scale pre-training. *Chemical Science*, 15(4):1449–1471, 2024.

Duanhua Cao, Geng Chen, Jiaxin Jiang, Jie Yu, Runze Zhang, Mingan Chen, Wei Zhang, Lifan Chen, Feisheng Zhong, Yingying Zhang, et al. Generic protein–ligand interaction scoring by integrating physical prior knowledge and data augmentation modelling. *Nature Machine Intelligence*, pp. 1–13, 2024.

Fabrice Carles, Stéphane Bourg, Christophe Meyer, and Pascal Bonnet. Pkidb: a curated, annotated and updated database of protein kinase inhibitors in clinical trials. *Molecules*, 23(4):908, 2018.

Wei Chu and Zoubin Ghahramani. Preference learning with gaussian processes. In *Proceedings of the 22nd International Conference on Machine Learning*, ICML '05, pp. 137–144, New York, NY, USA, 2005. Association for Computing Machinery. ISBN 1595931805. doi: 10.1145/1102351.1102369. URL https://doi.org/10.1145/1102351.1102369.

David Clark. Virtual screening: Is bigger always better? or can small be beautiful? *Journal of Chemical Information and Modeling*, 60, 05 2020. doi: 10.1021/acs.jcim.0c00101.

Philip Cohen, Darren Cross, and Pasi A Jänne. Kinase drug discovery 20 years after imatinib: progress and future directions. *Nature reviews drug discovery*, 20(7):551–569, 2021.

Gabriele Corso, Arthur Deng, Benjamin Fry, Nicholas Polizzi, Regina Barzilay, and Tommi Jaakkola. Deep confident steps to new pockets: Strategies for docking generalization. *arXiv preprint arXiv:2402.18396*, 2024.

Ivo Couckuyt, Dirk Deschrijver, and Tom Dhaene. Towards efficient multiobjective optimization: Multiobjective statistical criterions, 2012.

Samuel Daulton, Maximilian Balandat, and Eytan Bakshy. Parallel bayesian optimization of multiple noisy objectives with expected hypervolume improvement. *Advances in Neural Information Processing Systems*, 34:2187–2200, 2021.

Samuel Daulton, David Eriksson, Maximilian Balandat, and Eytan Bakshy. Multi-objective bayesian optimization over high-dimensional search spaces. In *Uncertainty in Artificial Intelligence*, pp. 507–517. PMLR, 2022.

Janani Durairaj, Yusuf Adeshina, Zhonglin Cao, Xuejin Zhang, Vladas Oleinikovas, Thomas Duignan, Zachary McClure, Xavier Robin, Daniel Kovtun, Emanuele Rossi, et al. Plinder: The protein-ligand interactions dataset and evaluation resource. *bioRxiv*, pp. 2024–07, 2024.

Jerome Eberhardt, Diogo Santos-Martins, Andreas F Tillack, and Stefano Forli. Autodock vina 1.2. 0: New docking methods, expanded force field, and python bindings. *Journal of chemical information and modeling*, 61(8):3891–3898, 2021.

Michael Emmerich, André Deutz, and Jan-Willem Klinkenberg. The computation of the expected improvement in dominated hypervolume of pareto front approximations. Technical Report LIACS-TR 4-2008, Leiden Institute for Advanced Computer Science (LIACS), Leiden, Netherlands, 09 2008.

Florian Flachsenberg, Christiane Ehrt, Torben Gutermuth, and Matthias Rarey. Redocking the pdb. *Journal of Chemical Information and Modeling*, 64(1):219–237, 2023.

Jenna C Fromer, David E Graff, and Connor W Coley. Pareto optimization to accelerate multi-objective virtual screening. *Digital Discovery*, 3(3):467–481, 2024.

Miguel García-Ortegón, Gregor NC Simm, Austin J Tripp, José Miguel Hernández-Lobato, Andreas Bender, and Sergio Bacallado. Dockstring: easy molecular docking yields better benchmarks for ligand design. *Journal of chemical information and modeling*, 62(15):3486–3502, 2022.

Francesco Gentile, Vibudh Agrawal, Michael Hsing, Anh-Tien Ton, Fuqiang Ban, Ulf Norinder, Martin E Gleave, and Artem Cherkasov. Deep docking: a deep learning platform for augmentation of structure based drug discovery. *ACS central science*, 6(6):939–949, 2020.

Christoph Gorgulla, Andras Boeszoermenyi, Zi-Fu Wang, Patrick D. Fischer, Paul W. Coote, Krishna M. Padmanabha Das, Yehor S. Malets, Dmytro S. Radchenko, Yurii S. Moroz, David A. Scott, Konstantin Fackeldey, Moritz Hoffmann, Iryna Iavniuk, Gerhard Wagner, and Haribabu Arthanari. An open-source drug discovery platform enables ultra-large virtual screens. *Nature*, 580(7805):663–668, 2020.

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. doi: 10.1039/D0SC06805E.

Michael M. Hann and György M. Keserü. Finding the sweet spot: the role of nature and nurture in medicinal chemistry. *Nature Reviews Drug Discovery*, 11(5):355–365, 2012.

Daniel Hernández-Lobato, Jose Hernandez-Lobato, Amar Shah, and Ryan Adams. Predictive entropy search for multi-objective bayesian optimization. In *International conference on machine learning*, pp. 1492–1501. PMLR, 2016.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

Aapo Hyvärinen and Peter Dayan. Estimation of non-normalized statistical models by score matching. *Journal of Machine Learning Research*, 6(4), 2005.

Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. *Advances in neural information processing systems*, 35:26565–26577, 2022.

Douglas B. Kitchen, Hélène Decornez, John R. Furr, and Jürgen Bajorath. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature Reviews Drug Discovery*, 3 (11):935–949, 2004.

J. Knowles. Parego: a hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE Transactions on Evolutionary Computation*, 10(1):50–66, 2006. doi: 10.1109/TEVC.2005.851274.

Craig Knox, Mike Wilson, Christen M Klinger, Mark Franklin, Eponine Oler, Alex Wilson, Allison Pon, Jordan Cox, Na Eun Chin, Seth A Strawbridge, et al. Drugbank 6.0: the drugbank knowledgebase for 2024. *Nucleic acids research*, 52(D1):D1265–D1275, 2024.

David Ryan Koes, Matthew P. Baumgartner, and Carlos J. Camacho. Lessons learned in empirical scoring with smina from the csar 2011 benchmarking exercise. *Journal of Chemical Information and Modeling*, 53(8):1893–1904, 08 2013. doi: 10.1021/ci300604z. URL https://doi.org/10.1021/ci300604z.

Harold J Kushner. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. 1964.

Evanthia Lionta, George Spyrou, Demetrios K Vassilatis, and Zoe Cournia. Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Current topics in medicinal chemistry*, 14(16):1923–1938, 2014.

Zhihai Liu, Minyi Su, Li Han, Jie Liu, Qifan Yang, Yan Li, and Renxiao Wang. Forging the basis for developing protein–ligand interaction scoring functions. *Accounts of chemical research*, 50(2): 302–309, 2017.

Jiankun Lyu, Sheng Wang, Trent E. Balius, Isha Singh, Anat Levit, Yurii S. Moroz, Matthew J. O'Meara, Tao Che, Enkhjargal Algaa, Kateryna Tolmachova, Andrey A. Tolmachev, Brian K. Shoichet, Bryan L. Roth, and John J. Irwin. Ultra-large library docking for discovering new chemotypes. *Nature*, 566(7743):224–229, 2019.

Jiankun Lyu, John J Irwin, and Brian K Shoichet. Modeling the expansion of virtual screening libraries. *Nature chemical biology*, 19(6):712–718, 2023.

Andrew McNutt, Paul Francoeur, Rishal Aggarwal, Tomohide Masuda, Rocco Meli, Matthew Ragoza, Jocelyn Sunseri, and David Koes. Gnina 1.0: molecular docking with deep learning. *Journal of Cheminformatics*, 13, 06 2021. doi: 10.1186/s13321-021-00522-2.

Seokhyun Moon, Wonho Zhung, Soojung Yang, Jaechang Lim, and Woo Youn Kim. Pignet: a physics-informed deep learning model toward generalized drug–target interaction predictions. *Chemical Science*, 13(13):3661–3673, 2022.

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.

RDKit, online. RDKit: Open-source cheminformatics. http://www.rdkit.org. [Online; accessed 11-April-2013].

Thomas Seidel. Chemical data processing toolkit source code repository. https://github.com/molinfo-vienna/CDPKit, 2023.

Thomas Seidel, Christian Permann, Oliver Wieder, Stefan M Kohlbacher, and Thierry Langer. High-quality conformer generation with conforge: algorithm and performance assessment. *Journal of Chemical Information and Modeling*, 63(17):5549–5570, 2023.

Brian K. Shoichet. Virtual screening of chemical libraries. *Nature*, 432(7019):862–865, 2004.

Yang Song and Stefano Ermon. Generative modeling by estimating gradients of the data distribution. *Advances in neural information processing systems*, 32, 2019.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020.

Niranjan Srinivas, Andreas Krause, Sham M Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. *arXiv preprint arXiv:0912.3995*, 2009.

Kyle Swanson, Parker Walther, Jeremy Leitz, Souhrid Mukherjee, Joseph C Wu, Rabindra V Shivnaraine, and James Zou. Admet-ai: a machine learning admet platform for evaluation of large-scale chemical libraries. *Bioinformatics*, 40(7):btae416, 2024.

William R Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3-4):285–294, 1933.

Jose Antonio Garrido Torres, Sii Hong Lau, Pranay Anchuri, Jason M Stevens, Jose E Tabora, Jun Li, Alina Borovika, Ryan P Adams, and Abigail G Doyle. A multi-objective active learning platform and web app for reaction optimization. *Journal of the American Chemical Society*, 144 (43):19999–20007, 2022.

Mikhail Volkov, Joseph-André Turk, Nicolas Drizard, Nicolas Martin, Brice Hoffmann, Yann Gaston-Mathé, and Didier Rognan. On the frustration to predict binding affinities from protein–ligand structures with deep neural networks. *Journal of medicinal chemistry*, 65(11):7946–7958, 2022.

Swapnil Wagle, Richard D Smith, Anthony J Dominic III, Debarati DasGupta, Sunil Kumar Tripathi, and Heather A Carlson. Sunsetting binding moad with its last data update and the addition of 3d-ligand polypharmacology tools. *Scientific Reports*, 13(1):3008, 2023.

wwPDB consortium. Protein data bank: the single global archive for 3d macromolecular structure data. *Nucleic acids research*, 47(D1):D520–D528, 2019.

Chengxin Zhang, Xi Zhang, Peter L Freddolino, and Yang Zhang. Biolip2: an updated structure database for biologically relevant ligand–protein interactions. *Nucleic Acids Research*, 52(D1): D404–D412, 2024.

Guangfeng Zhou, Domnita-Valeria Rusnac, Hahnbeom Park, Daniele Canzani, Hai Minh Nguyen, Lance Stewart, Matthew F. Bush, Phuong Tran Nguyen, Heike Wulff, Vladimir Yarov-Yarovoy, Ning Zheng, and Frank DiMaio. An artificial intelligence accelerated virtual screening platform for drug discovery. *Nature Communications*, 15(1):7761, 2024.

Yiheng Zhu, Jialu Wu, Chaowen Hu, Jiahuan Yan, Tingjun Hou, Jian Wu, et al. Sample-efficient multi-objective molecular optimization with gflownets. *Advances in Neural Information Processing Systems*, 36, 2024.

A NOTATION

Table 1: Notation Glossary

Symbol	Description
$\mathcal{L}$	Ligand library used for VS.
l_i	Ligand i in the ligand library.
f	Affinity GP model mapping ligand fingerprints to binding affinity.
g	Utility GP model learning from preference data.
x	Multi-objective properties of a ligand, including physicochemical and ADMET properties.
t	Fingerprint representation of the ligand's structure.
α	Acquisition function in BO for ligand selection.
a	GP model mapping fingerprints to binding affinity.
R	Regret metric quantifying the gap between the best possible and selected ligand.
U	Utility function ranking and prioritizing ligands.
k	Used for selecting the top-k compounds.
P	Protein target for VS.
θ	Docking model (typically a diffusion model) predicting ligand-protein binding.

B RELATED WORK

Multi-Objective Bayesian Optimization Multi-objective BO (MOBO) (Couckuyt et al., 2012) tackles the challenge of optimizing multiple, potentially conflicting objectives. A common approach uses the Expected Hypervolume Improvement (EHVI) (Emmerich et al., 2008; Daulton et al., 2021), while other strategies include Predictive Entropy Search, Max-value Entropy Search, and the Uncertainty-Aware Search Framework (Hernández-Lobato et al., 2016; Belakaria et al., 2020a;b). ParEGO (Knowles, 2006) addresses computationally expensive problems using landscape approximations. Recent work extends MOBO to high-dimensional spaces (Daulton et al., 2022), accelerates VS, molecular optimization, and reaction optimization (Fromer et al., 2024; Zhu et al., 2024; Torres et al., 2022). However, many MOBO methods still lack mechanisms to effectively incorporate general domain expert insights during the search process, which is a critical need, especially in applications like VS. *CheapVS* builds on this MOBO foundation (Couckuyt et al., 2012; Chu & Ghahramani, 2005; Brochu et al., 2010) by introducing a preference-learning framework that guides optimization towards solutions aligned with expert knowledge in drug discovery, enabling more targeted and efficient chemical space exploration.

Active Virtual Screening VS (Lionta et al., 2014; Kitchen et al., 2004) is a computational strategy for identifying promising molecules from large chemical libraries. Traditional high-throughput VS (HTVS) often employs computationally expensive molecular docking methods (McNutt et al., 2021; Koes et al., 2013; Eberhardt et al., 2021; Lyu et al., 2019). While the effectiveness of ultra-large libraries is debated (Clark, 2020), their use in structure-based drug design has seen an increase in popularity (Gorgulla et al., 2020; Acharya et al., 2020). However, docking billions of compounds is computationally demanding (Gorgulla et al., 2020). Active learning strategies, such as MolPAL (Graff et al., 2021), improve efficiency by integrating BO with docking. By training a machine learning model on initial docking scores, MolPAL predicts binding affinities on the entire set and strategically selects subsequent compounds, significantly reducing the number of docking calculations while maintaining accuracy.

Scalable Diffusion Training Diffusion-based generative models have gained significant attention for their ability to model complex data distributions through iterative refinements of noisy inputs. Grounded on denoising score-matching (Hyvärinen & Dayan, 2005; Song & Ermon, 2019; Song

et al., 2020), these models leverage a governing ODE, $\mathrm{d}x = -\dot{\sigma}(t)\sigma(t)\nabla_x \log p(x;\sigma(t))\,\mathrm{d}t$, where x represents the ligand's coordinates, $\sigma(t)$ is the noise level, and $\nabla_x \log p(x;\sigma(t))$ is the score function (Song et al., 2020; Ho et al., 2020). The denoiser $D(x;\sigma)$ minimizes the L_2 distance: $\mathbb{E}_{y\sim p_{\text{data}}}\mathbb{E}_{\epsilon\sim\mathcal{N}(0,\sigma^2 I)}\|D(y+\epsilon;\sigma)-y\|_2^2$, leading to $\nabla_x \log p(x;\sigma) = \frac{D(x;\sigma)-x}{\sigma^2}$. Drawing inspiration from NCSN (Song & Ermon, 2019; Song et al., 2020), DDPM (Ho et al., 2020), and EDM (Karras et al., 2022), this paradigm emphasizes scalable training and robust score estimation for generative modeling.

Large-scale Datasets Curation for Molecular Docking Modern machine-learning models for molecular docking often rely on experimentally verified protein-ligand structures from the Protein Data Bank (PDB). Although the PDB offers thousands of protein structures, it contains only around 40k distinct ligands. To broaden coverage, researchers often generate additional data. For example, PDBScreen (Cao et al., 2024) introduces "decoy" ligands presumed not to bind the protein, while PigNet (Moon et al., 2022) and CarsiDock (Cai et al., 2024) use techniques like re-docking, cross-docking, and random docking of molecules from large commercial libraries. These approaches greatly expand the range of protein-ligand poses, helping machine learning models learn from positive (likely binding) and negative (likely non-binding) examples. Ultimately, these large, varied datasets lead to better generalizing models across different proteins and chemical compounds.

C Acquisition Functions

In this paper, we utilize the following acquisition functions to guide our optimization process:

- **qExpected Improvement (qEI)** (Boender, 1991): Evaluates the expected gain in model performance across multiple candidates, emphasizing exploration where improvement potential is high.
- **qProbability of Improvement (qPI)** (Kushner, 1964): Computes the likelihood that a set of candidate samples will surpass the current best performance.
- **qUpper Confidence Bound (qUCB)** (Srinivas et al., 2009): Balances exploration and exploitation by selecting candidates with both high uncertainty and high predicted performance based on their upper confidence bounds.
- **qThompson Sampling (qTS)** (Thompson, 1933): Approximates the posterior distribution of the model and sample candidates to maximize predicted utility, promoting diverse exploration.
- **qExpected Utility of the Best Option (qEUBO)** (Astudillo et al., 2023): A decision-theoretic acquisition function for preferential BO (PBO) that maximizes the expected utility of the best option. It is computationally efficient, robust under noise, and offers superior performance with guaranteed regret convergence.

D Preliminary Analysis on the Data

As noted in Figure 5, the number of data points in the PDBScan training data is roughly four times as large as the data points in the PDBbind training data. Furthermore, the training data utilized covers 18 different protein groups.

D.1 Diversity of Data: Proteins

Understanding the diversity of protein classes in the dataset is essential for evaluating its coverage and potential biases in molecular docking tasks. Figure 5 illustrates the distribution of protein classes across different datasets, highlighting variations in data availability. The left panel compares the protein class distributions between PDBScan and PDBbind, showing that PDBScan contains a significantly larger number of data points across all protein categories, particularly in "Unclassified proteins" and "Kinases." This discrepancy suggests that PDBScan provides broader protein coverage, which may enhance model generalization.

The right panel further details the absolute counts of protein classes, emphasizing their relative abundance. The dataset is dominated by enzymatic proteins, including Oxidoreductases, Transferases, and Hydrolases, which are frequently studied in drug discovery. However, certain categories such as Toll-like and IL-1 receptors, Transporters, and Cytochrome P450 remain underrepresented, potentially

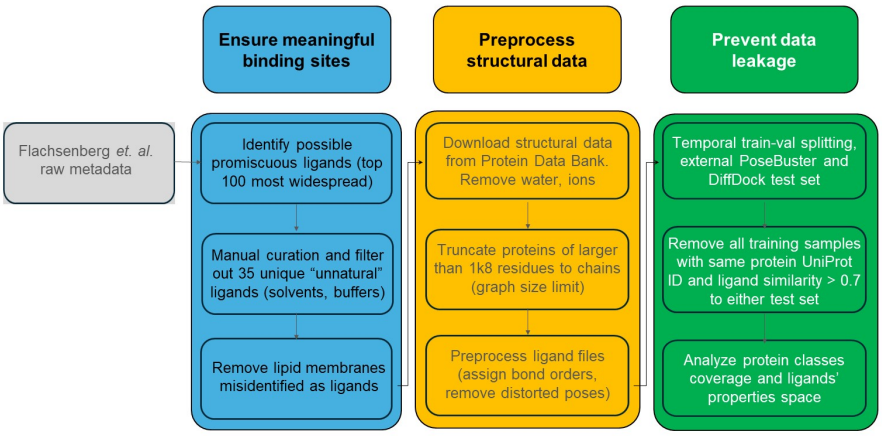


Figure 4: PDBScan22 data curation workflow. The process consists of three main steps: (1) Ensuring meaningful binding sites by filtering promiscuous ligands, removing unnatural molecules such as solvents and buffers, and eliminating misidentified lipid membranes. (2) Preprocessing structural data by downloading structures from the Protein Data Bank (PDB), removing water and ions, truncating proteins exceeding 1,800 residues, and refining ligand files by assigning bond orders and eliminating distorted poses. (3) Preventing data leakage through temporal train-validation splitting, external PoseBuster, and DiffDock test sets, and removing training samples with proteins sharing UniProt IDs and highly similar ligands (>0.7 similarity) with test sets. Additionally, protein class distributions and ligand properties are analyzed to ensure a well-balanced dataset.

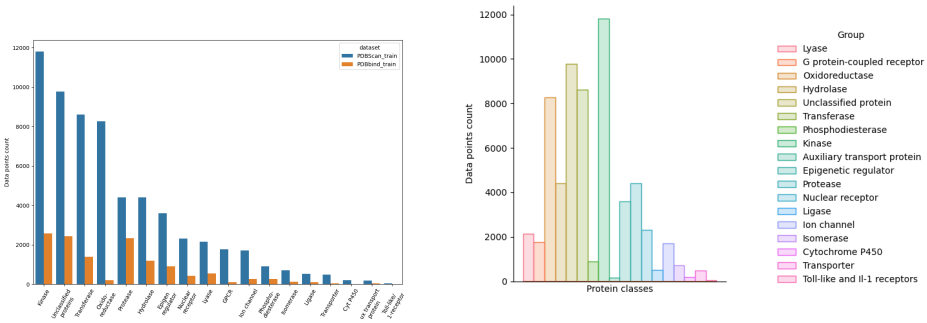


Figure 5: Analysis of protein properties: Protein classes distribution, Protein classes data point count

impacting model performance on these classes. These insights highlight the importance of data augmentation techniques to balance protein representation and improve downstream learning.

D.2 Diversity of Data: Ligands

The number of ligands with a higher QED drug-likeliness score is also significantly larger in the PDBScan training dataset when compared to the PDBbind training dataset. The PDBScan dataset notably contains a higher count of unique ligands with rotatable bonds. The PDBScan training dataset encompasses a greater quantity of ligands with a wider range of unique properties in comparison to the PDBbind training dataset, allowing for a wider analysis of ligand properties.

E More Results on Preferential Multi-Objective Bayesian Optimization

E.1 Synthetic Experiment

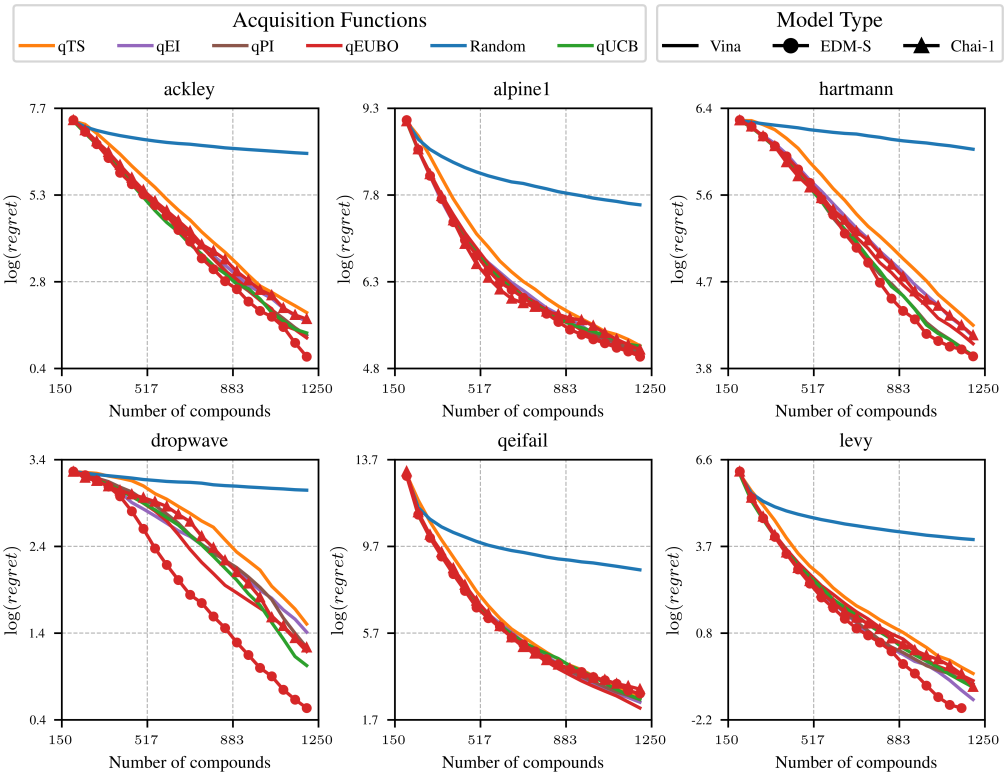


Figure 8: *CheapVS* results with all acquisition functions. The y-axis shows $\log -regret$

We examine how well *CheapVS* identifies high-utility solutions under *synthetic* conditions. We create complex utility landscapes by modeling multi-dimensional molecular designs with benchmark functions: Ackley, Alpine1, Hartmann, Dropwave, Qeifail, and Levy. Each benchmark outputs a scalar "utility," and we generate initial pairwise preference labels by comparing these function values. In addition, we simulate four main objectives relevant to drug discovery: binding affinity, rotatable bonds, molecular weight, and LogP. For computational feasibility, we used a 20k-ligand subset.

Figure 8 displays the logarithm of the regret ($\log(\text{optimum} - \text{current best})$) versus the number of compounds screened, focusing on `qEUBO`, `qUCB`, and `Random` acquisition strategy. All methods reduce regret over time, but advanced acquisitions converge faster than random baselines. EDM-S consistently achieves the lowest regret across most benchmarks, except for a slight underperformance on Qeifail. Chai-1 shows the highest regret, likely due to its heavier computation and reduced sampling efficiency, whereas Vina falls in between. These findings confirm *RQ1*: preferential BO can effectively learn multi-objective trade-offs in synthetic benchmarks, and an efficient docking model (EDM-S) accelerates convergence significantly. The results highlight that both the choice of acquisition function and the docking model's computational overhead can substantially influence optimization quality and speed.

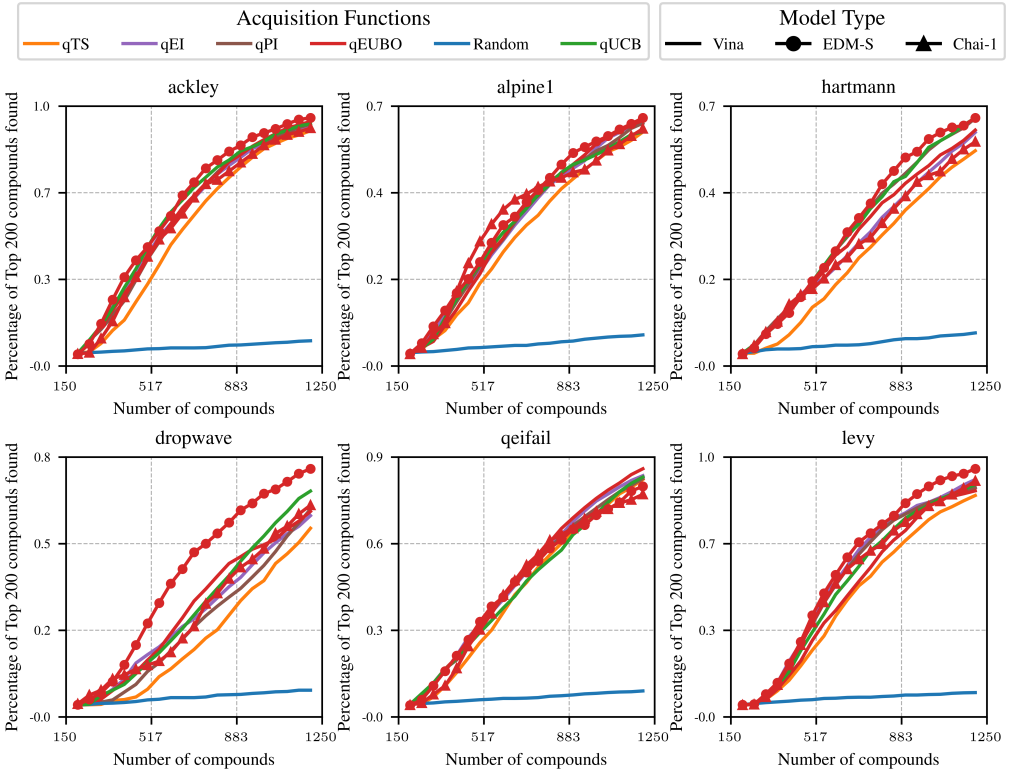


Figure 9: *CheapVS* results with all acquisition functions. The y-axis shows accuracy

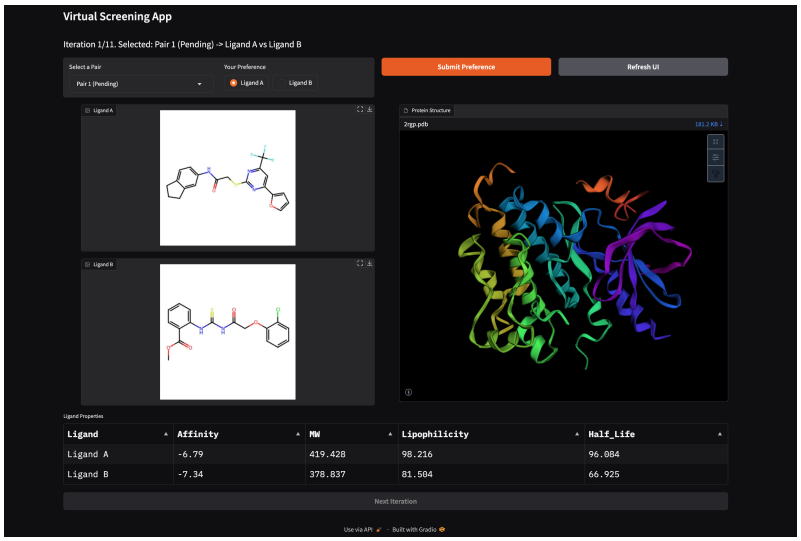


Figure 10: Virtual Screening (VS) App built with Gradio for seamless interaction with chemists, enabling expert-driven ligand selection and molecular analysis

E.2 HUMAN-IN-THE-LOOP EXPERIMENT

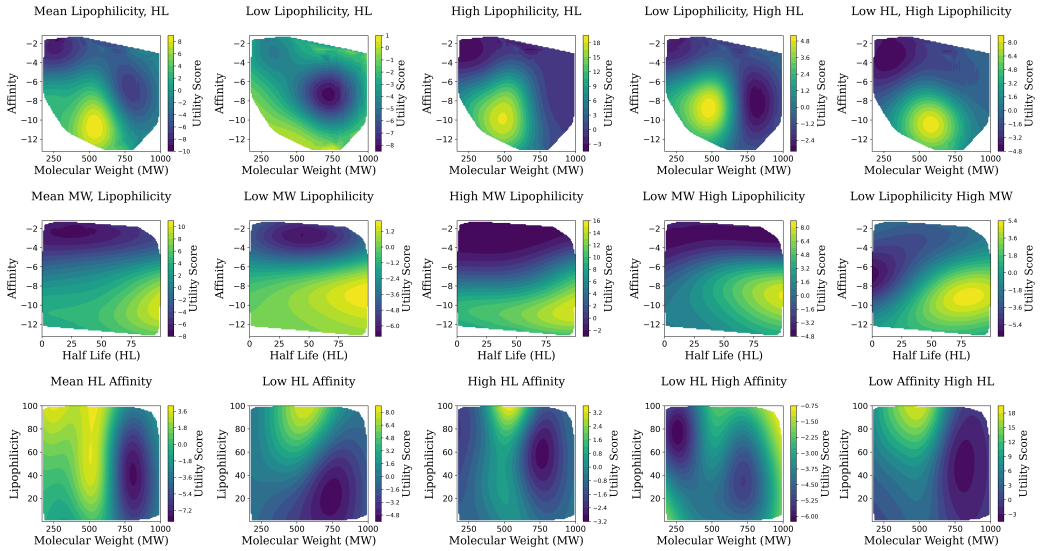


Figure 11: More on Gaussian process (GP) utility surfaces learned from expert preference data, illustrating the interplay among molecular weight (MW), half-life (HL), affinity, and lipophilicity. Each heatmap shows the predicted utility (color scale) over two of these variables while holding the others fixed at the levels indicated in each title. Higher (yellow) regions correspond to more favorable trade-offs according to the elicited expert preferences, providing insights for optimizing lead compounds in drug discovery.

E.3 DOCKING TIME-REGRET ANALYSIS

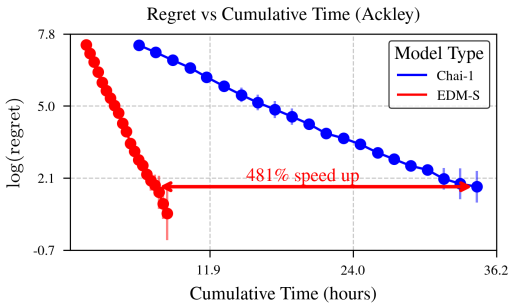


Figure 12: Regret reduction over cumulative runtime for the Ackley function using `qEUBO`. The y-axis represents $\log(\text{regret})$, while the x-axis shows cumulative time in hours. EDM-S (red) significantly outperforms Chai (blue) in convergence and speed, achieving the same regret level 481% faster.

Lastly, we tackle *RQ3* by investigating how the computational overhead of docking models affects convergence speed. We focus on the multi-modal **Ackley** function, plotting $\log(\text{regret})$ against total wall-clock hours. We compare EDM-S and Chai-1 under the `qEUBO` acquisition function, noting that EDM-S generates 128 poses per docking run, whereas Chai-1 is limited to a single pose. Figure 12 shows that **EDM-S** converges to lower regret up to four times faster than Chai-1. This outcome emphasizes the importance of computational efficiency in iterative BO pipelines.

These results address *RQ3*: *slower, computationally expensive docking methods impede the exploration-exploitation cycle*, while lighter, faster models (EDM-S) significantly boost throughput and convergence speed. By minimizing docking overhead, scientists can explore a larger chemical space more effectively within the same time budget.

F SURROGATE MODEL PERFORMACE

Model Type	MSE Loss	NLPD
Fully-connected Neural Network	1.0568±0.0437	1.4629±0.0198
Decision Tree	1.9785±0.2227	1.7572±0.0548
Gaussian Process (Tanimoto kernel)	**0.8549±0.0689**	**1.3389±0.0404**

Table 3: Comparison of model performance in predicting binding affinity values based on ligand fingerprints. The table reports the Mean Squared Error (MSE) Loss and Negative Log Predictive Density (NLPD) for different model types. Each model is trained on 6,000 samples using an 80/20 train/test split, and results are averaged over 20 random trials. The Gaussian Process (Tanimoto kernel) achieves the best performance with the lowest MSE and NLPD, indicating superior predictive accuracy and uncertainty quantification.

Model Type	Accuracy (%)	ROC AUC
Fully-connected Neural Net	0.9195±0.0199	0.9809±0.0081
Decision Tree	0.7853±0.0285	0.7858±0.029
Pairwise Gaussian Process	**0.9563±0.0146**	**0.9724±0.0161**

Table 4: Comparison of utility model performance in predicting preference-based rankings from ligand fingerprints. The table reports the classification accuracy of different model types. Each model is trained on 1,000 samples using an 80/20 train/test split, and results are averaged over 20 random trials. The Pairwise Gaussian Process achieves the highest classification accuracy and ROC-AUC, demonstrating superior performance in modeling pairwise preferences and learning utility functions from ligand physicochemical properties.

G DIFFUSION MODEL TRAINING: HYPERPARAMETERS AND PERFORMANCE RESULTS

We enhance our dataset by incorporating experimental binding poses and generating a synthetic dataset based on pharmacophore models. Learning-based docking models often fail to generalize to new protein classes, partly because standard datasets—such as PDBBind2020, which contains around 16,000 protein-ligand complexes across 4,000 unique proteins—are highly biased toward well-studied targets (e.g., kinases, proteases) (Corso et al., 2024; Buttenschoen et al., 2024; Volkov et al., 2022). While binding-affinity data is crucial for scoring models (which predict how strongly a ligand binds), it is not strictly necessary for generative docking frameworks, which learn docking poses directly from protein and ligand coordinates. This opens the door to leveraging broader datasets from the Protein Data Bank (PDB) (wwPDB consortium, 2019), such as PDBScan22 (Flachsenberg et al., 2023), thereby expanding protein diversity beyond the narrow range in PDBBind.

The original PDBScan22 contains 322,051 complexes but includes many "unnatural ligands" (solvents, surfactants, buffers) that don't represent drug-like molecules. We curated the dataset, removing entries with these ligands (see Appendix D), and retained biologically relevant molecules like nucleotides and amino acids because they often represent natural binding sites. This filtering ensures the dataset focuses on drug-like ligands in realistic protein environments. We detailed our curation process in Figure 4.

While strategies like redocking and cross-docking augment PDB ligands to nearby poses, they fail to address the low ligand diversity. We leverage the Papyrus dataset (Béquignon et al., 2023), containing 260,000 active ligands across 1300 protein UniProt IDs to tackle this. Molecules with reliable and good activity data (pChemBL > 5) matching curated PDBScan22 structures were retained. Using Conforge (Seidel et al., 2023), 50 conformers per molecule were generated (with RMSD > 0.2 Å between conformers). Binding sites from PDBScan22 were modeled into pharmacophores with CDPKit (Seidel, 2023), incorporating exclusion volumes and filtered for models with more than 5 features. Ligands were aligned to pharmacophores of matching UniProt IDs based on shape and

electrostatic properties. The aligned poses are further minimized of binding affinity using smina (Koes et al., 2013), enabling fast generation of large-scale protein-ligand structures without explicit docking.

For the Human-in-the-loop experiment in Section 3.1, focusing on the EGFR protein, we further fine-tune the EDM-S model on a set of 10,000 ligands sampled from the ChemDiv kinase inhibitor library. Any overlap between the training data and the Dockstring screening library is removed. Training poses are generated using smina (Koes et al., 2013) and the protein structure from PDB ID 2RGP. Only the top-scored pose by the Vinardo scoring (Quiroga & Villarreal, 2016) is retained to train the diffusion model.

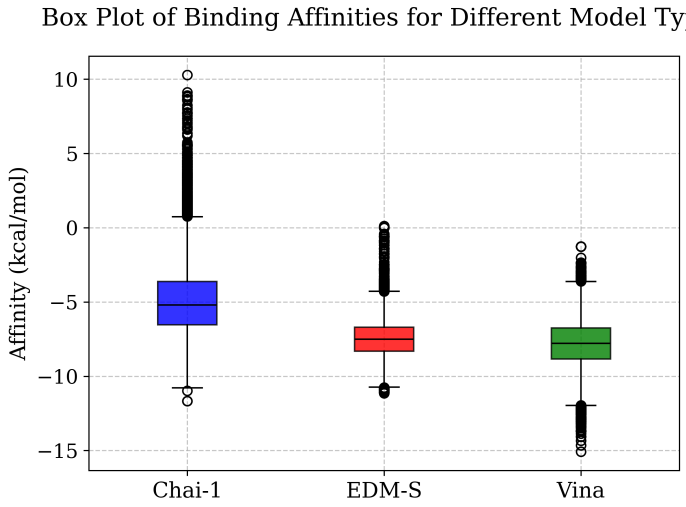


Figure 13: Box plot of binding affinities (kcal/mol) for different docking models. Vina achieves the lowest median binding affinity, followed by EDM-S, while Chai exhibits the weakest binding.

Model	DockScan22	EDM-S (Pre-train)	EDM-S (Fine-tune)	EDM-S (EGFR)
Parameters initialized from	Random	Random	EDM-S (Pre-train)	EDM-S (Fine-tune)
Batch Size	256	256	256	64
Number of Epochs	150	2.32	140	640
Dataset train on	PDBScan22	11M synthetic data	PDBScan22	ChemDiv 10k
Learning Rate	1.8×10^{-3}	1.8×10^{-3}	1×10^{-3}	2×10^{-3}
Diffusion steps	20	20	20	10
σ_{data}	32	32	32	5

Table 5: Hyperparameters for training EDM-S and DockScan22

Metrics	PoseBuster V1 Top-1 RMSD (Å) % < 2Å	% < 5Å	PoseBuster V2 Top-1 RMSD (Å) % < 2Å	% < 5Å	PDBBind Top-1 RMSD (Å) % < 2Å	% < 5Å	Inference time on 1 A100 seconds
DIFFDOCK-S (40)	24	45.1	-	-	31.1	-	**10**
DIFFDOCK (40)	37.9	49.3	-	-	**38.2**	**62**	30
AlphaFold 3 (25)	76.4	-	**80.5**	-	-	-	340
Chai-1 (25)	**77.05**	-	-	-	-	-	340
DockScan22 (40)	54.1	77.8	58.8	81.4	34.1	56	**10**
EDM-S (40)	30	**91**	32.2	**92.1**	-	-	**10**

Table 6: Performance comparison on PDBBind and PoseBuster benchmarks, with models sampling 40 or 25 ligand poses per protein-ligand pair. Highlighted rows show our proposed methods, offering competitive accuracy with significantly lower runtime.

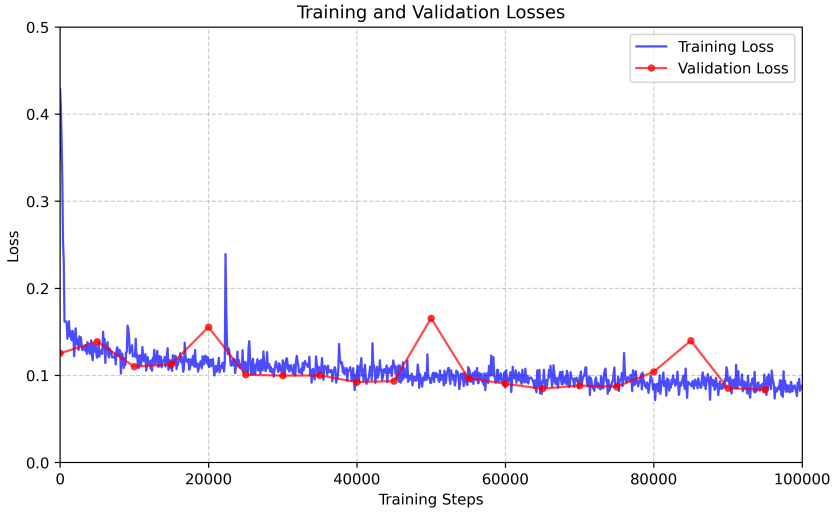
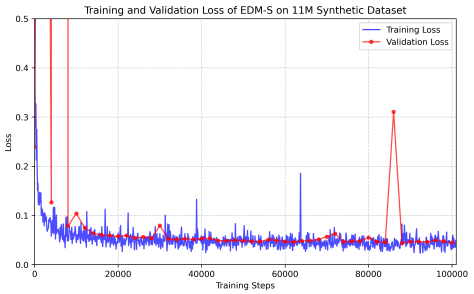
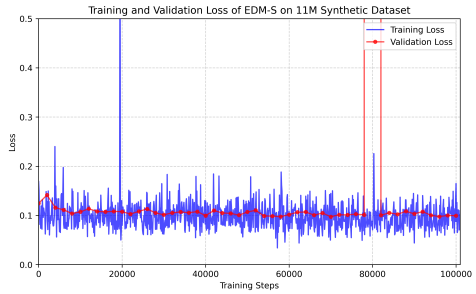


Figure 14: Training and validation loss of EDM-S on EGFR protein with 10k ligands.



(a) Training and validation loss of EDM-S on 11M synthetic data.



(b) Training and validation loss of EDM-S on 180k PDBScan22.

Figure 15: Comparison of EDM-S training and validation losses on different datasets: (a) 11M synthetic data and (b) 180k PDBScan22.

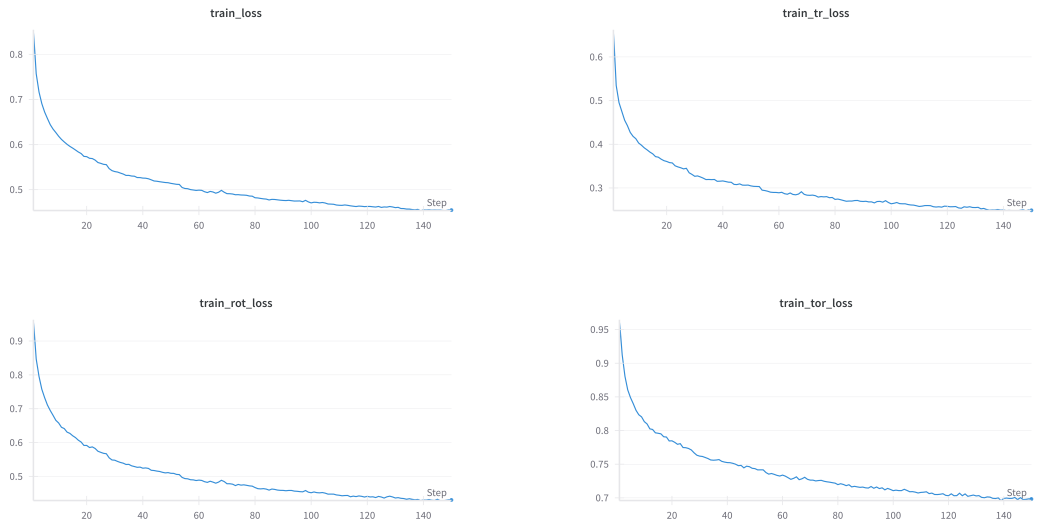


Figure 16: Training loss for DockScan22 on PDBScan22 across three components: translation, rotation, and torsion angle.

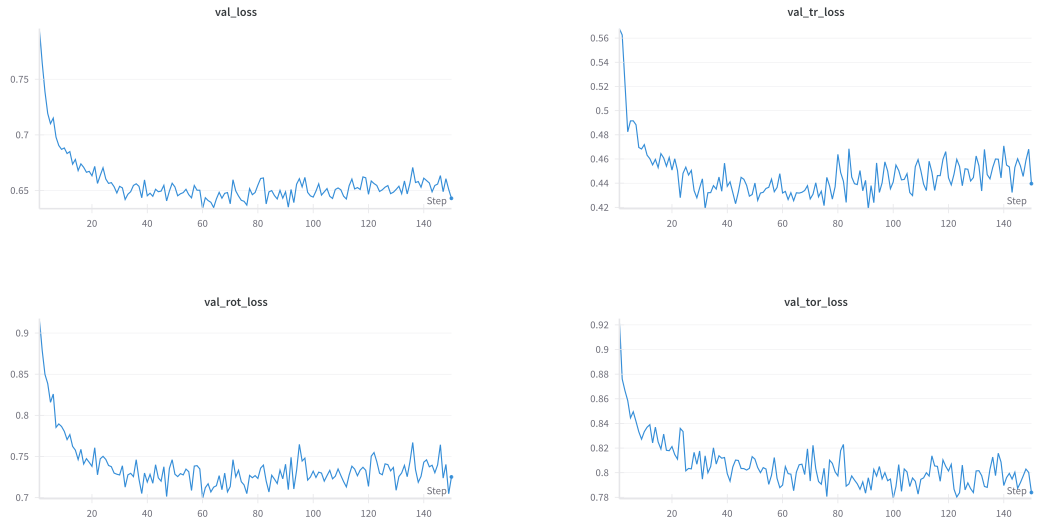


Figure 17: Validation loss for DockScan22 on PDBScan22 across three components: translation, rotation, and torsion angle.