PREFERENTIAL MULTI-OBJECTIVE BAYESIAN OPTIMIZATION FOR DRUG DISCOVERY

Tai Dang[*1], **Long-Hung Pham**[*2], **Sang T. Truong**[*1], **Ari Glenn**[3], **Wendy Nguyen**[4], **Edward A Pham**[3], **Jeffrey S. Glenn**[3], **Sanmi Koyejo**[1], **Thang Luong**[1]

[1]Stanford Computer Science, [2]Department of Chemistry, Imperial College London, [3]Stanford Medicine, [4]RHF.AI

ABSTRACT

Despite decades of advancements in automated ligand screening, large-scale drug discovery 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 named *CheapVS* that allows chemists to guide the ligand selection process by providing preferences regarding the trade-offs between drug properties via pairwise comparison. Our framework combines preferential multi-objective Bayesian optimization with a docking model for measuring binding affinity to capture human chemical intuition for improving hit identification. Specifically, on a library of 100K chemical candidates targeting EGFR and DRD2, *CheapVS* outperforms state-of-the-art screening methods in identifying drugs within a limited computational budget. Notably, our method can recover up to 16/37 EGFR and 37/58 DRD2 known drugs while screening only 6% of the library, showcasing its potential to significantly advance drug discovery.

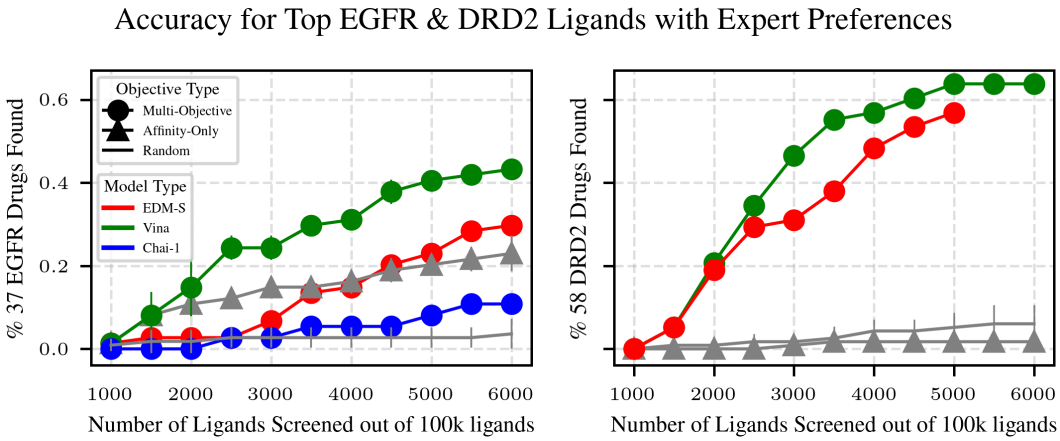

Figure 1: Chemist-guided Active Preferential Virtual Screening performance in identifying EGFR and DRD2 drugs. The search is conducted on a 100K ligand library, screened for a maximum of 6% of the library. The plot compares different methods for structure-based binding affinity measurement (Vina, EDM-S, Chai-1) and objective types. The y-axis shows the percentage of the top number of approved drugs identified, while the x-axis represents the number of ligands screened. Multi-objective optimization (circles) across all methods for affinity measures outperforms affinity-only selection (triangles) and random screening (gray line). Error bars indicate 1 standard deviation.

[*]Equal contribution. Correspondence to taitdang@stanford.edu, l.pham23@imperial.ac.uk, sttruong@cs.stanford.edu, lmthang@stanford.edu

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 docking 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 compounds potential. Despite its centrality to drug discovery, VS and hit selection remain both resource-intensive. 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 results to finalize which hits are worthy of costly experimental validation. In large-scale campaigns with millions of compounds, this process quickly becomes bottlenecked by both the computational costs and the limited bandwidth of experts. On top of the issue, vast computational resources are spent on characterizing 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 binding affinity prediction 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, 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 interdependencies Hann & Keserü (2012). For instance, adding bulky functional groups can enhance binding affinity but simultaneously lower solubility or increase off-target toxicity, complicating the search for high-potential candidates. Balancing competing properties is key to robust drug leads. While single-objective active-learning approaches Graff et al. (2021); Gentile et al. (2020) have shown promise in efficiently identifying top-scored molecules from large-scale libraries, the screening process still overlooks important considerations that medicinal chemists weigh in real-world pipelines, such as synthetic accessibility, stability, toxicity, etc. Thus, much computational power is still wasted on molecules with poor profiles other than binding affinity. This disconnection also highlights the critical role of domain expertise, balancing multiple factors that purely physics-based methods often fail to capture. Unfortunately, while invaluable, the expert-driven hit selection process is labor-intensive when scaled to large candidate pools.

To address these limitations, we present *CheapVS* (CHEmist-guided Active Preferential Virtual Screening) to assist 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 toxicityour framework ensures that computational optimization captures subtle trade-offs that purely physics-based methods often overlook. This expert-guided approach refines the VS process, prioritizing candidates based on 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.

Preference ranking relies on the availability of a good measurement of ligand properties. An important measurement is the binding affinity between the ligand and the target protein. Recent breakthroughs such as AlphaFold3 Abramson et al. (2024) and Chai-1 Boitreaud et al. (2024) have promised better measurement of binding affinity on a wide range of targeted proteins. Unfortunately, these methods are expensive, and the lack of understanding of their accuracy-efficiency trade-off has made it difficult for practitioners to select a suitable method for large scale VS. We compare the accuracy and efficiency of the physics-based and diffusion-based approaches, showing that while existing diffusion docking models are promising, their efficiency is far away from practical VS. Introducing a lightweight diffusion model on large-scale data, we significantly improve the efficiency of these tools while maintaining high performance, suggesting a path toward making deep learning models practical for VS.

Code and data are at github.com/vietai/cheapvs.

In summary, our key contributions are:

- **Eliciting Expert Preference for Efficient Virtual Screening:** We optimize trade-offs among interdependent drug properties by leveraging chemists' intuition through preference learning, translating domain knowledge into a latent utility function for more efficient VS.
- **Understanding Accuracy-Efficiency Trade-Off in Docking Models:** We evaluate the accuracy-efficiency trade-off of the physics-based and diffusion-based approaches and use data augmentation to significantly improve the efficiency of the diffusion docking model.
- **Efficient Multi-Objective Virtual Screening:** *CheapVS* considers various candidates' properties to simultaneously optimize them, such as binding affinity and toxicity, moving beyond single-objective paradigms.

2 Related Work

Efficient Decision Making in Virtual Screening VS Lionta et al. (2014); Kitchen et al. (2004) is a computational strategy for selecting promising molecules from large chemical libraries. Traditional high-throughput VS (HTVS) often employs computationally expensive structured-based binding affinity measurement 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). Therefore, active learning strategies, such as MolPAL Graff et al. (2021), improve efficiency by integrating optimal model-based sequential decision-making with docking. By training a machine learning model on initial binding affinities, MolPAL predicts binding affinities on the entire set and strategically selects subsequent compounds, significantly reducing the number of docking calculations while ensuring reliable performance.

Expert Preference in Virtual Screening 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 domain expert insights during the search process, which is a critical need in 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 VS. While prior work Choung et al. (2023) explores expert preferences via SMILES-based rankings, they do not consider ligand properties, which limits the optimization process

Measurement of Binding Affinity via Diffusion Model 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), these models leverage a governing ordinary differential equation, $\mathrm{d}x = -\dot{\sigma}(t)\sigma(t)\nabla_x \log p(x;\sigma(t))\,\mathrm{d}t$, where x represents the inputs, $\sigma(t)$ is the noise level, and $\nabla_x \log p(x;\sigma(t))$ is the score function. The denoiser $D(x;\sigma)$ minimizes the mean squared error loss. Modern machine-learning models for structure-based binding affinity prediction often rely on experimentally verified structures from the Protein Data Bank (PDB)wwPDB consortium (2019). Although the PDB offers thousands of structures, it contains only around 40k ligands. To broaden coverage, researchers often generate additional data, e.g., PDBScreen Cao et al. (2024) introduces "decoy" ligands presumed not to bind the protein, while PigNet and CarsiDock Moon et al. (2022); Cai et al. (2024) use techniques like re-docking, cross-docking, and random docking from large commercial libraries. These methods expand protein-ligand diversity, enabling models to improve their generalization for docking.

3 PRELIMINARY

We briefly describe the VS setup. Given a target protein ρ and a screening library $\mathcal{L} = \{\ell_1, \ldots, \ell_N\}$, the goal is to select the top k ligands with the highest potential to succeed as drugs. A value is assigned to each ligand toward this goal so they can be ranked. It is common in practice to use a molecular property vector of ligand ℓ, denoted as x_ℓ, as a proxy of drug likeliness. The vector can include various properties, such as binding affinity, toxicity, and solubility: $x_\ell = [x_\ell^{\text{aff}}, x_\ell^{\text{toc}}, x_\ell^{\text{sol}}]$. However, evaluating these properties can be costly and time-consuming for large molecular libraries. To address this, active screening methods are used to explore the chemical space efficiently. These methods start with a small, randomly chosen fraction of the ligand library as the initial training set. The properties of this small set of ligands, often binding affinity, are measured, which are then used to train a surrogate model. Once the initial model is established, the optimization proceeds iteratively. In each cycle, newly measured ligands update the model with their latest data. The updated surrogate model evaluates the remaining compounds, and an acquisition function α ranks them by balancing exploration and exploitation. The top-ranked candidates are selected, their properties are measured, and the resulting data is used to update the surrogate model further. This process continues until the termination criteria is reached. For evaluation, regret and top-k accuracy are the main metrics used in the screening procedure. *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 at iteration i. Importantly, U^* can only be determined if affinity computations are carried out for the entire library, making it a post-hoc evaluation metric rather than a run-time metric. Lower regret indicates a better screening strategy. *Top-k Accuracy* measures the proportion of correctly identified compounds within the *top-k* set, where the top-k corresponds to the compounds with the highest utility value.

Measuring molecular properties is essential for eliciting expert preference. One of the most important properties is binding affinity, denoted as $x_{\ell,\rho}^{\text{aff}}$. Unfortunately, this objective is computationally expensive to estimate because it requires *searching for an energetically optimal 3D structure of the ligand inside the protein binding pocket*:

$$x_{\ell,\rho}^{\text{aff}} = \min_{\ell_{\text{3D}} \in^{3\times N_\ell}} h(\ell_{\text{3D}}, \rho) \quad \forall \ell \in \mathcal{L} \tag{1}$$

where h is the physics-based affinity scoring function based on the atomic interaction between the ligand and the target protein, N_ℓ is the number of atoms in ligand ℓ (which is typically several dozen), and ℓ_{3D} is the 3D coordinate vector of ligand ℓ. Traditional docking methods use heuristics to search through the vast conformation space $^{3\times N_\ell}$, rendering an intractable procedure when applied at scale to all ligands in the library. Diffusion docking models have been introduced to bias the above optimization toward a 3D structure that geometrically fits the binding pocket. Given a protein target ρ, a ligand, and a corresponding experimentally obtained 3D binding pose, the training objective of diffusion docking model p_θ is to find the 3D pose that best fits the binding pocket in terms of mean squared error (MSE) loss. After training, the optimal 3D pose could be obtained rapidly without exhaustive search through sampling process $\ell_{\text{3D}} \sim p_\theta(\ell, \rho)$. The optimal pose is then used with the scoring function h to measure binding affinity.

4 METHODS

We introduce a comprehensive methodology comprised of three components: preference modeling, active ligand selection, and ligand property measurement. An overview of the complete pipeline is presented in Figure 2.

Preference Modeling Balancing multiple objectives, such as affinity, toxicity, and solubility, is necessary to identify drug candidates. This is challenging as the optimal trade-offs are unknown. We elicit expert preferences to guide the optimization process. These preferences are gathered through pairwise comparisons of ligand properties, where experts indicate which of the two ligands, ℓ_1 and ℓ_2, is more desirable based on multiple criteria. We assume the probability of preferring ligand ℓ_1 over ℓ_2 follows the Bradley-Terry model, generated via a utility function f mapping from ligand property vector x_ℓ to a utility scalar:

$$p(\ell_1 \succ \ell_2 \mid x_{\ell_1}, x_{\ell_2}) = \sigma(f(x_{\ell_1}) - f(x_{\ell_2})) \tag{2}$$

where σ is the logistic function. We model the distribution of the latent utility function with a Gaussian process (GP), assuming $f \sim \mathcal{GP}(\mu(X), k(X, X'))$, where $\mu(X)$ and $k(X, X')$ denote the mean and kernel functions. We train f using a dataset $_f = \{(x_{\ell_i}, x_{\ell_j}, y_{e_{ij}})\}$, where each pair

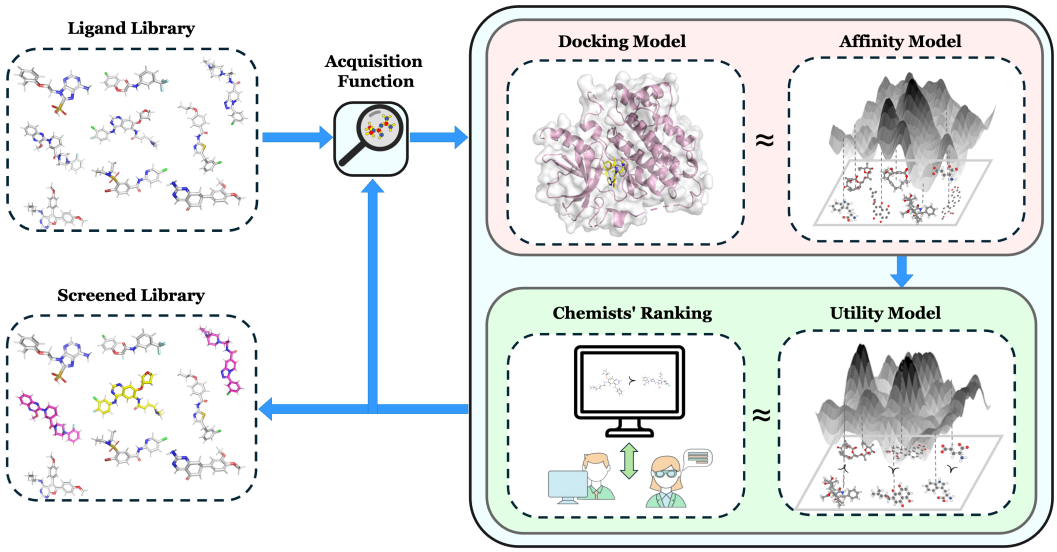

Figure 2: Overview of Chemist-guided Active Preferential Virtual Screening (*CheapVS*). Ligands from a large library are selected using an acquisition function and evaluated through structure-based affinity models. Chemists provide preference rankings, which inform a utility model to refine the selection process. The screened library iteratively improves, prioritizing ligands with desirable properties. Yellow-colored ligands represent found drug compounds, while purple ligands indicate screened compounds.

(x_{ℓ_i}, x_{ℓ_j}) consists of ligand property vectors, and $y_{e_{ij}}$ is the expert preference label indicating whether ℓ_i is preferred over ℓ_j. For posterior estimation, we employ the Laplace approximation. The choice of GPs for surrogate models is motivated by their superior performance over neural networks, as detailed in Appendix E.

Measurement of Binding Affinity via Diffusion Model A large-scale diffusion docking model holds the promise to accelerate finding optimal binding poses in comparison to traditional searches. Unfortunately, sampling from the diffusion model is still an expensive process, especially for a large model, making it impractical to use diffusion for large-scale screening. Here, we aim to train a lightweight diffusion model that is highly computationally tractable for a large-scale library while aiming for a minimal reduction in binding affinity estimation. Toward this goal, we curate a large, diverse diffusion training dataset:

1. From PDBScan22 (Flachsenberg et al., 2023), we remove non-drug-like ligands (e.g., solvents) to retain biologically relevant molecules (e.g., amino acids), ensuring the dataset focuses on drug-like ligands in realistic protein environments (see Appendix C), yielding 180,000 high-quality protein-ligand pairs.
2. To address low ligand diversity, we leverage the Papyrus dataset Béquignon et al. (2023), containing 260,000 active ligands across about 1,300 UniProt IDs. Molecules with reliable and good activity data (pChemBL > 5) matching curated PDBScan22 structures were retained. Using Conforge Seidel et al. (2023), we generate 50 conformers per molecule (RMSD > 0.2 Å). Pharmacophores from PDBScan22 are extracted with CDPKit Seidel (2023) and ligands are aligned based on shape and electrostatic properties. The aligned poses are minimized for binding affinity with Smina Koes et al. (2013).

After training the diffusion on this curated data, for each ligand, we generate 128 candidate conformation and greedily select the one with the lowest binding affinity to solve Equation 1 by inference-time best-of-N search. One can train a diffusion policy to solve this equation directly via reinforcement learning, and we defer this to future work.

Even with a highly efficient diffusion model, measuring binding affinity on a vast ligand library remains intractable. We further tackle this problem by introducing a protein-specific surrogate model

g_ρ that directly predicts binding affinity for a given ligand. Here, we use GP to model the affinity distribution of g_ρ with a Gaussian likelihood following Brochu et al. (2010). We featurize the ligand ℓ by its Morgan fingerprint ℓ, a fixed-length vector encoding its substructural features: $\hat{x}^{\text{aff}}_{\ell,\rho} = g_\rho(\ell_)$, where g_ρ is supervised trained by leveraging data from diffusion model $_{g_\rho} = \{(\ell_{,i}, h(\ell_{\text{3D},i}) : \ell_{\text{3D},i} \sim p_\theta(\ell_i, \rho))\}_i$

Active Virtual Screening with Expert Preference We have introduced a latent utility function and a scalable method for measuring the binding affinity of a large ligand library, which can provide a measure for drug likeliness $f \circ g_\rho(\ell)$ of ligand ℓ that balances various objectives aligning with human preference. Since various approximations have been made in favor of computational tractability, one should not greedily optimize for $f \circ g_\rho$. The composition of utility and affinity functions allows uncertainty in g_ρ to propagate through and impact decision quality if not handled carefully. Here, the optimal ligand for further measurement is the maximizer of the expected value of the acquisition function α, marginalizing over posterior predictive distribution induced by g_ρ for a given acquisition function α and a posterior distribution over f:

$$\ell^* = \arg\max_{\ell \in \mathcal{L}} \mathbb{E}_{p(x^{\text{aff}}_{\ell,\rho} | \ell, g_\rho)} \alpha(p(f|_f), x_\ell). \tag{3}$$

where the expectation is approximated using Monte Carlo sampling, and ligand pairs for expert preference queries are randomly chosen among the top-k candidates with the highest acquisition values. The full pseudocode for CheapVS's algorithm can be found in Appendix D.1.

5 Experiments

We conduct the experiments in three main stages. First, we investigate how well the utility model can learn from preference data, using both synthetic benchmark functions and real human-labeled preferences. Second, we explore the accuracy-speed tradeoff in affinity measurement by comparing our lightweight diffusion model (EDM-S) against Chai-1 and Vina, analyzing how computational efficiency impacts optimization performance. Finally, we integrate preference learning, molecular docking, and virtual screening into a comprehensive drug discovery pipeline, primarily targeting the EGFR and DRD2 proteins. For our main study, we utilize `qEUBO` as the main acquisition function. A full analysis of additional strategies can be found in Appendix B and D.5. Additionally, we compute physicochemical properties using RDKit, online and incorporate ADMET predictions (absorption, distribution, metabolism, excretion, and toxicity) Swanson et al. (2024) to account for realistic multi-objective trade-offs in drug discovery. Our study focuses on 3 Research Questions (RQ):

- **RQ1**: How effectively can the latent utility function learn and approximate the underlying true utility from both synthetic and expert pairwise comparisons?
- **RQ2**: How does the accuracy-efficiency trade off in diffusion docking model impact the virtual screening process and what is a promising path to improve the efficiency of diffusion docking model?
- **RQ3**: How effectively does *CheapVS* identify clinically relevant drug ligands using multi-objective optimization compared to affinity-only baseline?

5.1 Preference Elicitation from Pairwise Comparisons

To answer *RQ1*, we examine how well preference learning can correctly identify preferences using ligand properties (binding affinity, lipophilicity, molecular weight, and half-life) as input. For synthetic data, we generate 1,200 pairwise labels via functions–Ackley, Alpine1, Hartmann, Dropwave, Qeifail, and Levy. In contrast, for real human data, experts provide rankings on the EGFR target to form pairwise comparisons. All experiments are conducted under an 80/20 split and 20-fold cross-validation and evaluate model performance with accuracy and ROC AUC.

Table 1 indicates that our preference learning framework consistently achieves high accuracy and ROC AUC, demonstrating robust recovery of the latent utility function. While there is some variability across different synthetic functions, the overall trend confirms strong performance. Similarly, preliminary results on real human data show competitive performance, with an accuracy of approximately 80% and a ROC AUC of around 90%. These findings suggest that our approach effectively captures the underlying utility function from pairwise comparisons, supporting its potential for practical applications in drug candidate screening.

	Accuracy (%)	**ROC AUC**
ackley	95.75 ± 1.97	0.99 ± 0.01
alpine1	79.73 ± 2.99	0.88 ± 0.03
hartmann	90.52 ± 2.18	0.98 ± 0.01
levy	94.22 ± 1.31	0.99 ± 0.01
dropwave	66.32 ± 3.89	0.72 ± 0.05
qeifail	95.95 ± 1.35	0.99 ± 0.01
human	80.40 ± 0.03	0.90 ± 0.02

Table 1: GP Utility Performance across 20 trials on synthetic and human data using 80/20 split of 1200 pair comparisons, demonstrating that the model effectively captures the latent utility.

Summary: Preferential learning robustly recovers the latent utility function with high accuracy and AUC on both synthetic and human data.

5.2 Accuracy-Efficiency Trade-off in Measurement

We tackle *RQ2* by investigating how the computational overhead of diffusion models affects convergence speed. We focus on the EGFR target, plotting accuracy against total FLOPs (FLoating OPerations). We first compare Chai-1 and Vina under the ϵ-greedy acquisition function, noting that Chai produces five poses and Vina produces ten. Figure 3 shows that Vina requires fewer FLOPs to reach the highest accuracy of around 0.43, while Chai-1 is FLOP-intensive and ends with the lowest accuracy, around 0.10. These results show that slower, computationally expensive docking methods impede the exploration-exploitation cycle, while faster models (Vina) significantly boost throughput and convergence speed. Minimizing docking overhead enables the exploration of a larger chemical space.

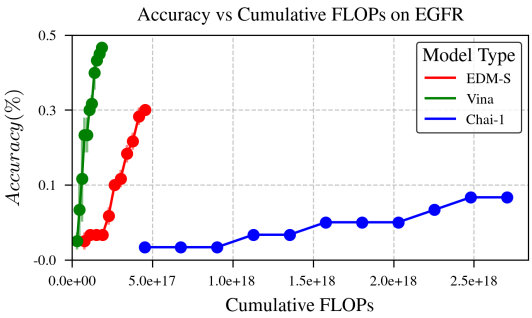

Figure 3: Accuracy over cumulative FLOPs on EGFR under the same screening settings. Vina achieves the highest accuracy with the fewest FLOPs, EDM-S is in between, and Chai uses the most FLOPs with the lowest accuracy.

Understanding the importance of efficiency in VS, we aim to train a highly efficient diffusion model while maintaining its accuracy using data augmentation. Our model (EDM-S) is based on Karras's diffusion model Karras et al. (2022). Training occurs in two phases: We first pre-train on 11 million synthetic pairs generated via pharmacophore alignment to capture diverse docking patterns. This enables the model to learn generalizable features from a broad chemical space that would be difficult to obtain solely from experimental data. We then fine-tune on PDBScan22, a high-quality dataset of approximately 180,000 experimentally determined complexes, to refine the models understanding of biologically relevant interactions. For targeted applications in (5.3), EDM-S is further fine-tuned on 10,000 pairs from García-Ortegón et al. (2022), ensuring robust binding affinity predictions. EDM-S achieves a final accuracy of about 0.30, drastically improving over Chai-1 but still lagging behind Vina. This result shows that diffusion models can be efficient via data augmentation, and future research should investigate methods to make these models even more practical for VS.

We further discuss the diffusion training result to deepen the understanding of the link between the training process and the downstream VS performance. Binding affinity (measured in kilocalories per mole) is evaluated using EGFR as the target, and results from all VS experiments are collected using the Vinardo scoring function. As shown in Figure 16, EDM-S outperforms Chai in binding affinity measurements, 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 accuracy and speed of measuring binding affinity. However, these findings are specific to the EGFR target; further experiments on diverse protein systems are necessary to assess broader generalizability.

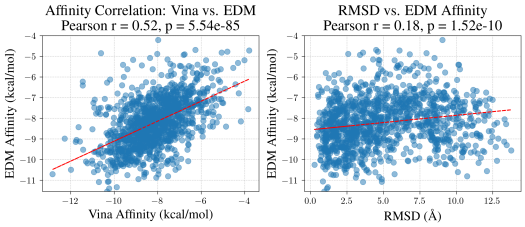

Figure 4: Scatter plots comparing EDM Affinity with Vina Affinity (left) and RMSD (right). A moderate correlation is observed between EDM and Vina affinities (r = 0.52), while no meaningful correlation exists between RMSD and EDM Affinity (r = 0.18).

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 (8) 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, running in 10s on an A100 GPU, demonstrating their practicality for large-scale applications. Most current methods train diffusion models on atomic coordinates, optimizing for low RMSD. However, RMSD only quantifies geometric similarity to a reference structure and is not sensitive to steric clashes or energetically unfavorable interactions: a ligand with RMSD $<$ 2Å may still exhibit steric clashes that disrupt binding, making it a poor candidate despite its structural similarity. In drug discovery, binding strength is more important than geometric accuracy. While binding affinity reflects a ligand's ability to bind to a protein and regulate its function, geometric accuracy only indicates how closely the predicted pose aligns with a reference structure, offering little insight into the drug's regulatory potential. The Vinardo scoring function provides a more meaningful measure of binding affinity by incorporating both energetic and steric factors. Figure 4 shows the weak RMSD-affinity correlation, emphasizing the need for affinity-based scoring as a better measure of drug quality.

Summary: Diffusion models show promise in binding affinity prediction, though physics-based methods demonstrate greater efficiency and accuracy.

5.3 Eliciting Expert Preference for Efficient Screening

To address *RQ3*, we focus on two targets, Epidermal Growth Factor Receptor (EGFR) and Dopamine D2 Receptor (DRD2) proteins, due to their clinical importance and the availability of multiple FDA-approved drugs Cohen et al. (2021). We collect 37 and 58 FDA-approved or late-stage clinical candidates from the PKIDB and DrugBank Carles et al. (2018); Knox et al. (2024) for EGFR and DRD2, respectively, treating them as "goal-optimal" molecules. The screening library comprises $260,000$ 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, defining nuanced *multi-objective* utility functions. In each BO iteration, these experts complete 200 pairwise comparisons via an interactive app, yielding a total of $2,200$ pairs over one full experiment of the study. Overall, the ranking process took a chemist roughly 10 hours to complete. During the process, they have access to SMILES visualizations and ligand properties to inform their decisions (see Figure 5). Our current experiment involves only one chemist; we leave bias mitigation for future work by incorporating feedback from multiple chemists. Finally, We also evaluate an affinity-optimization baseline to determine whether multi-property feedback yields more meaningful optimization for VS, using only Vina, as it typically provides the best performance for affinity-based docking.

As DRD2 is a protein located inside the Central Nervous System (CNS), its drug candidates require substantial considerations for brain penetration. To reflect the distinct pharmacological considerations, we select new objectives for DRD2 that differ from the previous target EGFR. For EGFR, we optimize

affinity, molecular weight (MW), lipophilicity, and half-life, aligning with key properties of kinase inhibitors. For EGFR, these properties are vital as they enable potent target binding, efficient cell penetration, and sustained drug activity-key traits for effective kinase inhibition. Meanwhile, for DRD2, we instead optimize affinity, MW, topological polar surface area (TPSA), predicted drug-induced liver injury (DILI), and predicted blood-brain barrier permeability (BBB). MW, TPSA, and BBB reflect important parameters allowing brain permeability, while DILI provides a standard toxicity indicator; for details on target-specific objective selection, see Appendix D.4.

To validate our objective selection, we compare these properties between drugs and non-drug molecules, confirming that the drug-like molecules exhibit characteristics consistent with our assumptions (see Figures 10 and 12). These choices ensure that our preference optimization aligns with real-world drug design. 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 6% of the library). All experiments are run on an A100 GPU with two seeds for EGFR and one for DRD2. Chai-1 requires 180 GPU-hours to complete 6000 dockings, making it computationally expensive for high-throughput tasks. In contrast, EDM-S finished in 17 GPU-hours. The BO computation–integrated within the cheapvs process–takes around 12 GPU-hours, resulting in an overall process time of about 3 days. Meanwhile, Vina requires no docking as affinities are precomputed from García-Ortegón et al. (2022). However, for a fair comparison, we refer to the measurements from Ding et al. (2023) for Vina on GPU, which we estimate a runtime of approximately 2.4 GPU-hours for 6,000 docking runs.

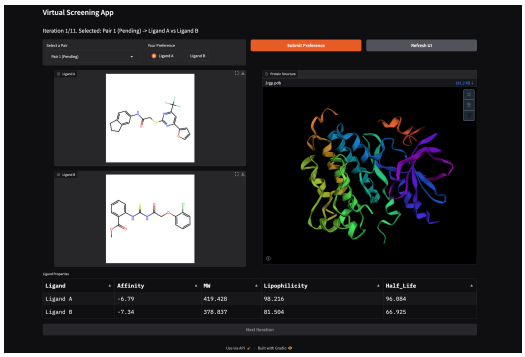

Figure 5: Virtual Screening (VS) App built with Gradio for seamless interaction with chemists

Figure 1 shows how effectively each approach (Multi-Objective, Affinity-Only, and Random) with different dock models (Vina, Chai, EDM-S) identifies the known EGFR and DRD2 ligands. For EGFR experiments, using Vina strategy, guided by expert preferences, attains about 42% accuracy in retrieving these known drugs, substantially surpassing the 22% accuracy of the best affinity-only approach. EDM-S reaches up to 30%. Chai-1 performs poorly due to its high affinity, highlighting that affinity still remains a crucial component in multi-objective optimization. Random screening performs poorly. We observe that two docking models 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. The results for DRD2 show that our multi-objective approach identifies a greater fraction of the 58 known DRD2 drugs compared to the affinity-only model and random selection. After screening about 1200 ligands, its accuracy quickly rises above 60%, while the best affinity-only model remains at zero. These findings *address RQ3*: leveraging expert preference leads to more clinically relevant molecules than relying solely on affinity, and reinforces our hypothesis that incorporating expert-defined preferences leads to more effective VS

Model	ROC AUC	Accuracy
No Interactions	0.67 ± 0.21	0.61 ± 0.16
Second-Order	0.72 ± 0.14	0.68 ± 0.11
Third-Order	0.77 ± 0.11	0.71 ± 0.09
Fourth-Order	$\mathbf{0.79 \pm 0.08}$	$\mathbf{0.74 \pm 0.05}$

Table 2: Linear regression performance across interaction orders. No Interaction uses individual features, while Pairwise, Triple, and Quadruple add second-, third-, and fourth-order interactions. Results averaged over 20 trials with 1200 pair-wise expert preferences (80/20 split), suggest trade-offs in ligand properties.

To understand why multi-objective optimization is more effective, we analyze how the utility model captures expert preferences and the resulting trade-offs in ligand selection. Figure **??** illustrates that the utility model captures expert preferences on EGFR: the box plot shows higher mean utility scores for drug-like compounds, while the heatmap highlights the *trade-offs* of multi-objective optimization. This interplay between pharmacokinetic properties reinforces how the model balances trade-offs to

identify clinically relevant candidates. Understanding the trade-off nature of drug discovery requires understanding how optimizing one objective impacts others. Toward this goal, we use simple linear regression to test whether high-order interaction is necessary for out-of-sample fit. The null model consists of individual effects, while the alternative models incorporate higher-order interactions to capture the interdependence among ligand properties. Table 2 shows that models with higher-order interactions generalize better, indicating that ligand properties exhibit complex interdependencies that influence predictive performance. GP, our main predictive model, generalizes this ideal to infinite dimensional feature space, capturing high-order interaction terms through kernel functions Schölkopf & Smola (2002); Mercer (1909); Williams (1998), allowing it to naturally model intricate dependencies among ligand properties without explicitly defining interaction orders. The superior out-of-sample fit of the alternative models and the complex utility landscape conclude that optimizing one objective does not necessarily improve overall drug potential, highlighting the shortcomings of single-objective screening.

Summary: Incorporating expert preferences outperforms affinity-only methods, emphasizing the critical role of chemical intuition in drug discovery.

6 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. 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. Specifically, *CheapVS* successfully identified up to 16 out of 37 known drugs for EGFR and 36 out of 57 for DRD2 targets. 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 and further accelerate the drug discovery process.

7 Acknowledgement

This work was supported by the 2024 HAI-Google Cloud Credits Grant on "Proactive Pandemic Preparedness: Accelerating Antiviral Drug Discovery by Empowering Chemists with Deep Generative Models." LH.P. acknowledges support from the Vingroup Science and Technology Scholarship for Doctoral Degrees and EPSRC 2886971. SK acknowledges support by NSF 2046795 and 2205329, IES R305C240046, ARPA-H, Stanford HAI, RAISE Health, OpenAI, and Google.

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.

Jacques Boitreaud, Jack Dent, Matthew McPartlon, Joshua Meier, Vinicius Reis, Alex Rogozhonikov, and Kevin Wu. Chai-1: Decoding the molecular interactions of life. *bioRxiv*, 2024. doi: 10.1101/2024.10.10.615955. URL https://www.biorxiv.org/content/early/2024/10/11/2024.10.10.615955.

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.

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.

Oh-Hyeon Choung, Riccardo Vianello, Marwin Segler, Nikolaus Stiefl, and José Jiménez-Luna. Extracting medicinal chemistry intuition via preference machine learning. *Nature Communications*, 14(1):6651, 2023.

Wei Chu and Zoubin Ghahramani. Preference learning with gaussian processes. In *Proceedings of the 22nd International Conference on Machine Learning*, ICML '05, pp. 137144, 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.

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.

Ji Ding, Shidi Tang, Zheming Mei, Lingyue Wang, Qinqin Huang, Haifeng Hu, Ming Ling, and Jiansheng Wu. Vina-gpu 2.0: further accelerating autodock vina and its derivatives with graphics processing units. *Journal of chemical information and modeling*, 63(7):1982–1998, 2023.

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, Andr 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.

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. *J. Basic Eng.*, 1964.

Tze Leung Lai and Herbert Robbins. Asymptotically efficient adaptive allocation rules. *Advances in applied mathematics*, 6(1):4–22, 1985.

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.

JXVI Mercer. Functions of positive and negative type, and their connection the theory of integral equations. philos. *Trans. Roy. Soc. London*, pp. 415–446, 1909.

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.

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

Bernhard Schölkopf and Alexander J Smola. *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press, 2002.

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.

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.

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.

Christopher KI Williams. Computation with infinite neural networks. *Neural Computation*, 10(5): 1203–1216, 1998.

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

We summarize the notation used in our paper in Table 3.

Symbol	Description
$\mathcal{L}$	Ligand library used for VS.
ℓ_i	Ligand i in the ligand library $\mathcal{L}$.
ℓ_{3D}	3D coordinate vector of ligand ℓ.
g_P	Affinity model mapping ligand fingerprints to binding affinity.
f	Latent Utility model learning from preference data.
h	Physics-based affinity scoring function.
x	Ligand properties, including physicochemical and ADMET.
ℓ	Morgan Fingerprint representation of the ligand's structure.
α	Acquisition function in BO for ligand selection.
R	Regret, quantifying the gap between the best possible and selected ligand.
U	Utility values of ligands.
k	Used for selecting the top-k compounds.
ρ	Protein target for VS.
p_θ	Docking diffusion model, predicting ligand-protein binding.
g_ρ	Datasets acquired to train affinity model.
f	Datasets acquired to train utility model.

Table 3: Notation

B 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.
- **Greedy**: Selects the candidate with the highest predicted performance at each step. This purely exploits the current model estimates and does not explicitly encourage exploration.
- **ϵ-Greedy** Lai & Robbins (1985): With probability ϵ (typically 5%), selects a candidate at random (exploration), and otherwise selects the best predicted candidate (exploitation). This method is a simple yet effective way to balance exploration and exploitation.

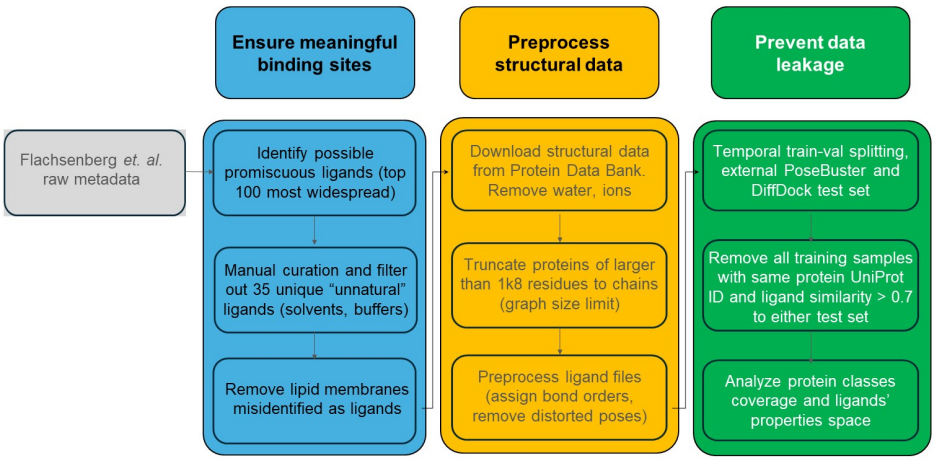

Figure 6: PDBScan22 data curation workflow. The process consists of three main steps: (1) Ensuring meaningful binding sites by filtering promiscuous ligands, removing unnatural molecules such as solvents and buffers, and eliminating misidentified lipid membranes. (2) Preprocessing structural data by downloading structures from 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, and removing training samples with proteins sharing UniProt IDs and highly similar ligands (>0.7 similarity) with test sets.

C Preliminary Analysis on the Data

As noted in Figure 7, 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. Additionally, we also perform a similar comparison on the Plinder dataset Durairaj et al. (2024) to further evaluate the differences in data distribution and model performance across diverse datasets.

C.1 Diversity of Data: Proteins

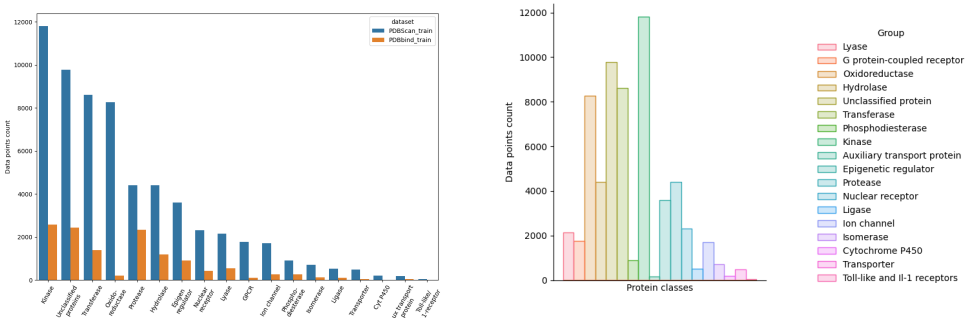

Figure 7: Analysis of protein properties: Protein classes distribution, Protein classes data point count

Understanding the diversity of protein classes in the dataset is essential for evaluating its coverage and potential biases in molecular docking tasks. Figure 7 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 impacting model performance on these classes. These insights highlight the importance of data augmentation techniques to balance protein representation and improve downstream learning.

C.2 DIVERSITY OF DATA: LIGANDS

Figure 9 compares the distribution of key molecular interactions across the PDBScan++ and PLINDER datasets, including hydrogen bonds, salt bridges, pi-stacking, hydrophobic interactions, and halogen bonds. PDBScan++ consistently contains more ligand-protein interactions than PLINDER, reflecting its larger dataset size. Hydrogen bonds and hydrophobic interactions are the most prevalent, while halogen bonds are the least common. Notably, PDBScan++ includes PDBScan22 along with 250k synthetic pharmacophore-ligand pairs with the lowest affinity, added to match the number of compounds in PLINDER, ensuring a balanced comparison.

Figure 8 presents the distribution of key physicochemical properties across the PDBScan++ and PLINDER datasets, including QED drug-likeness, molecular weight, Wildman-Crippen LogP, hydrogen bond donors and acceptors, polar surface area, rotatable bonds, and aromatic rings. Across all properties, PDBScan++ exhibits a broader and more diverse range of molecular characteristics compared to PLINDER, reflecting its larger dataset size. The QED scores and molecular weights of compounds in both datasets follow similar distributions, but PDBScan++ has a wider spread. The Wildman-Crippen LogP distribution indicates that PDBScan++ includes more hydrophobic molecules. Additionally, PDBScan++ contains a higher number of hydrogen bond donors and acceptors, as well as greater structural flexibility (rotatable bonds) and aromaticity (aromatic rings), highlighting its increased chemical diversity.

Together with the molecular interaction distributions in Figure 9, these results emphasize that PDBScan++ encompasses a broader and more chemically diverse set of compounds than PLINDER, ensuring a comprehensive representation of molecular properties. Notably, PDBScan++ includes PDBScan22 along with 250k synthetic pharmacophore-ligand pairs with the lowest affinity, introduced to match the number of compounds in PLINDER.

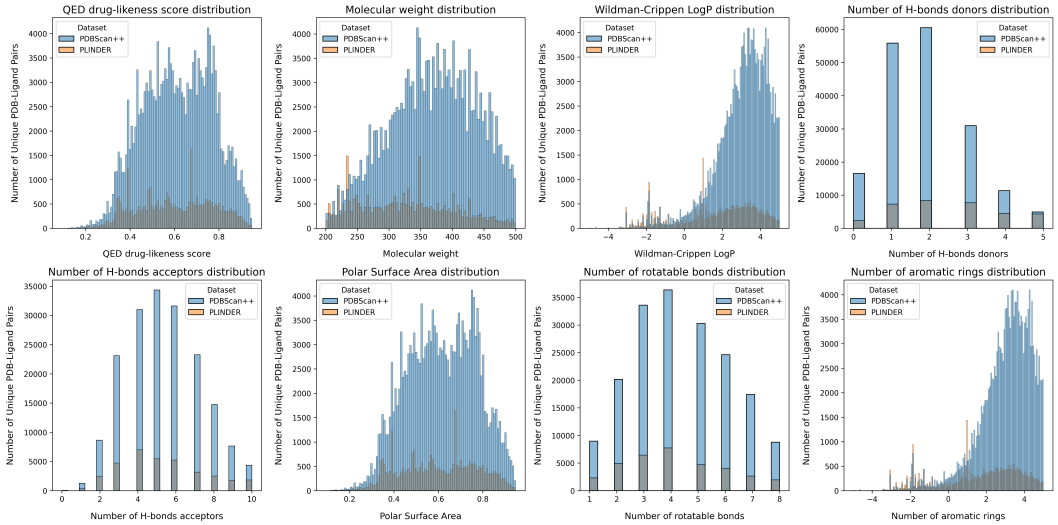

Figure 8: Comparison of physicochemical properties between PDBScan++ and PLINDER, including QED drug-likeness, molecular weight, LogP, hydrogen bond donors/acceptors, polar surface area, rotatable bonds, and aromatic rings. PDBScan++ shows greater diversity across all properties.

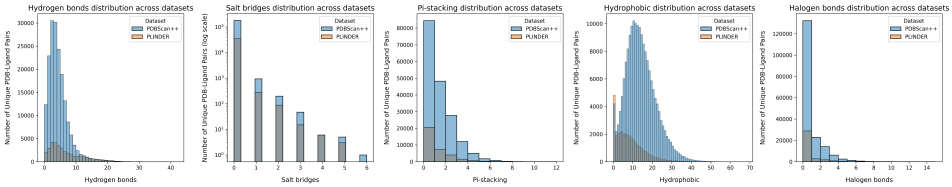

Figure 9: Distribution of molecular interactions in PDBScan++ and PLINDER, including hydrogen bonds, salt bridges, pi-stacking, hydrophobic interactions, and halogen bonds. PDBScan++ exhibits a broader range of interactions due to its larger dataset size.

C.3 Diversity of Protein-Ligand Dataset

Dataset	Dataset size	Main approaches	Limitations
PDBBind (Liu et al., 2017)	Approx. 20k complexes	Human-curated experimental structures with experimental binding affinity	Limited number of data points
BindingMOAD (Wagle et al., 2023)	Approx. 40k complexes	Human-curated experimental structures (with or without binding affinity)	Limited number of data points
BioLiP2 (Zhang et al., 2024)	Approx. 470k organic ligand complexes	Semi-manual-curated experimental structures (with or without binding affinity)	Include non-specific ligands (anions, crystal artefacts, solvents, etc)
PDBScreen (Cao et al., 2024)	True ligands: 23k unique ligands Generated decoys: approx. 110k	- Automated filtering from the PDB, excluding endogenous ligand (ATP, ADP, etc.) - Data augmentation by redocking and cross-docking - Also include artificially generated decoy ligands	- Purposely built for screening and scoring, but excluded endogenous ligands which represent important pockets - Redocked and cross-docked poses do not enrich protein or ligand diversity
PLINDER Durairaj et al. (2024)	Approx. 450k unique (organic) ligands	- Automated annotation of PDB structures to retrieve broad-termed ligands - Graph-based multiple-similarity data splitting to debias the train-test split	- Include covalent modifications of the proteins as surrogate ligands (glycosylation) - Include ions on the broadly defined ligand category
PapyrusScan	Approx. 11 million	- Synthetic data from 2D binding information	- Synthetic data - Unbalanced number of data points between proteins depending on 2D data

Table 4: Comparison of community-available protein-ligand structural datasets.

Table 4 compares various protein-ligand structural datasets, with a focus on PDBScreen and PapyrusScan in contrast to existing community datasets. PDBScreen refines structural data by filtering endogenous ligands and augmenting diversity through redocking and cross-docking, making it well-suited for screening and scoring tasks. However, its exclusion of endogenous ligands may overlook important binding pockets. In contrast, PapyrusScan is the largest dataset, containing approximately 11 million protein-ligand interactions derived from 2D binding data, offering extensive coverage but relying on synthetic data, leading to potential biases and imbalances across proteins. Compared to PDBBind, BindingMOAD, and BioLiP2, which primarily rely on human-curated or semi-curated experimental structures, PDBScreen and PapyrusScan emphasize data augmentation and large-scale synthetic generation, respectively. While PLINDER provides a broad dataset with automated annotation and debiased train-test splitting, it includes covalent modifications and ions as ligands, introducing potential noise. This comparison highlights the complementary nature of PDBScreen

and PapyrusScan, balancing curated experimental data with large-scale synthetic augmentation to enhance ligand-protein modeling.

D MORE RESULTS ON PREFERENTIAL MULTI-OBJECTIVE BAYESIAN OPTIMIZATION

D.1 CHEAPVS'S PSEUDOCODE

Algorithm 1 CheapVS's Algorithm

Require: Ligand library $\mathcal{L} = \{\ell_1, \ldots, \ell_N\}$, target protein ρ, docking model p_θ, acquisition function α

Ensure: Top-k drug ligands for target ρ

$\mathcal{D} \leftarrow \emptyset$, $\mathcal{D}_{g_\rho} \leftarrow \emptyset$, $\mathcal{D}_f \leftarrow \emptyset$, $\mathcal{F} \leftarrow \emptyset$, $\mathcal{L}_{\text{tox}} \leftarrow \emptyset$, $\mathcal{L}_{\text{sol}} \leftarrow \emptyset$, $X^{\text{aff}}_{\ell,\rho} \leftarrow \emptyset$

$\mathcal{D}_i \leftarrow \{\ell \in \mathcal{L} \mid U(0,1) < 0.01\}$ // Select 1% of $\mathcal{L}$ at random

for each ligand $\ell_i \in \mathcal{L}$ **do**

  $\ell_{i,\mathcal{M}} \leftarrow \text{MORGANFINGERPRINT}(\ell_i)$

  $x^{\text{tox}}_{\ell_i} \leftarrow \text{RDKITTOXICITY}(\ell_i)$

  $x^{\text{sol}}_{\ell_i} \leftarrow \text{RDKITSOLUBILITY}(\ell_i)$

  $\mathcal{F} \leftarrow \mathcal{F} \cup \{\ell_{i,\mathcal{M}}\}$, $\mathcal{L}_{\text{toc}} \leftarrow \mathcal{L}_{\text{toc}} \cup \{x^{\text{tox}}_{\ell_i}\}$, $\mathcal{L}_{\text{sol}} \leftarrow \mathcal{L}_{\text{sol}} \cup \{x^{\text{sol}}_{\ell_i}\}$

end for

while computational budget not reached **do**

  $g_P \sim \mathcal{GP}(\mu, k)$ // Initialize affinity model with Gaussian likelihood

  $f \sim \mathcal{GP}(\mu, k)$ // Initialize utility model with pairwise likelihood:

  $\mathcal{D} \leftarrow \mathcal{D} \cup \mathcal{D}_i$ // Add selected ligands to the dataset

  $\mathcal{L} \leftarrow \mathcal{L} \setminus \mathcal{D}_i$ // Remove selected ligands from the library

  For each ligand $\ell_i \in \mathcal{D}_i$:

    $\ell_{i;\text{3D}} \sim p_\theta(\ell_i, \rho)$

    $x^{\text{aff}}_{\ell_i,\rho} \leftarrow \min_{\ell_{i;\text{3D}} \in \mathbb{R}^{3\times N_{\ell_i}}} h(\ell_{i;\text{3D}}, \rho)$

    $X^{\text{aff}}_{\ell,\rho} \leftarrow X^{\text{aff}}_{\ell,\rho} \cup \{x^{\text{aff}}_{\ell_i,\rho}\}$

  $\mathcal{D}_{g_\rho} \leftarrow \mathcal{D}_{g_\rho} \cup \{(\mathcal{F}(\mathcal{D}_i),\ X^{\text{aff}}_{\ell,\rho})\}$

  Fit g_P on $\mathcal{D}_{g_\rho}$

  $\mathcal{I} \leftarrow \text{random pairs}(\mathcal{D}_i)$ // $\mathcal{I}$: set of index pairs from $\mathcal{D}_i$

  $X_{\text{train}} \leftarrow \text{concat}\Big(X^{\text{aff}}_{\ell,\rho}, \mathcal{L}_{\text{tox}}(\mathcal{D}_i),\ \mathcal{L}_{\text{sol}}(\mathcal{D}_i)\Big)$

  $Y_e \leftarrow \text{chemists_ranking}(X_{\text{train}}, \mathcal{I})$

  $\mathcal{D}_f \leftarrow \mathcal{D}_f \cup \{(X_{\text{train}},\ Y_e)\}$

  Fit f on $\mathcal{D}_f$

  $\hat{X}^{\text{aff}}_{\ell,\rho} \leftarrow g_\rho(\mathcal{F}(\mathcal{L}))$ // Posterior inference: $\hat{x}^{\text{aff}}_{\ell,\rho} = g_\rho(\ell_{\mathcal{M}})\ \forall\ \ell \in \mathcal{L}$

  $\hat{X} \leftarrow \text{concat}(\hat{X}^{\text{aff}}_{\ell,\rho},\ \mathcal{L}_{\text{sol}}(\mathcal{L}),\ \mathcal{L}_{\text{tox}}(\mathcal{L}))$

  $\mathcal{D}_i \leftarrow \text{Top}_k\Big\{\ell \in \mathcal{L} \mid \mathbb{E}_{p(x^{\text{aff}}_{\ell,\rho}|\ell,g_\rho)}\alpha(f(\hat{X}))\Big\}$

end while

Return $\mathcal{D}$

D.2 DRD2 EXPERIMENTS

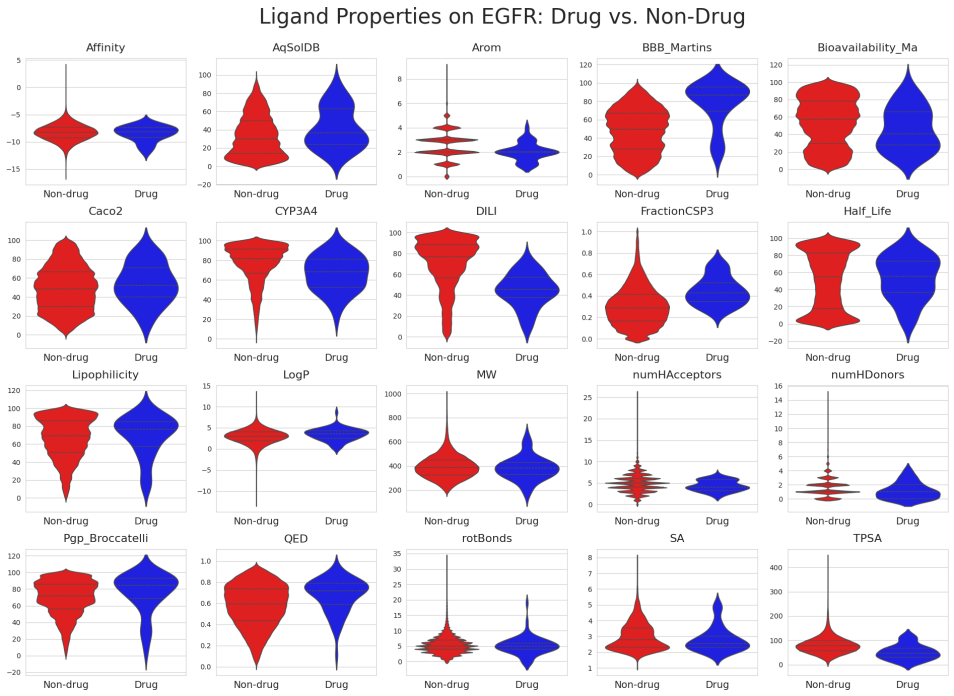

Figure 10: Comparison of physicochemical and pharmacokinetic properties between known DRD2-targeting drugs (blue) and non-drug molecules (red) within 100,000-compound screening library. Violin plots illustrate key attributes such as affinity, molecular weight (MW), topological polar surface area (TPSA), blood-brain barrier permeability (BBB), and drug-induced liver injury (DILI), among others. The observed differences validate our objective selection, showing that drug-like molecules generally align with expected characteristics for CNS-active compounds, such as lower MW, optimized BBB permeability, and favorable toxicity profiles.

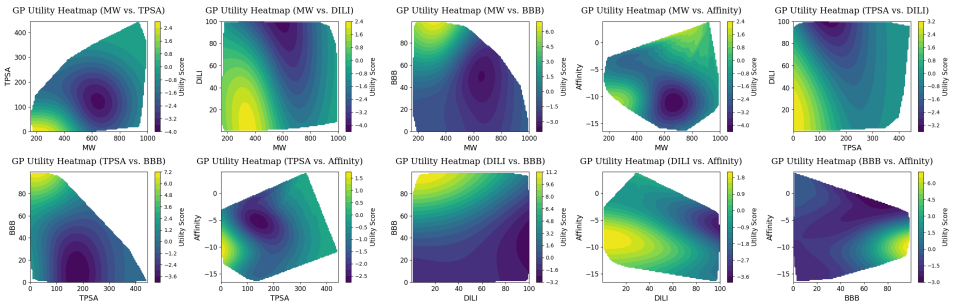

Figure 11: Predictive utility scores after BO on expert preference elicitation on DRD2. Heatmaps illustrate utility over two objectives while keeping others three at their mean. Results align well with established medicinal chemistry ranges, favoring optimal MW (200-400), TPSA (below 140), while maximizing BBB and minimizing DILI and binding affinity.

D.3 EGFR EXPERIMENTS

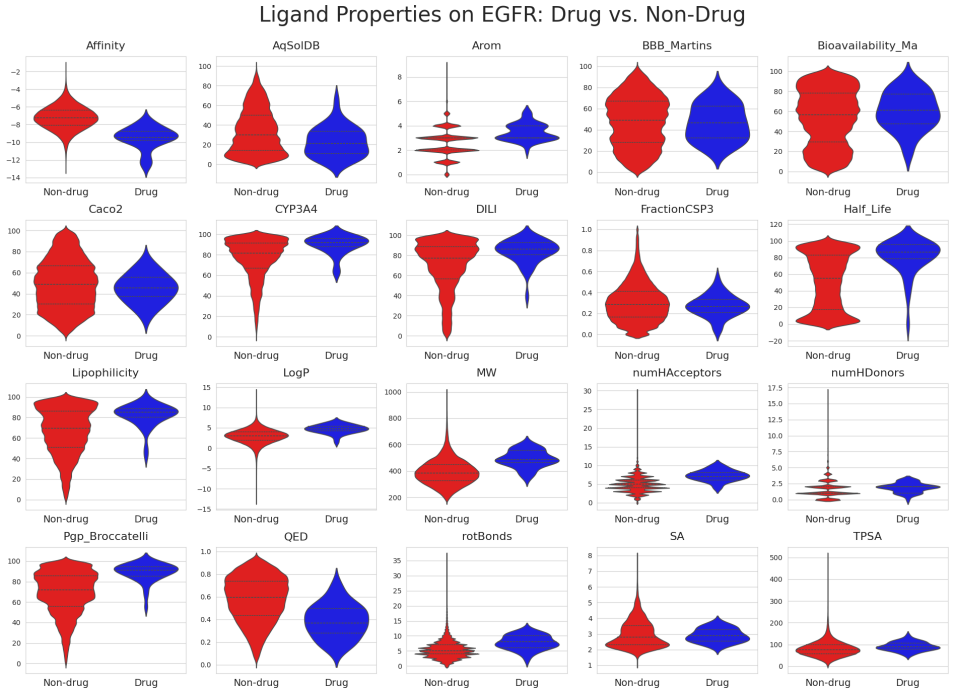

Figure 12: Comparison of physicochemical and pharmacokinetic properties between known EGFR-targeting drugs (blue) and non-drug molecules (red) within the 100,000-compound screening library. The observed differences confirm that drug-like molecules generally exhibit characteristics favorable for kinase inhibition, including higher MW and optimized lipophilicity.

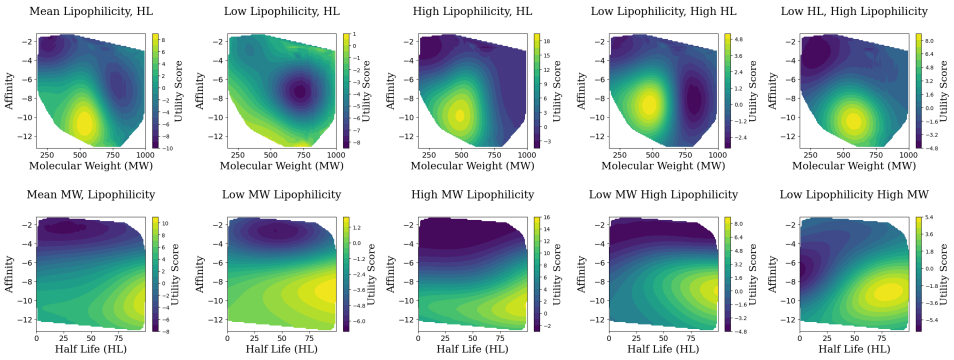

Figure 13: More on Gaussian process (GP) utility surfaces learned from expert preference data, illustrating the interplay among molecular weight (MW), halflife (HL), affinity, and lipophilicity. Each heatmap shows the predicted utility (color scale) over two of these variables while holding the others fixed at the levels indicated in each title. Higher (yellow) regions correspond to more favorable tradeoffs according to the elicited expert preferences, providing insights for optimizing lead compounds in drug discovery.

D.4 GUIDELINES FOR CHEMISTS

The virtual screening app 5 assists chemists in evaluating and comparing ligands by providing key molecular properties such as binding affinity, molecular weight (MW), lipophilicity, and half-life. It integrates SMILES-based molecular visualizations alongside numerical data, enabling users to analyze structural and chemical characteristics effectively. Chemists select their preferred ligand

based on predefined criteria, and their selections contribute to refining the models predictive capabilities, improving its ability to identify promising drug-like candidates over time. However, the selection process is highly dependent on the biological target, as different proteins require distinct pharmacokinetic and pharmacodynamic considerations.

For example, targeting DRD2 in neuropharmacology necessitates prioritizing blood-brain barrier (BBB) permeability, as compounds must effectively penetrate the central nervous system while maintaining an appropriate balance between molecular weight and topological polar surface area (TPSA). Additionally, potential toxicity, such as predicted drug-induced liver injury (DILI), should be considered to ensure safety. In contrast, when designing inhibitors for EGFR in cancer therapy, selectivity, and affinity becomes paramount, as high target specificity minimizes off-target interactions and reduces systemic toxicity. A well-structured screening approach should reflect these protein-specific requirements, allowing chemists to weigh molecular properties appropriately when ranking compounds. Effective use of CheapVS requires some degree of expertise in medicinal chemistry, as misprioritizing criteria may lead to the selection of ineffective molecules.

D.5 Synthetic Experiments

We examine how well *CheapVS* identifies high-utility solutions under various synthetic utility functions. Before running on real human preference data, we first test on synthetic functions. 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 based on the corresponding utility values. In addition, we simulate four main objectives relevant to drug discovery: binding affinity, rotatable bonds, molecular weight, and LogP. For computational feasibility, we use a 20k-ligand subset sampled from the Dockstring library García-Ortegón et al. (2022). Since the docking affinity values have already been computed for all compounds, we can determine both regret and accuracy. To ensure robustness, we repeat all experiments across five random seeds and report mean and standard deviation across runs. Furthermore, we evaluate a range of acquisition functions, including `qEUBO`, `qTS`, `qEI`, `qPI`, `qUCB`, `Greedy`, ϵ`-Greedy`, and `Random`. Figure 14 and Figure 15 displays the log regret and accuracy versus the number of compounds screened, illustrating the effectiveness of different acquisition strategies. The results show that regret consistently decreases and accuracy improves across all acquisition functions, with more advanced methods converging significantly faster than random baselines. These findings demonstrate that preferential BO effectively learns multi-objective trade-offs in synthetic benchmarks.

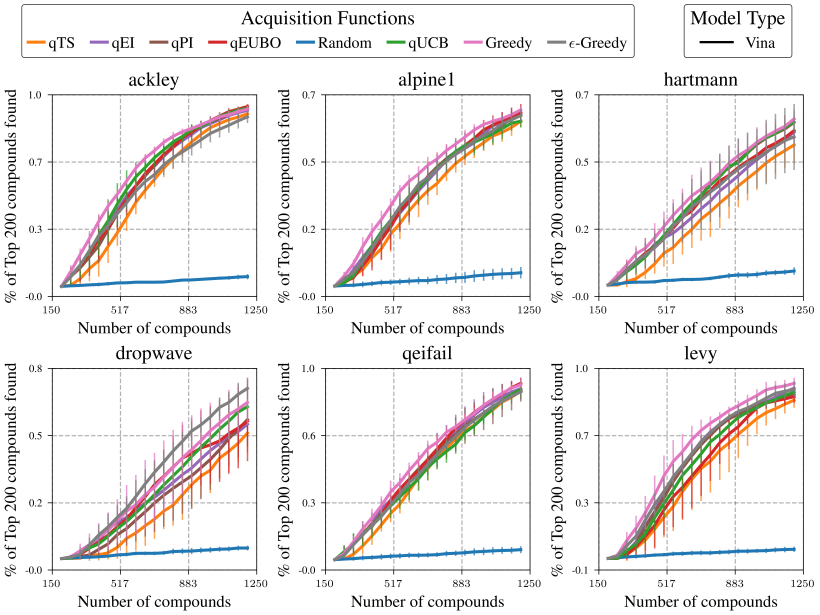

Figure 14: Preferential Multi-Objective Optimization results on multiple synthetic functions. The y-axis shows $\log(\textit{regret})$. The results compare multiple acquisition functions across various benchmark functions. Error bars indicate standard deviations across five seeds.

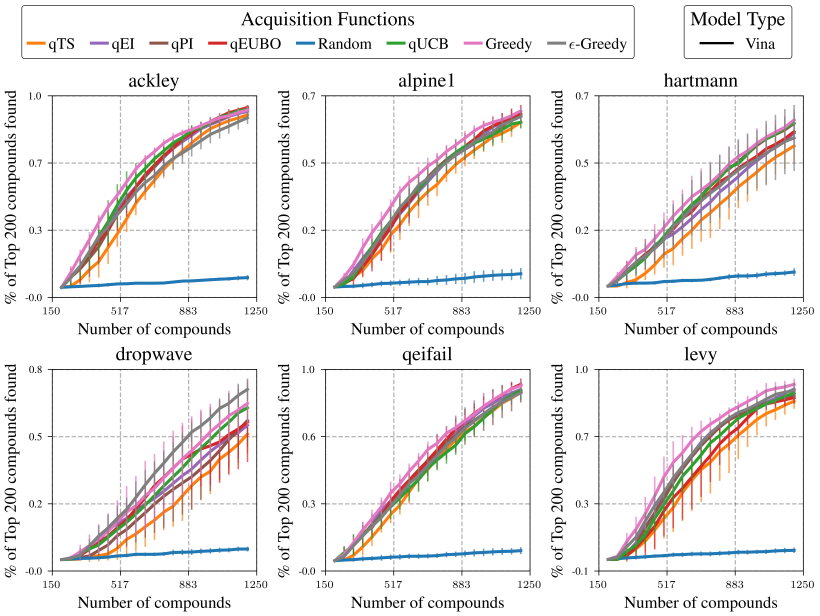

Figure 15: Preferential Multi-Objective Optimization results on multiple synthetic functions. The y-axis shows accuracy. The results compare multiple acquisition functions across various benchmark functions. Error bars indicate standard deviations across five seeds.

E 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 5: 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.

Model Type	Accuracy (%)	ROC AUC
Fully-connected Neural Net	0.9505±0.0146	**0.9913±0.0081**
Decision Tree	0.7853±0.0285	0.7858±0.029
Pairwise Gaussian Process	**0.9563±0.0146**	0.9724±0.0161

Table 6: Comparison of utility model performance in predicting preference-based rankings from ligand properties on Ackley function. The table reports the classification accuracy and ROC-AUC 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 second highest ROC-AUC, demonstrating superior performance in modeling pairwise preferences and learning utility functions from ligand physicochemical properties.

F DIFFUSION MODEL TRAINING: HYPERPARAMETERS AND PERFORMANCE RESULTS

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 7: Hyperparameters for training EDM-S and DockScan22

Metrics	PoseBuster V1 Top-1 RMSD () % < 2	PoseBuster V1 Top-1 RMSD () % < 5	PoseBuster V2 Top-1 RMSD () % < 2	PoseBuster V2 Top-1 RMSD () % < 5	PDBBind Top-1 RMSD () % < 2	PDBBind Top-1 RMSD () % < 5	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 8: 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.

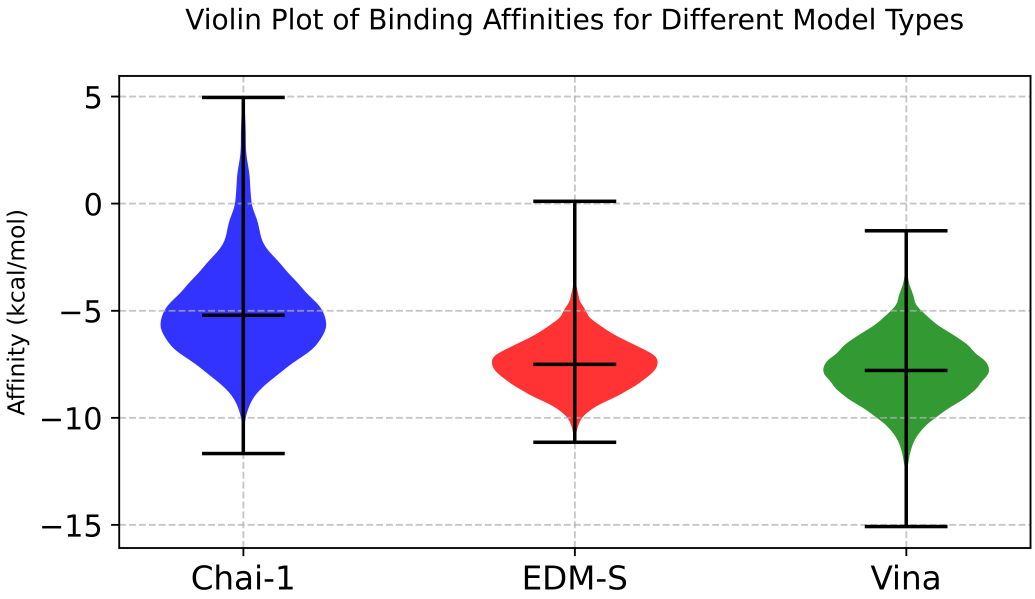
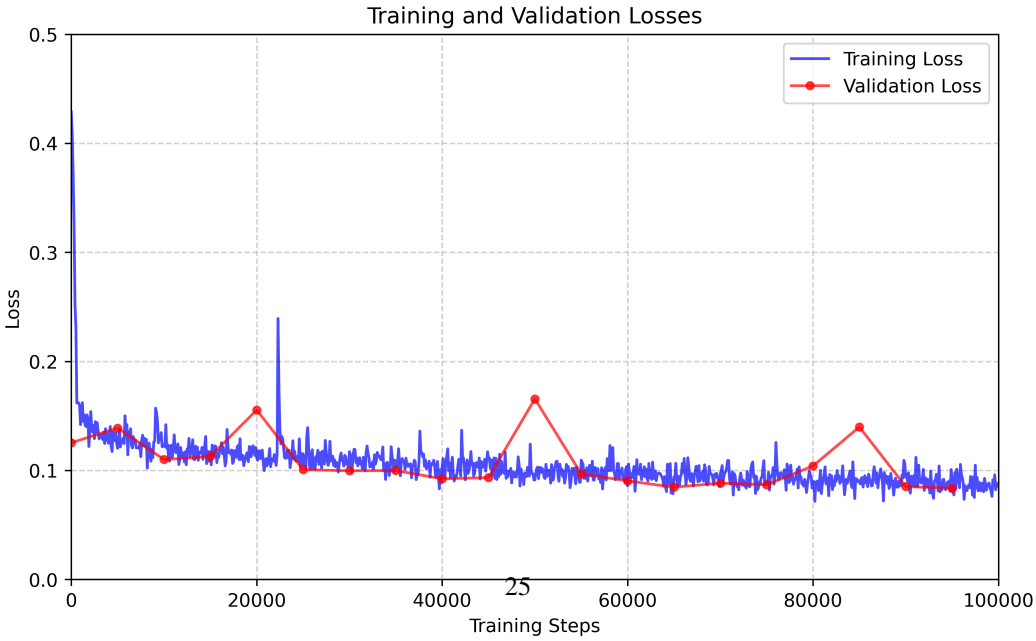
Figure 16: Violin plot of binding affinities (kcal/mol) for different docking models on the EGFR protein with 6000 ligands. Vina achieves the lowest median binding affinity, followed by EDM-S, while Chai exhibits the weakest binding.