Multi-Objective Optimization to Improve Structure-Based Virtual Screening at Large Scale

Stasa Skorupan

Thesis submitted to the University of Ottawa
in partial fulfillment of the requirements for the
Master of Science in Chemistry

Department of Chemistry and Biomolecular Sciences
Faculty of Science
University of Ottawa

Abstract

The recent expansion of commercial, chemical databases to billions of make-on demand molecules has restructured hit identification methods in early-stage drug discovery. The exploration of these ultra-large chemical databases represents a new frontier for the identification of novel, potent, and selective drug candidates. Molecular docking is a computational technique used to predict the binding modes and affinities of small molecules to a protein binding site, enabling fast and cost-effective virtual screening (VS) of large chemical libraries to identify novel hit compounds. Machine learning (ML)-augmented molecular docking methods were proposed as a new paradigm to screen ultra-large, commercial, chemical libraries and expand VS to billions of molecules. These models are generally trained on a small subset of docked protein-ligand complexes to predict binding affinities of the remainder of the dataset, thereby streamlining the identification of promising top-scoring candidates at reasonable computational costs.

Molecular docking continues to be crucial in the acceleration and cost-effectiveness of hit identification. However, the simplified modelling of protein-ligand interactions introduces a significant number of artifact molecules, especially for ultra-large libraries, with the visual inspection of the top-ranked docking hits being the standard protocol to remove these artifacts. Common criteria considered in the assessment of modelled protein-ligand binding includes strain and unsatisfied polar ligand or protein heteroatoms and thus, various computational tools have been developed to automate this filtering process at scale. Moreover, ML-accelerated docking models can implicitly learn and propagate these inherent molecular docking artifacts. Importantly, no current model addresses the effect of artifacts inherent to molecular docking predictions performed at large scale.

The research herein developed new multi-objective optimization (MOO), ML-accelerated molecular docking models to aid in artifact filtering and selection of promising candidates in the early stages of drug discovery. These models incorporate selected three-dimensional medicinal chemistry properties that were thoroughly evaluated for their potential to improve early enrichment in VS. Chapter 2 presents a retrospective evaluation of ligand strain and unsatisfied hydrogen bonds as filters in the post-processing of molecular dockings to assess the impact on early enrichment. We found their effect on enrichment to be highly system-dependent: there was

no single threshold that led to an enrichment of all protein-ligand datasets explored, but several proteins showed significant enrichment when using strain, unsatisfied hydrogen bonds, or both as filters at specific thresholds. Chapter 3 presents simulated large-scale prospective VS campaigns with the developed MOO ML-accelerated molecular docking models and selected three-dimensional medicinal chemistry filters. Multi-task learning (MTL) was explored to this end. In most systems studied, there was an improvement in early enrichment with filtering compared to molecular docking score predictions alone. MTL models showed the potential to improve early enrichment in large-scale VS, while accelerating runtime (3-4x) and significantly reducing (90-99%) the chemical database size. This research contributes an open-source, medicinal chemistry-informed ML-accelerated molecular docking model towards the development of new drug discovery tools.

Acknowledgements

I would like to express immense gratitude to my supervisor, Francesco Gentile – you are the example of an excellent supervisor and researcher. Your exceptional leadership, knowledge, encouragement, and humble character made this experience invaluable. I thank you for the time to learn new skills in the field of computational drug discovery under your supervision.

Thank you to my thesis advisory committee, Professors Tom Woo and Jeffrey Keillor, for your excellent teaching, time, and advice in the fulfillment of this masters.

To those involved in this research – Thank you to David Kouvchinov and Jean-Paul Kazzi for your help in the earlier stages of this research, and Yasaman Shahrasbi for your help in machine learning.

Thank you Sumin Park, who led us in this masters together – I am so grateful to be one of the two first graduate students in this research group, but moreover that you were the other one.

Thank you to all other members of the Gentile Lab, Rahul Ravichandran, Thanawat Thaingtamtanha, and Kaitlyn Bessette.

Kian Mansour, thank you for upholding me throughout this masters. Thank you to the Zoom algorithm in communications class for pairing me with my soon-to-be husband every single week of the course – haha it worked.

Thank you to my family, Zoran Skorupan, Jadranka Skorupan, Marko Skorupan, and my dog Jack Skorupan. I am eternally grateful for everything.

Table of Contents

List of Figures

List of Tables

Abbreviations

AI – Artificial Intelligence
AUC – Area Under Curve
CNN – Convolutional Neural Network
CASP – Critical Assessment of Structure Prediction
CE – Cross Entropy
Cryo-EM – Cryo-Electron Microscopy
DUD –Database of Useful Decoys
DUD-E – Database of Useful Decoys: Enhanced
DD – Deep Docking
DL – Deep-Learning
DNN – Deep Neural Network
FPR – False Positive Rate
GPCR – G Protein-Coupled Receptor
GPU – Graphical Processing Unit
HTS – High Throughput Screening
LGA – Lamarckian Genetic Algorithm
ML – Machine learning
MSE – Mean Squared Error
MOE – Molecular Operating Environment
MLP – Multi-Layer Perceptron
MOO – Multi-Objective Optimization
MTL – Multi-Task Learning
NMR – Nuclear Magnetic Resonance spectroscopy
PDB – Protein Data Bank
r – Pearson Correlation Coefficient
R^2 – R-squared
RMSD – Root Mean Square Deviation
RMSE – Root Mean Squared Error
ROC – Receiver Operating Characteristic
STL – Single-Task Learning
SBDD – Structure-Based Drug Discovery
SBVS – Structure-Based Virtual Screening
3D – Three-Dimensional
TEU – Torsional Energy Units
TVT – Training Validation Testing
TPR – True Positive Rate
VS – Virtual Screening

Preface

Section 1.2.2 [Artificial Intelligence-Augmented Docking] was adapted from the following publication:

Ravichandran, R.; Park, S.; Skorupan, S.; Bessette, K.; Gentile, F. Artificial Intelligence in Early Stages of Structure-Based Drug Discovery. In *Drug Discovery Stories*; Elsevier, 2025; pp 3–24. https://doi.org/10.1016/B978-0-443-23932-8.00001-7.

This chapter section was written by Stasa Skorupan under the supervision of Dr. Francesco Gentile and in collaboration with co-authors Rahul Ravichandran, Sumin Park, and Kaitlyn Bessette.

Chapter 1: Introduction

1.1 Structure-Based Drug Discovery

Drug discovery is the multi-step and multi-disciplinary process of identifying and characterizing potential disease modulators[1] – a complex task with a failure rate of 90% in clinical development.[2] The drug discovery and development process costs over 1 billion CAD and takes 10-15 years before market release.[3] The early stages of drug discovery focus on identifying diversified chemical starting points, through the multistep processes of target identification, target validation, lead identification, and lead optimization (Figure 1.1). Target-based drug discovery has served as the main approach to lead generation in the pharmaceutical industry for the past three decades.[4] Modern hit identification methods are categorized into biochemical and cellular high throughput screening (HTS), biophysical screening, fragment-based screening, or in-silico screening.[5] HTS is a conventional tool used in the investigation of druggable protein families[6] and has contributed to the discovery of most novel scaffolds of recent clinical candidates[7] – however this method is expensive, time-intensive, and limited to chemical screenings of hundreds of thousands to few million compounds.[5] The chemical search space is restricted to chemical vendors' and in-house screening databases containing several million on-shelf compounds.[7] Moreover, most clinical candidates are more complex and structurally different from their starting points[8] – driving the investigation of new chemical spaces. In-silico methods have been proposed to mitigate the limitations of HTS in the early stages of drug discovery. Structure-based drug discovery (SBDD) uses the three-dimensional (3D) structure of a protein target to characterize the development of new disease modulators. The integration of artificial intelligence (AI) in the earlier stages of SBDD are discussed herein.

	Target identification	Target validation	Hit identification	Lead generation
Objective	The identification of a biological target involved in a particular disease mechanism, that a therapeutic can modulate	The experimental confirmation that a particular biological target can be engaged to produce the intended therapeutic effect	The identification of chemical compounds that show measurable activity against the validated target	The optimization and selection of hits into lead compounds, with drug-like properties suitable for further development

Figure 1.1. Pre-clinical drug discovery stages. The objective of each pre-clinical drug discovery stage.

1.1.1 Target Structure Determination

The first step in the early stages of drug discovery is target identification – the identification of biological targets involved in a particular disease mechanism that can be modulated with a drug to produce a therapeutic effect.[9,10] The majority of pharmaceutical drug targets are part of five protein families: G protein-coupled receptors (GPCRs), ion channels, kinases, nuclear hormone receptors, and proteases.[11] GPCRs are the largest protein family targeted in drug development – implicated in a multitude of diseases.[12] Target identification methods include experimental, multiomic, and computational methods or a combination thereof to characterize on-target and off-target effects within the proposed therapeutic mechanism.[13,14] Recent AI methods have been introduced to integrate multiomic data and analyze biological networks in the search of modulable targets.[15] Subsequent to target identification, target validation is the experimental confirmation that a particular biological target can be engaged to produce the intended therapeutic effect.[16] Modern techniques include initial in-silico target screenings with potential drugs prior to in vivo/vitro validation.[17] SBDD uses resolved, 3D protein target structures determined through techniques like X-ray crystallography, Nuclear Magnetic Resonance spectroscopy (NMR), or Cryo-Electron Microscopy (Cryo-EM);[18] the majority of protein structures, in state or complexes relevant to its biological function, of clinical relevance are available in the Protein Data Bank (PDB)[19].[7] The PDB database now (2025) contains more than 235,000 3D macromolecules and complexed structures resolved through experimental methods.[20,21] While structural biology continues to resolve new target structures in the advancement of drug discovery, challenges persist in experimental limitations, biological complexity, and resource investment;[22] the number of known protein sequences (> 200 million) vastly exceeds the number of resolved protein structures in PDB.[23] AI-based sequence-to-structure predictive methods are a consequential advancement in structural biology.[24] AlphaFold is a Deep-learning (DL) model with a central convolutional neural network component trained on PDB structures to predict protein structure. AlphaFold exceeded state-of-the art conventional structure prediction methods in the Critical Assessment of Structure Prediction (CASP) 13 competition[25] and the improved version (AlphaFold2) with a new equivariant transformer architecture prevailed again in the subsequent CASP competition.[26] The newest version (AlphaFold3) is a diffusion-based model with updated complex prediction for most molecular

types present in the PDB.[27] Other DL sequence-based protein structure prediction models include RoseTTAFold[28], OpenFold[29], RGN2[30], ESMfold[31], OmegaFold[32], and EMBER2[33].[34]

1.1.2 Binding Site Elucidation

The characterization of protein-ligand binding sites is crucial in the investigation of orthosteric or allosteric sites that can be targeted by potential disease modulators.[35] Protein targets with known, resolved protein-ligand complex structures may have additional allosteric binding sites that merit exploration.[36] Traditional structure-based binding site prediction methods are categorized into template and pocket-based.[37] Template-based methods identify putative binding sites based on similar protein templates with known structures and functions;[38] these methods have shown the best performance across conventional methods where applicable.[39] Pocket-based methods utilize geometric and physiochemical properties to locate and characterize potential binding sites. Recent structure-based binding site prediction methods have integrated and augmented conventional physics-based methods with AI. Deep learning (DL) methods[40–45] and point cloud algorithms[46,47] have been used to extract geometrical patterns in protein-ligand binding sites and hybrid methods[48–50] that integrate large-scale structural data in the identification of potential druggable target sites.[51]

1.1.3 Chemical Databases

Commercial databases have been developed to facilitate the efficient exploration of potential disease modulators. Early chemical databases contained pre-synthesized molecules with structural scaffolds, functional groups, and chemical properties representative of biomolecules, available in inventories at pharmaceutical vendors.[52] The recent expansion of commercial, chemical databases to billions of molecules has restructured hit identification methods – predefined building blocks and chemical reactions enable on-demand synthesis of small molecules. The exploration of ultra-large, chemical databases represents a new frontier for the identification of novel, potent, and selective drug candidates.[53,54] Reported trends indicate that increasing the scale of virtual screenings improves molecular docking results, and the identification of new chemotypes and higher-affinity binders.[55,56] Several commercial databases have been developed to this end. The newest ZINC version (ZINC-22) contains over 37 billion enumerated, available compounds derived from multiple multi-billion scale make-on-demand

libraries. ZINC has contributed to the identification of numerous, potential inhibitors across different protein families.[57–63] Enamine REAL[64] has over 76 billion make-on-demand compounds and numerous subsets such as the REAL diversity set (drug likeness set). REAL has been explored in the identification of novel GPCR chemotypes[65–67], SARS-CoV-2 Mpro binders[68,69], and more[70–73]. Other commercial chemical databases include GalaXi, CHEMriya, eXplore, Freedom Space, ULTIMATE, and SaVI[74]. These resources have advanced chemical space exploration with a focus on easily synthesizable compounds, but neglect druglike molecules derived from more abstract and complex chemical reactions. Academic research groups have made efforts to develop commercial databases based on innovative chemistry invented in academic laboratories. The Pan-Canadian Chemical Library is an open-source database with more than 148 billion compounds derived from chemical reactions developed in academic laboratories and compatible ZINC reagents, which have minimal intersection with other commercial databases.[75]

Bioactivity databases compile well-annotated biological activity data of chemical compounds, curated from scientific literature, patents, and experimental repositories. Virtual screening (VS) campaigns depend on bioactivity databases to source known, active compounds that can serve as starting points and/or as benchmarks to evaluate VS performances in drug discovery programs.[76] PubChem[77] is a public chemical database at the U.S National Institutes of Health with chemical compounds (>120 million) and biological activities tested in assay experiments[78], and other chemical information (properties, pharmacology, toxicology, drug target, metabolism, safety and handling, etc.[79]). LIT-PCBA is a benchmark dataset designed for VS/ML applications, founded on 149 PubChem dose-response bioassays further processed to remove false positives and assay artifacts, and maintain similar active and inactive molecular property distributions; the dataset contains 15 targets representing protein families of pharmaceutical interest, confirmed actives, and confirmed inactive compounds.[80] ChEMBL[81] is a large, open-access bioactivity database (>2.5 million) at the European Bioinformatics Institute with information curated from assays in several core medicinal chemistry journals.[82] Other bioactivity databases include PDBbind[83], CSAR[84], BindingMOAD[85], and BindingBD[86].

Decoy databases are commonly used in retrospective screens to select an appropriate VS protocol that can separate known, active compounds and inactive or decoy compounds. Inactive compounds are scarcely described and documented[87]; computational decoys are generated to simulate inactive compounds based on property match to known, active compounds while maintaining topological distinction. Decoy databases are designed to minimize bias between active and decoy characteristics (structural complexity and diversity, chemical space coverage, and latent actives in the decoy set).[87] The Database of Useful Decoys: Enhanced (DUD-E)[88] is an optimized version of the Database of Useful Decoys (DUD)[89], which has been widely used in VS benchmarking. DUD-E contains 102 drug targets representing the major pharmaceutical protein families, ChEMBL clustered actives each with 50 property-matched ZINC decoys, and a tool to generate improved, matched decoys for user-supplied actives of alternative drug targets.[88] The latest version (DUDE-Z[90]) is further optimized to address unintended biases identified in former versions: the addition of decoys that represent charge extrema and characteristics of the prospective dataset to mitigate overoptimization for electrostatic interactions and artificial enrichment, respectively.[90] Other decoy databases include MUV[91], DEKOIS 2.0[92], D-COID[93], and LUDe[94]. ML-based decoy generation tools are a more recent method: DeepCoy[95] and TocoDecoy.[96]

VS benchmarking datasets have been developed specific to major protein families in drug discovery. These datasets provide standardized collections for protein targets of pharmaceutical interest to assess and optimize VS protocols.[97] GPCR-Bench is a benchmark dataset with 24 GPCR co-structures, ChEMBL19 actives, and DUD-E-generated decoys[98]. Kinase-Bench is a another focused protein family benchmark dataset.[99]

1.2 Structure-Based Virtual Screening

Structure-based virtual screening (SBVS) is a computational technique in SBDD that uses molecular docking to predict the binding affinity of small molecules to the binding site of a protein target. This enables fast and cost-effective screening of millions of molecules in the identification of virtual hit compounds. AI-augmented molecular docking methods were proposed to screen ultra-large, commercial, chemical libraries – expanding screening to billions of molecules. The rapid advancement of affordable computing resources has made it feasible to

implement these methods at scale.[7] Machine learning (ML)- and deep learning (DL)-accelerated docking models are commonly used to predict binding affinities for ultra-large chemical libraries. These methods, their current limitations, and recent discoveries implementing these methods are discussed herein.

1.2.1 Molecular Docking

Molecular docking is a computational technique used to predict the binding mode and affinity of small molecules (ligands) to the binding site of a protein target – expanding the scope of small molecule exploration at reduced time and cost compared to traditional experimental methods.[100] Molecular docking uses algorithmic sampling to explore protein-ligand binding modes and a scoring function to assess conformations in the predicted binding mode.[101] Popular molecular docking programs include AutoDock[102], AutoDock-Vina[103], Glide[104], GNINA[105], and more.[18]

Rigid docking treats both the protein and ligand as rigid bodies that do not undergo conformational change during the docking process to simplify the modelled binding interaction – this treatment ignores the dynamic nature of protein-ligand binding (induced-fit) as it relies on pre-computed ligand conformations and rigid protein structures.[106] Flexible docking treats the ligand as flexible, and some programs also consider aspects of the protein as flexible during the docking process; flexible protein treatment can involve flexible modeling of key protein sidechains or the use of molecular dynamics simulations to consider multiple protein conformations.

Molecular docking search algorithms are used to predict the most energetically favourable ligand binding mode within the protein binding site. These search algorithms must be efficient and comprehensive in conformational sampling – exhaustive exploration of the entire search space is not feasible. Molecular docking search algorithms are classified into systematic or stochastic methods. Systematic search algorithms (iterative[18], fragmentation[107], and database[108] methods) produce protein-ligand binding conformations through the exploration of slight variations in ligand structural parameters to probe the energy landscape of the conformational space and converge to a minimum energy solution.[109] Stochastic search algorithms (Monte Carlo[101], genetic[109], Tabu[106], and swarm[110] optimization methods) produce protein-ligand binding

conformations through the generation of conformation ensembles to populate a wide range energy landscape and converge to a minimum energy solution; this search method increases the probability of finding a global minimum energy solution at the consequent increase in computational cost.[109]

Molecular docking scoring functions are mathematical functions used to predict the absolute free binding energy of protein-ligand complex formation – entropic and enthalpic terms are incorporated to model the binding interaction. These scoring functions must be efficient and accurate in scoring binding modes, reliable in the identification of active compounds, and discriminative in binding to the native target and off-targets.[111] Classical scoring functions are classified into physics, empirical, and knowledge-based. Physics, force field-based scoring functions use molecular mechanics force-field terms to estimate binding energy as the sum of enthalpic interactions between protein-ligand atom pairs and neglect entropic effects.[112] Empirical-based scoring functions estimate protein-ligand binding affinity as the sum of weighted energetic terms (hydrophobicity, electrostatic, desolvation, hydrogen bond, steric clashes, etc.);[113] the weights are optimized using linear regression analysis on a training set with known binding affinities.[114] Knowledge-based scoring functions are founded on the statistical analysis of protein-ligand complex atom pair interactions from experimental structural information.[115]ML scoring functions incorporate non-linear regression. Common algorithms include random forest, support vector machine, neural networks, etc. These ML models are trained on limited, experimental protein-ligand structural data and therefore are commonly used to rescore docked protein-ligand complexes.[114,116,117]

Protein-ligand binding is dependent on intermolecular interactions, solvent effects, and dynamics.[118] The inaccuracies in scoring functions produce implausible molecules (torsional strain, absent intermolecular interactions, steric clashes, etc.).[119–121] Computational protein-ligand binding can model molecules with strained internal torsions to increase contact with the static protein representation.[119] These irregular torsions are not present in the torsion distribution of molecules in experimental structural databases.[122] Moreover, hydrogen bonding is crucial component of protein-ligand binding in their aqueous environments. New protein-ligand interactions should compensate the disruption of protein- and ligand-water interactions in

binding.[123] The implicit solvation model used in common scoring functions is an oversimplified treatment of solvent interactions.[124] Implicit solvent models treat the solvent as a continuous medium to consider macroscopic properties and speed up simulations – this neglects the representation of specific solvent-solute interactions (hydrogen bonding or the solvent effect on the protein-ligand solvation shell). Computational protein-ligand binding can also be modelled with buried, unsatisfied hydrogen bond donors and acceptors in the binding site – unsatisfied polar heteroatoms in protein binding sites are very rare.[125] The high intrinsic energetic cost of these modelled molecules is not accounted for in their scoring. These artifacts are a product of simplified molecular models, force fields, and search algorithms used in the modelling of protein-ligand binding.

Molecular docking continues to be crucial in the acceleration and cost-effectiveness of the early stages of drug discovery.[126] Though additional post-processing techniques or computational tools are required to select hit candidates for experimental testing.

1.2.2 Artificial Intelligence-Augmented Docking

Recent computational techniques have been introduced to capacitate ultra-large scale VS: molecular docking with high-performance computing cluster parallelization, Graphical processing unit (GPU)-accelerated molecular docking, hierarchical VS, and AI-augmented molecular docking.[127] AI-augmented molecular docking methods (Figure 1.2) were proposed as a new paradigm to screen ultra-large, commercial, chemical libraries. ML/DL-accelerated models are trained on a small subset of docked protein-ligand complexes to predict binding affinities. These algorithms can learn complex patterns and interactions to predict more accurate binding affinities.[100]

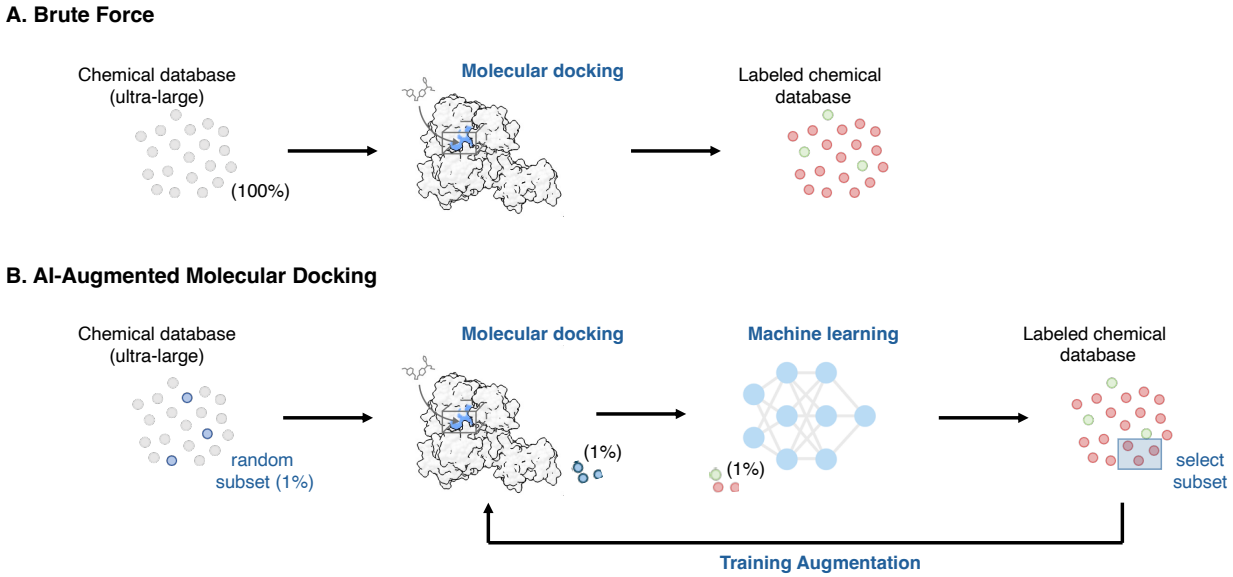


Figure 1.2. Brute force vs. AI-augmented molecular docking. (A) Brute force molecular docking. The entire chemical database is docked. (B) AI-augmented molecular docking docks a small, random subset (1%) of the ultra-large chemical database to train a machine learning model to predict the molecular docking scores for the rest of the chemical database. The top machine learning virtual hits (1-10%) are prioritized for molecular docking to confirm the molecular docking scores. Active learning can be incorporated into the machine learning workflow (training augmentation), wherein the model identifies a small subset (1%) of the most informative, unlabeled data to label in an iterative manner and improve model performance, before selecting the top virtual hits for the final molecular docking round.

Recent advancements in DL and the development of large, commercial chemical databases have increased the adoption of these methods that originated in the early 2000s. Surrogate Docking introduced a novel VS method that requires docking a representative subset of a chemical database to train a quantitative structure–activity relationship (QSAR) model to predict the docking scores of the rest of the database.[128] Progressive Docking presented a hybrid QSAR/docking method that leverages inductive, 3D-sensitive QSAR descriptors to estimate the binding potential of compounds in an iterative process and eliminate predicted non-binders – this resulted in up to a sixfold reduction in docking simulations while maintaining 80-90% hit recovery.[129] The combination of molecular docking and conformal predictions was proposed to further improve iterative screenings.[130,131] Lean-Docking presented a linear support regression model trained on molecular fingerprints to predict docking scores with a significant speed increase.[132] These methods use shallower ML models suitable for smaller scale VS to provide 2 to 4- fold acceleration in contrast to brute force docking.[133] Novel docking accelerators were then introduced based on nascent DL architectures to accommodate commercial databases of

hundreds of millions to billions of small molecules. Deep Docking (DD) is a deep neural network (DNN) trained on a subset of Morgan fingerprints and docking scores from a chemical database to predict the scores of the rest of the database in an iterative manner with an active learning training loop[134] (Figure 1.2) – this model demonstrated up to a 100-fold database reduction and 6000-fold enrichment in VS 12 protein targets with the 1.36 billion molecule ZINC15 database.[135] DD identified novel chemotypes for different targets of clinical relevance: Lin28 inhibitors for the suppression of prostate cancer cell stemness[136], $A_{2A}R$ antagonists for cancer immunotherapy[137], SARS-CoV-2 papain-like protease inhibitors[61], and others[68,138–140]. Several similar methods were proposed afterwards with graph neural network architectures.[141,142] MolPAL is an active learning model-guided method that uses Bayesian optimization to prioritize molecules – a retrospective screen of 100 million compounds with a directed-message passing neural network identified most top scoring molecules with considerable cost reductions.[143] Design space pruning is an extension to the model-guided optimization method which iteratively prunes and select molecules for docking to narrow the design space at each iteration.[144] SPDF is a ML-based surrogate docking model that incorporates virtual hit and dataset size thresholds to prefilter and enrich chemical libraries prior to docking.[145] MEMES is a Bayesian optimization-driven method designed to accelerate high-throughput VS.[146] MO-MEMES is a multi-objective extension of MEMES that identifies virtual hits with optimal predicted binding affinity, LogP, and synthetic accessibility – most high docking scoring molecules violate other drug properties.[147] These DL architectures provide the speed acceleration required to address the limitations of SBVS in ultra-large scale settings, but no current models address the artifacts inherent to molecular docking predictions in their predictions.

1.2.3 Post-Processing Molecular Docking

Molecular docking serves to differentiate putative binders and non-binders.[119] The simplified modelling of protein-ligand interactions introduces artifacts (99%) in the top predictions.[133] Molecular docking artifacts are false positive molecules that cheat scoring functions, but are artificial or unrealistic and would not exhibit binding in experimental assays.[148] These artifacts are further magnified in the top-ranked dockings as the chemical database size increases.[55,56] Top-ranked docking molecules are further processed with other criteria to select leads for more

expensive computational or experimental analysis.[149] Post-processing methods include consensus docking[150] or scoring[151], rescoring[152], filtering[119], and visual inspection[120].[153,154]

Consensus docking or scoring techniques are used to combine results from multiple docking programs or scoring functions to improve the accuracy of predicted binding affinities – improvement can be achieved in the selection of binding poses or scoring. These methods were introduced to negate program- or scoring function-specific problems, but can conflict if the selected programs or functions suffer from the same problems (training-set dependencies and scoring function parameterization[155]).[156]

Rescoring is the re-evaluation of the top-ranked docked molecules with a more expensive computational technique. Common rescoring methods include protein-ligand pose refinement with molecular dynamics and rescoring with molecular mechanics and/or quantum mechanics.[151] These rescoring methods are time-intensive and expensive, which restrict their use to a few thousand molecules after other, initial filtration strategies (pharmacophore, shape similarity, in-silico ADME prediction, etc.).[157]

3D medicinal chemistry properties are often considered in the post-processing of molecular dockings.[158] Pharmacophore, strain, and hydrogen bond filtering computational tools have been developed to automate this process at large scale. Pharmacophore filtering isolates molecules that share the same features, 3D location, and spatial orientation as the native ligand(s) in complex. These features include hydrogen-bond donors, hydrogen-bond acceptors, charged functional groups, hydrophobic regions, aromatics, etc.[159] Pharmacophore modelling programs include Catalyst[160], MOE[161], Phase[162], and LigandScout[163], etc.[164] Moreover, strain is an important and well-studied component in protein-ligand binding[165–167] that is considered in the post-processing of molecular dockings. Numerous computational strain calculation tools are available: statistics-based[168–171], physics-based[172–174] or hybrid[175–177] (physics- and ML-based). Statistics-based tools use experimental torsion angle distributions found in structural databases to calculate strain energies in a high-throughput manner; these statistics-based methods provide the speeds required to handle large-scale VS. The relaxation of biomolecule steric strains has shown both substantial improvement in molecular docking scores and change composition (>50%) in the top 1% molecules.[178] Hydrogen bonding is another important component in protein-ligand

binding that is considered in the post-processing of molecular dockings. Computational protein-ligand interaction tools have been developed to accelerate the inspection of modelled protein-ligand binding. These tools can be used to characterize present protein-ligand interactions (hydrogen bonding, ionic interactions, π- π stacking and π-cation interactions, hydrophobic contacts, etc.) and select molecules that exhibit specific interactions of interest. Available tools include PLIP[179,180], LigGrep[181], Rosetta[125], and IChem[182]. The presence of specific protein-ligand hydrogen bonds has been crucial in the discovery of novel inhibitors.[63,183] Good practices recommend tools to remove virtual hit compounds with strained torsions and unsatisfied hydrogen bond donors and acceptors[184] – these practices have been incorporated separately and cooperatively in numerous VS.

The standard protocol to remove artifact contamination is the visual inspection of top-ranked dockings performed by expert medicinal chemists – a practice that is highly subjective and limited in its application to large chemical libraries.[153] The top criterion considered in the assessment of modelled protein-ligand binding by experts in academia and the pharmaceutical industry includes analogy to crystal structure, specific protein-ligand interactions, ligand strain, unsatisfied polar ligand or protein heteroatoms, etc.[120] Computational tools are recommended to narrow down or prioritize molecules for subsequent visual inspection.[153] Visual inspection is commonly conducted as the final step in candidate prioritization before experimental testing[120], with few thousand compounds inspected.[119,185] Recent efforts to expand visual inspection have been made with ML-automation. Autoparty is a ML-guided tool developed for the visual inspection of molecular dockings.[186] The active-learning model identifies the most useful molecules for the user to grade, then the trained model grades the rest of the docked database – this expands the docked database that can be inspected to remove artifacts and recover lower-ranked molecules. Visual inspection post-docking is integral in the selection of hit candidate molecules for experimentation.

1.3 Multi-Objective Optimization

Drug discovery is a multi-objective optimization (MOO) problem – potential disease modulators must satisfy multiple pharmaceutical objectives (pharmacodynamic, pharmacokinetic, clinical,

etc.).[187] MOO techniques have been reported in QSAR modelling, molecular docking[188], de novo drug design, and chemical database design.[189]

1.3.1 Multi-Objective AI-Augmented Molecular Docking Models

Several MOO AI-augmented molecular docking models with active learning have been developed. The multi-objective extension of MolPAL[143] is designed to accommodate custom objectives and facilitate the optimization of user-selected properties related to drug design. This architecture demonstrated acceleration of docking-based multi-objective VS in retrospective studies: three two-objective studies that identified selective compounds based on docking scores to on- and off-targets and one three-objective study that identified dual inhibitor compounds based on docking scores, with additional selectivity over another target. MO-MEMES[147] is the multi-objective Bayesian optimization extension to Enhanced MolEcular Screening (MEMES[146]). The MO-MEMES architecture uses a Pareto optimization to select molecules that balance binding affinity, octanol partition coefficient (LogP), and synthetic accessibility predictions. Molecular docking is not a standalone method – most high docking scoring compounds violate drug properties' constraints (quantitative estimate of drug likeness (QED) and LogP).[147] This multi-objective architecture identified over 90% of the top compounds with respect to all the required properties considered.

1.3.2 Multi-Task Learning

Single-task learning (STL) is a ML method to train a model to predict the relationship between an instance and particular task (Figure 1.3) – this requires training independent models for each task in a repetitive, parallel process.[190] Multi-task learning (MTL) is a form of MOO[191] in which a single model is trained to learn multiple tasks simultaneously[192] (Figure 1.3) – shared representation across multiple, related tasks can improve model sample efficiency, learning speed, generalization, and performance.[192–194] The obstacle in implementation is negative transfer; destructive interference is the result of increased model performance on one task and decreased performance on other task(s) with different needs.[195] These gradient conflicts can be addressed with gradient modification[196,197], dynamic weighting[198–200], and other methods. There are two common parameter sharing mechanisms to implement MTL: hard parameter sharing (weight sharing to optimize multiple loss functions) and soft parameter sharing (weight

regularization to optimize a joint objective function).[192,201] The shared layers learn a common representation across related tasks. The task-specific layers learn task-dependent information.

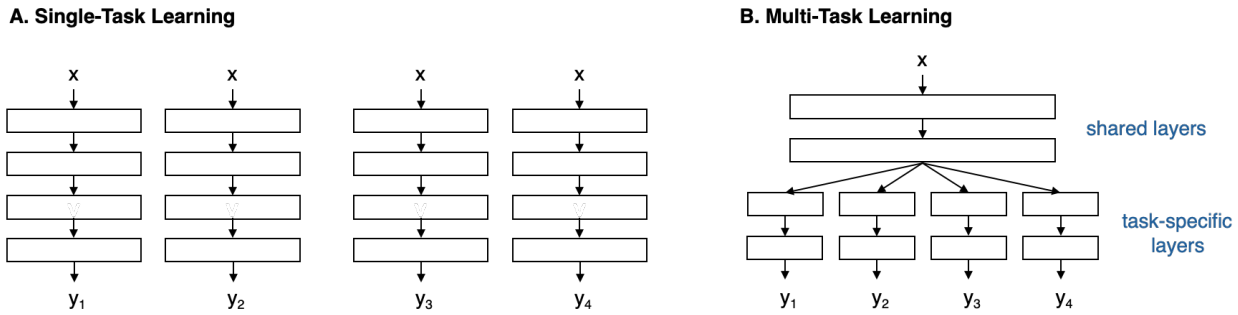


Figure 1.3. Single-task learning (STL) and multi-task learning (MTL) architectures. (A) STL trains separate models per task: 4 separate models and tasks (y) depicted. (B) MTL (hard-parameter sharing shown) trains a single model to predict multiple related tasks: 1 multi-objective model and 4 tasks (y) depicted.

MTL has been leveraged to improve molecular properties[202–206], bioactivities[207–214], ADMET[215–219] predictions, and combinations thereof. No current model aims to improve AI-accelerated molecular docking predictions through the incorporation of select 3D medicinal chemistry properties used in the classification of drug candidates or explores MTL to this end.

1.4 Research Objective

As discussed earlier in this chapter, the interest in exploring ultra-large commercial, chemical databases lies in the identification of novel drug candidates. The exploration of these spaces has been made feasible with the introduction of AI.[7] AI-accelerated molecular docking models trained on a small subset of docked protein-ligand complexes can predict binding affinities for millions or billions of molecules[145] – but these models do not account for the artifacts inherent to molecular docking that are further amplified at large scale.[148] Artifacts are introduced in the modelling of most (99%) molecules in the top docking-ranked predictions.[133] The standard protocol to remove artifact contamination is the visual inspection of top-ranked dockings.[120] This practice is constrained to the top few thousand molecules[119] – but the number of molecules with good docking scores increases with chemical database size.[55] Molecules further down in the docking-ranking can be promising candidates which are not explored due to the time-intensive nature of visual inspection.[186] The top criteria considered in the visual inspection of molecular dockings includes strain and unsatisfied heteroatoms.[120] These 3D medicinal chemistry

properties can be used to filter out molecular docking artifacts.[119,158,184] Several computational tools have been developed to automate the assessment of strain[168–173,175–177] and unsatisfied heteroatoms[179–182] in the protein binding site. There is no MOO model to date that integrates multiple, 3D medicinal chemistry properties not accounted or simplified in molecular docking programs into the prediction of AI-accelerated molecular docking.

The research herein aims to improve the prediction of AI-accelerated molecular docking through the development of a new MOO DL-model that incorporates the prediction of multiple, 3D medicinal chemistry properties used in the classification of drug candidates.

This thesis pursues the following objectives:

- Evaluate and test multiple, 3D medicinal chemistry filters used in the selection of virtual hit candidates in simulated VS campaigns.
- Develop and test improved MOO DL-accelerated molecular docking models in simulated VS campaigns.

These objectives intend to deliver an open-source, medicinal chemistry-informed DL-accelerated molecular docking model for high-performance VS of ultra-large chemical databases that can impact the development of next-generation drug discovery tools and therapeutics.

1.5 References

(1) Bateman, T. J. Drug Discovery. In *Atkinson's Principles of Clinical Pharmacology*; Elsevier, 2022; pp 563–572. https://doi.org/10.1016/B978-0-12-819869-8.00019-7.

(2) Sun, D.; Gao, W.; Hu, H.; Zhou, S. Why 90% of Clinical Drug Development Fails and How to Improve It? *Acta. Pharmaceutica. Sinica. B* **2022**, *12* (7), 3049–3062. https://doi.org/10.1016/j.apsb.2022.02.002.

(3) *Audit: Evaluation of Innovation, Science and Economic Development (ISED) Canada's funding to the Centre for Drug Research and Development (CDRD)*. https://ised-isde.canada.ca/site/audits-evaluations/en/evaluation/evaluation-innovation-science-and-economic-development-ised-canadas-funding-centre-drug-research-and.

(4) Jia, Z.-C.; Yang, X.; Wu, Y.-K.; Li, M.; Das, D.; Chen, M.-X.; Wu, J. The Art of Finding the Right Drug Target: Emerging Methods and Strategies. *Pharmacol. Rev.* **2024**, *76* (5), 896–914. https://doi.org/10.1124/pharmrev.123.001028.

(5) Holenz, J.; Stoy, P. Advances in Lead Generation. *Bioorg. Med. Chem. Lett.* **2019**, *29* (4), 517–524. https://doi.org/10.1016/j.bmcl.2018.12.001.

(6) Ashraf, S. N.; Blackwell, J. H.; Holdgate, G. A.; Lucas, S. C. C.; Solovyeva, A.; Storer, R. I.; Whitehurst, B. C. Hit Me with Your Best Shot: Integrated Hit Discovery for the next Generation of Drug Targets. *Drug Discov. Today* **2024**, *29* (10), 104143. https://doi.org/10.1016/j.drudis.2024.104143.

(7) The Atomwise AIMS Program; Wallach, I.; Bernard, D.; Nguyen, K.; Ho, G.; Morrison, A.; Stecula, A.; Rosnik, A.; O'Sullivan, A. M.; Davtyan, A.; Samudio, B.; Thomas, B.; Worley, B.; Butler, B.; Laggner, C.; Thayer, D.; Moharreri, E.; Friedland, G.; Truong, H.; Van Den Bedem, H.; Ng, H. L.; Stafford, K.; Sarangapani, K.; Giesler, K.; Ngo, L.; Mysinger, M.; Ahmed, M.; Anthis, N. J.; Henriksen, N.; Gniewek, P.; Eckert, S.; De Oliveira, S.; Suterwala, S.; PrasadPrasad, S. V. K.; Shek, S.; Contreras, S.; Hare, S.; Palazzo, T.; O'Brien, T. E.; Van Grack, T.; Williams, T.; Chern, T.-R.; Kenyon, V.; Lee, A. H.; Cann, A. B.; Bergman, B.; Anderson, B. M.; Cox, B. D.; Warrington, J. M.; Sorenson, J. M.; Goldenberg, J. M.; Young, M. A.; DeHaan, N.; Pemberton, R. P.; Schroedl, S.; Abramyan, T. M.; Gupta, T.; Mysore, V.; Presser, A. G.; Ferrando, A. A.; Andricopulo, A. D.; Ghosh, A.; Ayachi, A. G.; Mushtaq, A.; Shaqra, A. M.; Toh, A. K. L.; Smrcka, A. V.; Ciccia, A.; De Oliveira, A. S.; Sverzhinsky, A.; De Sousa, A. M.; Agoulnik, A. I.; Kushnir, A.; Freiberg, A. N.; Statsyuk, A. V.; Gingras, A. R.; Degterev, A.; Tomilov, A.; Vrielink, A.; Garaeva, A. A.; Bryant-Friedrich, A.; Caflisch, A.; Patel, A. K.; Rangarajan, A. V.; Matheeussen, A.; Battistoni, A.; Caporali, A.; Chini, A.; Ilari, A.; Mattevi, A.; Foote, A. T.; Trabocchi, A.; Stahl, A.; Herr, A. B.; Berti, A.; Freywald, A.; Reidenbach, A. G.; Lam, A.; Cuddihy, A. R.; White, A.; Taglialatela, A.; Ojha, A. K.; Cathcart, A. M.; Motyl, A. A. L.; Borowska, A.; D'Antuono, A.; Hirsch, A. K. H.; Porcelli, A. M.; Minakova, A.; Montanaro, A.; Müller, A.; Fiorillo, A.; Virtanen, A.; O'Donoghue, A. J.; Del Rio Flores, A.; Garmendia, A. E.; Pineda-Lucena, A.; Panganiban, A. T.; Samantha, A.; Chatterjee, A. K.; Haas, A. L.; Paparella, A. S.; John, A. L. St.; Prince, A.; ElSheikh, A.; Apfel, A. M.; Colomba, A.; O'Dea, A.; Diallo, B. N.; Ribeiro, B. M. R. M.; Bailey-Elkin, B. A.; Edelman, B. L.; Liou, B.; Perry, B.; Chua, B. S. K.; Kováts, B.; Englinger, B.; Balakrishnan, B.; Gong, B.; Agianian, B.; Pressly, B.; Salas, B. P. M.; Duggan, B. M.; Geisbrecht, B. V.; Dymock, B. W.; Morten, B. C.; Hammock, B. D.; Mota, B. E. F.; Dickinson, B. C.; Fraser, C.; Lempicki, C.; Novina, C. D.;

Torner, C.; Ballatore, C.; Bon, C.; Chapman, C. J.; Partch, C. L.; Chaton, C. T.; Huang, C.; Yang, C.-Y.; Kahler, C. M.; Karan, C.; Keller, C.; Dieck, C. L.; Huimei, C.; Liu, C.; Peltier, C.; Mantri, C. K.; Kemet, C. M.; Müller, C. E.; Weber, C.; Zeina, C. M.; Muli, C. S.; Morisseau, C.; Alkan, C.; Reglero, C.; Loy, C. A.; Wilson, C. M.; Myhr, C.; Arrigoni, C.; Paulino, C.; Santiago, C.; Luo, D.; Tumes, D. J.; Keedy, D. A.; Lawrence, D. A.; Chen, D.; Manor, D.; Trader, D. J.; Hildeman, D. A.; Drewry, D. H.; Dowling, D. J.; Hosfield, D. J.; Smith, D. M.; Moreira, D.; Siderovski, D. P.; Shum, D.; Krist, D. T.; Riches, D. W. H.; Ferraris, D. M.; Anderson, D. H.; Coombe, D. R.; Welsbie, D. S.; Hu, D.; Ortiz, D.; Alramadhani, D.; Zhang, D.; Chaudhuri, D.; Slotboom, D. J.; Ronning, D. R.; Lee, D.; Dirksen, D.; Shoue, D. A.; Zochodne, D. W.; Krishnamurthy, D.; Duncan, D.; Glubb, D. M.; Gelardi, E. L. M.; Hsiao, E. C.; Lynn, E. G.; Silva, E. B.; Aguilera, E.; Lenci, E.; Abraham, E. T.; Lama, E.; Mameli, E.; Leung, E.; Giles, E.; Christensen, E. M.; Mason, E. R.; Petretto, E.; Trakhtenberg, E. F.; Rubin, E. J.; Strauss, E.; Thompson, E. W.; Cione, E.; Lisabeth, E. M.; Fan, E.; Kroon, E. G.; Jo, E.; García-Cuesta, E. M.; Glukhov, E.; Gavathiotis, E.; Yu, F.; Xiang, F.; Leng, F.; Wang, F.; Ingoglia, F.; Van Den Akker, F.; Borriello, F.; Vizeacoumar, F. J.; Luh, F.; Buckner, F. S.; Vizeacoumar, F. S.; Bdira, F. B.; Svensson, F.; Rodriguez, G. M.; Bognár, G.; Lembo, G.; Zhang, G.; Dempsey, G.; Eitzen, G.; Mayer, G.; Greene, G. L.; Garcia, G. A.; Lukacs, G. L.; Prikler, G.; Parico, G. C. G.; Colotti, G.; De Keulenaer, G.; Cortopassi, G.; Roti, G.; Girolimetti, G.; Fiermonte, G.; Gasparre, G.; Leuzzi, G.; Dahal, G.; Michlewski, G.; Conn, G. L.; Stuchbury, G. D.; Bowman, G. R.; Popowicz, G. M.; Veit, G.; De Souza, G. E.; Akk, G.; Caljon, G.; Alvarez, G.; Rucinski, G.; Lee, G.; Cildir, G.; Li, H.; Breton, H. E.; Jafar-Nejad, H.; Zhou, H.; Moore, H. P.; Tilford, H.; Yuan, H.; Shim, H.; Wulff, H.; Hoppe, H.; Chaytow, H.; Tam, H.-K.; Van Remmen, H.; Xu, H.; Debonsi, H. M.; Lieberman, H. B.; Jung, H.; Fan, H.-Y.; Feng, H.; Zhou, H.; Kim, H. J.; Greig, I. R.; Caliandro, I.; Corvo, I.; Arozarena, I.; Mungrue, I. N.; Verhamme, I. M.; Qureshi, I. A.; Lotsaris, I.; Cakir, I.; Perry, J. J. P.; Kwiatkowski, J.; Boorman, J.; Ferreira, J.; Fries, J.; Kratz, J. M.; Miner, J.; Siqueira-Neto, J. L.; Granneman, J. G.; Ng, J.; Shorter, J.; Voss, J. H.; Gebauer, J. M.; Chuah, J.; Mousa, J. J.; Maynes, J. T.; Evans, J. D.; Dickhout, J.; MacKeigan, J. P.; Jossart, J. N.; Zhou, J.; Lin, J.; Xu, J.; Wang, J.; Zhu, J.; Liao, J.; Xu, J.; Zhao, J.; Lin, J.; Lee, J.; Reis, J.; Stetefeld, J.; Bruning, J. B.; Bruning, J. B.; Coles, J. G.; Tanner, J. J.; Pascal, J. M.; So, J.; Pederick, J. L.; Costoya, J. A.; Rayman, J. B.; Maciag, J. J.; Nasburg, J. A.; Gruber, J. J.; Finkelstein, J. M.; Watkins, J.; Rodríguez-Frade, J. M.; Arias, J. A. S.; Lasarte, J. J.; Oyarzabal, J.; Milosavljevic, J.; Cools, J.; Lescar, J.; Bogomolovas, J.; Wang, J.; Kee, J.-M.; Kee, J.-M.; Liao, J.; Sistla, J. C.; Abrahão, J. S.; Sishtla, K.; Francisco, K. R.; Hansen, K. B.; Molyneaux, K. A.; Cunningham, K. A.; Martin, K. R.; Gadar, K.; Ojo, K. K.; Wong, K. S.; Wentworth, K. L.; Lai, K.; Lobb, K. A.; Hopkins, K. M.; Parang, K.; Machaca, K.; Pham, K.; Ghilarducci, K.; Sugamori, K. S.; McManus, K. J.; Musta, K.; Faller, K. M. E.; Nagamori, K.; Mostert, K. J.; Korotkov, K. V.; Liu, K.; Smith, K. S.; Sarosiek, K.; Rohde, K. H.; Kim, K. K.; Lee, K. H.; Pusztai, L.; Lehtiö, L.; Haupt, L. M.; Cowen, L. E.; Byrne, L. J.; Su, L.; Wert-Lamas, L.; Puchades-Carrasco, L.; Chen, L.; Malkas, L. H.; Zhuo, L.; Hedstrom, L.; Hedstrom, L.; Walensky, L. D.; Antonelli, L.; Iommarini, L.; Whitesell, L.; Randall, L. M.; Fathallah, M. D.; Nagai, M. H.; Kilkenny, M. L.; Ben-Johny, M.; Lussier, M. P.; Windisch, M. P.; Lolicato, M.; Lolli, M. L.; Vleminckx, M.; Caroleo, M. C.; Macias, M. J.; Valli, M.; Barghash, M. M.; Mellado, M.; Tye, M. A.; Wilson, M. A.; Hannink, M.; Ashton, M. R.; Cerna, M. V. C. dela; Giorgis, M.; Safo, M. K.; Maurice, M. St.; McDowell, M. A.; Pasquali, M.; Mehedi, M.; Serafim, M. S. M.; Soellner, M. B.; Alteen, M. G.; Champion, M. M.; Skorodinsky, M.; O'Mara, M. L.; Bedi, M.; Rizzi, M.; Levin, M.; Mowat, M.; Jackson, M. R.; Paige, M.; Al-Yozbaki, M.;

Giardini, M. A.; Maksimainen, M. M.; De Luise, M.; Hussain, M. S.; Christodoulides, M.; Stec, N.; Zelinskaya, N.; Van Pelt, N.; Merrill, N. M.; Singh, N.; Kootstra, N. A.; Singh, N.; Gandhi, N. S.; Chan, N.-L.; Trinh, N. M.; Schneider, N. O.; Matovic, N.; Horstmann, N.; Longo, N.; Bharambe, N.; Rouzbeh, N.; Mahmoodi, N.; Gumede, N. J.; Anastasio, N. C.; Khalaf, N. B.; Rabal, O.; Kandror, O.; Escaffre, O.; Silvennoinen, O.; Bishop, O. T.; Iglesias, P.; Sobrado, P.; Chuong, P.; O'Connell, P.; Martin-Malpartida, P.; Mellor, P.; Fish, P. V.; Moreira, P. O. L.; Zhou, P.; Liu, P.; Liu, P.; Wu, P.; Agogo-Mawuli, P.; Jones, P. L.; Ngoi, P.; Toogood, P.; Ip, P.; Von Hundelshausen, P.; Lee, P. H.; Rowswell-Turner, R. B.; Balaña-Fouce, R.; Rocha, R. E. O.; Guido, R. V. C.; Ferreira, R. S.; Agrawal, R. K.; Harijan, R. K.; Ramachandran, R.; Verma, R.; Singh, R. K.; Tiwari, R. K.; Mazitschek, R.; Koppisetti, R. K.; Dame, R. T.; Douville, R. N.; Austin, R. C.; Taylor, R. E.; Moore, R. G.; Ebright, R. H.; Angell, R. M.; Yan, R.; Kejriwal, R.; Batey, R. A.; Blelloch, R.; Vandenberg, R. J.; Hickey, R. J.; Kelm, R. J.; Lake, R. J.; Bradley, R. K.; Blumenthal, R. M.; Solano, R.; Gierse, R. M.; Viola, R. E.; McCarthy, R. R.; Reguera, R. M.; Uribe, R. V.; Do Monte-Neto, R. L.; Gorgoglione, R.; Cullinane, R. T.; Katyal, S.; Hossain, S.; Phadke, S.; Shelburne, S. A.; Geden, S. E.; Johannsen, S.; Wazir, S.; Legare, S.; Landfear, S. M.; Radhakrishnan, S. K.; Ammendola, S.; Dzhumaev, S.; Seo, S.-Y.; Li, S.; Zhou, S.; Chu, S.; Chauhan, S.; Maruta, S.; Ashkar, S. R.; Shyng, S.-L.; Conticello, S. G.; Buroni, S.; Garavaglia, S.; White, S. J.; Zhu, S.; Tsimbalyuk, S.; Chadni, S. H.; Byun, S. Y.; Park, S.; Xu, S. Q.; Banerjee, S.; Zahler, S.; Espinoza, S.; Gustincich, S.; Sainas, S.; Celano, S. L.; Capuzzi, S. J.; Waggoner, S. N.; Poirier, S.; Olson, S. H.; Marx, S. O.; Van Doren, S. R.; Sarilla, S.; Brady-Kalnay, S. M.; Dallman, S.; Azeem, S. M.; Teramoto, T.; Mehlman, T.; Swart, T.; Abaffy, T.; Akopian, T.; Haikarainen, T.; Moreda, T. L.; Ikegami, T.; Teixeira, T. R.; Jayasinghe, T. D.; Gillingwater, T. H.; Kampourakis, T.; Richardson, T. I.; Herdendorf, T. J.; Kotzé, T. J.; O'Meara, T. R.; Corson, T. W.; Hermle, T.; Ogunwa, T. H.; Lan, T.; Su, T.; Banjo, T.; O'Mara, T. A.; Chou, T.; Chou, T.-F.; Baumann, U.; Desai, U. R.; Pai, V. P.; Thai, V. C.; Tandon, V.; Banerji, V.; Robinson, V. L.; Gunasekharan, V.; Namasivayam, V.; Segers, V. F. M.; Maranda, V.; Dolce, V.; Maltarollo, V. G.; Scoffone, V. C.; Woods, V. A.; Ronchi, V. P.; Van Hung Le, V.; Clayton, W. B.; Lowther, W. T.; Houry, W. A.; Li, W.; Tang, W.; Zhang, W.; Van Voorhis, W. C.; Donaldson, W. A.; Hahn, W. C.; Kerr, W. G.; Gerwick, W. H.; Bradshaw, W. J.; Foong, W. E.; Blanchet, X.; Wu, X.; Lu, X.; Qi, X.; Xu, X.; Yu, X.; Qin, X.; Wang, X.; Yuan, X.; Zhang, X.; Zhang, Y. J.; Hu, Y.; Aldhamen, Y. A.; Chen, Y.; Li, Y.; Sun, Y.; Zhu, Y.; Gupta, Y. K.; Pérez-Pertejo, Y.; Li, Y.; Tang, Y.; He, Y.; Tse-Dinh, Y.-C.; Sidorova, Y. A.; Yen, Y.; Li, Y.; Frangos, Z. J.; Chung, Z.; Su, Z.; Wang, Z.; Zhang, Z.; Liu, Z.; Inde, Z.; Artía, Z.; Heifets, A. AI Is a Viable Alternative to High Throughput Screening: A 318-Target Study. *Sci. Rep.* **2024**, *14* (1), 7526. https://doi.org/10.1038/s41598-024-54655-z.

(8) Brown, D. G.; Boström, J. Where Do Recent Small Molecule Clinical Development Candidates Come From? *J. Med. Chem.* **2018**, *61* (21), 9442–9468. https://doi.org/10.1021/acs.jmedchem.8b00675.

(9) Tabana, Y.; Babu, D.; Fahlman, R.; Siraki, A. G.; Barakat, K. Target Identification of Small Molecules: An Overview of the Current Applications in Drug Discovery. *BMC Biotechnol.* **2023**, *23* (1), 44. https://doi.org/10.1186/s12896-023-00815-4.

(10) Hoffman, H. I.; Guo, J. A.; Hwang, W. L. Drug Discovery. In *Translational Radiation Oncology*; Elsevier, 2023; pp 39–43. https://doi.org/10.1016/B978-0-323-88423-5.00065-0.

(11) Hauser, A. S.; Attwood, M. M.; Rask-Andersen, M.; Schiöth, H. B.; Gloriam, D. E. Trends in GPCR Drug Discovery: New Agents, Targets and Indications. *Nat. Rev. Drug. Discov.* **2017**, *16* (12), 829–842. https://doi.org/10.1038/nrd.2017.178.

(12) Chan, W.; Wu, J.; Bell, E.; Zhang, Y. Virtual Screening and Bioactivity Modeling for G Protein-Coupled Receptors. In *GPCRs as Therapeutic Targets*; Gilchrist, A., Ed.; Wiley, 2022; pp 388–423. https://doi.org/10.1002/9781119564782.ch12.

(13) Pun, F. W.; Ozerov, I. V.; Zhavoronkov, A. AI-Powered Therapeutic Target Discovery. *Trends Pharmacol. Sci.* **2023**, *44* (9), 561–572. https://doi.org/10.1016/j.tips.2023.06.010.

(14) Schenone, M.; Dančík, V.; Wagner, B. K.; Clemons, P. A. Target Identification and Mechanism of Action in Chemical Biology and Drug Discovery. *Nat. Chem. Biol.* **2013**, *9* (4), 232–240. https://doi.org/10.1038/nchembio.1199.

(15) Ocana, A.; Pandiella, A.; Privat, C.; Bravo, I.; Luengo-Oroz, M.; Amir, E.; Gyorffy, B. Integrating Artificial Intelligence in Drug Discovery and Early Drug Development: A Transformative Approach. *Biomark. Res.* **2025**, *13* (1), 45. https://doi.org/10.1186/s40364-025-00758-2.

(16) *Improving and Accelerating Therapeutic Development for Nervous System Disorders: Workshop Summary*; National Academies Press (US), 2014.

(17) Wenteler, A.; Cabrera, C. P.; Wei, W.; Neduva, V.; Barnes, M. R. AI Approaches for the Discovery and Validation of Drug Targets. *Camb. Prism. Precis. Med.* **2024**, *2*, e7. https://doi.org/10.1017/pcm.2024.4.

(18) Agu, P. C.; Afiukwa, C. A.; Orji, O. U.; Ezeh, E. M.; Ofoke, I. H.; Ogbu, C. O.; Ugwuja, E. I.; Aja, P. M. Molecular Docking as a Tool for the Discovery of Molecular Targets of Nutraceuticals in Diseases Management. *Sci. Rep.* **2023**, *13* (1), 13398. https://doi.org/10.1038/s41598-023-40160-2.

(19) Berman, H. M. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28* (1), 235–242. https://doi.org/10.1093/nar/28.1.235.

(20) Burley, S. K.; Bhatt, R.; Bhikadiya, C.; Bi, C.; Biester, A.; Biswas, P.; Bittrich, S.; Blaumann, S.; Brown, R.; Chao, H.; Chithari, V. R.; Craig, P. A.; Crichlow, G. V.; Duarte, J. M.; Dutta, S.; Feng, Z.; Flatt, J. W.; Ghosh, S.; Goodsell, D. S.; Green, R. K.; Guranovic, V.; Henry, J.; Hudson, B. P.; Joy, M.; Kaelber, J. T.; Khokhriakov, I.; Lai, J.-S.; Lawson, C. L.; Liang, Y.; Myers-Turnbull, D.; Peisach, E.; Persikova, I.; Piehl, D. W.; Pingale, A.; Rose, Y.; Sagendorf, J.; Sali, A.; Segura, J.; Sekharan, M.; Shao, C.; Smith, J.; Trumbull, M.; Vallat, B.; Voigt, M.; Webb, B.; Whetstone, S.; Wu-Wu, A.; Xing, T.; Young, J. Y.; Zalevsky, A.; Zardecki, C. Updated Resources for Exploring Experimentally-Determined PDB Structures and Computed Structure Models at the RCSB Protein Data Bank. *Nucleic Acids Res.* **2025**, *53* (D1), D564–D574. https://doi.org/10.1093/nar/gkae1091.

(21) Bertoline, L. M. F.; Lima, A. N.; Krieger, J. E.; Teixeira, S. K. Before and after AlphaFold2: An Overview of Protein Structure Prediction. *Front. Bioinform.* **2023**, *3*, 1120370. https://doi.org/10.3389/fbinf.2023.1120370.

(22) Timofeev, V.; Samygina, V. Protein Crystallography: Achievements and Challenges. *Crystals* **2023**, *13* (1), 71. https://doi.org/10.3390/cryst13010071.

(23) Bittrich, S.; Bhikadiya, C.; Bi, C.; Chao, H.; Duarte, J. M.; Dutta, S.; Fayazi, M.; Henry, J.; Khokhriakov, I.; Lowe, R.; Piehl, D. W.; Segura, J.; Vallat, B.; Voigt, M.; Westbrook, J. D.; Burley, S. K.; Rose, Y. RCSB Protein Data Bank: Efficient Searching and Simultaneous Access to One Million Computed Structure Models Alongside the PDB Structures Enabled by Architectural Advances. *J. Mol. Biol.* **2023**, *435* (14), 167994. https://doi.org/10.1016/j.jmb.2023.167994.

(24) Senior, A. W.; Evans, R.; Jumper, J.; Kirkpatrick, J.; Sifre, L.; Green, T.; Qin, C.; Žídek, A.; Nelson, A. W. R.; Bridgland, A.; Penedones, H.; Petersen, S.; Simonyan, K.; Crossan, S.; Kohli, P.; Jones, D. T.; Silver, D.; Kavukcuoglu, K.; Hassabis, D. Improved Protein Structure Prediction Using Potentials from Deep Learning. *Nature* **2020**, *577* (7792), 706–710. https://doi.org/10.1038/s41586-019-1923-7.

(25) Senior, A. W.; Evans, R.; Jumper, J.; Kirkpatrick, J.; Sifre, L.; Green, T.; Qin, C.; Žídek, A.; Nelson, A. W. R.; Bridgland, A.; Penedones, H.; Petersen, S.; Simonyan, K.; Crossan, S.; Kohli, P.; Jones, D. T.; Silver, D.; Kavukcuoglu, K.; Hassabis, D. Protein Structure Prediction Using Multiple Deep Neural Networks in the 13th Critical Assessment of Protein Structure Prediction (CASP13). *Proteins: Struct., Funct., and Bioinf.* **2019**, *87* (12), 1141–1148. https://doi.org/10.1002/prot.25834.

(26) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2.

(27) Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Ronneberger, O.; Willmore, L.; Ballard, A. J.; Bambrick, J.; Bodenstein, S. W.; Evans, D. A.; Hung, C.-C.; O'Neill, M.; Reiman, D.; Tunyasuvunakool, K.; Wu, Z.; Žemgulytė, A.; Arvaniti, E.; Beattie, C.; Bertolli, O.; Bridgland, A.; Cherepanov, A.; Congreve, M.; Cowen-Rivers, A. I.; Cowie, A.; Figurnov, M.; Fuchs, F. B.; Gladman, H.; Jain, R.; Khan, Y. A.; Low, C. M. R.; Perlin, K.; Potapenko, A.; Savy, P.; Singh, S.; Stecula, A.; Thillaisundaram, A.; Tong, C.; Yakneen, S.; Zhong, E. D.; Zielinski, M.; Žídek, A.; Bapst, V.; Kohli, P.; Jaderberg, M.; Hassabis, D.; Jumper, J. M. Accurate Structure Prediction of Biomolecular Interactions with AlphaFold 3. *Nature* **2024**, *630* (8016), 493–500. https://doi.org/10.1038/s41586-024-07487-w.

(28) Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G. R.; Wang, J.; Cong, Q.; Kinch, L. N.; Schaeffer, R. D.; Millán, C.; Park, H.; Adams, C.; Glassman, C. R.; DeGiovanni, A.; Pereira, J. H.; Rodrigues, A. V.; Van Dijk, A. A.; Ebrecht, A. C.; Opperman, D. J.; Sagmeister, T.; Buhlheller, C.; Pavkov-Keller, T.; Rathinaswamy, M. K.; Dalwadi, U.; Yip, C. K.; Burke, J. E.; Garcia, K. C.; Grishin, N. V.; Adams, P. D.; Read, R. J.; Baker, D. Accurate Prediction of Protein Structures and Interactions Using a Three-Track Neural Network. *Science* **2021**, *373* (6557), 871–876. https://doi.org/10.1126/science.abj8754.

(29) Ahdritz, G.; Bouatta, N.; Floristean, C.; Kadyan, S.; Xia, Q.; Gerecke, W.; O'Donnell, T. J.; Berenberg, D.; Fisk, I.; Zanichelli, N.; Zhang, B.; Nowaczynski, A.; Wang, B.; Stepniewska-Dziubinska, M. M.; Zhang, S.; Ojewole, A.; Guney, M. E.; Biderman, S.; Watkins, A. M.; Ra, S.; Lorenzo, P. R.; Nivon, L.; Weitzner, B.; Ban, Y.-E. A.; Chen, S.; Zhang, M.; Li, C.; Song, S.

L.; He, Y.; Sorger, P. K.; Mostaque, E.; Zhang, Z.; Bonneau, R.; AlQuraishi, M. OpenFold: Retraining AlphaFold2 Yields New Insights into Its Learning Mechanisms and Capacity for Generalization. *Nat. Methods.* **2024**, *21* (8), 1514–1524. https://doi.org/10.1038/s41592-024-02272-z.

(30) Chowdhury, R.; Bouatta, N.; Biswas, S.; Floristean, C.; Kharkar, A.; Roy, K.; Rochereau, C.; Ahdritz, G.; Zhang, J.; Church, G. M.; Sorger, P. K.; AlQuraishi, M. Single-Sequence Protein Structure Prediction Using a Language Model and Deep Learning. *Nat. Biotechnol.* **2022**, *40* (11), 1617–1623. https://doi.org/10.1038/s41587-022-01432-w.

(31) Lin, Z.; Akin, H.; Rao, R.; Hie, B.; Zhu, Z.; Lu, W.; Smetanin, N.; Verkuil, R.; Kabeli, O.; Shmueli, Y.; Dos Santos Costa, A.; Fazel-Zarandi, M.; Sercu, T.; Candido, S.; Rives, A. Evolutionary-Scale Prediction of Atomic-Level Protein Structure with a Language Model. *Science* **2023**, *379* (6637), 1123–1130. https://doi.org/10.1126/science.ade2574.

(32) Wu, R.; Ding, F.; Wang, R.; Shen, R.; Zhang, X.; Luo, S.; Su, C.; Wu, Z.; Xie, Q.; Berger, B.; Ma, J.; Peng, J. High-Resolution *de Novo* Structure Prediction from Primary Sequence. *bioRxiv*. July 22, 2022. https://doi.org/10.1101/2022.07.21.500999.

(33) Weissenow, K.; Heinzinger, M.; Rost, B. Protein Language-Model Embeddings for Fast, Accurate, and Alignment-Free Protein Structure Prediction. *Structure* **2022**, *30* (8), 1169-1177.e4. https://doi.org/10.1016/j.str.2022.05.001.

(34) Jänes, J.; Beltrao, P. Deep Learning for Protein Structure Prediction and Design—Progress and Applications. *Mol. Syst. Biol.* **2024**, *20* (3), 162–169. https://doi.org/10.1038/s44320-024-00016-x.

(35) Özçelik, R.; Van Tilborg, D.; Jiménez-Luna, J.; Grisoni, F. Structure-Based Drug Discovery with Deep Learning**. *ChemBioChem* **2023**, *24* (13), e202200776. https://doi.org/10.1002/cbic.202200776.

(36) Xia, Y.; Pan, X.; Shen, H.-B. A Comprehensive Survey on Protein-Ligand Binding Site Prediction. *Curr. Opin. in Struct. Biol.* **2024**, *86*, 102793. https://doi.org/10.1016/j.sbi.2024.102793.

(37) Liao, J.; Wang, Q.; Wu, F.; Huang, Z. In Silico Methods for Identification of Potential Active Sites of Therapeutic Targets. *Molecules* **2022**, *27* (20), 7103. https://doi.org/10.3390/molecules27207103.

(38) Petrey, D.; Chen, T. S.; Deng, L.; Garzon, J. I.; Hwang, H.; Lasso, G.; Lee, H.; Silkov, A.; Honig, B. Template-Based Prediction of Protein Function. *Curr. Opin. Struct. Biol.* **2015**, *32*, 33–38. https://doi.org/10.1016/j.sbi.2015.01.007.

(39) Macari, G.; Toti, D.; Polticelli, F. Computational Methods and Tools for Binding Site Recognition between Proteins and Small Molecules: From Classical Geometrical Approaches to Modern Machine Learning Strategies. *J. Comput. Aided Mol. Des.* **2019**, *33* (10), 887–903. https://doi.org/10.1007/s10822-019-00235-7.

(40) Meller, A.; Ward, M.; Borowsky, J.; Kshirsagar, M.; Lotthammer, J. M.; Oviedo, F.; Ferres, J. L.; Bowman, G. R. Predicting Locations of Cryptic Pockets from Single Protein Structures Using the PocketMiner Graph Neural Network. *Nat. Commun.* **2023**, *14* (1), 1177. https://doi.org/10.1038/s41467-023-36699-3.

(41) Jiménez, J.; Doerr, S.; Martínez-Rosell, G.; Rose, A. S.; De Fabritiis, G. DeepSite: Protein-Binding Site Predictor Using 3D-Convolutional Neural Networks. *Bioinformatics* **2017**, *33* (19), 3036–3042. https://doi.org/10.1093/bioinformatics/btx350.

(42) Kozlovskii, I.; Popov, P. Spatiotemporal Identification of Druggable Binding Sites Using Deep Learning. *Commun. Biol.* **2020**, *3* (1), 618. https://doi.org/10.1038/s42003-020-01350-0.

(43) Jeevan, K.; Palistha, S.; Tayara, H.; Chong, K. T. PUResNetV2.0: A Deep Learning Model Leveraging Sparse Representation for Improved Ligand Binding Site Prediction. *J. Cheminform.* **2024**, *16* (1), 66. https://doi.org/10.1186/s13321-024-00865-6.

(44) Evteev, S. A.; Ereshchenko, A. V.; Ivanenkov, Y. A. SiteRadar: Utilizing Graph Machine Learning for Precise Mapping of Protein–Ligand-Binding Sites. *J. Chem. Inf. Model.* **2023**, *63* (4), 1124–1132. https://doi.org/10.1021/acs.jcim.2c01413.

(45) Xia, Y.; Pan, X.; Shen, H.-B. LigBind: Identifying Binding Residues for Over 1000 Ligands with Relation-Aware Graph Neural Networks. *J. Mol. Biol.* **2023**, *435* (13), 168091. https://doi.org/10.1016/j.jmb.2023.168091.

(46) Yan, X.; Lu, Y.; Li, Z.; Wei, Q.; Gao, X.; Wang, S.; Wu, S.; Cui, S. PointSite: A Point Cloud Segmentation Tool for Identification of Protein Ligand Binding Atoms. *J. Chem. Inf. Model.* **2022**, *62* (11), 2835–2845. https://doi.org/10.1021/acs.jcim.1c01512.

(47) Krapp, L. F.; Abriata, L. A.; Cortés Rodriguez, F.; Dal Peraro, M. PeSTo: Parameter-Free Geometric Deep Learning for Accurate Prediction of Protein Binding Interfaces. *Nat. Commun.* **2023**, *14* (1), 2175. https://doi.org/10.1038/s41467-023-37701-8.

(48) Carbery, A.; Buttenschoen, M.; Skyner, R.; Von Delft, F.; Deane, C. M. Learnt Representations of Proteins Can Be Used for Accurate Prediction of Small Molecule Binding Sites on Experimentally Determined and Predicted Protein Structures. *J. Cheminform.* **2024**, *16* (1), 32. https://doi.org/10.1186/s13321-024-00821-4.

(49) Xia, Y.; Xia, C.; Pan, X.; Shen, H. BINDWEB : A Web Server for Ligand Binding Residue and Pocket Prediction from Protein Structures. *Protein Sci.* **2022**, *31* (12), e4462. https://doi.org/10.1002/pro.4462.

(50) Yuan, Q.; Tian, C.; Yang, Y. Genome-Scale Annotation of Protein Binding Sites via Language Model and Geometric Deep Learning. *eLife.* March 26, 2024. https://doi.org/10.7554/eLife.93695.2.

(51) Sim, J.; Kim, D.; Kim, B.; Choi, J.; Lee, J. Recent Advances in AI-Driven Protein-Ligand Interaction Predictions. *Curr. Opin. Struct. Biol.* **2025**, *92*, 103020. https://doi.org/10.1016/j.sbi.2025.103020.

(52) Suzuki, S. Chemical Libraries: Unlocking the Power of Molecules. *J. Med. Org. Chem* **2023**, *6* (3), 46–49. https://doi.org/10.37532/jmoc.2023.6(3).46-49.

(53) Roggia, M.; Natale, B.; Amendola, G.; Di Maro, S.; Cosconati, S. Streamlining Large Chemical Library Docking with Artificial Intelligence: The PyRMD2Dock Approach. *J. Chem. Inf. Model.* **2024**, *64* (7), 2143–2149. https://doi.org/10.1021/acs.jcim.3c00647.

(54) Sadybekov, A. V.; Katritch, V. Computational Approaches Streamlining Drug Discovery. *Nature* **2023**, *616* (7958), 673–685. https://doi.org/10.1038/s41586-023-05905-z.

(55) Lyu, J.; Irwin, J. J.; Shoichet, B. K. Modeling the Expansion of Virtual Screening Libraries. *Nat. Chem. Biol.* **2023**, *19* (6), 712–718. https://doi.org/10.1038/s41589-022-01234-w.

(56) Liu, F.; Mailhot, O.; Glenn, I. S.; Vigneron, S. F.; Bassim, V.; Xu, X.; Fonseca-Valencia, K.; Smith, M. S.; Radchenko, D. S.; Fraser, J. S.; Moroz, Y. S.; Irwin, J. J.; Shoichet, B. K. The Impact of Library Size and Scale of Testing on Virtual Screening. *Nat. Chem. Biol*. **2025**, *21*, 1039-1045. https://doi.org/10.1038/s41589-024-01797-w.

(57) Priyankha, S.; Polk, S.; Thilagavathi, R.; Prakash, M.; Kathiravan, M. K.; Lakshmana Prabhu, S.; Brunozi De Oliveira, I.; Selvam, C. The Virtual Screening of Compounds from the ZINC Database against PARP-1 in Triple-Negative Breast Cancer. *ChemistrySelect* **2024**, *9* (10), e202304729. https://doi.org/10.1002/slct.202304729.

(58) Poonia, P.; Sharma, M.; Jha, P.; Chopra, M. Pharmacophore-Based Virtual Screening of ZINC Database, Molecular Modeling and Designing New Derivatives as Potential HDAC6 Inhibitors. *Mol. Divers.* **2023**, *27* (5), 2053–2071. https://doi.org/10.1007/s11030-022-10540-3.

(59) Arba, M.; Nur-Hidayat, A.; Surantaadmaja, S. I.; Tjahjono, D. H. Pharmacophore-Based Virtual Screening for Identifying B5 Subunit Inhibitor of 20S Proteasome. *Comput. Biol. Chem.* **2018**, *77*, 64–71. https://doi.org/10.1016/j.compbiolchem.2018.08.009.

(60) Razzaghi-Asl, N.; Mirzayi, S.; Mahnam, K.; Sepehri, S. Identification of COX-2 Inhibitors via Structure-Based Virtual Screening and Molecular Dynamics Simulation. *J. Mol. Graph. Model.* **2018**, *83*, 138–152. https://doi.org/10.1016/j.jmgm.2018.05.010.

(61) Garland, O.; Ton, A.-T.; Moradi, S.; Smith, J. R.; Kovacic, S.; Ng, K.; Pandey, M.; Ban, F.; Lee, J.; Vuckovic, M.; Worrall, L. J.; Young, R. N.; Pantophlet, R.; Strynadka, N. C. J.; Cherkasov, A. Large-Scale Virtual Screening for the Discovery of SARS-CoV-2 Papain-like Protease (PLpro) Non-Covalent Inhibitors. *J. Chem. Inf. Model.* **2023**, *63* (7), 2158–2169. https://doi.org/10.1021/acs.jcim.2c01641.

(62) Luttens, A.; Gullberg, H.; Abdurakhmanov, E.; Vo, D. D.; Akaberi, D.; Talibov, V. O.; Nekhotiaeva, N.; Vangeel, L.; De Jonghe, S.; Jochmans, D.; Krambrich, J.; Tas, A.; Lundgren, B.; Gravenfors, Y.; Craig, A. J.; Atilaw, Y.; Sandström, A.; Moodie, L. W. K.; Lundkvist, Å.; Van Hemert, M. J.; Neyts, J.; Lennerstrand, J.; Kihlberg, J.; Sandberg, K.; Danielson, U. H.; Carlsson, J. Ultralarge Virtual Screening Identifies SARS-CoV-2 Main Protease Inhibitors with Broad-Spectrum Activity against Coronaviruses. *J. Am. Chem. Soc.* **2022**, *144* (7), 2905–2920. https://doi.org/10.1021/jacs.1c08402.

(63) Singh, I.; Li, F.; Fink, E. A.; Chau, I.; Li, A.; Rodriguez-Hernández, A.; Glenn, I.; Zapatero-Belinchón, F. J.; Rodriguez, M. L.; Devkota, K.; Deng, Z.; White, K.; Wan, X.; Tolmachova, N. A.; Moroz, Y. S.; Kaniskan, H. Ü.; Ott, M.; García-Sastre, A.; Jin, J.; Fujimori, D. G.; Irwin, J. J.; Vedadi, M.; Shoichet, B. K. Structure-Based Discovery of Inhibitors of the SARS-CoV-2 Nsp14 N7-Methyltransferase. *J. Med. Chem.* **2023**, *66* (12), 7785–7803. https://doi.org/10.1021/acs.jmedchem.2c02120.

(64) *Enamine REAL database*. Enamine.net. https://enamine.net/compound-collections/real-compounds/real-database.

(65) Alon, A.; Lyu, J.; Braz, J. M.; Tummino, T. A.; Craik, V.; O'Meara, M. J.; Webb, C. M.; Radchenko, D. S.; Moroz, Y. S.; Huang, X.-P.; Liu, Y.; Roth, B. L.; Irwin, J. J.; Basbaum, A. I.; Shoichet, B. K.; Kruse, A. C. Structures of the Σ2 Receptor Enable Docking for Bioactive

Ligand Discovery. *Nature* **2021**, *600* (7890), 759–764. https://doi.org/10.1038/s41586-021-04175-x.

(66) Fink, E. A.; Xu, J.; Hübner, H.; Braz, J. M.; Seemann, P.; Avet, C.; Craik, V.; Weikert, D.; Schmidt, M. F.; Webb, C. M.; Tolmachova, N. A.; Moroz, Y. S.; Huang, X.-P.; Kalyanaraman, C.; Gahbauer, S.; Chen, G.; Liu, Z.; Jacobson, M. P.; Irwin, J. J.; Bouvier, M.; Du, Y.; Shoichet, B. K.; Basbaum, A. I.; Gmeiner, P. Structure-Based Discovery of Nonopioid Analgesics Acting through the α_{2A} -Adrenergic Receptor. *Science* **2022**, *377* (6614), eabn7065. https://doi.org/10.1126/science.abn7065.

(67) Stein, R. M.; Kang, H. J.; McCorvy, J. D.; Glatfelter, G. C.; Jones, A. J.; Che, T.; Slocum, S.; Huang, X.-P.; Savych, O.; Moroz, Y. S.; Stauch, B.; Johansson, L. C.; Cherezov, V.; Kenakin, T.; Irwin, J. J.; Shoichet, B. K.; Roth, B. L.; Dubocovich, M. L. Virtual Discovery of Melatonin Receptor Ligands to Modulate Circadian Rhythms. *Nature* **2020**, *579* (7800), 609–614. https://doi.org/10.1038/s41586-020-2027-0.

(68) Gentile, F.; Fernandez, M.; Ban, F.; Ton, A.-T.; Mslati, H.; Perez, C. F.; Leblanc, E.; Yaacoub, J. C.; Gleave, J.; Stern, A.; Wong, B.; Jean, F.; Strynadka, N.; Cherkasov, A. Automated Discovery of Noncovalent Inhibitors of SARS-CoV-2 Main Protease by Consensus Deep Docking of 40 Billion Small Molecules. *Chem. Sci.* **2021**, *12* (48), 15960–15974. https://doi.org/10.1039/D1SC05579H.

(69) Rossetti, G. G.; Ossorio, M. A.; Rempel, S.; Kratzel, A.; Dionellis, V. S.; Barriot, S.; Tropia, L.; Gorgulla, C.; Arthanari, H.; Thiel, V.; Mohr, P.; Gamboni, R.; Halazonetis, T. D. Non-Covalent SARS-CoV-2 Mpro Inhibitors Developed from in Silico Screen Hits. *Sci. Rep.* **2022**, *12* (1), 2505. https://doi.org/10.1038/s41598-022-06306-4.

(70) Reis, J.; Gorgulla, C.; Massari, M.; Marchese, S.; Valente, S.; Noce, B.; Basile, L.; Törner, R.; Cox, H.; Viennet, T.; Yang, M. H.; Ronan, M. M.; Rees, M. G.; Roth, J. A.; Capasso, L.; Nebbioso, A.; Altucci, L.; Mai, A.; Arthanari, H.; Mattevi, A. Targeting ROS Production through Inhibition of NADPH Oxidases. *Nat. Chem. Biol.* **2023**, *19* (12), 1540–1550. https://doi.org/10.1038/s41589-023-01457-5.

(71) Tuerkova, A.; Bongers, B.; Norinder, U.; Ungvári, O.; Szekely, V.; Tarnovskiy, A.; Szakács, G.; Özvegy-Laczka, C.; Van Westen, G. J. P.; Zdrazil, B. Identifying Novel Inhibitors for Hepatic Organic Anion Transporting Polypeptides by Machine-Learning Based Virtual Screening. *J. Chem. Inf. Model.* **2022**, *62* (24), 6323-6335. https://doi.org/10.1021/acs.jcim.1c01460.

(72) Gorgulla, C.; Boeszoermenyi, A.; Wang, Z.-F.; Fischer, P. D.; Coote, P. W.; Padmanabha Das, K. M.; Malets, Y. S.; Radchenko, D. S.; Moroz, Y. S.; Scott, D. A.; Fackeldey, K.; Hoffmann, M.; Iavniuk, I.; Wagner, G.; Arthanari, H. An Open-Source Drug Discovery Platform Enables Ultra-Large Virtual Screens. *Nature* **2020**, *580* (7805), 663–668. https://doi.org/10.1038/s41586-020-2117-z.

(73) Michino, M.; Beautrait, A.; Boyles, N. A.; Nadupalli, A.; Dementiev, A.; Sun, S.; Ginn, J.; Baxt, L.; Suto, R.; Bryk, R.; Jerome, S. V.; Huggins, D. J.; Vendome, J. Shape-Based Virtual Screening of a Billion-Compound Library Identifies Mycobacterial Lipoamide Dehydrogenase Inhibitors. *ACS Bio. Med. Chem. Au* **2023**, *3* (6), 507–515. https://doi.org/10.1021/acsbiomedchemau.3c00046.

(74) Patel, H.; Ihlenfeldt, W.-D.; Judson, P. N.; Moroz, Y. S.; Pevzner, Y.; Peach, M. L.; Delannée, V.; Tarasova, N. I.; Nicklaus, M. C. SAVI, in Silico Generation of Billions of Easily Synthesizable Compounds through Expert-System Type Rules. *Sci. Data.* **2020**, *7* (1), 384. https://doi.org/10.1038/s41597-020-00727-4.

(75) Bedart, C.; Shimokura, G.; West, F. G.; Wood, T. E.; Batey, R. A.; Irwin, J. J.; Schapira, M. The Pan-Canadian Chemical Library: A Mechanism to Open Academic Chemistry to High-Throughput Virtual Screening. *Sci. Data.* **2024**, *11* (1), 597. https://doi.org/10.1038/s41597-024-03443-5.

(76) Tiikkainen, P.; Franke, L. Analysis of Commercial and Public Bioactivity Databases. *J. Chem. Inf. Model.* **2012**, *52* (2), 319–326. https://doi.org/10.1021/ci2003126.

(77) Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B. A.; Thiessen, P. A.; Yu, B.; Zaslavsky, L.; Zhang, J.; Bolton, E. E. PubChem 2025 Update. *Nucleic Acids Res.* **2025**, *53* (D1), D1516–D1525. https://doi.org/10.1093/nar/gkae1059.

(78) Kim, S.; Thiessen, P. A.; Bolton, E. E.; Chen, J.; Fu, G.; Gindulyte, A.; Han, L.; He, J.; He, S.; Shoemaker, B. A.; Wang, J.; Yu, B.; Zhang, J.; Bryant, S. H. PubChem Substance and Compound Databases. *Nucleic Acids Res.* **2016**, *44* (D1), D1202–D1213. https://doi.org/10.1093/nar/gkv951.

(79) Kim, S. Public Chemical Databases. In *Encyclopedia of Bioinformatics and Computational Biology*; Elsevier, 2019; pp 628–639. https://doi.org/10.1016/B978-0-12-809633-8.20192-1.

(80) Tran-Nguyen, V.-K.; Jacquemard, C.; Rognan, D. LIT-PCBA: An Unbiased Data Set for Machine Learning and Virtual Screening. *J. Chem. Inf. Model.* **2020**, *60* (9), 4263–4273. https://doi.org/10.1021/acs.jcim.0c00155.

(81) Zdrazil, B. Fifteen Years of ChEMBL and Its Role in Cheminformatics and Drug Discovery. *J. Cheminform.* **2025**, *17* (1), 32. https://doi.org/10.1186/s13321-025-00963-z.

(82) Mendez, D.; Gaulton, A.; Bento, A. P.; Chambers, J.; De Veij, M.; Félix, E.; Magariños, M. P.; Mosquera, J. F.; Mutowo, P.; Nowotka, M.; Gordillo-Marañón, M.; Hunter, F.; Junco, L.; Mugumbate, G.; Rodriguez-Lopez, M.; Atkinson, F.; Bosc, N.; Radoux, C. J.; Segura-Cabrera, A.; Hersey, A.; Leach, A. R. ChEMBL: Towards Direct Deposition of Bioassay Data. *Nucleic Acids Res.* **2019**, *47* (D1), D930–D940. https://doi.org/10.1093/nar/gky1075.

(83) Liu, Z.; Li, Y.; Han, L.; Li, J.; Liu, J.; Zhao, Z.; Nie, W.; Liu, Y.; Wang, R. PDB-Wide Collection of Binding Data: Current Status of the PDBbind Database. *Bioinformatics* **2015**, *31* (3), 405–412. https://doi.org/10.1093/bioinformatics/btu626.

(84) Dunbar, J. B.; Smith, R. D.; Damm-Ganamet, K. L.; Ahmed, A.; Esposito, E. X.; Delproposto, J.; Chinnaswamy, K.; Kang, Y.-N.; Kubish, G.; Gestwicki, J. E.; Stuckey, J. A.; Carlson, H. A. CSAR Data Set Release 2012: Ligands, Affinities, Complexes, and Docking Decoys. *J. Chem. Inf. Model.* **2013**, *53* (8), 1842–1852. https://doi.org/10.1021/ci4000486.

(85) Wagle, S.; Smith, R. D.; Dominic, A. J.; DasGupta, D.; Tripathi, S. K.; Carlson, H. A. Sunsetting Binding MOAD with Its Last Data Update and the Addition of 3D-Ligand Polypharmacology Tools. *Sci. Rep.* **2023**, *13* (1), 3008. https://doi.org/10.1038/s41598-023-29996-w.

(86) Liu, T.; Hwang, L.; Burley, S. K.; Nitsche, C. I.; Southan, C.; Walters, W. P.; Gilson, M. K. BindingDB in 2024: A FAIR Knowledgebase of Protein-Small Molecule Binding Data. *Nucleic Acids Res.* **2025**, *53* (D1), D1633–D1644. https://doi.org/10.1093/nar/gkae1075.

(87) Torres, P. H. M.; Sodero, A. C. R.; Jofily, P.; Silva-Jr, F. P. Key Topics in Molecular Docking for Drug Design. *IJMS* **2019**, *20* (18), 4574. https://doi.org/10.3390/ijms20184574.

(88) Mysinger, M. M.; Carchia, M.; Irwin, John. J.; Shoichet, B. K. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *J. Med. Chem.* **2012**, *55* (14), 6582–6594. https://doi.org/10.1021/jm300687e.

(89) Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking Sets for Molecular Docking. *J. Med. Chem.* **2006**, *49* (23), 6789–6801. https://doi.org/10.1021/jm0608356.

(90) Stein, R. M.; Yang, Y.; Balius, T. E.; O'Meara, M. J.; Lyu, J.; Young, J.; Tang, K.; Shoichet, B. K.; Irwin, J. J. Property-Unmatched Decoys in Docking Benchmarks. *J. Chem. Inf. Model.* **2021**, *61* (2), 699–714. https://doi.org/10.1021/acs.jcim.0c00598.

(91) Rohrer, S. G.; Baumann, K. Maximum Unbiased Validation (MUV) Data Sets for Virtual Screening Based on PubChem Bioactivity Data. *J. Chem. Inf. Model.* **2009**, *49* (2), 169–184. https://doi.org/10.1021/ci8002649.

(92) Bauer, M. R.; Ibrahim, T. M.; Vogel, S. M.; Boeckler, F. M. Evaluation and Optimization of Virtual Screening Workflows with DEKOIS 2.0 – A Public Library of Challenging Docking Benchmark Sets. *J. Chem. Inf. Model.* **2013**, *53* (6), 1447–1462. https://doi.org/10.1021/ci400115b.

(93) Adeshina, Y. O.; Deeds, E. J.; Karanicolas, J. Machine Learning Classification Can Reduce False Positives in Structure-Based Virtual Screening. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117* (31), 18477–18488. https://doi.org/10.1073/pnas.2000585117.

(94) Alberca, L. N.; Prada Gori, D. N.; Fallico, M. J.; Fassio, A. V.; Talevi, A.; Bellera, C. L. LIDEB's Useful Decoys (LUDe): A Freely Available Decoy-Generation Tool. Benchmarking and Scope. *Artif. Intell. Life Sci.* **2025**, *7*, 100129. https://doi.org/10.1016/j.ailsci.2025.100129.

(95) Imrie, F.; Bradley, A. R.; Deane, C. M. Generating Property-Matched Decoy Molecules Using Deep Learning. *Bioinformatics* **2021**, *37* (15), 2134–2141. https://doi.org/10.1093/bioinformatics/btab080.

(96) Zhang, X.; Shen, C.; Liao, B.; Jiang, D.; Wang, J.; Wu, Z.; Du, H.; Wang, T.; Huo, W.; Xu, L.; Cao, D.; Hsieh, C.-Y.; Hou, T. TocoDecoy: A New Approach to Design Unbiased Datasets for Training and Benchmarking Machine-Learning Scoring Functions. *J. Med. Chem.* **2022**, *65* (11), 7918–7932. https://doi.org/10.1021/acs.jmedchem.2c00460.

(97) Lagarde, N.; Zagury, J.-F.; Montes, M. Benchmarking Data Sets for the Evaluation of Virtual Ligand Screening Methods: Review and Perspectives. *J. Chem. Inf. Model.* **2015**, *55* (7), 1297–1307. https://doi.org/10.1021/acs.jcim.5b00090.

(98) Weiss, D. R.; Bortolato, A.; Tehan, B.; Mason, J. S. GPCR-Bench: A Benchmarking Set and Practitioners' Guide for G Protein-Coupled Receptor Docking. *J. Chem. Inf. Model.* **2016**, *56* (4), 642–651. https://doi.org/10.1021/acs.jcim.5b00660.

(99) Wei, T.-H.; Zhou, S.-S.; Jing, X.-L.; Liu, J.-C.; Sun, M.; Zhao, Z.-H.; Li, Q.-Q.; Wang, Z.-X.; Yang, J.; Zhou, Y.; Wang, X.; Ling, C.-X.; Ding, N.; Xue, X.; Yu, Y.-C.; Wang, X.-L.;

Yin, X.-Y.; Sun, S.-L.; Cao, P.; Li, N.-G.; Shi, Z.-H. Kinase-Bench: Comprehensive Benchmarking Tools and Guidance for Achieving Selectivity in Kinase Drug Discovery. *J. Chem. Inf. Model.* **2024**, *64* (24), 9528–9550. https://doi.org/10.1021/acs.jcim.4c01830.

(100) Aguiar, C.; Camps, I. Molecular Docking in Drug Discovery: Techniques, Applications, and Advancements. *Curr. Med. Chem.* **2025**, *32* (28), 5924-5938. https://doi.org/10.2174/0109298673325827240926081845.

(101) Meng, X.-Y.; Zhang, H.-X.; Mezei, M.; Cui, M. Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Curr. Comput. Aided Drug Des.* **2011**, *7* (2), 146–157. https://doi.org/10.2174/157340911795677602.

(102) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30* (16), 2785–2791. https://doi.org/10.1002/jcc.21256.

(103) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455–461. https://doi.org/10.1002/jcc.21334.

(104) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47* (7), 1739–1749. https://doi.org/10.1021/jm0306430.

(105) McNutt, A. T.; Francoeur, P.; Aggarwal, R.; Masuda, T.; Meli, R.; Ragoza, M.; Sunseri, J.; Koes, D. R. GNINA 1.0: Molecular Docking with Deep Learning. *J. Cheminform.* **2021**, *13* (1), 43. https://doi.org/10.1186/s13321-021-00522-2.

(106) Huang, S.-Y.; Zou, X. Advances and Challenges in Protein-Ligand Docking. *IJMS* **2010**, *11* (8), 3016–3034. https://doi.org/10.3390/ijms11083016.

(107) Nivatya, H. K.; Singh, A.; Kumar, N.; Sonam; Sharma, L.; Singh, V.; Mishra, R.; Gaur, N.; Mishra, A. K. Assessing Molecular Docking Tools: Understanding Drug Discovery and Design. *Futur. J. Pharm. Sci.* **2025**, *11* (1), 111. https://doi.org/10.1186/s43094-025-00862-y.

(108) Crampon, K.; Giorkallos, A.; Deldossi, M.; Baud, S.; Steffenel, L. A. Machine-Learning Methods for Ligand–Protein Molecular Docking. *Drug Discov. Today* **2022**, *27* (1), 151–164. https://doi.org/10.1016/j.drudis.2021.09.007.

(109) Ferreira, L. G.; Dos Santos, R. N.; Oliva, G.; Andricopulo, A. D. Molecular Docking and Structure-Based Drug Design Strategies. *Molecules* **2015**, *20* (7), 13384–13421. https://doi.org/10.3390/molecules200713384.

(110) Zhou, J.; Yang, Z.; He, Y.; Ji, J.; Lin, Q.; Li, J. A Novel Molecular Docking Program Based on a Multi-Swarm Competitive Algorithm. *Swarm Evol. Comput.* **2023**, *78*, 101292. https://doi.org/10.1016/j.swevo.2023.101292.

(111) Zheng, L.; Meng, J.; Jiang, K.; Lan, H.; Wang, Z.; Lin, M.; Li, W.; Guo, H.; Wei, Y.; Mu, Y. Improving Protein-Ligand Docking and Screening Accuracies by Incorporating a Scoring Function Correction Term. *Brief. Bioinform.* **2022**, *23* (3), bbac051. https://doi.org/10.1093/bib/bbac051.

(112) Meli, R.; Morris, G. M.; Biggin, P. C. Scoring Functions for Protein-Ligand Binding Affinity Prediction Using Structure-Based Deep Learning: A Review. *Front. Bioinform.* **2022**, *2*, 885983. https://doi.org/10.3389/fbinf.2022.885983.

(113) Shirali, A.; Stebliankin, V.; Karki, U.; Shi, J.; Chapagain, P.; Narasimhan, G. A Comprehensive Survey of Scoring Functions for Protein Docking Models. *BMC Bioinform.* **2025**, *26* (1), 25. https://doi.org/10.1186/s12859-024-05991-4.

(114) Li, J.; Fu, A.; Zhang, L. An Overview of Scoring Functions Used for Protein–Ligand Interactions in Molecular Docking. *Interdiscip. Sci. Comput. Life Sci.* **2019**, *11* (2), 320–328. https://doi.org/10.1007/s12539-019-00327-w.

(115) Guedes, I. A.; Pereira, F. S. S.; Dardenne, L. E. Empirical Scoring Functions for Structure-Based Virtual Screening: Applications, Critical Aspects, and Challenges. *Front. Pharmacol.* **2018**, *9*, 1089. https://doi.org/10.3389/fphar.2018.01089.

(116) Pellicani, F.; Dal Ben, D.; Perali, A.; Pilati, S. Machine Learning Scoring Functions for Drug Discovery from Experimental and Computer-Generated Protein-Ligand Structures: Towards Per-Target Scoring Functions. *Molecules* **2023**, *28* (4), 1661. https://doi.org/10.3390/molecules28041661.

(117) Velasquez-López, Y.; Tejera, E.; Perez-Castillo, Y. Can Docking Scoring Functions Guarantee Success in Virtual Screening? In *Annual Reports in Medicinal Chemistry*; Elsevier, 2022; Vol. 59, pp 1–41. https://doi.org/10.1016/bs.armc.2022.08.008.

(118) Pantsar, T.; Poso, A. Binding Affinity via Docking: Fact and Fiction. *Molecules* **2018**, *23* (8), 1899. https://doi.org/10.3390/molecules23081899.

(119) Bender, B. J.; Gahbauer, S.; Luttens, A.; Lyu, J.; Webb, C. M.; Stein, R. M.; Fink, E. A.; Balius, T. E.; Carlsson, J.; Irwin, J. J.; Shoichet, B. K. A Practical Guide to Large-Scale Docking. *Nat. Protoc.* **2021**, *16* (10), 4799–4832. https://doi.org/10.1038/s41596-021-00597-z.

(120) Fischer, A.; Smieško, M.; Sellner, M.; Lill, M. A. Decision Making in Structure-Based Drug Discovery: Visual Inspection of Docking Results. *J. Med. Chem.* **2021**, *64* (5), 2489–2500. https://doi.org/10.1021/acs.jmedchem.0c02227.

(121) Luttens, A.; Cabeza De Vaca, I.; Sparring, L.; Brea, J.; Martínez, A. L.; Kahlous, N. A.; Radchenko, D. S.; Moroz, Y. S.; Loza, M. I.; Norinder, U.; Carlsson, J. Rapid Traversal of Vast Chemical Space Using Machine Learning-Guided Docking Screens. *Nat. Comput. Sci.* **2025**, *5* (4), 301–312. https://doi.org/10.1038/s43588-025-00777-x.

(122) Xu, M.; Shen, C.; Yang, J.; Wang, Q.; Huang, N. Systematic Investigation of Docking Failures in Large-Scale Structure-Based Virtual Screening. *ACS Omega* **2022**, *7* (43), 39417–39428. https://doi.org/10.1021/acsomega.2c05826.

(123) Xie, A.; Zhao, G.; Liang, H.; Gao, T.; Gao, X.; Hou, N.; Wei, F.; Li, J.; Zhao, H.; Xu, X. LeScore: A Scoring Function Incorporating Hydrogen Bonding Penalty for Protein–Ligand Docking. *J. Mol. Model.* **2025**, *31* (4), 106. https://doi.org/10.1007/s00894-025-06328-5.

(124) Zhao, H.; Huang, D. Hydrogen Bonding Penalty upon Ligand Binding. *PLoS ONE* **2011**, *6* (6), e19923. https://doi.org/10.1371/journal.pone.0019923.

(125) Coventry, B.; Baker, D. Protein Sequence Optimization with a Pairwise Decomposable Penalty for Buried Unsatisfied Hydrogen Bonds. *PLoS Comput. Biol.* **2021**, *17* (3), e1008061. https://doi.org/10.1371/journal.pcbi.1008061.

(126) Muhammed, M. T.; Aki-Yalcin, E. Molecular Docking: Principles, Advances, and Its Applications in DrugDiscovery. *LDDD* **2024**, *21* (3), 480–495. https://doi.org/10.2174/1570180819666220922103109.

(127) Zhou, G.; Rusnac, D.-V.; Park, H.; Canzani, D.; Nguyen, H. M.; Stewart, L.; Bush, M. F.; Nguyen, P. T.; Wulff, H.; Yarov-Yarovoy, V.; Zheng, N.; DiMaio, F. An Artificial Intelligence Accelerated Virtual Screening Platform for Drug Discovery. *Nat. Commun.* **2024**, *15* (1), 7761. https://doi.org/10.1038/s41467-024-52061-7.

(128) Yoon, S.; Smellie, A.; Hartsough, D.; Filikov, A. Surrogate Docking: Structure-Based Virtual Screening at High Throughput Speed. *J. Comput. Aided Mol. Des.* **2005**, *19* (7), 483–497. https://doi.org/10.1007/s10822-005-9002-6.

(129) Cherkasov, A.; Ban, F.; Li, Y.; Fallahi, M.; Hammond, G. L. Progressive Docking: A Hybrid QSAR/Docking Approach for Accelerating In Silico High Throughput Screening. *J. Med. Chem.* **2006**, *49* (25), 7466–7478. https://doi.org/10.1021/jm060961+.

(130) Svensson, F.; Norinder, U.; Bender, A. Improving Screening Efficiency through Iterative Screening Using Docking and Conformal Prediction. *J. Chem. Inf. Model.* **2017**, *57* (3), 439–444. https://doi.org/10.1021/acs.jcim.6b00532.

(131) Ahmed, L.; Georgiev, V.; Capuccini, M.; Toor, S.; Schaal, W.; Laure, E.; Spjuth, O. Efficient Iterative Virtual Screening with Apache Spark and Conformal Prediction. *J. Cheminform.* **2018**, *10* (1), 8. https://doi.org/10.1186/s13321-018-0265-z.

(132) Berenger, F.; Kumar, A.; Zhang, K. Y. J.; Yamanishi, Y. Lean-Docking: Exploiting Ligands' Predicted Docking Scores to Accelerate Molecular Docking. *J. Chem. Inf. Model.* **2021**, *61* (5), 2341–2352. https://doi.org/10.1021/acs.jcim.0c01452.

(133) Kuan, J.; Radaeva, M.; Avenido, A.; Cherkasov, A.; Gentile, F. Keeping Pace with the Explosive Growth of Chemical Libraries with Structure-based Virtual Screening. *WIREs Comput. Mol. Sci.* **2023**, *13* (6), e1678. https://doi.org/10.1002/wcms.1678.

(134) Gentile, F.; Agrawal, V.; Hsing, M.; Ton, A.-T.; Ban, F.; Norinder, U.; Gleave, M. E.; Cherkasov, A. Deep Docking: A Deep Learning Platform for Augmentation of Structure Based Drug Discovery. *ACS Cent. Sci.* **2020**, *6* (6), 939–949. https://doi.org/10.1021/acscentsci.0c00229.

(135) Gentile, F.; Yaacoub, J. C.; Gleave, J.; Fernandez, M.; Ton, A.-T.; Ban, F.; Stern, A.; Cherkasov, A. Artificial Intelligence–Enabled Virtual Screening of Ultra-Large Chemical Libraries with Deep Docking. *Nat. Protoc.* **2022**, *17* (3), 672–697. https://doi.org/10.1038/s41596-021-00659-2.

(136) Radaeva, M.; Ho, C.-H.; Xie, N.; Zhang, S.; Lee, J.; Liu, L.; Lallous, N.; Cherkasov, A.; Dong, X. Discovery of Novel Lin28 Inhibitors to Suppress Cancer Cell Stemness. *Cancers* **2022**, *14* (22), 5687. https://doi.org/10.3390/cancers14225687.

(137) Tang, M.; Wen, C.; Lin, J.; Chen, H.; Ran, T. Discovery of Novel A2AR Antagonists through Deep Learning-Based Virtual Screening. *Artif. Intell. Life Sci.* **2023**, *3*, 100058. https://doi.org/10.1016/j.ailsci.2023.100058.

(138) Foo, J.; Gentile, F.; Massah, S.; Morin, H.; Singh, K.; Lee, J.; Smith, J.; Ban, F.; LeBlanc, E.; Young, R.; Strynadka, N.; Lallous, N.; Cherkasov, A. Characterization of Novel Small Molecule Inhibitors of Estrogen Receptor-Activation Function 2 (ER-AF2). *Breast Cancer Res.* **2024**, *26* (1), 168. https://doi.org/10.1186/s13058-024-01926-2.

(139) Vendruscolo, M.; Brezinova, M.; Brotzakis, Z. F.; Horne, R.; Chowdhury, V. R.; Gregory, R.; Gentile, F. Identification of High-Affinity Secondary Nucleation Inhibitors of Aβ42 Aggregation from an Ultra-Large Chemical Library Using Deep Docking. *Res. Sq.* June 10, 2024. https://doi.org/10.21203/rs.3.rs-4512167/v1.

(140) Gutkin, E.; Gusev, F.; Gentile, F.; Ban, F.; Koby, S. B.; Narangoda, C.; Isayev, O.; Cherkasov, A.; Kurnikova, M. G. *In Silico* Screening of LRRK2 WDR Domain Inhibitors Using Deep Docking and Free Energy Simulations. *Chem. Sci.* **2024**, *15* (23), 8800–8812. https://doi.org/10.1039/D3SC06880C.

(141) Jastrzębski, S.; Szymczak, M.; Pocha, A.; Mordalski, S.; Tabor, J.; Bojarski, A. J.; Podlewska, S. Emulating Docking Results Using a Deep Neural Network: A New Perspective for Virtual Screening. *J. Chem. Inf. Model.* **2020**, *60* (9), 4246–4262. https://doi.org/10.1021/acs.jcim.9b01202.

(142) Yang, Y.; Yao, K.; Repasky, M. P.; Leswing, K.; Abel, R.; Shoichet, B. K.; Jerome, S. V. Efficient Exploration of Chemical Space with Docking and Deep Learning. *J. Chem. Theory Comput.* **2021**, *17* (11), 7106–7119. https://doi.org/10.1021/acs.jctc.1c00810.

(143) Graff, D. E.; Shakhnovich, E. I.; Coley, C. W. Accelerating High-Throughput Virtual Screening through Molecular Pool-Based Active Learning. *Chem. Sci.* **2021**, *12* (22), 7866–7881. https://doi.org/10.1039/D0SC06805E.

(144) Graff, D. E.; Aldeghi, M.; Morrone, J. A.; Jordan, K. E.; Pyzer-Knapp, E. O.; Coley, C. W. Self-Focusing Virtual Screening with Active Design Space Pruning. *J. Chem. Inf. Model.* **2022**, *62* (16), 3854–3862. https://doi.org/10.1021/acs.jcim.2c00554.

(145) Clyde, A.; Liu, X.; Brettin, T.; Yoo, H.; Partin, A.; Babuji, Y.; Blaiszik, B.; Mohd-Yusof, J.; Merzky, A.; Turilli, M.; Jha, S.; Ramanathan, A.; Stevens, R. AI-Accelerated Protein-Ligand Docking for SARS-CoV-2 Is 100-Fold Faster with No Significant Change in Detection. *Sci. Rep.* **2023**, *13* (1), 2105. https://doi.org/10.1038/s41598-023-28785-9.

(146) Mehta, S.; Laghuvarapu, S.; Pathak, Y.; Sethi, A.; Alvala, M.; Priyakumar, U. D. MEMES: Machine Learning Framework for Enhanced MolEcular Screening. *Chem. Sci.* **2021**, *12* (35), 11710–11721. https://doi.org/10.1039/D1SC02783B.

(147) Mehta, S.; Goel, M.; Priyakumar, U. D. MO-MEMES: A Method for Accelerating Virtual Screening Using Multi-Objective Bayesian Optimization. *Front. Med.* **2022**, *9*, 916481. https://doi.org/10.3389/fmed.2022.916481.

(148) Wu, Y.; Liu, F.; Glenn, I.; Fonseca-Valencia, K.; Paris, L.; Xiong, Y.; Jerome, S. V.; Brooks, C. L.; Shoichet, B. K. Identifying Artifacts from Large Library Docking. *J. Med. Chem.* **2024**, *67* (18), 16796–16806. https://doi.org/10.1021/acs.jmedchem.4c01632.

(149) Yang, X.; Wang, Y.; Byrne, R.; Schneider, G.; Yang, S. Concepts of Artificial Intelligence for Computer-Assisted Drug Discovery. *Chem. Rev.* **2019**, *119* (18), 10520–10594. https://doi.org/10.1021/acs.chemrev.8b00728.

(150) Blanes-Mira, C.; Fernández-Aguado, P.; De Andrés-López, J.; Fernández-Carvajal, A.; Ferrer-Montiel, A.; Fernández-Ballester, G. Comprehensive Survey of Consensus Docking for High-Throughput Virtual Screening. *Molecules* **2022**, *28* (1), 175. https://doi.org/10.3390/molecules28010175.

(151) Rastelli, G.; Pinzi, L. Refinement and Rescoring of Virtual Screening Results. *Front. Chem.* **2019**, *7*, 498. https://doi.org/10.3389/fchem.2019.00498.

(152) Gkeka, P.; Svensson, F.; Magadán, C. R.; De Groot, M. J.; Jerome, S. V. Computational Hit Finding: An Industry Perspective. *J. Med. Chem.* **2025**, *68* (11), 10507–10519. https://doi.org/10.1021/acs.jmedchem.4c03087.

(153) Podlewska, S.; Bojarski, A. J. Post-Processing of Docking Results: Tools and Strategies. In *Molecular Docking for Computer-Aided Drug Design*; Elsevier, 2021; pp 57–74. https://doi.org/10.1016/B978-0-12-822312-3.00004-7.

(154) Zhu, H.; Zhang, Y.; Li, W.; Huang, N. A Comprehensive Survey of Prospective Structure-Based Virtual Screening for Early Drug Discovery in the Past Fifteen Years. *IJMS* **2022**, *23* (24), 15961. https://doi.org/10.3390/ijms232415961.

(155) Palacio-Rodríguez, K.; Lans, I.; Cavasotto, C. N.; Cossio, P. Exponential Consensus Ranking Improves the Outcome in Docking and Receptor Ensemble Docking. *Sci. Rep.* **2019**, *9* (1), 5142. https://doi.org/10.1038/s41598-019-41594-3.

(156) Varela-Rial, A.; Majewski, M.; De Fabritiis, G. Structure Based Virtual Screening: Fast and Slow. *WIREs Comput. Mol. Sci.* **2022**, *12* (2), e1544. https://doi.org/10.1002/wcms.1544.

(157) Tuccinardi, T. What Is the Current Value of MM/PBSA and MM/GBSA Methods in Drug Discovery? *Expert Opin. Drug Discov.* **2021**, *16* (11), 1233–1237. https://doi.org/10.1080/17460441.2021.1942836.

(158) Mahgoub, R. E.; Atatreh, N.; Ghattas, M. A. Using Filters in Virtual Screening: A Comprehensive Guide to Minimize Errors and Maximize Efficiency. In *Annual Reports in Medicinal Chemistry*; Elsevier, 2022; Vol. 59, pp 99–136. https://doi.org/10.1016/bs.armc.2022.09.002.

(159) Daisy, P.; Singh, S. K.; Vijayalakshmi, P.; Selvaraj, C.; Rajalakshmi, M.; Suveena, S. A Database for the Predicted Pharmacophoric Features of Medicinal Compounds. *Bioinformation* **2011**, *6* (4), 167–168. https://doi.org/10.6026/97320630006167.

(160) CATALYST, Version 4.7 Software; Copyright Accelrys Inc. 2002.

(161) Molecular Operating Environment, v.2010, The Chemical Computing Group, Montreal, QC, Canada.

(162) Dixon, S. L.; Smondyrev, A. M.; Knoll, E. H.; Rao, S. N.; Shaw, D. E.; Friesner, R. A. PHASE: A New Engine for Pharmacophore Perception, 3D QSAR Model Development, and 3D Database Screening: 1. Methodology and Preliminary Results. *J. Comput. Aided Mol. Des.* **2006**, *20* (10–11), 647–671. https://doi.org/10.1007/s10822-006-9087-6.

(163) Wolber, G.; Langer, T. LigandScout: 3-D Pharmacophores Derived from Protein-Bound Ligands and Their Use as Virtual Screening Filters. *J. Chem. Inf. Model.* **2005**, *45* (1), 160–169. https://doi.org/10.1021/ci049885e.

(164) Aparoy, P.; Kumar Reddy, K.; Reddanna, P. Structure and Ligand Based Drug Design Strategies in the Development of Novel 5- LOX Inhibitors. *CMC* **2012**, *19* (22), 3763–3778. https://doi.org/10.2174/092986712801661112.

(165) Smola, M.; Gutten, O.; Dejmek, M.; Kožíšek, M.; Evangelidis, T.; Tehrani, Z. A.; Novotná, B.; Nencka, R.; Birkuš, G.; Rulíšek, L.; Boura, E. Ligand Strain and Its Conformational Complexity Is a Major Factor in the Binding of Cyclic Dinucleotides to STING Protein. *Angew. Chem. Int. Ed.* **2021**, *60* (18), 10172–10178. https://doi.org/10.1002/anie.202016805.

(166) Foloppe, N.; Chen, I.-J. The Reorganization Energy of Compounds upon Binding to Proteins, from Dynamic and Solvated Bound and Unbound States. *Bioorg. Med. Chem.* **2021**, *51*, 116464. https://doi.org/10.1016/j.bmc.2021.116464.

(167) Mobley, D. L.; Dill, K. A. Binding of Small-Molecule Ligands to Proteins: "What You See" Is Not Always "What You Get." *Structure* **2009**, *17* (4), 489–498. https://doi.org/10.1016/j.str.2009.02.010.

(168) Penner, P.; Guba, W.; Schmidt, R.; Meyder, A.; Stahl, M.; Rarey, M. The Torsion Library: Semiautomated Improvement of Torsion Rules with SMARTScompare. *J. Chem. Inf. Model.* **2022**, *62* (7), 1644–1653. https://doi.org/10.1021/acs.jcim.2c00043.

(169) Gu, S.; Smith, M. S.; Yang, Y.; Irwin, J. J.; Shoichet, B. K. Ligand Strain Energy in Large Library Docking. *J. Chem. Inf. Model.* **2021**, *61* (9), 4331–4341. https://doi.org/10.1021/acs.jcim.1c00368.

(170) Jain, A. N.; Brueckner, A. C.; Cleves, A. E.; Reibarkh, M.; Sherer, E. C. A Distributional Model of Bound Ligand Conformational Strain: From Small Molecules up to Large Peptidic Macrocycles. *J. Med. Chem.* **2023**, *66* (3), 1955–1971. https://doi.org/10.1021/acs.jmedchem.2c01744.

(171) Peach, M. L.; Cachau, R. E.; Nicklaus, M. C. Conformational Energy Range of Ligands in Protein Crystal Structures: The Difficult Quest for Accurate Understanding. *J. Mol. Recognit.* **2017**, *30* (8). https://doi.org/10.1002/jmr.2618.

(172) Tong, J.; Zhao, S. Large-Scale Analysis of Bioactive Ligand Conformational Strain Energy by *Ab Initio* Calculation. *J. Chem. Inf. Model.* **2021**, *61* (3), 1180–1192. https://doi.org/10.1021/acs.jcim.0c01197.

(173) Rai, B. K.; Sresht, V.; Yang, Q.; Unwalla, R.; Tu, M.; Mathiowetz, A. M.; Bakken, G. A. Comprehensive Assessment of Torsional Strain in Crystal Structures of Small Molecules and Protein–Ligand Complexes Using Ab Initio Calculations. *J. Chem. Inf. Model.* **2019**, *59* (10), 4195–4208. https://doi.org/10.1021/acs.jcim.9b00373.

(174) SZYBKI 2.9.0.1. OpenEye, Cadence Molecular Sciences, Santa Fe, NM. http://www.eyesopen.com.

(175) Berenger, F.; Tsuda, K. An ANI -2 Enabled Open-source Protocol to Estimate Ligand Strain after Docking. *J. Comput. Chem.* **2025**, *46* (1), e27478. https://doi.org/10.1002/jcc.27478.

(176) Wallace, E. R. S.; Frey, N. C.; Rackers, J. A. Strain Problems Got You in a Twist? Try StrainRelief: A Quantum-Accurate Tool for Ligand Strain Calculations. *J. Chem. Inf. Model.* **2025**, *65* (13), 6613–6620. https://doi.org/10.1021/acs.jcim.5c00586.

(177) Rai, B. K.; Sresht, V.; Yang, Q.; Unwalla, R.; Tu, M.; Mathiowetz, A. M.; Bakken, G. A. TorsionNet: A Deep Neural Network to Rapidly Predict Small-Molecule Torsional Energy Profiles with the Accuracy of Quantum Mechanics. *J. Chem. Inf. Model.* **2022**, *62* (4), 785–800. https://doi.org/10.1021/acs.jcim.1c01346.

(178) Rumyantseva, V. K.; Morozkina, S. N.; Uspenskaya, M. V.; Petukhov, M. G. Relaxation of Steric Strains of TTR-Type Amyloid Fibril Inhibitors Radically Changes the Results of Their Virtual Screening. *Cell Tiss. Biol.* **2024**, *18* (5), 561–569. https://doi.org/10.1134/S1990519X24700433.

(179) Salentin, S.; Schreiber, S.; Haupt, V. J.; Adasme, M. F.; Schroeder, M. PLIP: Fully Automated Protein–Ligand Interaction Profiler. *Nucleic Acids Res.* **2015**, *43* (W1), W443–W447. https://doi.org/10.1093/nar/gkv315.

(180) Schake, P.; Bolz, S. N.; Linnemann, K.; Schroeder, M. PLIP 2025: Introducing Protein–Protein Interactions to the Protein–Ligand Interaction Profiler. *Nucleic Acids Res.* **2025**, *53* (W1), W463–W465. https://doi.org/10.1093/nar/gkaf361.

(181) Ha, E. J.; Lwin, C. T.; Durrant, J. D. LigGrep: A Tool for Filtering Docked Poses to Improve Virtual-Screening Hit Rates. *J. Cheminform.* **2020**, *12* (1), 69. https://doi.org/10.1186/s13321-020-00471-2.

(182) Marcou, G.; Rognan, D. Optimizing Fragment and Scaffold Docking by Use of Molecular Interaction Fingerprints. *J. Chem. Inf. Model.* **2007**, *47* (1), 195–207. https://doi.org/10.1021/ci600342e.

(183) Beroza, P.; Crawford, J. J.; Ganichkin, O.; Gendelev, L.; Harris, S. F.; Klein, R.; Miu, A.; Steinbacher, S.; Klingler, F.-M.; Lemmen, C. Chemical Space Docking Enables Large-Scale Structure-Based Virtual Screening to Discover ROCK1 Kinase Inhibitors. *Nat. Commun.* **2022**, *13* (1), 6447. https://doi.org/10.1038/s41467-022-33981-8.

(184) Sindt, F.; Bret, G.; Rognan, D. On the Difficulty to Rescore Hits from Ultralarge Docking Screens. *J. Chem. Inf. Model.* **2025**, *65* (11), 5553–5566. https://doi.org/10.1021/acs.jcim.5c00730.

(185) Lionta, E.; Spyrou, G.; Vassilatis, D. K.; Cournia, Z. Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. *Curr. Top. Med. Chem.* **2014**, *14* (16), 1923–1938. https://doi.org/10.2174/1568026614666140929124445.

(186) Shub, L.; Korczynska, M.; Muir, D. F.; Lin, F.-Y.; Hall, B. W.; Mathiowetz, A. M.; Keiser, M. J. Autoparty: Machine Learning-Guided Visual Inspection of Molecular Docking Results. *J. Chem. Inf. Model.* **2025**, *65* (15), 7817–7826. https://doi.org/10.1021/acs.jcim.5c00850.

(187) Luukkonen, S.; Van Den Maagdenberg, H. W.; Emmerich, M. T. M.; Van Westen, G. J. P. Artificial Intelligence in Multi-Objective Drug Design. *Curr. Opin. Struct. Biol.* **2023**, *79*, 102537. https://doi.org/10.1016/j.sbi.2023.102537.

(188) Zhang, X.; Gao, H.; Wang, H.; Chen, Z.; Zhang, Z.; Chen, X.; Li, Y.; Qi, Y.; Wang, R. PLANET: A Multi-Objective Graph Neural Network Model for Protein–Ligand Binding Affinity Prediction. *J. Chem. Inf. Model.* **2024**, *64* (7), 2205–2220. https://doi.org/10.1021/acs.jcim.3c00253.

(189) Nicolaou, C. A.; Brown, N. Multi-Objective Optimization Methods in Drug Design. *Drug Discov. Today Technol.* **2013**, *10* (3), e427–e435. https://doi.org/10.1016/j.ddtec.2013.02.001.

(190) Allenspach, S.; Hiss, J. A.; Schneider, G. Neural Multi-Task Learning in Drug Design. *Nat. Mach. Intell.* **2024**, *6* (2), 124–137. https://doi.org/10.1038/s42256-023-00785-4.

(191) Sener, O.; Koltun, V. Multi-Task Learning as Multi-Objective Optimization. *arXiv*. 2018. https://doi.org/10.48550/ARXIV.1810.04650.

(192) Crawshaw, M. Multi-Task Learning with Deep Neural Networks: A Survey. *arXiv*. 2020. https://doi.org/10.48550/ARXIV.2009.09796.

(193) Yu, J.; Dai, Y.; Liu, X.; Huang, J.; Shen, Y.; Zhang, K.; Zhou, R.; Adhikarla, E.; Ye, W.; Liu, Y.; Kong, Z.; Zhang, K.; Yin, Y.; Namboodiri, V.; Davison, B. D.; Moore, J. H.; Chen, Y. Unleashing the Power of Multi-Task Learning: A Comprehensive Survey Spanning Traditional, Deep, and Pretrained Foundation Model Eras. *arXiv*. 2024. https://doi.org/10.48550/ARXIV.2404.18961.

(194) Caruana, R. Multitask Learning. *Mach. Learn.* **1997**, *28* (1), 41–75. https://doi.org/10.1023/A:1007379606734.

(195) Guo, X.; Ha, M.; Tao, X.; Li, S.; Li, Y.; Zhu, Z.; Shen, Z.; Ma, L. Multi-Task Learning with Sequential Dependence Toward Industrial Applications: A Systematic Formulation. *ACM Trans. Knowl. Discov. Data* **2024**, *18* (5), 1–29. https://doi.org/10.1145/3640468.

(196) Yu, T.; Kumar, S.; Gupta, A.; Levine, S.; Hausman, K.; Finn, C. Gradient Surgery for Multi-Task Learning. *arXiv*. 2020. https://doi.org/10.48550/ARXIV.2001.06782.

(197) Liu, B.; Liu, X.; Jin, X.; Stone, P.; Liu, Q. Conflict-Averse Gradient Descent for Multi-Task Learning. *arXiv*. 2021. https://doi.org/10.48550/ARXIV.2110.14048.

(198) Kendall, A.; Gal, Y.; Cipolla, R. Multi-Task Learning Using Uncertainty to Weigh Losses for Scene Geometry and Semantics. *arXiv*. 2017. https://doi.org/10.48550/ARXIV.1705.07115.

(199) Liu, S.; Johns, E.; Davison, A. J. End-to-End Multi-Task Learning with Attention. *arXiv*. 2018. https://doi.org/10.48550/ARXIV.1803.10704.

(200) Chen, Z.; Badrinarayanan, V.; Lee, C.-Y.; Rabinovich, A. GradNorm: Gradient Normalization for Adaptive Loss Balancing in Deep Multitask Networks. *arXiv*. 2017. https://doi.org/10.48550/ARXIV.1711.02257.

(201) Ruder, S. An Overview of Multi-Task Learning in Deep Neural Networks. *arXiv*. 2017. https://doi.org/10.48550/ARXIV.1706.05098.

(202) Liu, S.; Qu, M.; Zhang, Z.; Cai, H.; Tang, J. Structured Multi-Task Learning for Molecular Property Prediction. *arXiv*. 2022. https://doi.org/10.48550/ARXIV.2203.04695.

(203) Lee, C.; Jeong, D.-W.; Ko, S. M.; Lee, S.; Kim, H.; Yim, S.; Han, S.; Kim, S.; Lim, S. Scalable Multi-Task Transfer Learning for Molecular Property Prediction. *arXiv*. 2024. https://doi.org/10.48550/ARXIV.2410.00432.

(204) Javaid, M. B.; Gervens, T.; Mitsos, A.; Grohe, M.; Rittig, J. G. Exploring Data Augmentation: Multi-Task Methods for Molecular Property Prediction. *Comput. Chem. Eng.* **2025**, *201*, 109253. https://doi.org/10.1016/j.compchemeng.2025.109253.

(205) Han, X.; Cai, J.; Bai, C.; Wu, Z. Triview Molecular Representation Learning Combined with Multitask Optimization for Enhanced Molecular Property Prediction. *J. Chem. Inf. Model.* **2025**, *65* (10), 5163–5175. https://doi.org/10.1021/acs.jcim.5c00436.

(206) Yang, X.; Wang, Y.; Lin, Y.; Zhang, M.; Liu, O.; Shuai, J.; Zhao, Q. A Multi-Task Self-Supervised Strategy for Predicting Molecular Properties and FGFR1 Inhibitors. *Adv. Sci.* **2025**, *12* (13), 2412987. https://doi.org/10.1002/advs.202412987.

(207) Rodríguez-Pérez, R.; Bajorath, J. Multitask Machine Learning for Classifying Highly and Weakly Potent Kinase Inhibitors. *ACS Omega* **2019**, *4* (2), 4367–4375. https://doi.org/10.1021/acsomega.9b00298.

(208) Tan, M. Prediction of Anti-Cancer Drug Response by Kernelized Multi-Task Learning. *Artif. Intell. Med.* **2016**, *73*, 70–77. https://doi.org/10.1016/j.artmed.2016.09.004.

(209) Lin, S.; Shi, C.; Chen, J. GeneralizedDTA: Combining Pre-Training and Multi-Task Learning to Predict Drug-Target Binding Affinity for Unknown Drug Discovery. *BMC Bioinform.* **2022**, *23* (1), 367. https://doi.org/10.1186/s12859-022-04905-6.

(210) Lee, K.; Kim, D. In-Silico Molecular Binding Prediction for Human Drug Targets Using Deep Neural Multi-Task Learning. *Genes* **2019**, *10* (11), 906. https://doi.org/10.3390/genes10110906.

(211) Zhao, Z.; Qin, J.; Gou, Z.; Zhang, Y.; Yang, Y. Multi-Task Learning Models for Predicting Active Compounds. *J. Biomed. Inform.* **2020**, *108*, 103484. https://doi.org/10.1016/j.jbi.2020.103484.

(212) Shah, P. M.; Zhu, H.; Lu, Z.; Wang, K.; Tang, J.; Li, M. DeepDTAGen: A Multitask Deep Learning Framework for Drug-Target Affinity Prediction and Target-Aware Drugs Generation. *Nat. Commun.* **2025**, *16* (1), 5021. https://doi.org/10.1038/s41467-025-59917-6.

(213) Santos, E. S. D. A.; Lemos, J. M.; Dos Santos Carvalho, A. M.; Mendonça De Melo, F. D. S.; Pereira, E. D. S.; Moreira-Filho, J. T.; Braga, R. D. C.; Muratov, E. N.; Grellier, P.; Charneau, S.; Bastos, I. M. D.; Neves, B. J. Deep Multitask Learning-Driven Discovery of New Compounds Targeting *Leishmania Infantum*. *ACS Omega* **2024**, *9* (52), 51271–51284. https://doi.org/10.1021/acsomega.4c07994.

(214) Li, X.; Xu, Y.; Lai, L.; Pei, J. Prediction of Human Cytochrome P450 Inhibition Using a Multitask Deep Autoencoder Neural Network. *Mol. Pharmaceutics* **2018**, *15* (10), 4336–4345. https://doi.org/10.1021/acs.molpharmaceut.8b00110.

(215) Wenzel, J.; Matter, H.; Schmidt, F. Predictive Multitask Deep Neural Network Models for ADME-Tox Properties: Learning from Large Data Sets. *J. Chem. Inf. Model.* **2019**, *59* (3), 1253–1268. https://doi.org/10.1021/acs.jcim.8b00785.

(216) Walter, M.; Borghardt, J. M.; Humbeck, L.; Skalic, M. Multi-Task ADME/PK Prediction at Industrial Scale: Leveraging Large and Diverse Experimental Datasets. *Mol. Inf.* **2024**, *43* (10), e202400079. https://doi.org/10.1002/minf.202400079.

(217) Du, B.-X.; Xu, Y.; Yiu, S.-M.; Yu, H.; Shi, J.-Y. ADMET Property Prediction via Multi-Task Graph Learning under Adaptive Auxiliary Task Selection. *iScience* **2023**, *26* (11), 108285. https://doi.org/10.1016/j.isci.2023.108285.

(218) Shahid, S.; Maity, D.; Chakrabarty, S. MTAN-ADMET: A Multi-Task Adaptive Neural Network for Efficient and Accurate Prediction of ADMET Properties. *chemRxiv.* August 12, 2025. https://doi.org/10.26434/chemrxiv-2025-zhrsk.

(219) Goossens, K.; Tricarico, G.; Hofmans, J.; Dréanic, M.-P.; De Cesco, S.; Lenselink, E. B. ChemProp Multi-Task Models for Predicting ADME Properties in the Polaris Challenge. *chemRxiv.* June 10, 2025. https://doi.org/10.26434/chemrxiv-2025-q12vh.

Chapter 2: Retrospective Virtual Screening: Benchmarking Enrichment with Three-Dimensional Medicinal Chemistry Filters Post-Molecular Docking

As discussed earlier in the previous chapter, there is an interest in improving AI-accelerated docking with the incorporation of select 3D medicinal chemistry properties that are simplified or omitted in conventional docking and contribute to common docking artifacts. Computational tools have been developed to characterize these 3D medicinal chemistry properties (geometry, features, interactions, etc.) in protein-ligand binding.[1–13] These tools can assist in the post-processing of molecular dockings to help prioritize candidates for experimental testing. The most common criteria considered in the visual inspection of molecular dockings includes strain and unsatisfied heteroatoms;[14] virtual hits with highly strained conformations and multiple unsatisfied hydrogen bonds are filtered out.[15] These filters have been incorporated separately[16–20] or in tandem[15,21–25] in VS to improve hit enrichment. Filter thresholds found in the literature for strain and unsatisfied hydrogen bond donors and acceptors are often inconsistent between reports. Gu et al.'s statistics-based strain filter[2] is a method that is well suited for large-scale VS. Their method benchmark of 40 DUD-E datasets at different strain thresholds in torsional energy units (TEU) shows the highest average enrichment at 4 TEU and most proteins showing improvement at 7.0 and 7.5 TEUs. The positive average enrichment across all explored thresholds (4-8 TEUs) suggests that strain filtering improves virtual hit lists. The most reported strain threshold in VS campaigns is 8 TEU[18,19,22] – this selection is not the most supported threshold in the method benchmark, but is the least stringent threshold explored and therefore retains the most hits. Moreover, other strain thresholds are reported[20] that serve no consensus in threshold selection. Similar inconsistencies are found in the thresholds selected to filter virtual hits based on the number of unsatisfied hydrogen bonds. Numerous thresholds are recommended or reported for unsatisfied hydrogen bond donors (0-2) and acceptors (1-4, 6).[15–17,21–23,26] Some of these thresholds may be selected based on specific binding site knowledge where available. There are no current, public benchmarks to support the selection of filter thresholds for unsatisfied hydrogen bonds. Therefore, this chapter aims to benchmark strain and unsatisfied hydrogen bond donor and acceptor filter thresholds with retrospective VS datasets to assess whether their application improves enrichment compared with molecular docking scores alone and observe whether there is a general threshold consensus for prospective studies.

2.1 Methods

2.1.1 Binding Databases

To test multiple 3D medicinal chemistry filters used in the selection of virtual hit candidates, we considered datasets of available protein-ligand bioactivities for retrospective VS. GPCRs are the largest protein family targeted in drug development, representing an active area of interest.[27] GPCR-Bench[28] is a retrospective validation database developed to evaluate molecular docking protocols. The GPCR-Bench database contains twenty-four GPCR protein-ligand complexes, ChEMBL19 actives, and DUD-E generated decoys – developed with the objective to simulate a diverse, retrospective VS campaign. Developers collected protein-ligand complex X-ray crystal structures from the PBD and selected the highest-resolution and/or most complete representative structure for each protein. Experimental active compounds were searched for in the ChEMBL19 database and further filtered based on annotated K_i, K_d, IC_{50}, or EC_{50} values and drug-likeness (molecular weight, hydrogen bonds, rotatable bonds, ring, reactive non-metals, etc.). An average of sixty computational decoys per experimental active were generated with the DUD-E tool. The GPCR-Bench database was downloaded and used as available. LIT-PCBA[29] is another retrospective validation database developed to evaluate VS protocols. This database contains fifteen protein-ligand complexes representative of protein families of pharmaceutical interest, 7844 confirmed actives, and 407,381 confirmed inactive compounds – designed to mimic typical HTS hit rates and potency distribution. Active and inactive compounds were collected from 149 PubChem dose-response bioassays and further processed to remove false positives, frequent hitters, assay artifacts, etc. The LIT-PCBA database was downloaded and used as available.

2.1.2 Molecular Docking

2.1.2.1 Molecular Docking Programs

AutoDock-GPU and GNINA 1.0 are among common molecular docking programs used.[30–32] GNINA was pronounced the winning submission of the 2024 CACHE challenge.[33] These programs were selected for their relative speed and ease of use.

2.1.2.1.1 AutoDock

AutoDock-GPU (AutoDock) is an OpenCL implementation of the AutoDock4[34] engine developed to accelerate the search for optimal protein-ligand binding in large scale VS.[35] AutoDock uses the Lamarckian Genetic Algorithm (LGA) and semi-empirical free energy scoring function. The LGA generates random ligand conformations which are further evolved through iterative, genetic operations to search the conformational space.[36] Then, the semi-empirical scoring function estimates the binding free energy of the generated simulation conformations – including implicit dispersion/repulsion, hydrogen bonding, electrostatic, and desolvation contributions.[37] Protein-ligand complexes with known structure and inhibition constants (K_i) were used to parameterize the force field.[38] The force field assesses binding in two steps: the intramolecular energetics involved in the transition of the unbound ligand state to its respective protein bound conformation, and the intermolecular energetics of their interaction in complex.[39] Protein-ligand simulations output the predicted binding mode and affinity (kcal/mol).[34]

The parameters for AutoDock were set to 100 independent docking runs per ligand using the LGA local search method. Heuristic optimization was enabled to accelerate convergence with the adaptive tuning of LGA parameters (population size, mutation rate, local search frequency, etc.) based on system size and GPU architecture. Docking grids were precomputed based on the native protein-ligand complex. The binding mode with the lowest docking score was selected to be the best predictive binding mode for each ligand docking. AutoDock molecular dockings output in PDBQT format were converted to PDB, then SDF format with Open Babel[40].

2.1.2.1.2 GNINA

GNINA 1.0[41] (GNINA) is a DL-enhanced fork of smina[42], further derived from AutoDock Vina[43], with an integrated convolutional neural network (CNN) to support geometry optimization and scoring. GNINA uses Monte Carlo sampling to search the conformational space, empirical scoring to assess the generated conformations, and CNN refinement to rescore the binding modes. The Monte Carlo chain selects a random translation, rotation, or torsional operation to mutate the ligand's geometry. Then, an approximate energy minimization optimizes the binding conformations and an empirical scoring function scores the conformations. Minimized

conformations are accepted or rejected based on the Metropolis criteria; this prioritizes lower energy conformations but enables some uphill movement to escape local minima. The global top binding conformations from all Monte Carlo chains are further refined. After refinement, the CNN rescores each final binding mode and ranks the binding modes with the CNNscore (probabilistic binding mode confidence score) and the CNNaffinity score predicts the pK value. In the retrospective screenings herein, the CNN_VS score (CNNaffinity \times CNNscore) was used.

2.1.2.2 Protein Preparation

The protein structures were prepared in Molecular Operating Environment (MOE[44]). Proteins were visualized to remove solvent, water, ions, and bound ligand, where present. The Structure Preparation module was used to correct structural issues and Protonate3D was used to add hydrogen atoms. The prepared protein structures were saved in PDB format. Proteins prepared for AutoDock were converted from PDB to PDBQT format with AutoDockTools.

2.1.2.3 Binding Site

GPCR-Bench and LIT-PCBA datasets include the native protein-ligand structures. The binding site was set and centred on the native ligand bound to each respective protein and grid maps (0.375 Å spacing[39]) were generated with AutoDockTools. The same grid generated for AutoDock was used in GNINA docking; grid spacing was converted to angstroms with a conversion factor of 0.375 Å.

2.1.2.4 Ligand Preparation

OpenEye's Flipper module[45] was used to enumerate stereocentres with unspecified stereochemistry (maxcentres were set to 3) and QUACPAC module[46] to calculate one dominant tautomeric and protonation form of each isomer at physiological pH (7.4).[47]

2.1.3 Strain Calculations

The selected strain tool[2] is a statistical-based method that calculates strain through the matching of torsion patterns in a molecule to patterns in a torsion database. Developers used the Cambridge Structure Database to convert (Boltzmann) observed torsion angle distributions into

artificial torsion energy units (TEUs) in the construction of the torsion database. This tool was selected for its speed in support of ultra-large scale VS – takes less than 0.04 seconds to assess a conformation on a standard core.[2]

The strain filter tool was downloaded from http://tldr.docking.org. AutoDock and GNINA molecular dockings in SDF format were converted to the required input file format (mol2) using OpenEye Scientific OEChem Toolkit and run through the strain tool. The total strain energy of the molecule was used. Total strain energy is the sum of all individual torsion energies in the molecule.

2.1.4 Hydrogen Bonds Detection

IChem[48] is a free suite of software designed to analyze protein-ligand interactions. Three-dimensional molecular information is converted into fingerprints or graphs to enable high-throughput analysis and feed ML models in the prediction of important molecular features and interactions.

IChem was downloaded from http://bioinfo-pharma.u-strasbg.fr/labwebsite/download.html. The developers were contacted to get an in-house script that parses IChem output to count the number of ligand polar atoms that do not interact with the protein binding site.[15] AutoDock and GNINA molecular dockings in SDF format were converted to the required input file format (mol2) using OpenEye Scientific OEChem Toolkit. LIT-PCBA proteins were provided in the required mol2 format and prepared in MOE; GPCR-Bench proteins in PDB format were prepared and converted to mol2 format in MOE. IChem was run both with and without the intramolecular hydrogen bond parameter.

2.1.5 Virtual Screening Metrics

We considered widely adopted VS metrics such as the area under (AUC) the receiver operating characteristic (ROC) curve or variations thereof and enrichment factor.[26,49,50] All these metrics consider VS as a classification task between active versus inactive molecules.

ROC is a plot of the true positive rate (TPR) versus the false positive rate (FPR) at all classification thresholds (Figure 2.1, A). The ROC AUC is a binary classification performance

metric that measures a model's ability to differentiate between classes. An AUC of 1 represents a perfect classification model: the model will predict a higher rank for a random positive example than negative example, and an AUC of 0.5 represents a random classification: the model does not predict better than random selection. The early recognition of active molecules is critical in prospective VS because only the top-ranked molecules are selected for experimental testing[49] – ROC AUC fails to differentiate between methods with the same AUC. Therefore, several ROC AUC variations have been proposed to address the early recognition of active molecules.

We evaluated VS with the log-weighted area under the ROC curve (LogAUC), Figure 2.1, B. The x-axis (FPR) is plotted on a logarithmic scale to emphasize early enrichment; preferential weight is given to the early behaviour of the function.[51] The standard x-axis offset parameter (a) is 1.0 x 10^{-3}, which sets the LogAUC of a random classifier to 0.145.[52] LogAUC herein is scaled for better readability; the reference LogAUC (random classifier) becomes 1.45.

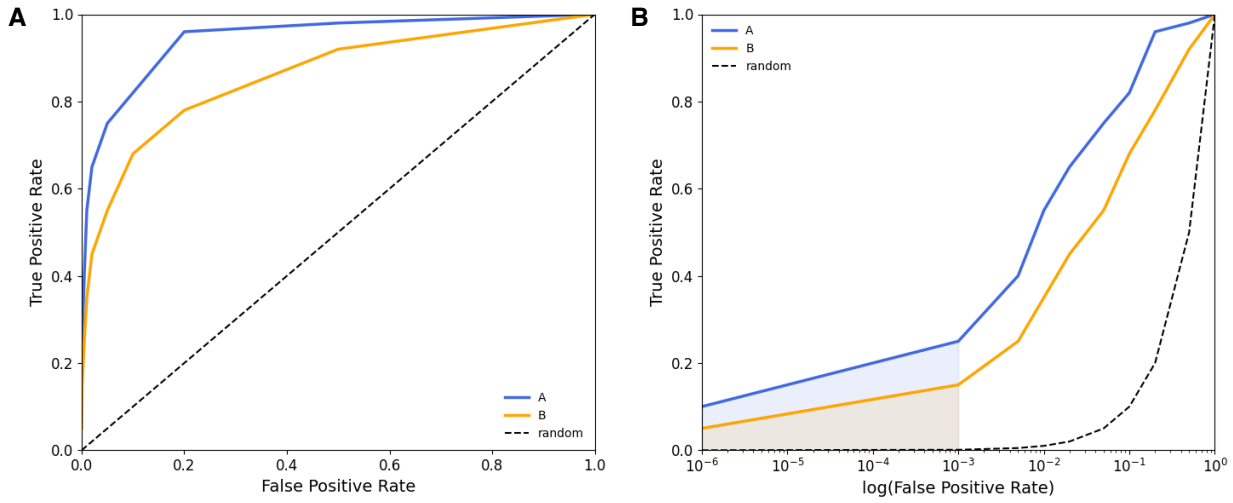


Figure 2.1. Receiver operating characteristic (ROC) and log-weighted area under the ROC (LogAUC) curves. (A) ROC curve. The diagonal (dashed, black) represents a random classifier. Curves above the diagonal perform better than a random classifier and curves below the diagonal perform worse than a random classifier. A (blue) performs better than B (orange). **(B) LogAUC curve.** The x-axis plotted on a logarithmic scale (log10) emphasizes early enrichment (shaded region. A (blue) shows better early enrichment with greater AUC (shaded blue region) than B (shaded orange region).

2.2 Results and Discussion

All datasets in the retrospective databases were docked with AutoDock and GNINA. Docked molecules were run through the strain and IChem programs to calculate strain and the number of unsatisfied hydrogen bond donors and acceptors, respectively. Threshold-based filtering was used. AutoDock and GNINA dockings were ranked with their respective docking scores (ADSCOR and CNN_VS) and then separately filtered at various strain and unsatisfied hydrogen bond thresholds to assess enrichment with LogAUC.

The IChem program used to count the number of unsatisfied hydrogen bond donors and acceptors was run both with and without the intramolecular hydrogen bond parameter. When selected, this parameter accounts for intramolecular hydrogen bonding in the count of unsatisfied hydrogen bonds. These polar heteroatoms that participate in intramolecular hydrogen bonding are considered satisfied. There was no difference observed in enrichment when filtering molecules with or without the intramolecular hydrogen bond parameter considered in the count of unsatisfied hydrogen bonds. Moreover, intramolecular hydrogen bonds in protein-ligand binding do not guarantee improved binding activity.[53] Inclusion of this parameter may also further complicate the prediction of unsatisfied hydrogen bonds in the machine learning portion (chapter 3) of this research – increases the complexity of the task. Therefore, the simulated VS campaigns herein did not consider intramolecular hydrogen bonds in the count of unsatisfied hydrogen bonds.

Some examples of molecular docking artifacts that were filtered out are illustrated (Figure 2.2). These examples come from the simulated VS campaigns; shown are molecules ranked within the top 35 virtual hits against their respective targets. Figure 2.2, A shows the top predicted molecule in the MAPK1 (LIT-PCBA dataset) simulated VS campaign with AutoDock; this is an example of a molecule with a strained dihedral that exploited the simplifications in common molecular docking scoring functions. Figure 2.2, B shows the top predicted molecule in the KAT2A (LIT-PCBA dataset) simulated VS campaign with GNINA; this is an example of a molecule with numerous unsatisfied molecule heteroatoms in the protein binding site: 3 unsatisfied donors and 5 unsatisfied acceptors.

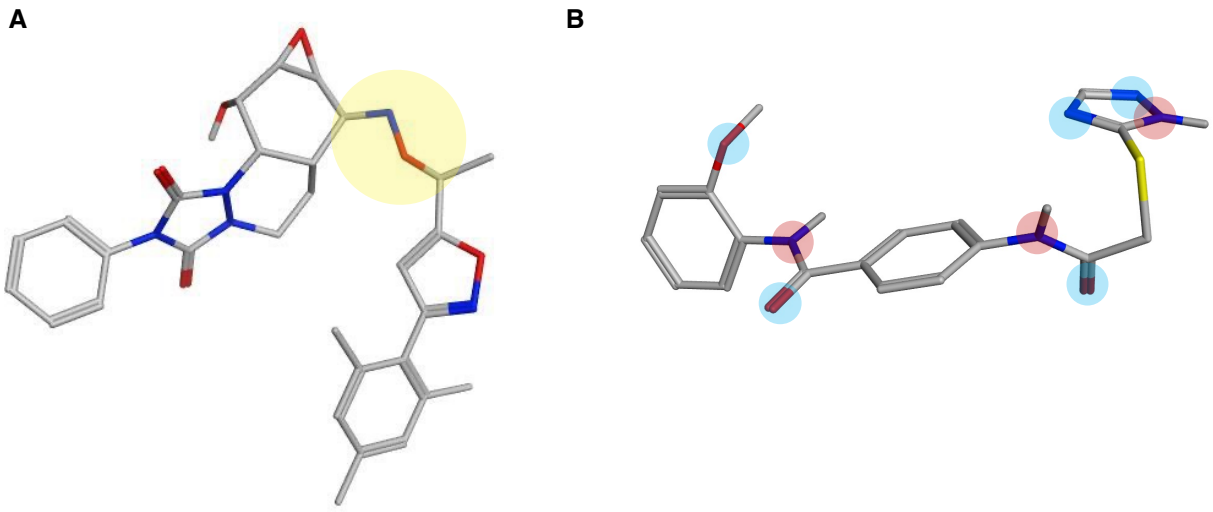


Figure 2.2. Select molecular docking artifact examples. (A) Molecule 26751022_5 docked to MAPK1 (LIT-PCBA dataset) with AutoDock. The modelled dihedral (yellow highlight) is highly strained (3.69 TEU), as calculated with the strain program. (B) Molecule 24271103_1 docked to KAT2A (LIT-PCBA dataset) with GNINA. Ligand unsatisfied hydrogen bond donors (red highlight) and unsatisfied hydrogen bond acceptors (blue highlight) are shown, as identified with the IChem program. The protein binding site was omitted in both images for clear visualization. Note that only polar hydrogens are visualized.

The methods discussed above were implemented to test the impact of strain, unsatisfied hydrogen bond donors, and unsatisfied hydrogen bond acceptors as filters in simulated VS campaigns and assess their impact on early enrichment. Figure 2.3, A-F shows the plots of delta LogAUC versus the selected medicinal chemistry filters at different thresholds, where the baseline reference is ranking with docking score alone.

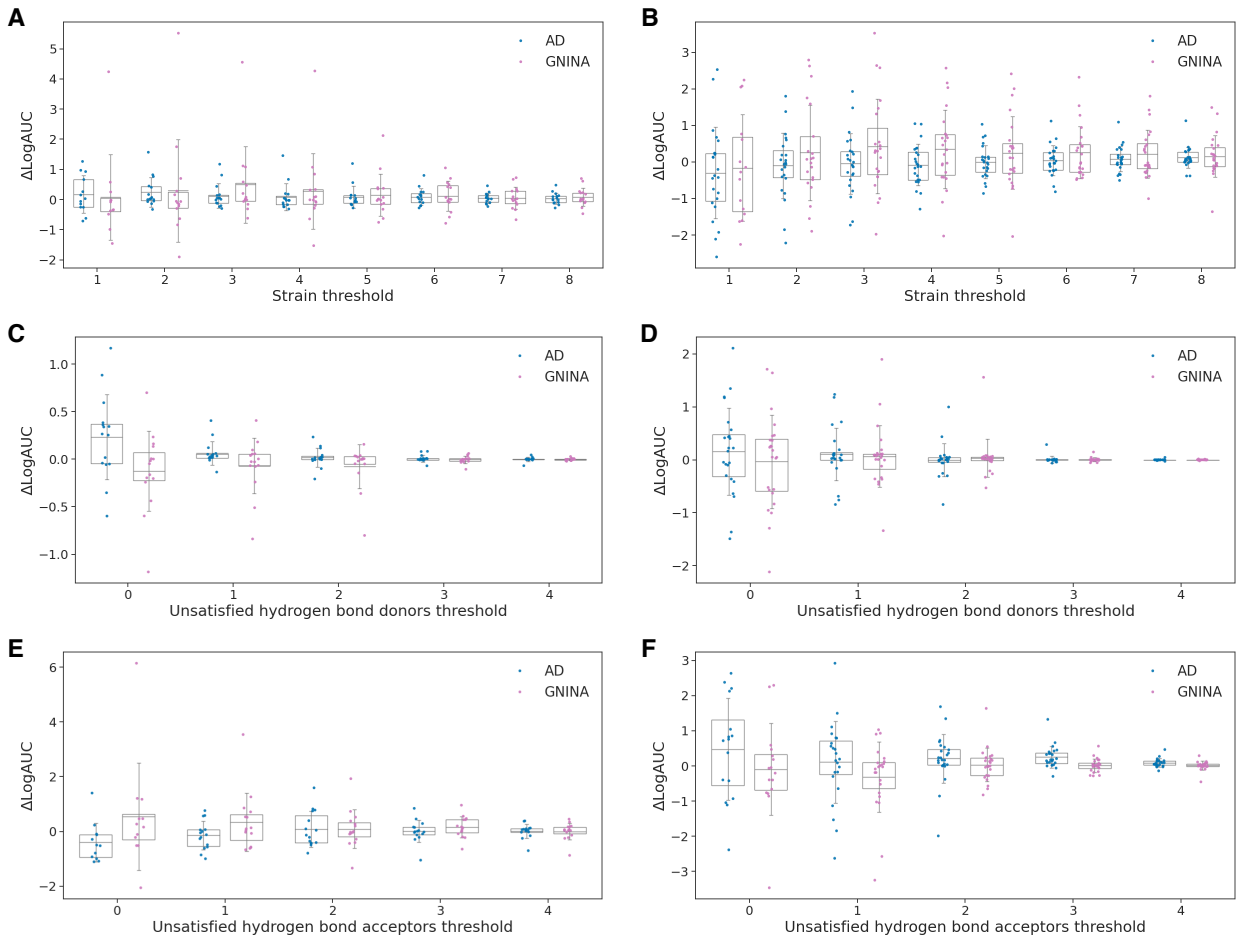


Figure 2.3 Enrichment in simulated virtual screening campaigns with selected retrospective databases. Upper panel (A-B): strain filtering. Middle panel (C-D): unsatisfied hydrogen bond donors filtering. Lower panel (E-F): unsatisfied hydrogen bond acceptors filtering. Left panel (A, C, E): LIT-PCBA datasets. Right panel (B, D, F): GPCR-Bench datasets. GPCR-Bench and LIT-PCBA databases were docked with AutoDock (blue) and GNINA (pink). ΔLogAUC is difference in LogAUC between the baseline (molecular docking scores) and filtered dataset. Each datapoint on the plots represents the ΔLogAUC for a particular protein-ligand dataset. The mean ΔLogAUC (horizontal grey lines), standard deviations (vertical grey lines), and interquartile range (grey boxes) are shown.

Figure 2.3, A-F shows that no single threshold led to a noticeable enrichment of all protein-ligand datasets in the two examined databases. This observation suggests that the selected filters are not applicable in general and need to be retrospectively assessed on a specific protein of interest before prospective VS, at least for the tested molecular docking programs shown. Other independent studies which evaluated different strain methods on the same LIT-PCBA database have suggested the same system-dependence.[7]

Moreover, we also conducted an exhaustive search of filter threshold combinations, Figure 2.4. This showed the same target-dependence: no specific filter combination led to a generalized enrichment across all protein-ligand datasets.

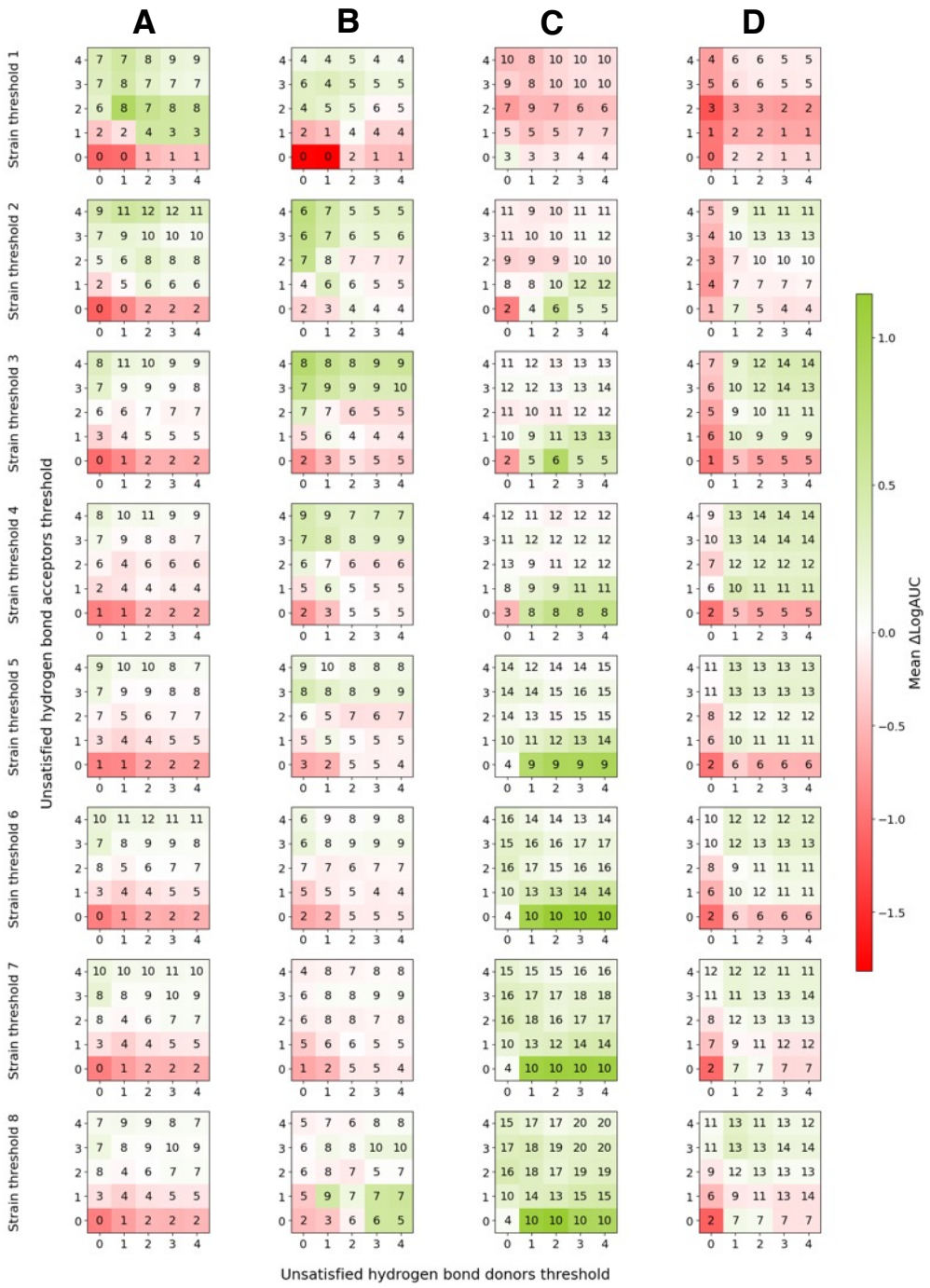


Figure 2.4. An exhaustive search of all filter threshold combinations with selective retrospective databases. (A) LIT-PCBA database docked with AutoDock. (B) LIT-PCBA database docked with GNINA. (C) GPCR-Bench database docked with AutoDock. (D) GPCR-Bench database docked with GNINA. Each row of plots represents a strain threshold. The x-axis is the unsatisfied hydrogen bond donors threshold and the y-axis is unsatisfied hydrogen bond acceptors threshold on each plot. The heatmap indicates the mean ΔLogAUC of a database (A-D) at a specific threshold combination. The count inside each cell indicates the number of proteins in each database that had a positive ΔLogAUC at that threshold combination. Total count proteins in database: LIT-PCBA (15) and GPCR-Bench (24).

Several proteins in the explored databases clearly benefitted from post-processing molecular docking with the selected filters, Figure 2.5, A-D. Most protein-ligand datasets at their best filter combination showed an improvement in enrichment with the selected filters.

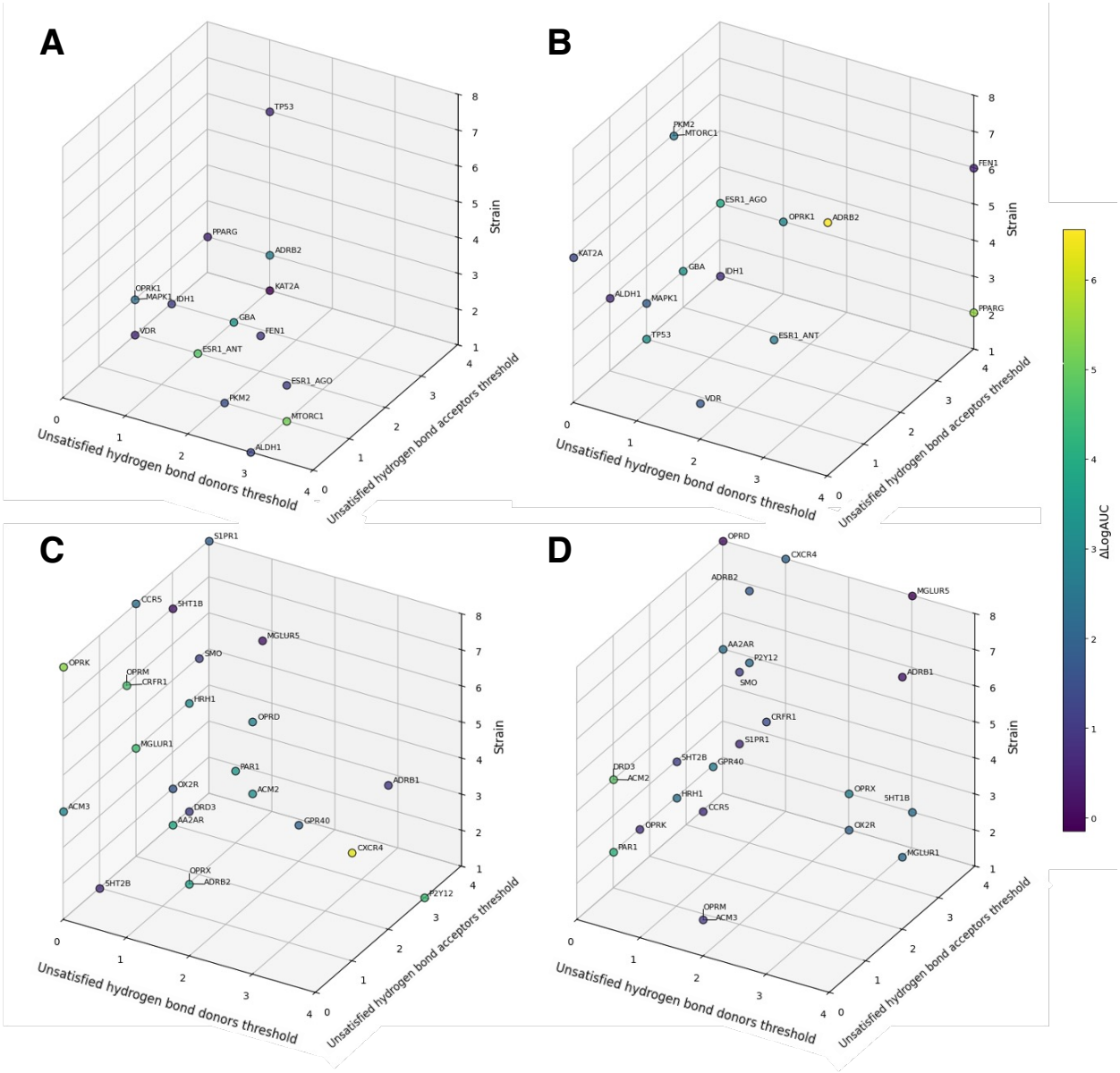


Figure 2.5. Best filter threshold combination for each protein-ligand dataset in selected retrospective databases. (A) LIT-PCBA database docked with AutoDock. (B) LIT-PCBA database docked with GNINA. (C) GPCR-Bench database docked with AutoDock. (D) GPCR-Bench database docked with GNINA. The x-axis is the unsatisfied hydrogen bond donors threshold; y-axis is unsatisfied hydrogen bond acceptors threshold; z-axis is the strain threshold. The heatmap indicates the ΔLogAUC for a protein-ligand dataset at specific threshold combination.

2.3 Conclusions

The results presented in this chapter strongly suggest that retrospective VS studies, where data is available, is a crucial step before a prospective campaign. While strain and unsatisfied hydrogen bond filters are broadly used in VS campaigns, we showed their effect on enrichment to be highly system-dependent. Therefore, the choice of these parameters should be carefully calibrated prior to prospective studies. There was no single threshold that led to an enrichment of all protein-ligand datasets explored, but several proteins showed significant enrichment when using strain, unsatisfied hydrogen bonds, or both as filters at specific thresholds. Thus, these filters can be effective post-processing tools in VS when retrospectively confirmed to benefit the selection of candidates for the protein target of interest.

2.4 References

(1) Penner, P.; Guba, W.; Schmidt, R.; Meyder, A.; Stahl, M.; Rarey, M. The Torsion Library: Semiautomated Improvement of Torsion Rules with SMARTScompare. *J. Chem. Inf. Model.* **2022**, *62* (7), 1644–1653. https://doi.org/10.1021/acs.jcim.2c00043.

(2) Gu, S.; Smith, M. S.; Yang, Y.; Irwin, J. J.; Shoichet, B. K. Ligand Strain Energy in Large Library Docking. *J. Chem. Inf. Model.* **2021**, *61* (9), 4331–4341. https://doi.org/10.1021/acs.jcim.1c00368.

(3) Jain, A. N.; Brueckner, A. C.; Cleves, A. E.; Reibarkh, M.; Sherer, E. C. A Distributional Model of Bound Ligand Conformational Strain: From Small Molecules up to Large Peptidic Macrocycles. *J. Med. Chem.* **2023**, *66* (3), 1955–1971. https://doi.org/10.1021/acs.jmedchem.2c01744.

(4) Peach, M. L.; Cachau, R. E.; Nicklaus, M. C. Conformational Energy Range of Ligands in Protein Crystal Structures: The Difficult Quest for Accurate Understanding. *J. Mol. Recognit.* **2017**, *30* (8). https://doi.org/10.1002/jmr.2618.

(5) Tong, J.; Zhao, S. Large-Scale Analysis of Bioactive Ligand Conformational Strain Energy by *Ab Initio* Calculation. *J. Chem. Inf. Model.* **2021**, *61* (3), 1180–1192. https://doi.org/10.1021/acs.jcim.0c01197.

(6) Rai, B. K.; Sresht, V.; Yang, Q.; Unwalla, R.; Tu, M.; Mathiowetz, A. M.; Bakken, G. A. Comprehensive Assessment of Torsional Strain in Crystal Structures of Small Molecules and Protein–Ligand Complexes Using Ab Initio Calculations. *J. Chem. Inf. Model.* **2019**, *59* (10), 4195–4208. https://doi.org/10.1021/acs.jcim.9b00373.

(7) Berenger, F.; Tsuda, K. An ANI -2 Enabled Open-source Protocol to Estimate Ligand Strain after Docking. *J. Comput. Chem.* **2025**, *46* (1), e27478. https://doi.org/10.1002/jcc.27478.

(8) Wallace, E. R. S.; Frey, N. C.; Rackers, J. A. Strain Problems Got You in a Twist? Try StrainRelief: A Quantum-Accurate Tool for Ligand Strain Calculations. *J. Chem. Inf. Model.* **2025**, *65* (13), 6613–6620. https://doi.org/10.1021/acs.jcim.5c00586.

(9) Rai, B. K.; Sresht, V.; Yang, Q.; Unwalla, R.; Tu, M.; Mathiowetz, A. M.; Bakken, G. A. TorsionNet: A Deep Neural Network to Rapidly Predict Small-Molecule Torsional Energy Profiles with the Accuracy of Quantum Mechanics. *J. Chem. Inf. Model.* **2022**, *62* (4), 785–800. https://doi.org/10.1021/acs.jcim.1c01346.

(10) Salentin, S.; Schreiber, S.; Haupt, V. J.; Adasme, M. F.; Schroeder, M. PLIP: Fully Automated Protein–Ligand Interaction Profiler. *Nucleic Acids Res.* **2015**, *43* (W1), W443–W447. https://doi.org/10.1093/nar/gkv315.

(11) Schake, P.; Bolz, S. N.; Linnemann, K.; Schroeder, M. PLIP 2025: Introducing Protein–Protein Interactions to the Protein–Ligand Interaction Profiler. *Nucleic Acids Res.* **2025**, *53* (W1), W463–W465. https://doi.org/10.1093/nar/gkaf361.

(12) Ha, E. J.; Lwin, C. T.; Durrant, J. D. LigGrep: A Tool for Filtering Docked Poses to Improve Virtual-Screening Hit Rates. *J. Cheminform.* **2020**, *12* (1), 69. https://doi.org/10.1186/s13321-020-00471-2.

(13) Marcou, G.; Rognan, D. Optimizing Fragment and Scaffold Docking by Use of Molecular Interaction Fingerprints. *J. Chem. Inf. Model.* **2007**, *47* (1), 195–207. https://doi.org/10.1021/ci600342e.

(14) Fischer, A.; Smieško, M.; Sellner, M.; Lill, M. A. Decision Making in Structure-Based Drug Discovery: Visual Inspection of Docking Results. *J. Med. Chem.* **2021**, *64* (5), 2489–2500. https://doi.org/10.1021/acs.jmedchem.0c02227.

(15) Sindt, F.; Seyller, A.; Eguida, M.; Rognan, D. Protein Structure-Based Organic Chemistry-Driven Ligand Design from Ultralarge Chemical Spaces. *ACS Cent. Sci.* **2024**, *10* (3), 615–627. https://doi.org/10.1021/acscentsci.3c01521.

(16) Zhou, G.; Rusnac, D.-V.; Park, H.; Canzani, D.; Nguyen, H. M.; Stewart, L.; Bush, M. F.; Nguyen, P. T.; Wulff, H.; Yarov-Yarovoy, V.; Zheng, N.; DiMaio, F. An Artificial Intelligence Accelerated Virtual Screening Platform for Drug Discovery. *Nat. Commun.* **2024**, *15* (1), 7761. https://doi.org/10.1038/s41467-024-52061-7.

(17) Liu, F.; Mailhot, O.; Glenn, I. S.; Vigneron, S. F.; Bassim, V.; Xu, X.; Fonseca-Valencia, K.; Smith, M. S.; Radchenko, D. S.; Fraser, J. S.; Moroz, Y. S.; Irwin, J. J.; Shoichet, B. K. The Impact of Library Size and Scale of Testing on Virtual Screening. *Nat. Chem. Biol*. **2025**, *21*, 1039-1045. https://doi.org/10.1038/s41589-024-01797-w.

(18) Liu, F.; Wu, C.-G.; Tu, C.-L.; Glenn, I.; Meyerowitz, J.; Kaplan, A. L.; Lyu, J.; Cheng, Z.; Tarkhanova, O. O.; Moroz, Y. S.; Irwin, J. J.; Chang, W.; Shoichet, B. K.; Skiniotis, G. Large Library Docking Identifies Positive Allosteric Modulators of the Calcium-Sensing Receptor. *Science* **2024**, *385* (6715), eado1868. https://doi.org/10.1126/science.ado1868.

(19) Lyu, J.; Kapolka, N.; Gumpper, R.; Alon, A.; Wang, L.; Jain, M. K.; Barros-Álvarez, X.; Sakamoto, K.; Kim, Y.; DiBerto, J.; Kim, K.; Glenn, I. S.; Tummino, T. A.; Huang, S.; Irwin, J. J.; Tarkhanova, O. O.; Moroz, Y.; Skiniotis, G.; Kruse, A. C.; Shoichet, B. K.; Roth, B. L. AlphaFold2 Structures Guide Prospective Ligand Discovery. *Science* **2024**, *384* (6702), eadn6354. https://doi.org/10.1126/science.adn6354.

(20) Fink, E. A.; Xu, J.; Hübner, H.; Braz, J. M.; Seemann, P.; Avet, C.; Craik, V.; Weikert, D.; Schmidt, M. F.; Webb, C. M.; Tolmachova, N. A.; Moroz, Y. S.; Huang, X.-P.; Kalyanaraman, C.; Gahbauer, S.; Chen, G.; Liu, Z.; Jacobson, M. P.; Irwin, J. J.; Bouvier, M.; Du, Y.; Shoichet, B. K.; Basbaum, A. I.; Gmeiner, P. Structure-Based Discovery of Nonopioid Analgesics Acting through the α_{2A} -Adrenergic Receptor. *Science* **2022**, *377* (6614), eabn7065. https://doi.org/10.1126/science.abn7065.

(21) Chakrabarti, M.; Tan, Y. S.; Balius, T. E. Considerations Around Structure-Based Drug Discovery for KRAS Using DOCK. In *KRAS*; Stephen, A. G., Esposito, D., Eds.; Methods in Molecular Biology; Springer US: New York, NY, 2024; Vol. 2797, pp 67–90. https://doi.org/10.1007/978-1-0716-3822-4_6.

(22) Alon, A.; Lyu, J.; Braz, J. M.; Tummino, T. A.; Craik, V.; O'Meara, M. J.; Webb, C. M.; Radchenko, D. S.; Moroz, Y. S.; Huang, X.-P.; Liu, Y.; Roth, B. L.; Irwin, J. J.; Basbaum, A. I.; Shoichet, B. K.; Kruse, A. C. Structures of the Σ2 Receptor Enable Docking for Bioactive Ligand Discovery. *Nature* **2021**, *600* (7890), 759–764. https://doi.org/10.1038/s41586-021-04175-x.

(23) Sindt, F.; Bret, G.; Rognan, D. On the Difficulty to Rescore Hits from Ultralarge Docking Screens. *J. Chem. Inf. Model.* **2025**, *65* (11), 5553–5566. https://doi.org/10.1021/acs.jcim.5c00730.

(24) Singh, I.; Li, F.; Fink, E. A.; Chau, I.; Li, A.; Rodriguez-Hernández, A.; Glenn, I.; Zapatero-Belinchón, F. J.; Rodriguez, M. L.; Devkota, K.; Deng, Z.; White, K.; Wan, X.; Tolmachova, N. A.; Moroz, Y. S.; Kaniskan, H. Ü.; Ott, M.; García-Sastre, A.; Jin, J.; Fujimori, D. G.; Irwin, J. J.; Vedadi, M.; Shoichet, B. K. Structure-Based Discovery of Inhibitors of the SARS-CoV-2 Nsp14 N7-Methyltransferase. *J. Med. Chem.* **2023**, *66* (12), 7785–7803. https://doi.org/10.1021/acs.jmedchem.2c02120.

(25) Singh, I.; Seth, A.; Billesbølle, C. B.; Braz, J.; Rodriguiz, R. M.; Roy, K.; Bekele, B.; Craik, V.; Huang, X.-P.; Boytsov, D.; Pogorelov, V. M.; Lak, P.; O'Donnell, H.; Sandtner, W.; Irwin, J. J.; Roth, B. L.; Basbaum, A. I.; Wetsel, W. C.; Manglik, A.; Shoichet, B. K.; Rudnick, G. Structure-Based Discovery of Conformationally Selective Inhibitors of the Serotonin Transporter. *Cell* **2023**, *186* (10), 2160-2175.e17. https://doi.org/10.1016/j.cell.2023.04.010.

(26) Bender, B. J.; Gahbauer, S.; Luttens, A.; Lyu, J.; Webb, C. M.; Stein, R. M.; Fink, E. A.; Balius, T. E.; Carlsson, J.; Irwin, J. J.; Shoichet, B. K. A Practical Guide to Large-Scale Docking. *Nat. Protoc.* **2021**, *16* (10), 4799–4832. https://doi.org/10.1038/s41596-021-00597-z.

(27) Sriram, K.; Insel, P. A. G Protein-Coupled Receptors as Targets for Approved Drugs: How Many Targets and How Many Drugs? *Mol. Pharmacol.* **2018**, *93* (4), 251–258. https://doi.org/10.1124/mol.117.111062.

(28) Weiss, D. R.; Bortolato, A.; Tehan, B.; Mason, J. S. GPCR-Bench: A Benchmarking Set and Practitioners' Guide for G Protein-Coupled Receptor Docking. *J. Chem. Inf. Model.* **2016**, *56* (4), 642–651. https://doi.org/10.1021/acs.jcim.5b00660.

(29) Tran-Nguyen, V.-K.; Jacquemard, C.; Rognan, D. LIT-PCBA: An Unbiased Data Set for Machine Learning and Virtual Screening. *J. Chem. Inf. Model.* **2020**, *60* (9), 4263–4273. https://doi.org/10.1021/acs.jcim.0c00155.

(30) Wang, J.-C.; Lin, J.-H.; Chen, C.-M.; Perryman, A. L.; Olson, A. J. Robust Scoring Functions for Protein-Ligand Interactions with Quantum Chemical Charge Models. *J. Chem. Inf. Model.* **2011**, *51* (10), 2528–2537. https://doi.org/10.1021/ci200220v.

(31) Domínguez-Ramírez, L.; Anaya-Ruiz, M.; Cortés-Hernández, P. Quality over Quantity: How to Get the Best Results When Using Docking for Repurposing. *Front. Bioinform.* **2025**, *5*, 1536504. https://doi.org/10.3389/fbinf.2025.1536504.

(32) Buccheri, R.; Rescifina, A. High-Throughput, High-Quality: Benchmarking GNINA and AutoDock Vina for Precision Virtual Screening Workflow. *Molecules* **2025**, *30* (16), 3361. https://doi.org/10.3390/molecules30163361.

(33) Dunn, I.; Pirhadi, S.; Wang, Y.; Ravindran, S.; Concepcion, C.; Koes, D. R. CACHE Challenge #1: Docking with GNINA Is All You Need. *J. Chem. Inf. Model.* **2024**, *64* (24), 9388–9396. https://doi.org/10.1021/acs.jcim.4c01429.

(34) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30* (16), 2785–2791. https://doi.org/10.1002/jcc.21256.

(35) Santos-Martins, D.; Solis-Vasquez, L.; Tillack, A. F.; Sanner, M. F.; Koch, A.; Forli, S. Accelerating AutoDock4 with GPUs and Gradient-Based Local Search. *J. Chem. Theory. Comput.* **2021**, *17* (2), 1060–1073. https://doi.org/10.1021/acs.jctc.0c01006.

(36) Sarkar, A.; Concilio, S.; Sessa, L.; Marrafino, F.; Piotto, S. Advancements and Novel Approaches in Modified AutoDock Vina Algorithms for Enhanced Molecular Docking. *Results Chem.* **2024**, *7*, 101319. https://doi.org/10.1016/j.rechem.2024.101319.

(37) Hill, A. D.; Reilly, P. J. Scoring Functions for AutoDock. In *Glycoinformatics*; Lütteke, T., Frank, M., Eds.; Methods in Molecular Biology; Springer New York: New York, NY, 2015; Vol. 1273, pp 467–474. https://doi.org/10.1007/978-1-4939-2343-4_27.

(38) Cosconati, S.; Forli, S.; Perryman, A. L.; Harris, R.; Goodsell, D. S.; Olson, A. J. Virtual Screening with AutoDock: Theory and Practice. *Expert Opin. Drug Discov.* **2010**, *5* (6), 597–607. https://doi.org/10.1517/17460441.2010.484460.

(39) Forli, S.; Botta, M. Lennard-Jones Potential and Dummy Atom Settings to Overcome the AUTODOCK Limitation in Treating Flexible Ring Systems. *J. Chem. Inf. Model.* **2007**, *47* (4), 1481–1492. https://doi.org/10.1021/ci700036j.

(40) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *J. Cheminform.* **2011**, *3* (1), 33. https://doi.org/10.1186/1758-2946-3-33.

(41) McNutt, A. T.; Francoeur, P.; Aggarwal, R.; Masuda, T.; Meli, R.; Ragoza, M.; Sunseri, J.; Koes, D. R. GNINA 1.0: Molecular Docking with Deep Learning. *J. Cheminform.* **2021**, *13* (1), 43. https://doi.org/10.1186/s13321-021-00522-2.

(42) Koes, D. R.; Baumgartner, M. P.; Camacho, C. J. Lessons Learned in Empirical Scoring with Smina from the CSAR 2011 Benchmarking Exercise. *J. Chem. Inf. Model.* **2013**, *53* (8), 1893–1904. https://doi.org/10.1021/ci300604z.

(43) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455–461. https://doi.org/10.1002/jcc.21334.

(44) Molecular Operating Environment, v. 2010, The Chemical Computing Group, Montreal, QC, Canada.

(45) FLIPPER, Ver. 2025.2.1; OMEGA Application Ver. 6.1.1.1; OpenEye, Cadence Molecular Sciences, Santa Fe, NM, USA, 2025. http://www.eyesopen.com.

(46) QUACPAC Application, Ver. 2.2.7.1; OpenEye, Cadence Molecular Sciences, Santa Fe, NM, USA, 2025. http://www.eyesopen.com.

(47) Gentile, F.; Yaacoub, J. C.; Gleave, J.; Fernandez, M.; Ton, A.-T.; Ban, F.; Stern, A.; Cherkasov, A. Artificial Intelligence–Enabled Virtual Screening of Ultra-Large Chemical Libraries with Deep Docking. *Nat. Protoc.* **2022**, *17* (3), 672–697. https://doi.org/10.1038/s41596-021-00659-2.

(48) Da Silva, F.; Desaphy, J.; Rognan, D. IChem: A Versatile Toolkit for Detecting, Comparing, and Predicting Protein–Ligand Interactions. *ChemMedChem* **2018**, *13* (6), 507–510. https://doi.org/10.1002/cmdc.201700505.

(49) Truchon, J.-F.; Bayly, C. I. Evaluating Virtual Screening Methods: Good and Bad Metrics for the “Early Recognition” Problem. *J. Chem. Inf. Model.* **2007**, *47* (2), 488–508. https://doi.org/10.1021/ci600426e.

(50) Tian, T.; Li, S.; Zhang, Z.; Chen, L.; Zou, Z.; Zhao, D.; Zeng, J. Benchmarking Compound Activity Prediction for Real-World Drug Discovery Applications. *Commun. Chem.* **2024**, *7* (1), 127. https://doi.org/10.1038/s42004-024-01204-4.

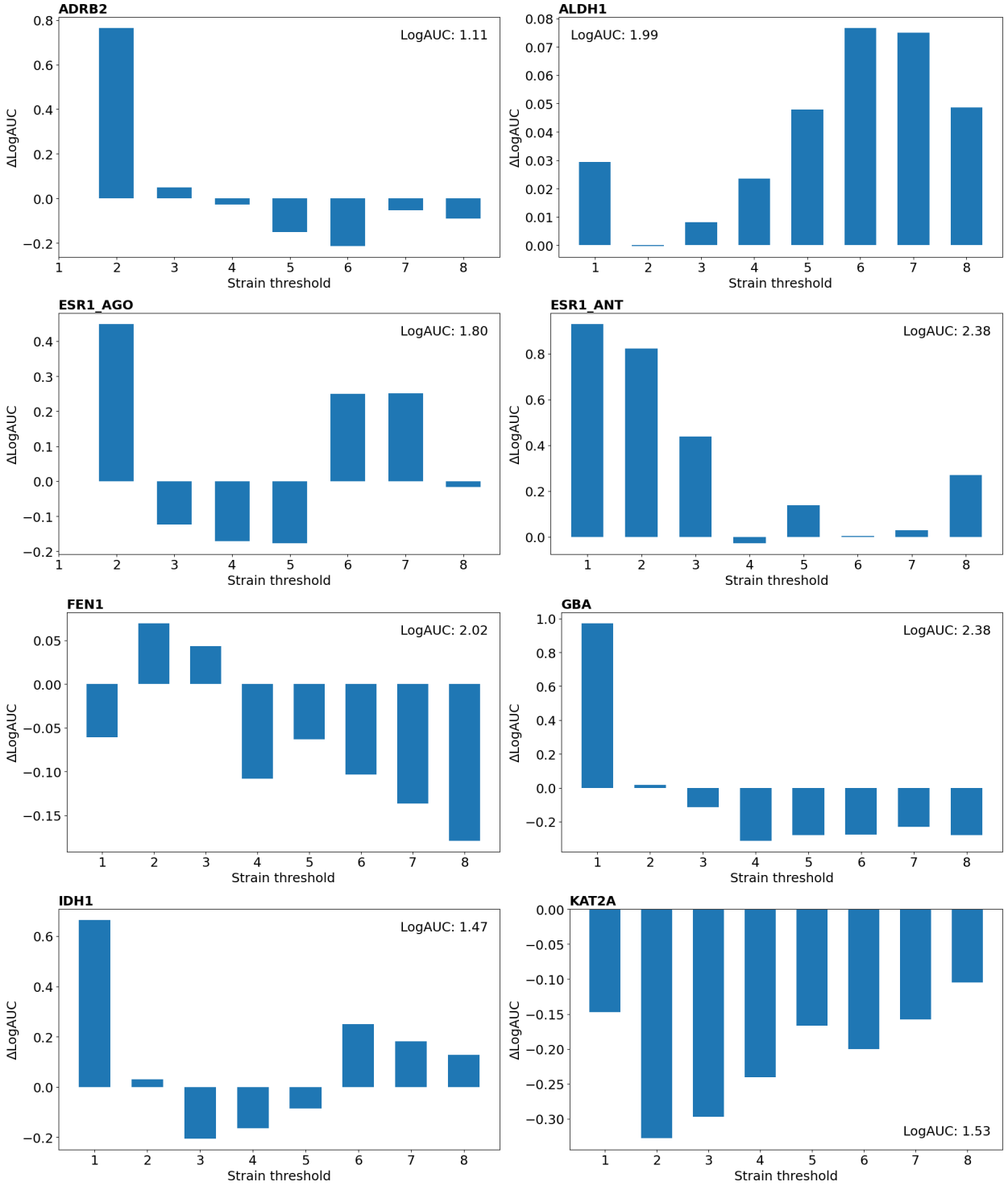
(51) Knight, I. S.; Naprienko, S.; Irwin, J. J. Enrichment Score: A Better Quantitative Metric for Evaluating the Enrichment Capacity of Molecular Docking Models. *arXiv*. 2022. https://doi.org/10.48550/ARXIV.2210.10905.

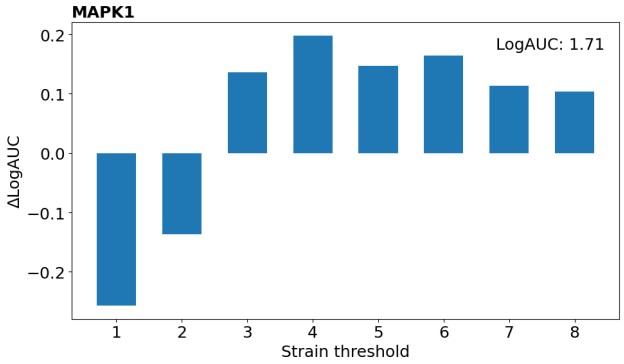
(52) Stein, R. M.; Yang, Y.; Balius, T. E.; O’Meara, M. J.; Lyu, J.; Young, J.; Tang, K.; Shoichet, B. K.; Irwin, J. J. Property-Unmatched Decoys in Docking Benchmarks. *J. Chem. Inf. Model.* **2021**, *61* (2), 699–714. https://doi.org/10.1021/acs.jcim.0c00598.

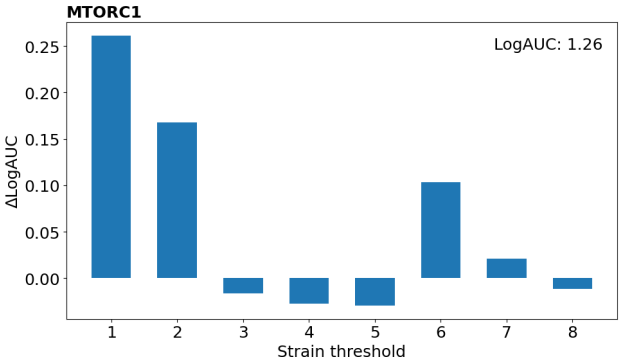
(53) Giordanetto, F.; Tyrchan, C.; Ulander, J. Intramolecular Hydrogen Bond Expectations in Medicinal Chemistry. *ACS Med. Chem. Lett.* **2017**, *8* (2), 139–142. https://doi.org/10.1021/acsmedchemlett.7b00002.

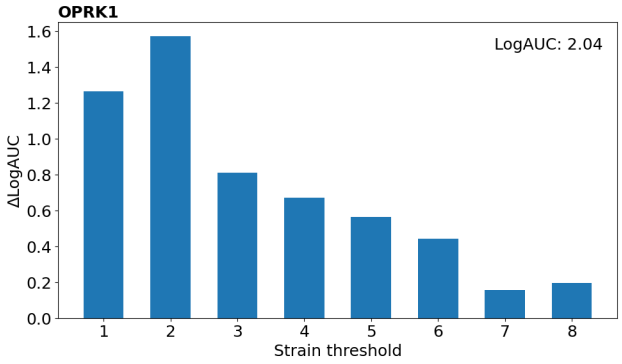
2.5 Appendix

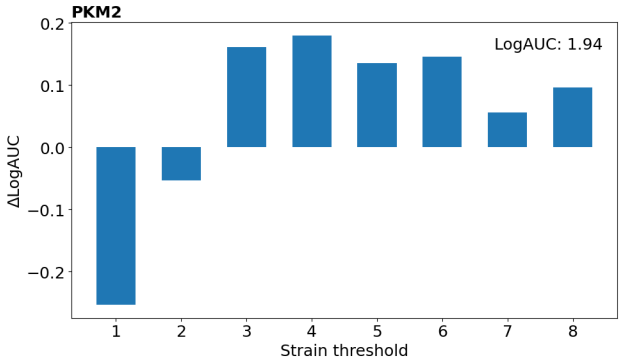
Supplementary figure series 2.5.1. Retrospective virtual screening results: benchmarking enrichment with strain filtering on the LIT-PCBA database docked with AutoDock.

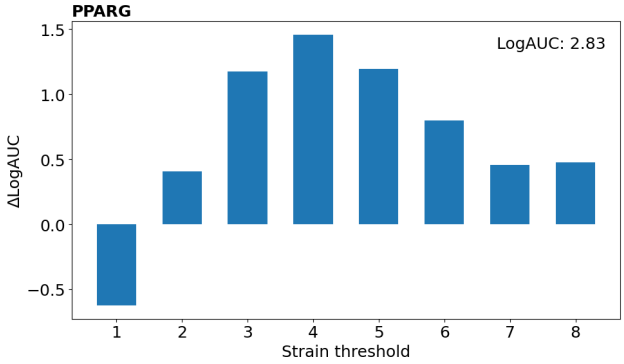


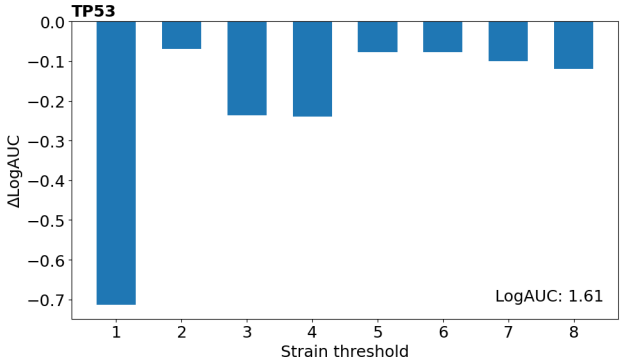

MAPK1
LogAUC: 1.71
ΔLogAUC
Strain threshold

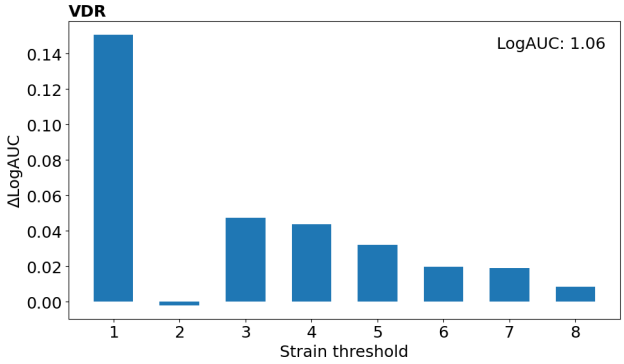

MTORC1
LogAUC: 1.26
ΔLogAUC
Strain threshold


OPRK1
LogAUC: 2.04
ΔLogAUC
Strain threshold

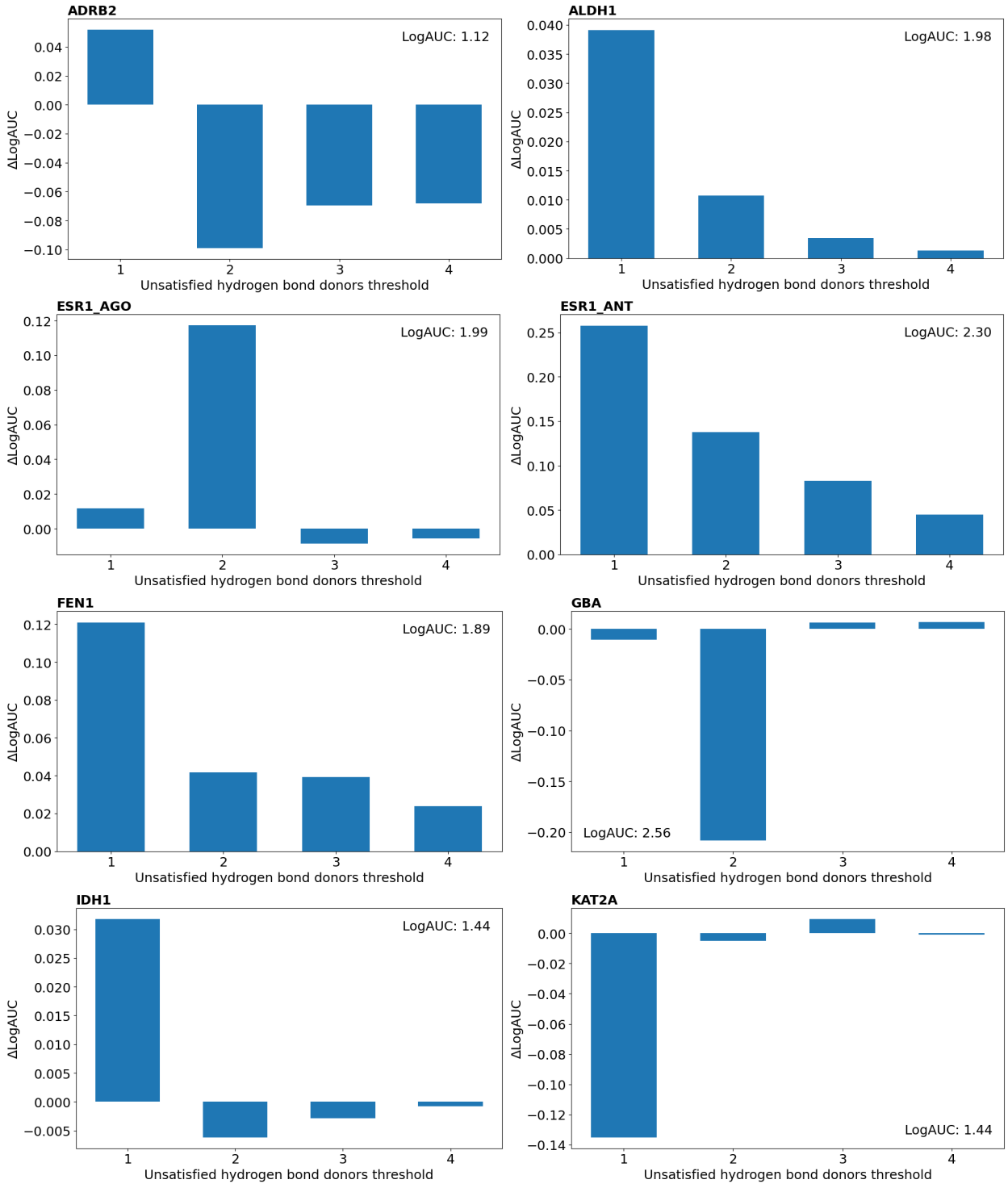

PKM2
LogAUC: 1.94
ΔLogAUC
Strain threshold

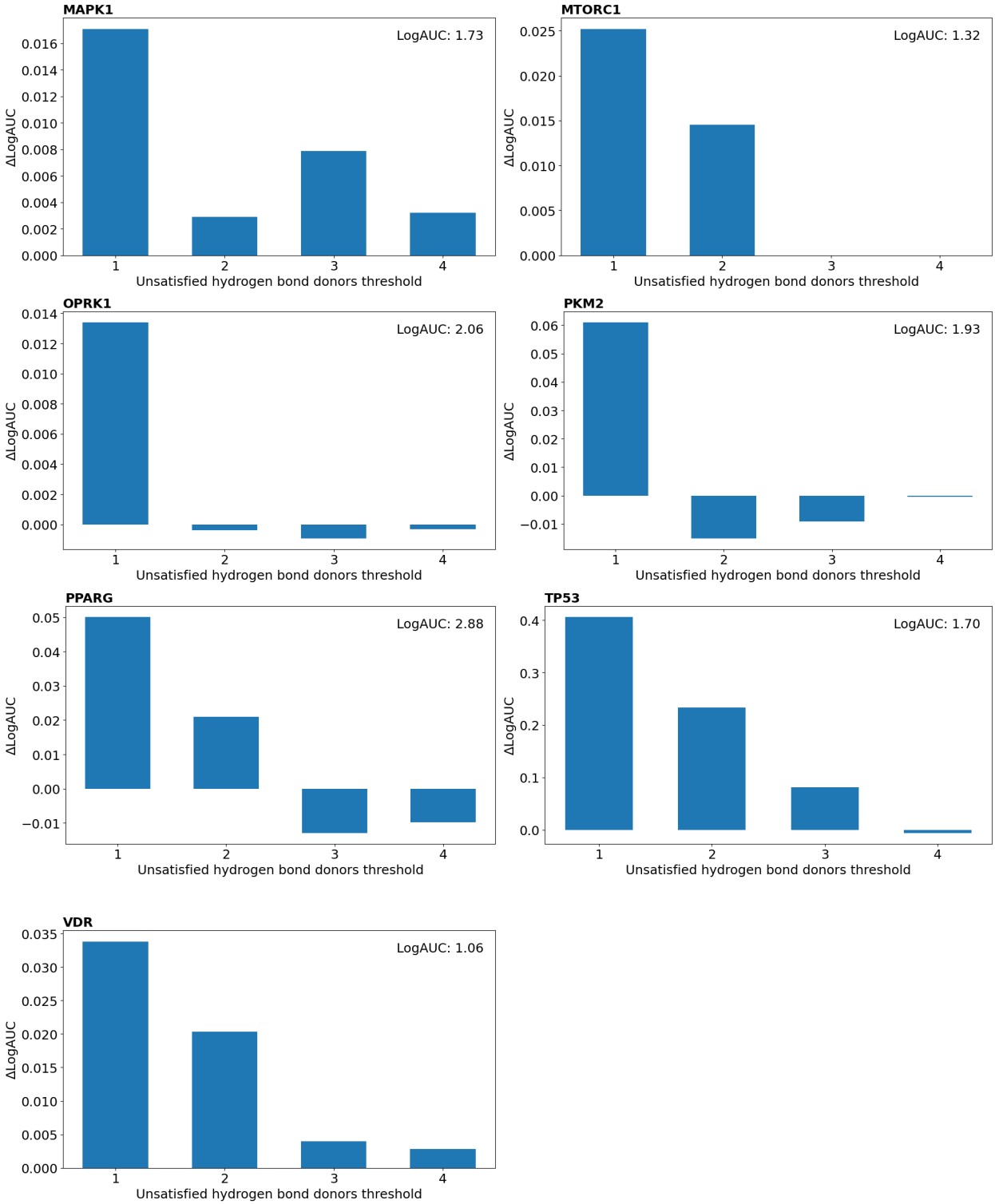

PPARG
LogAUC: 2.83
ΔLogAUC
Strain threshold


TP53
LogAUC: 1.61
ΔLogAUC
Strain threshold

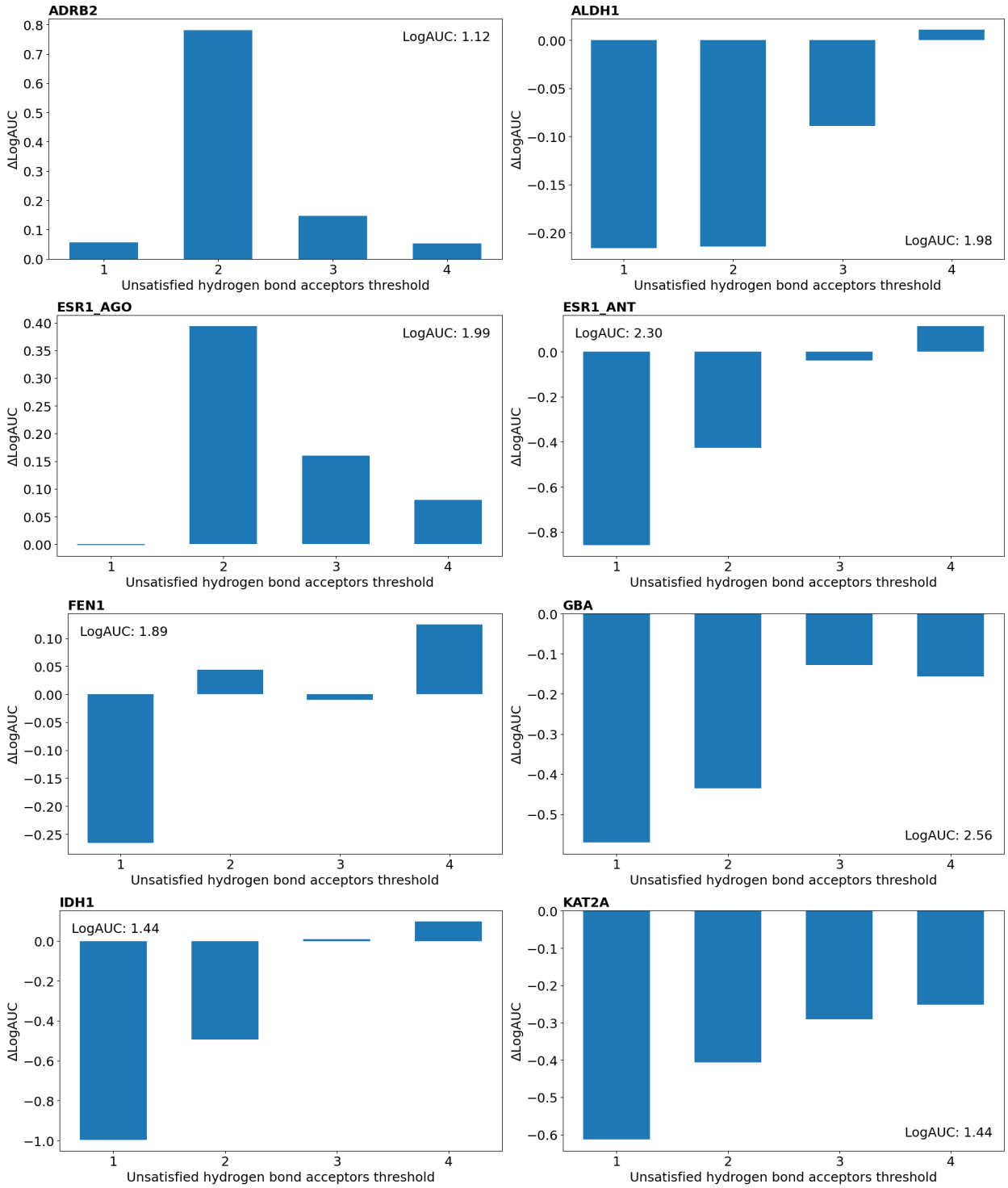

VDR
LogAUC: 1.06
ΔLogAUC
Strain threshold

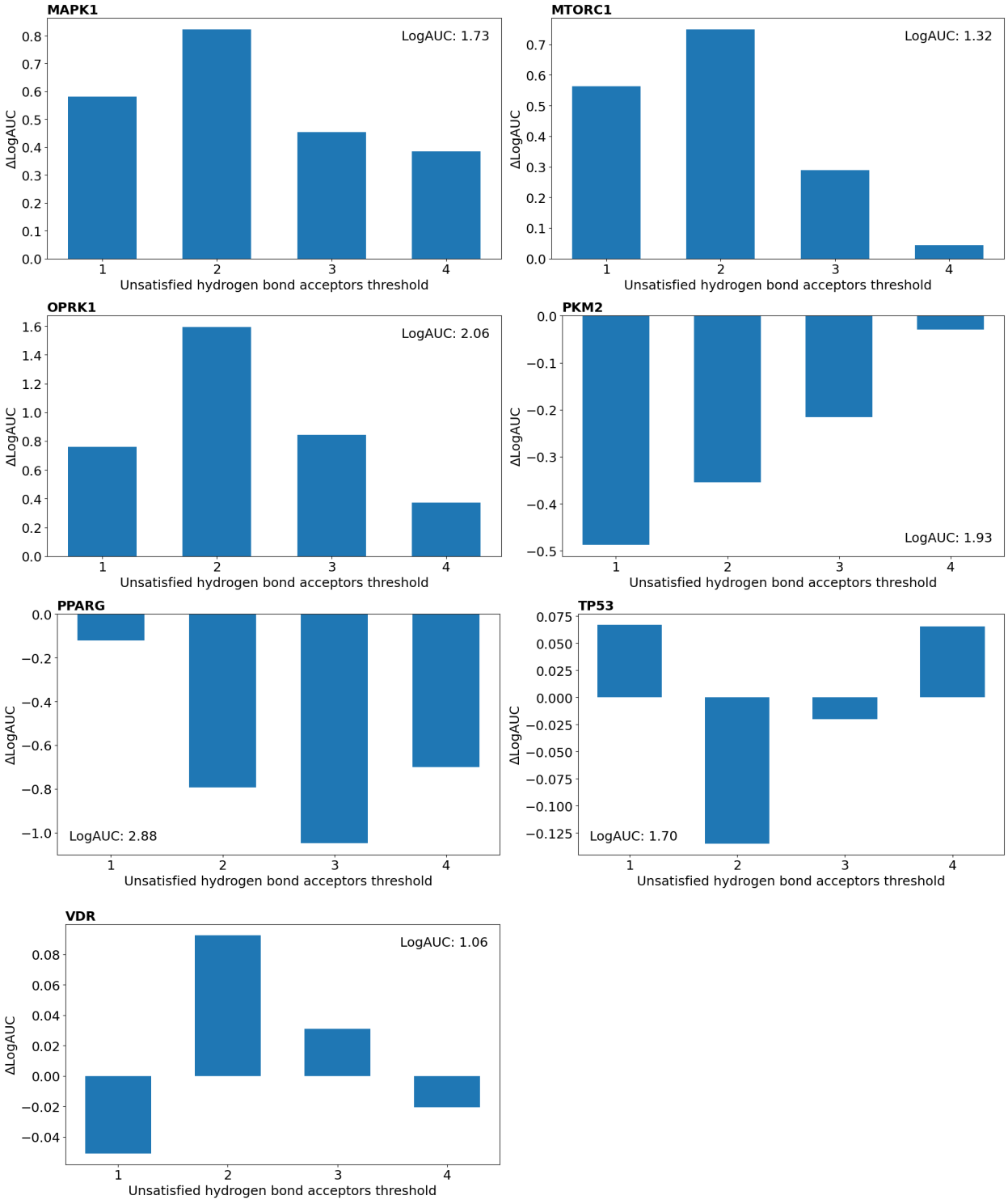
Supplementary figure series 2.5.2. Retrospective virtual screening results: benchmarking enrichment with unsatisfied hydrogen bond donors filtering on the LIT-PCBA database docked with AutoDock.



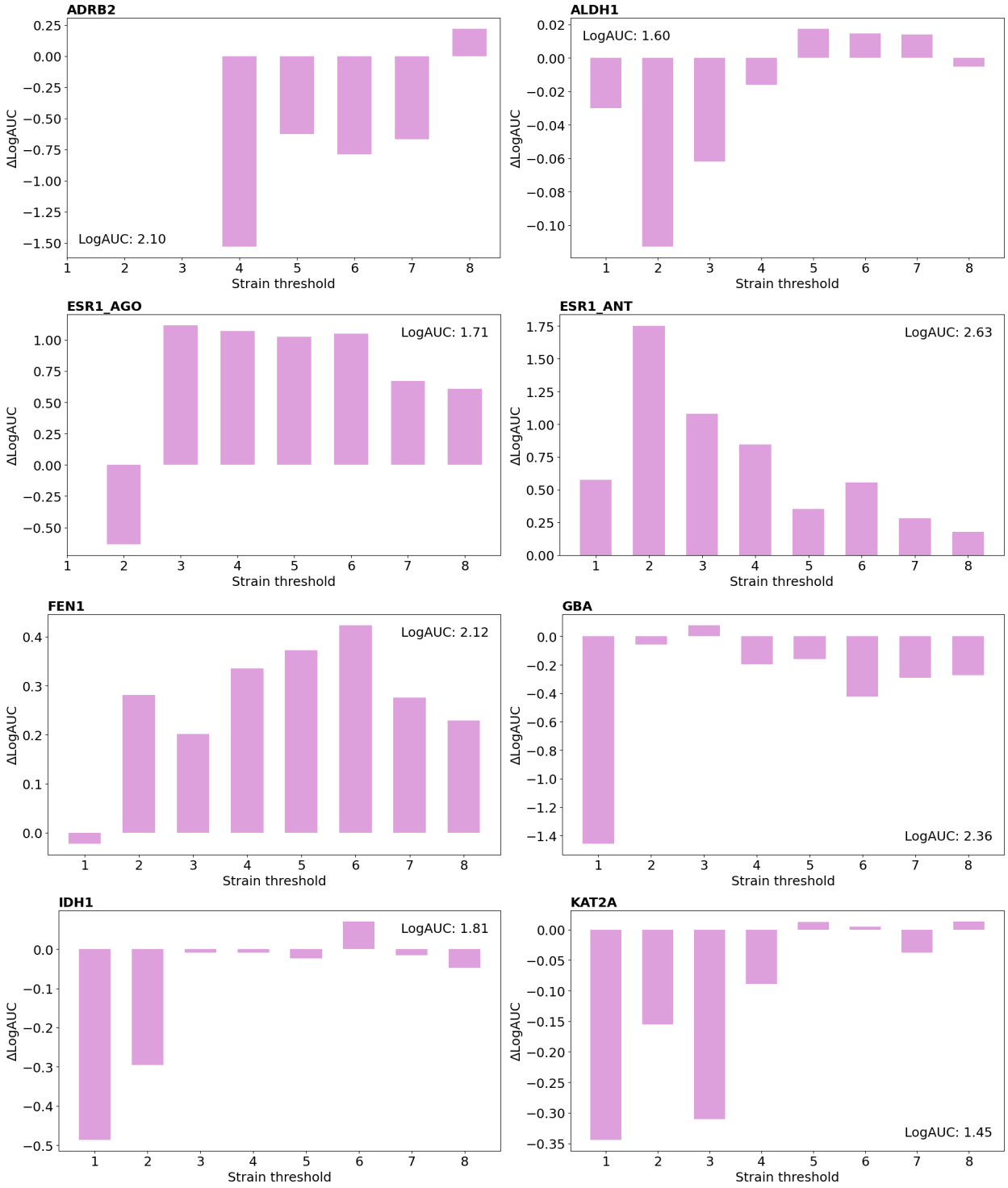

MAPK1
LogAUC: 1.73
ΔLogAUC
Unsatisfied hydrogen bond donors threshold
MTORC1
LogAUC: 1.32
ΔLogAUC
Unsatisfied hydrogen bond donors threshold
OPRK1
LogAUC: 2.06
ΔLogAUC
Unsatisfied hydrogen bond donors threshold
PKM2
LogAUC: 1.93
ΔLogAUC
Unsatisfied hydrogen bond donors threshold
PPARG
LogAUC: 2.88
ΔLogAUC
Unsatisfied hydrogen bond donors threshold
TP53
LogAUC: 1.70
ΔLogAUC
Unsatisfied hydrogen bond donors threshold
VDR
LogAUC: 1.06
ΔLogAUC
Unsatisfied hydrogen bond donors threshold

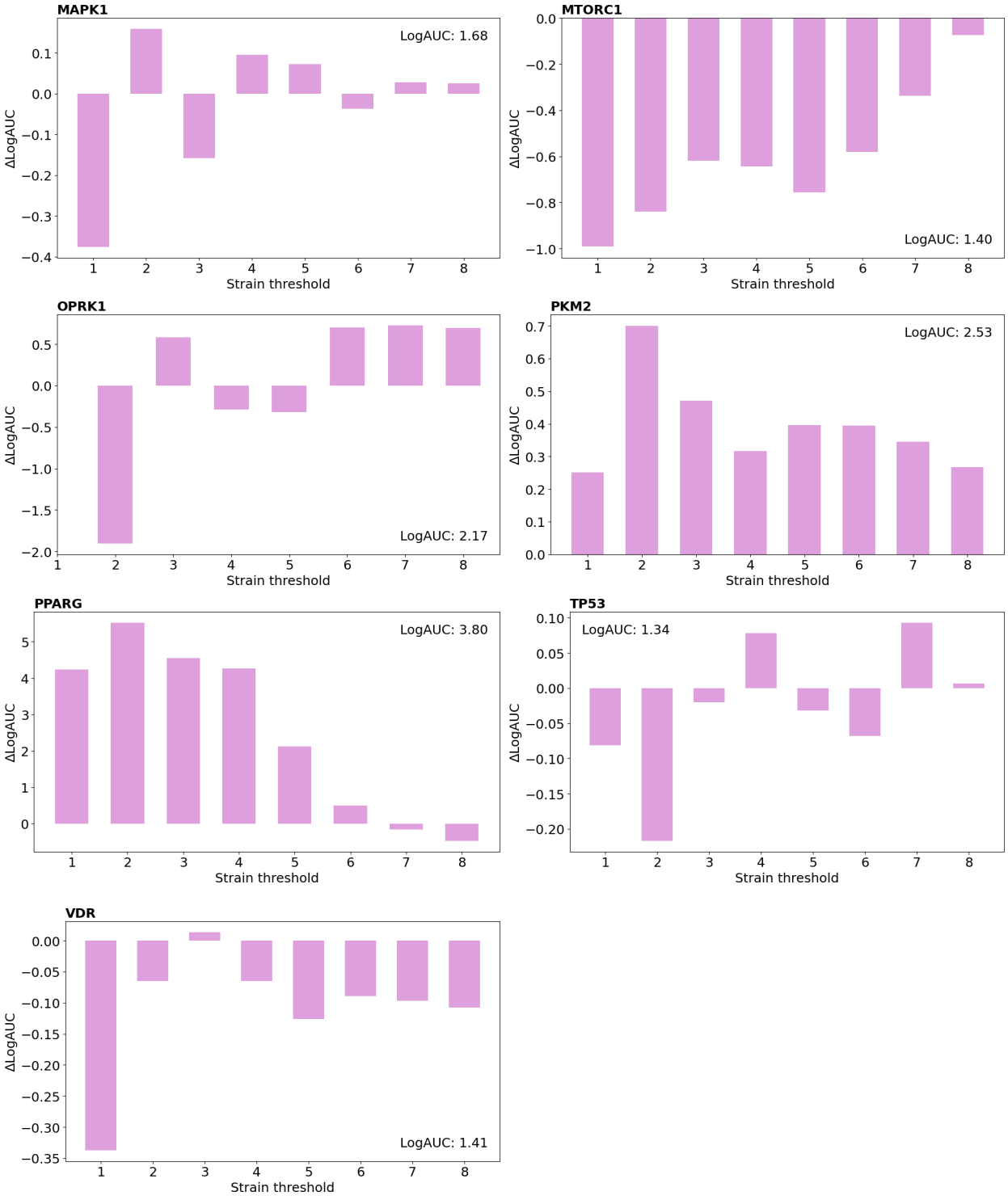
Supplementary figure series 2.5.3. Retrospective virtual screening results: benchmarking enrichment with unsatisfied hydrogen bond acceptors filtering on the LIT-PCBA database docked with AutoDock.



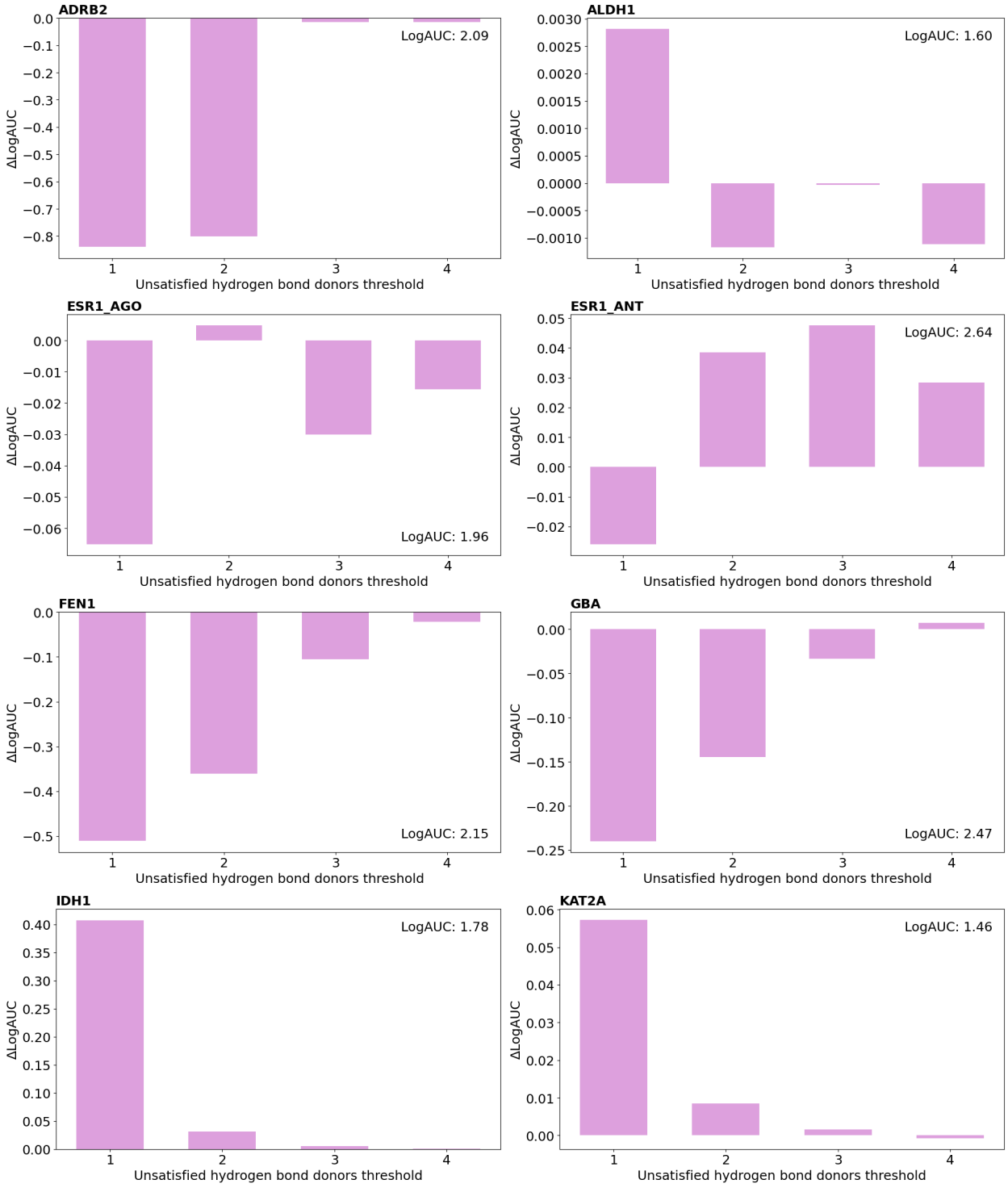

MAPK1
LogAUC: 1.73
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold
MTORC1
LogAUC: 1.32
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold
OPRK1
LogAUC: 2.06
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold
PKM2
LogAUC: 1.93
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold
PPARG
LogAUC: 2.88
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold
TP53
LogAUC: 1.70
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold
VDR
LogAUC: 1.06
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold

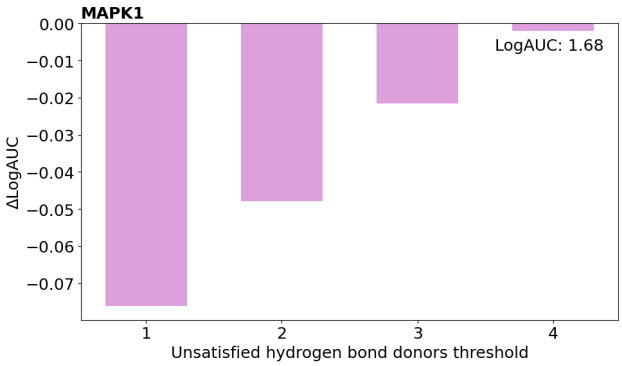
Supplementary figure series 2.5.4. Retrospective virtual screening results: benchmarking enrichment with strain filtering on the LIT-PCBA database docked with GNINA.

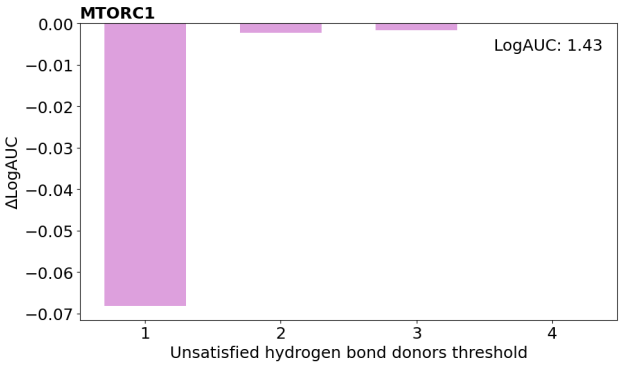


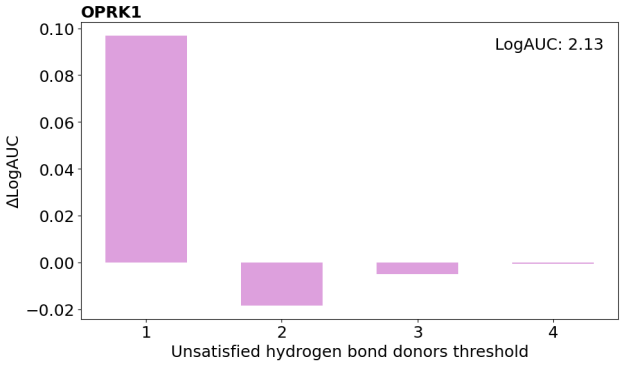

MAPK1
LogAUC: 1.68
MTORC1
LogAUC: 1.40
OPRK1
LogAUC: 2.17
PKM2
LogAUC: 2.53
PPARG
LogAUC: 3.80
TP53
LogAUC: 1.34
VDR
LogAUC: 1.41
ΔLogAUC
Strain threshold

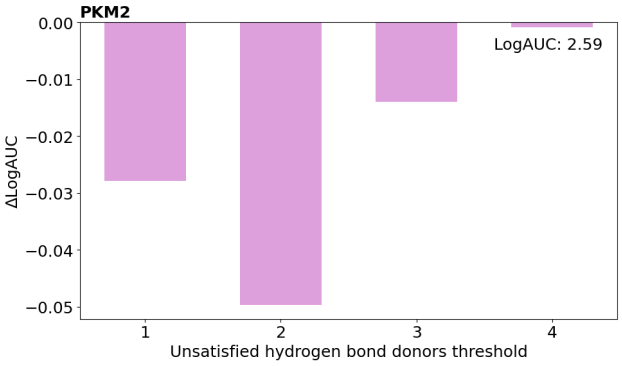
Supplementary figure series 2.5.5. Retrospective virtual screening results: benchmarking enrichment with unsatisfied hydrogen bond donors filtering on the LIT-PCBA database docked with GNINA.

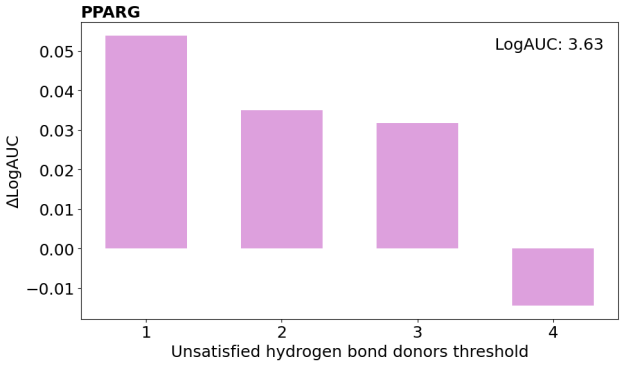


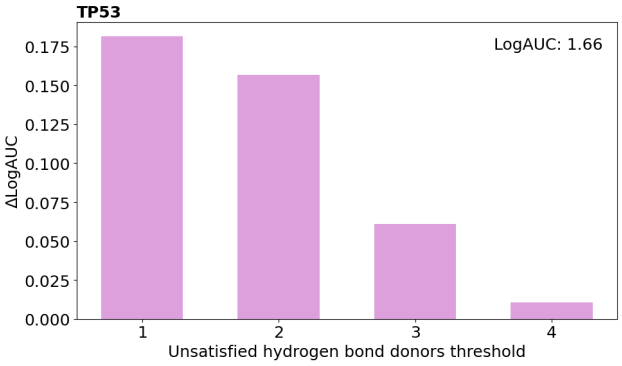

MAPK1
LogAUC: 1.68
ΔLogAUC
Unsatisfied hydrogen bond donors threshold

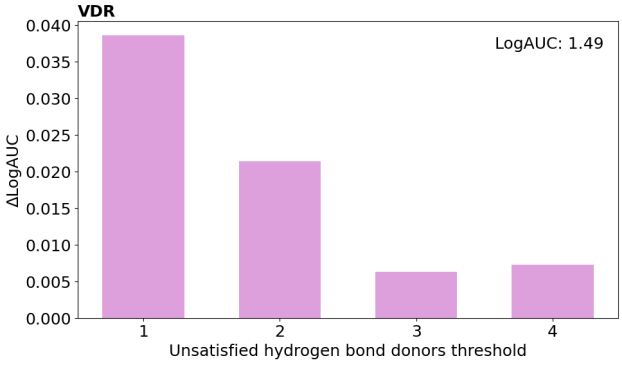

MTORC1
LogAUC: 1.43
ΔLogAUC
Unsatisfied hydrogen bond donors threshold


OPRK1
LogAUC: 2.13
ΔLogAUC
Unsatisfied hydrogen bond donors threshold

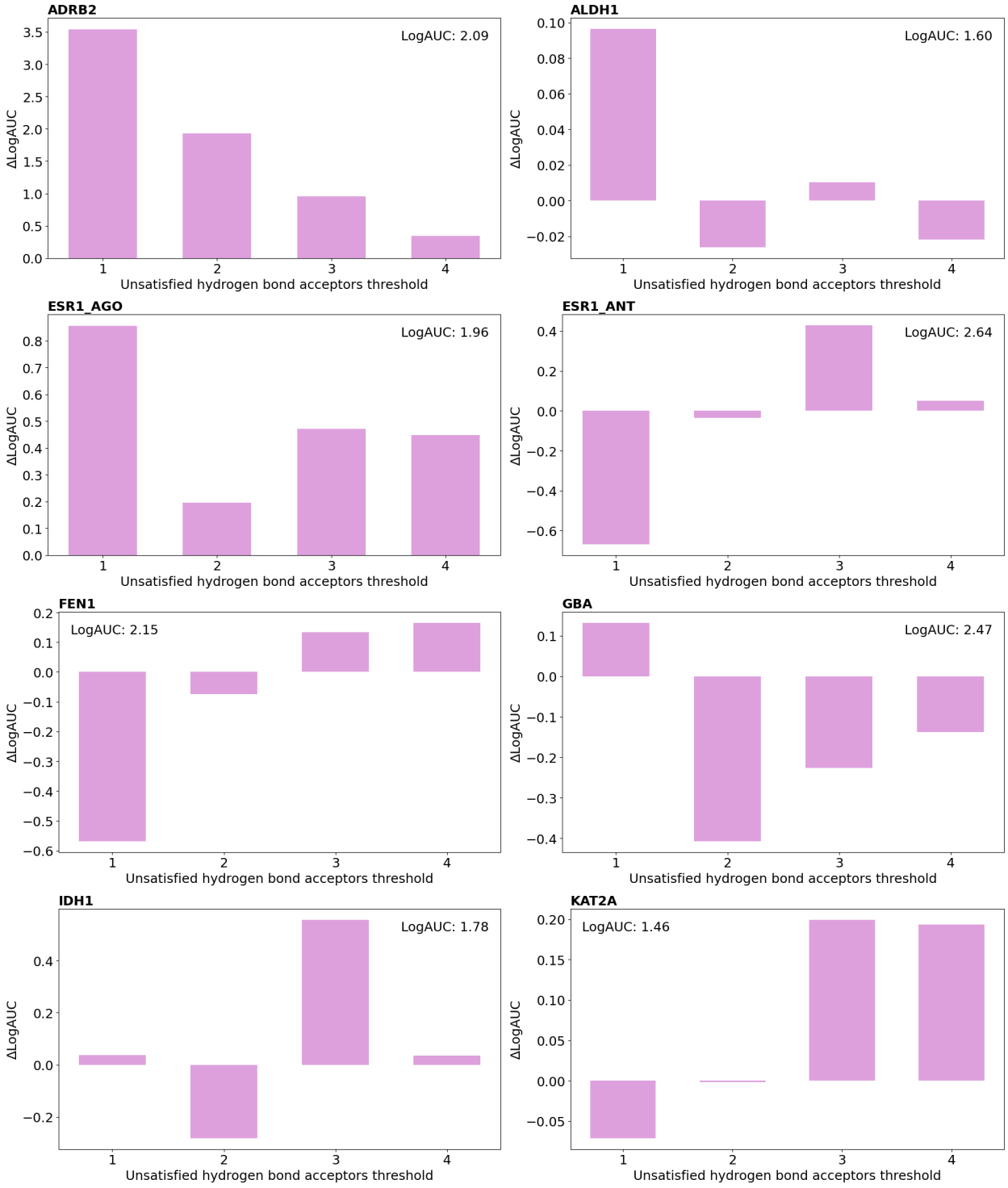

PKM2
LogAUC: 2.59
ΔLogAUC
Unsatisfied hydrogen bond donors threshold

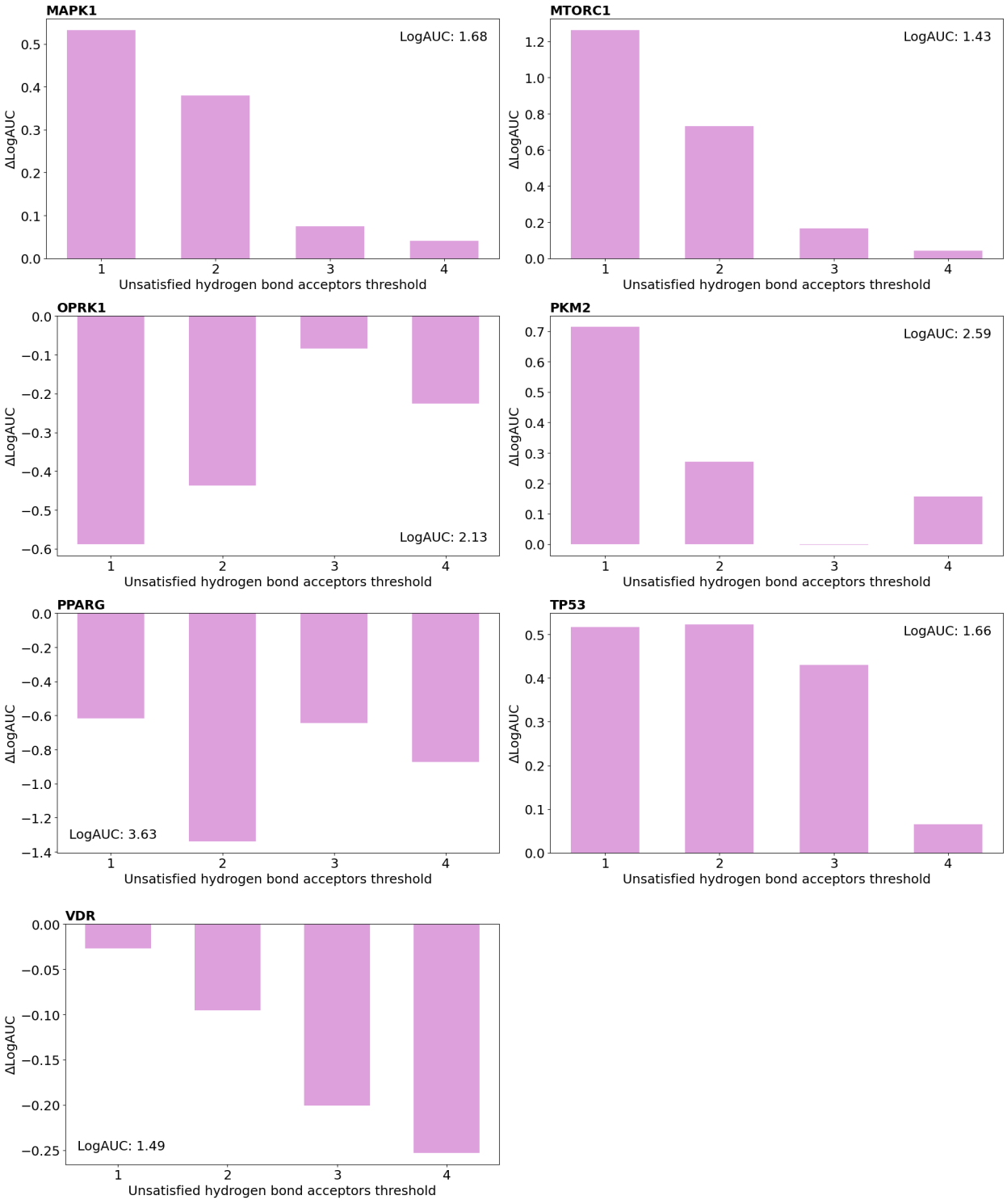

PPARG
LogAUC: 3.63
ΔLogAUC
Unsatisfied hydrogen bond donors threshold


TP53
LogAUC: 1.66
ΔLogAUC
Unsatisfied hydrogen bond donors threshold

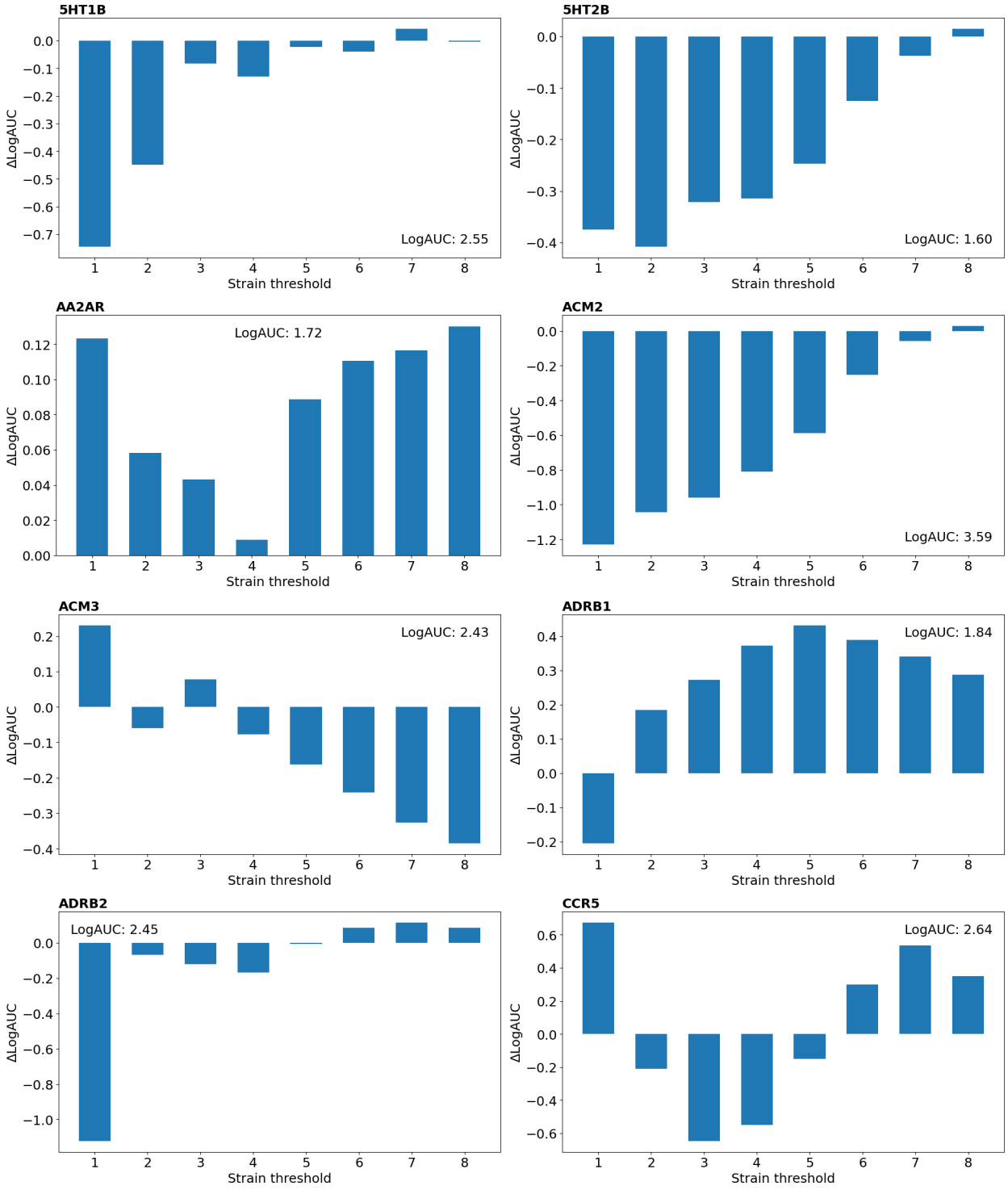

VDR
LogAUC: 1.49
ΔLogAUC
Unsatisfied hydrogen bond donors threshold

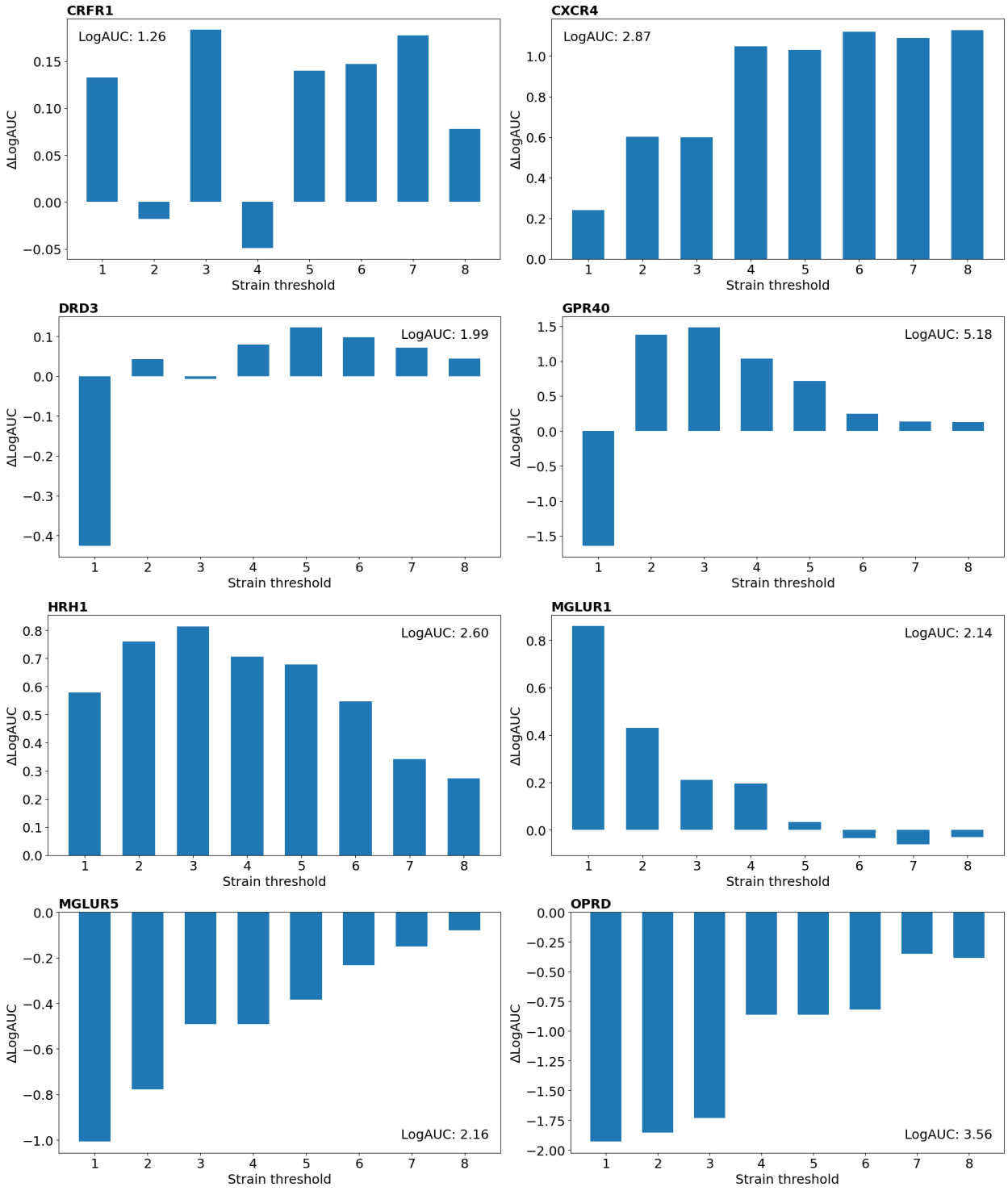
Supplementary figure series 2.5.6. Retrospective virtual screening results: benchmarking enrichment with unsatisfied hydrogen bond acceptors filtering on the LIT-PCBA database docked with GNINA.

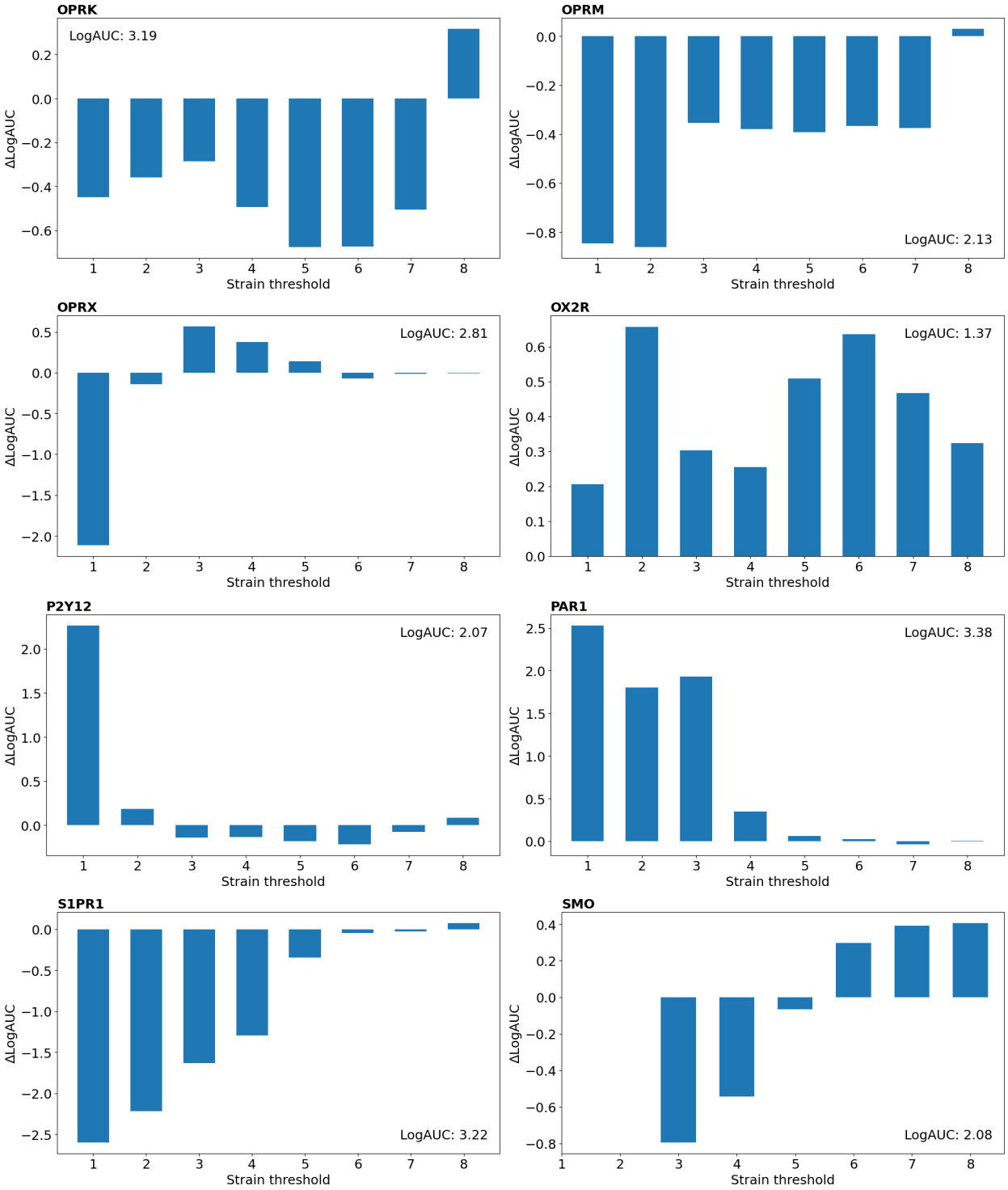



MAPK1
LogAUC: 1.68
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold
MTORC1
LogAUC: 1.43
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold
OPRK1
LogAUC: 2.13
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold
PKM2
LogAUC: 2.59
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold
PPARG
LogAUC: 3.63
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold
TP53
LogAUC: 1.66
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold
VDR
LogAUC: 1.49
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold

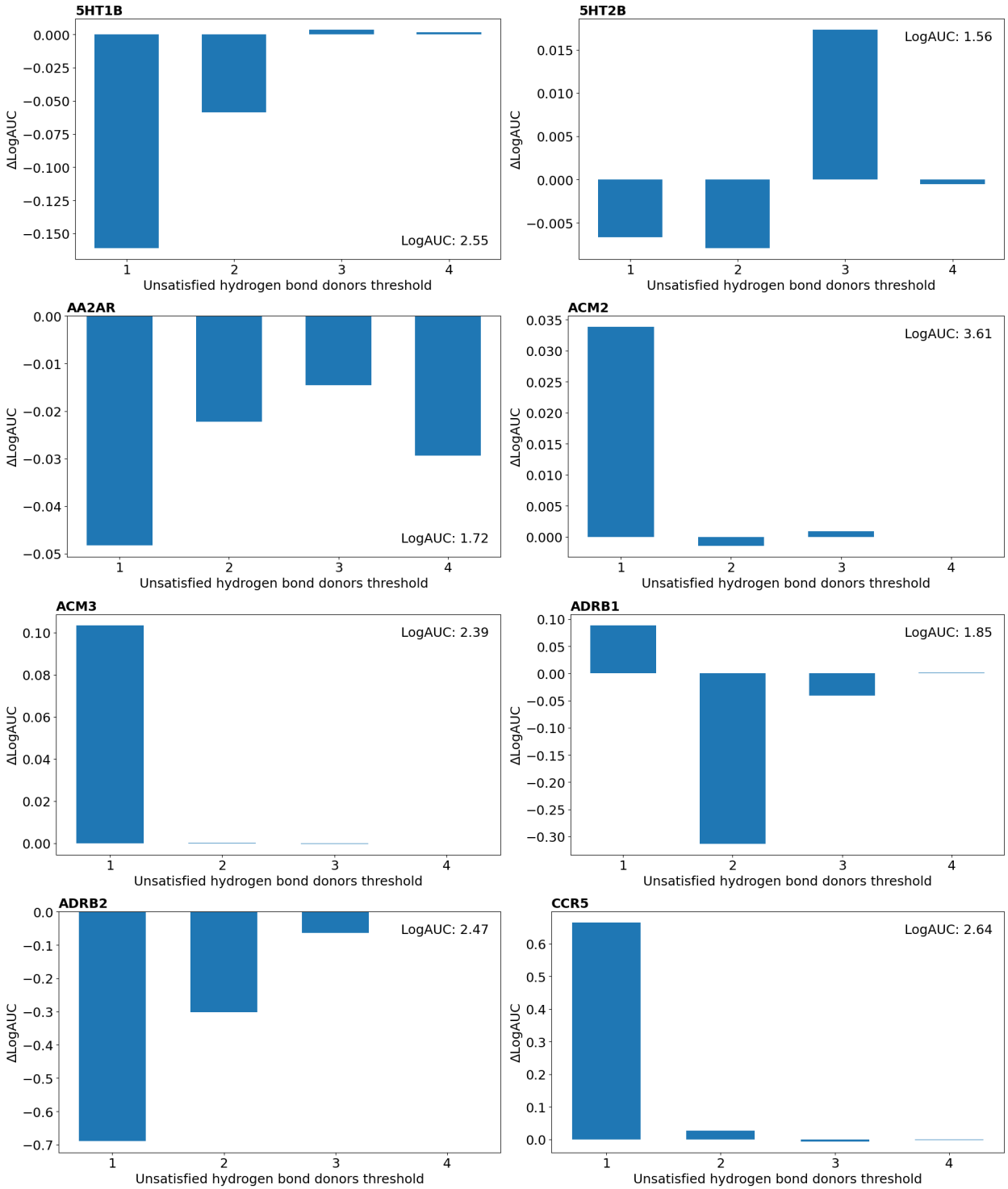
Supplementary figure series 2.5.7. Retrospective virtual screening results: benchmarking enrichment with strain filtering on the GPCR-Bench database docked with AutoDock.

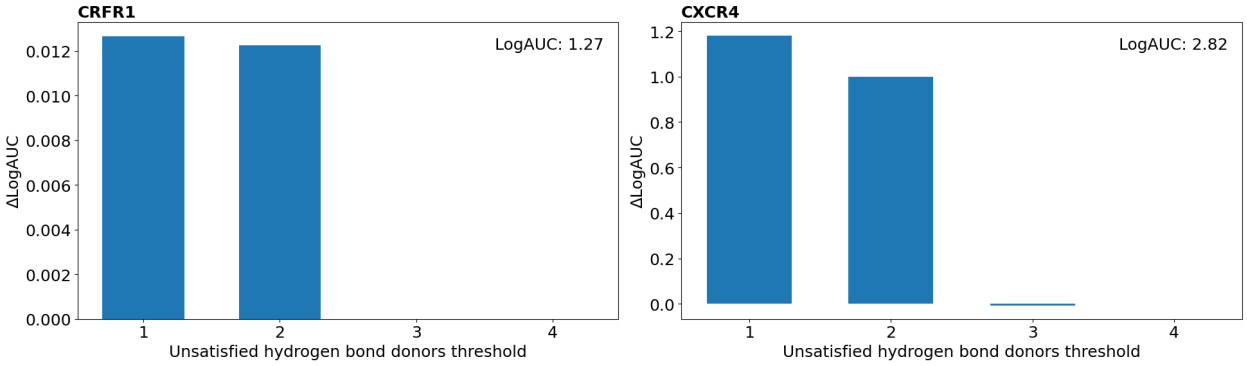



CRFR1
LogAUC: 1.26
CXCR4
LogAUC: 2.87
DRD3
LogAUC: 1.99
GPR40
LogAUC: 5.18
HRH1
LogAUC: 2.60
MGLUR1
LogAUC: 2.14
MGLUR5
LogAUC: 2.16
OPRD
LogAUC: 3.56
ΔLogAUC
Strain threshold

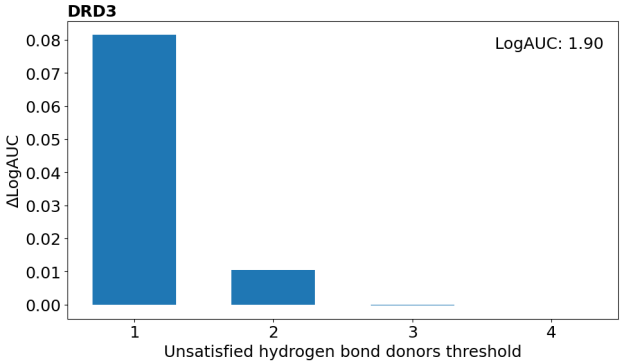

OPRK
LogAUC: 3.19
OPRM
LogAUC: 2.13
OPRX
LogAUC: 2.81
OX2R
LogAUC: 1.37
P2Y12
LogAUC: 2.07
PAR1
LogAUC: 3.38
S1PR1
LogAUC: 3.22
SMO
LogAUC: 2.08
ΔLogAUC
Strain threshold

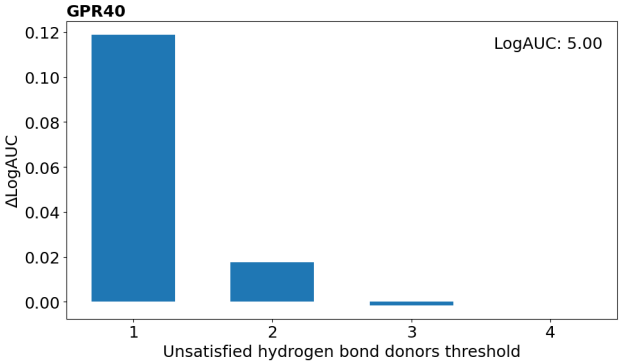
Supplementary figure series 2.5.8. Retrospective virtual screening results: benchmarking enrichment with unsatisfied hydrogen bond donors filtering on the GPCR-Bench database docked with AutoDock.

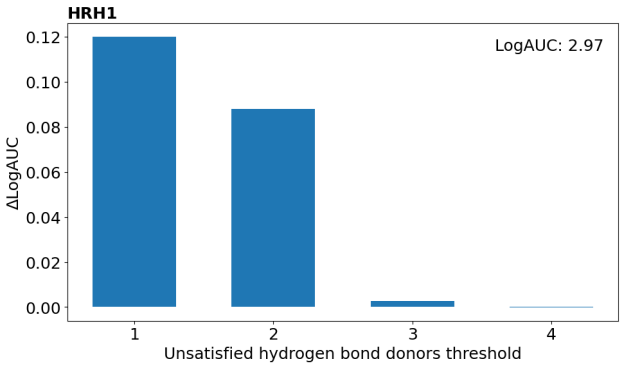


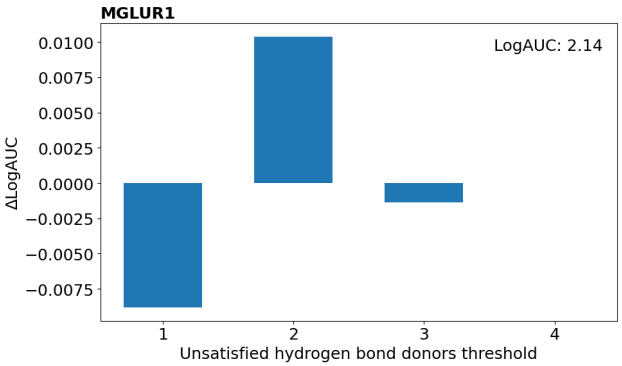

CRFR1
LogAUC: 1.27
ΔLogAUC
Unsatisfied hydrogen bond donors threshold

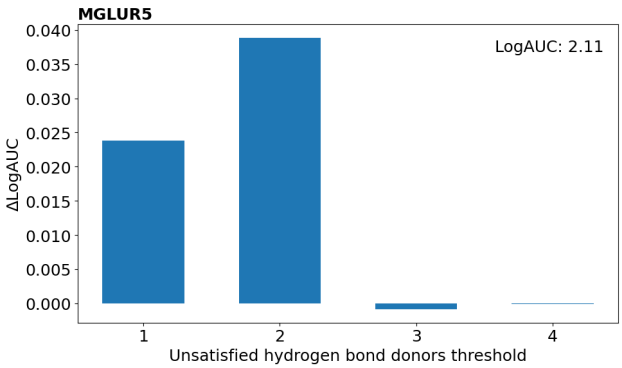
CXCR4
LogAUC: 2.82
ΔLogAUC
Unsatisfied hydrogen bond donors threshold

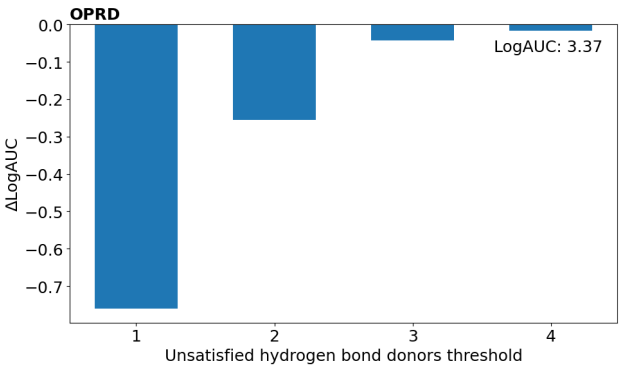

DRD3
LogAUC: 1.90
ΔLogAUC
Unsatisfied hydrogen bond donors threshold

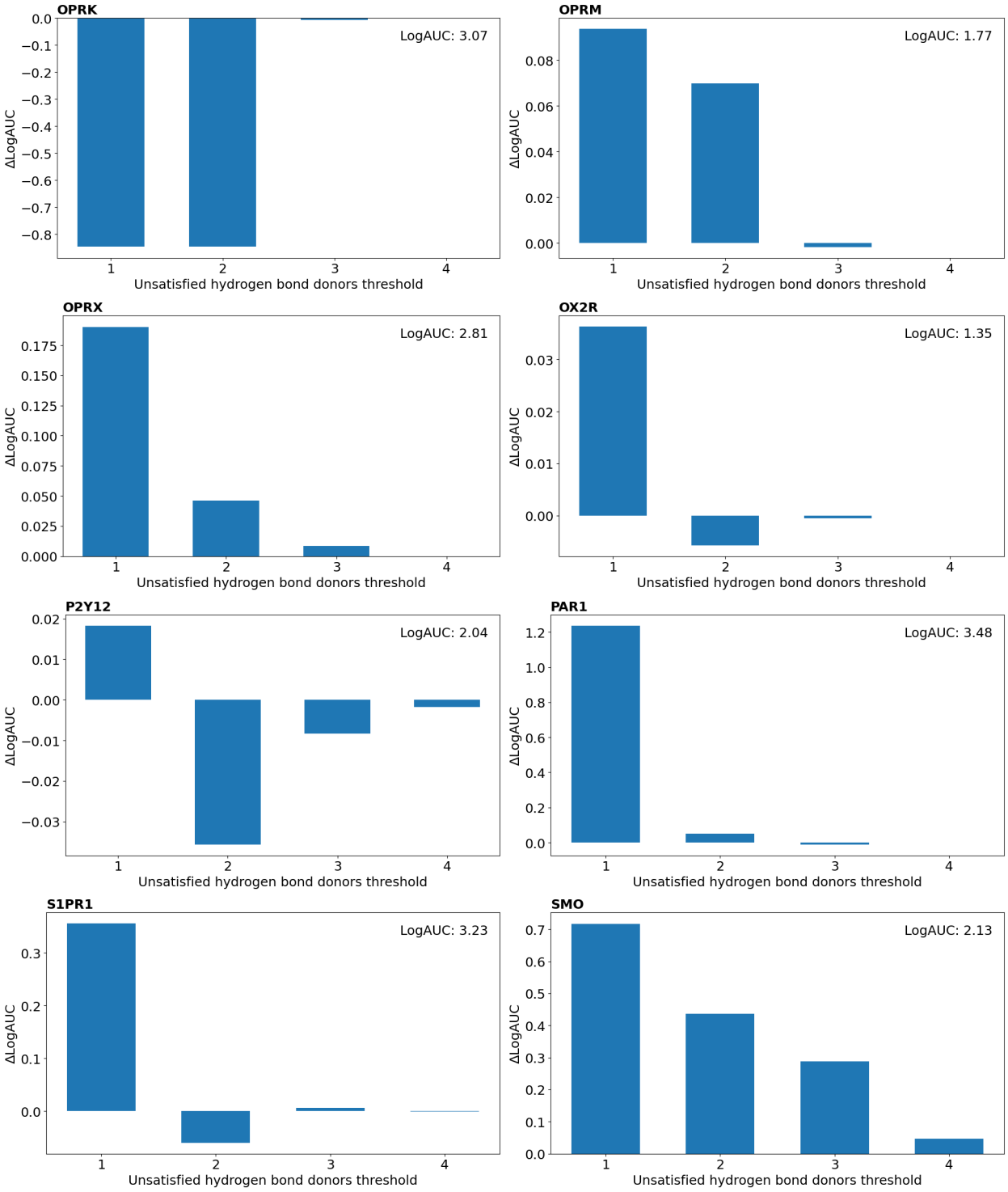

GPR40
LogAUC: 5.00
ΔLogAUC
Unsatisfied hydrogen bond donors threshold


HRH1
LogAUC: 2.97
ΔLogAUC
Unsatisfied hydrogen bond donors threshold

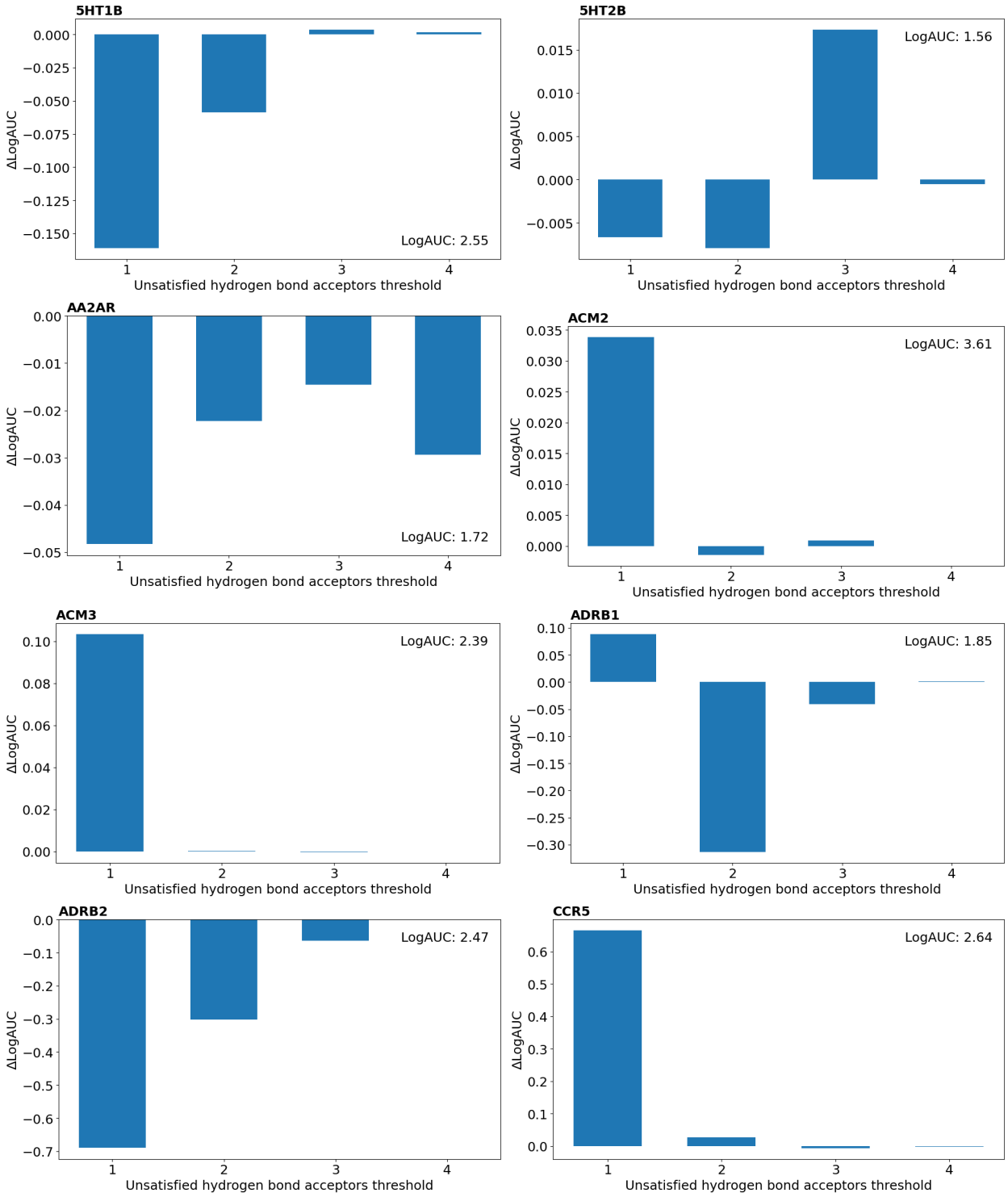

MGLUR1
LogAUC: 2.14
ΔLogAUC
Unsatisfied hydrogen bond donors threshold

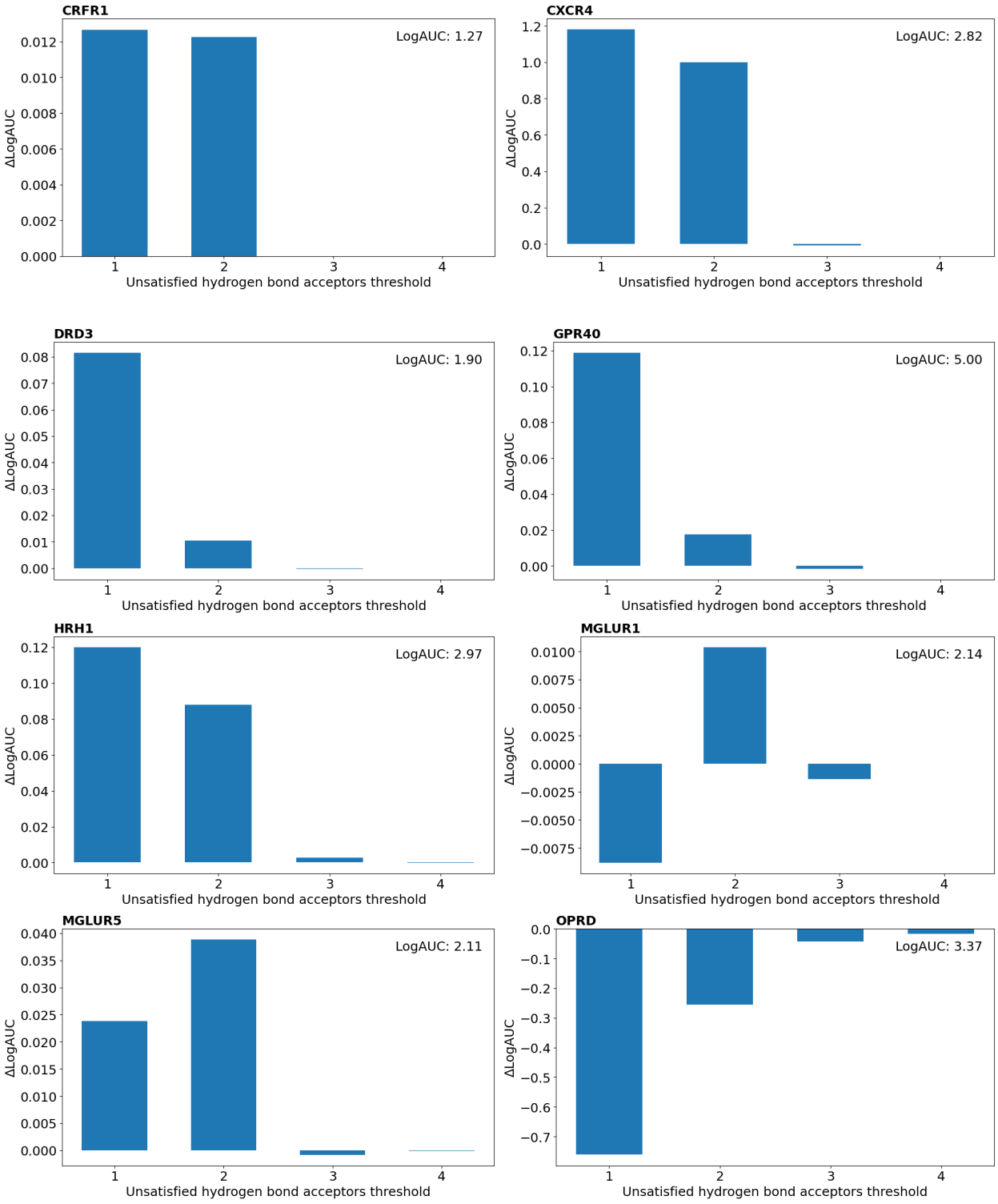

MGLUR5
LogAUC: 2.11
ΔLogAUC
Unsatisfied hydrogen bond donors threshold

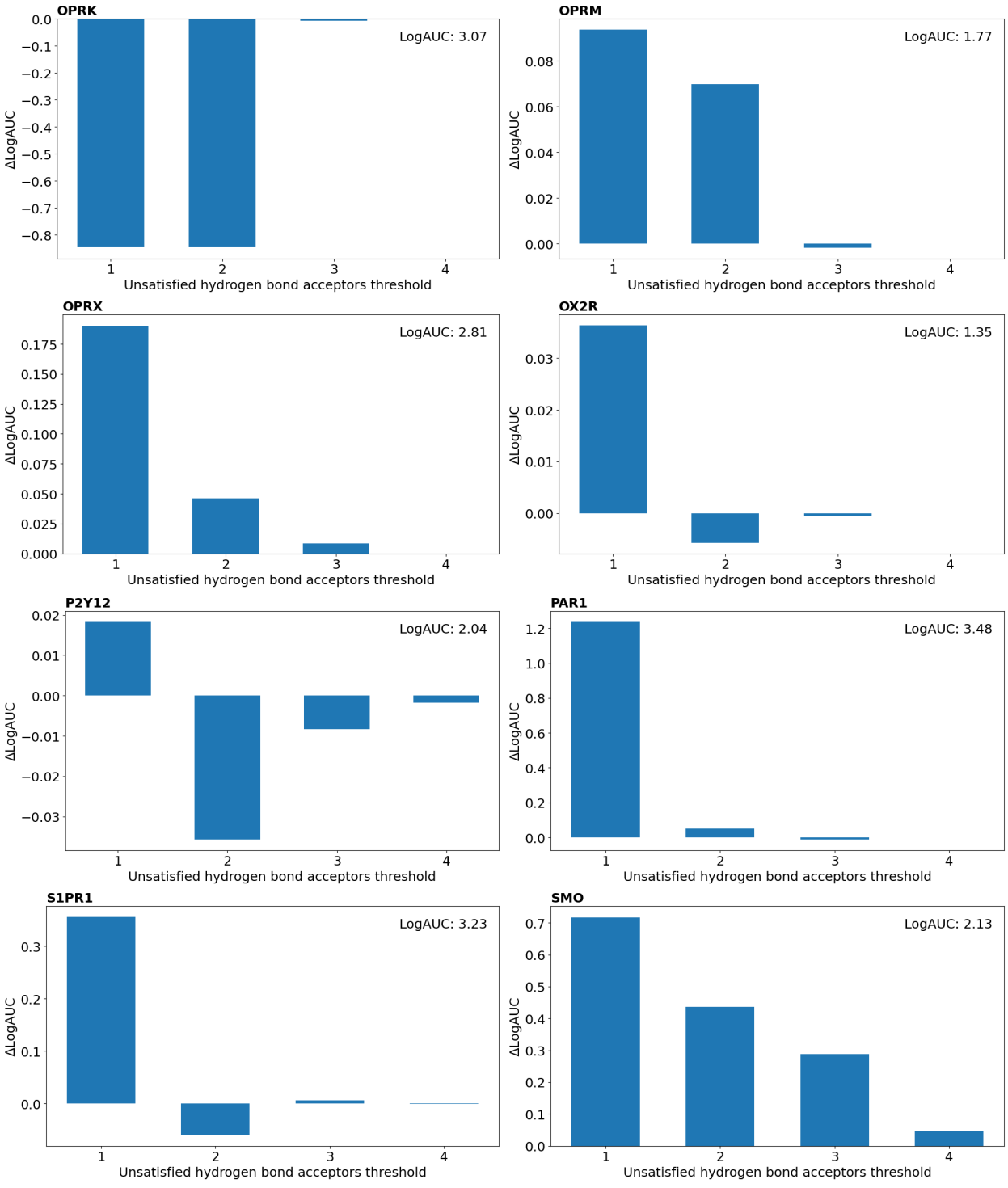

OPRD
LogAUC: 3.37
ΔLogAUC
Unsatisfied hydrogen bond donors threshold


OPRK
LogAUC: 3.07
OPRM
LogAUC: 1.77
OPRX
LogAUC: 2.81
OX2R
LogAUC: 1.35
P2Y12
LogAUC: 2.04
PAR1
LogAUC: 3.48
S1PR1
LogAUC: 3.23
SMO
LogAUC: 2.13
ΔLogAUC
Unsatisfied hydrogen bond donors threshold

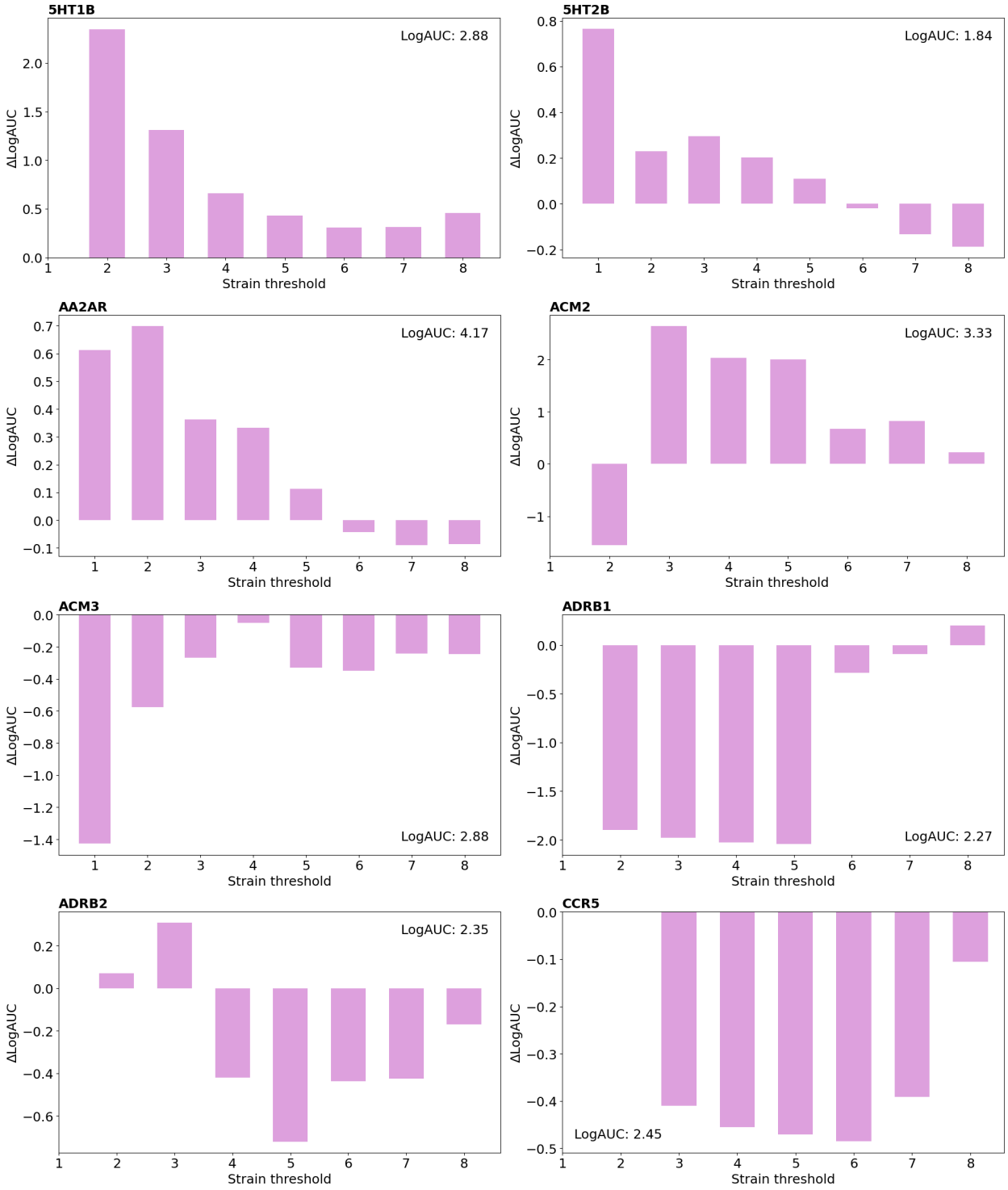
Supplementary figure series 2.5.9. Retrospective virtual screening results: benchmarking enrichment with unsatisfied hydrogen bond acceptors filtering on the GPCR-Bench database docked with AutoDock.

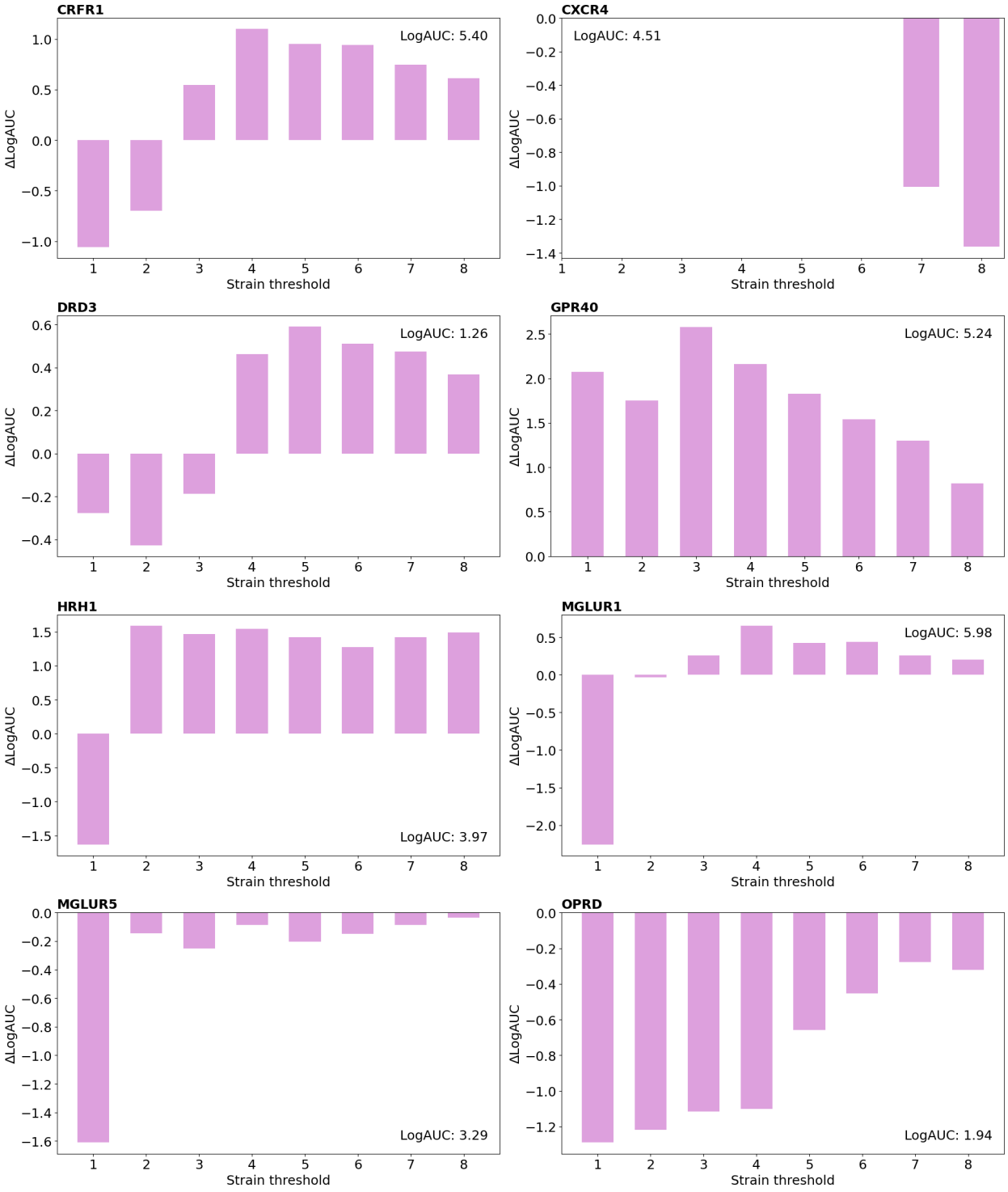


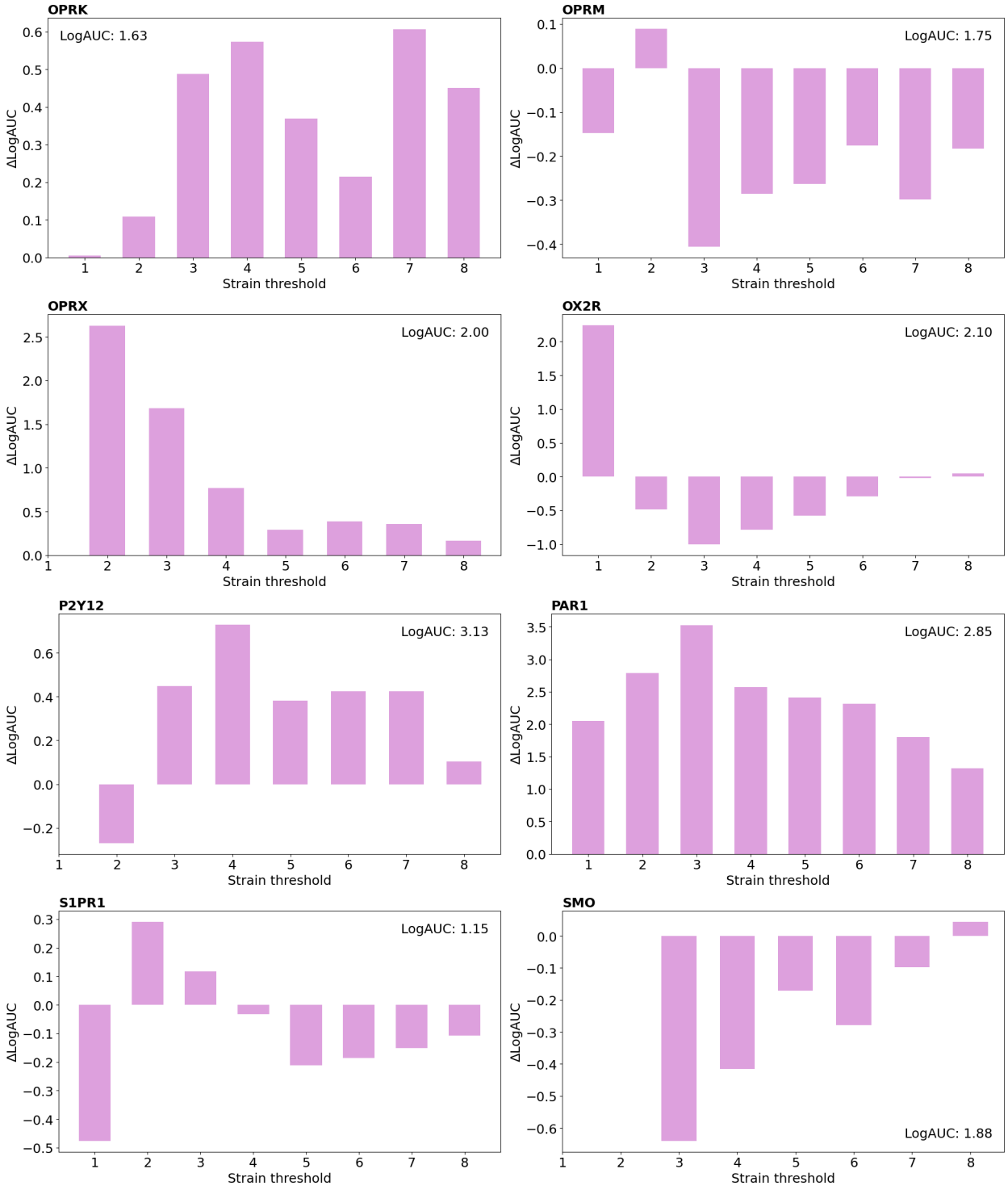

CRFR1
LogAUC: 1.27
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold
CXCR4
LogAUC: 2.82
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold
DRD3
LogAUC: 1.90
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold
GPR40
LogAUC: 5.00
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold
HRH1
LogAUC: 2.97
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold
MGLUR1
LogAUC: 2.14
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold
MGLUR5
LogAUC: 2.11
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold
OPRD
LogAUC: 3.37
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold


OPRK
LogAUC: 3.07
OPRM
LogAUC: 1.77
OPRX
LogAUC: 2.81
OX2R
LogAUC: 1.35
P2Y12
LogAUC: 2.04
PAR1
LogAUC: 3.48
S1PR1
LogAUC: 3.23
SMO
LogAUC: 2.13
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold

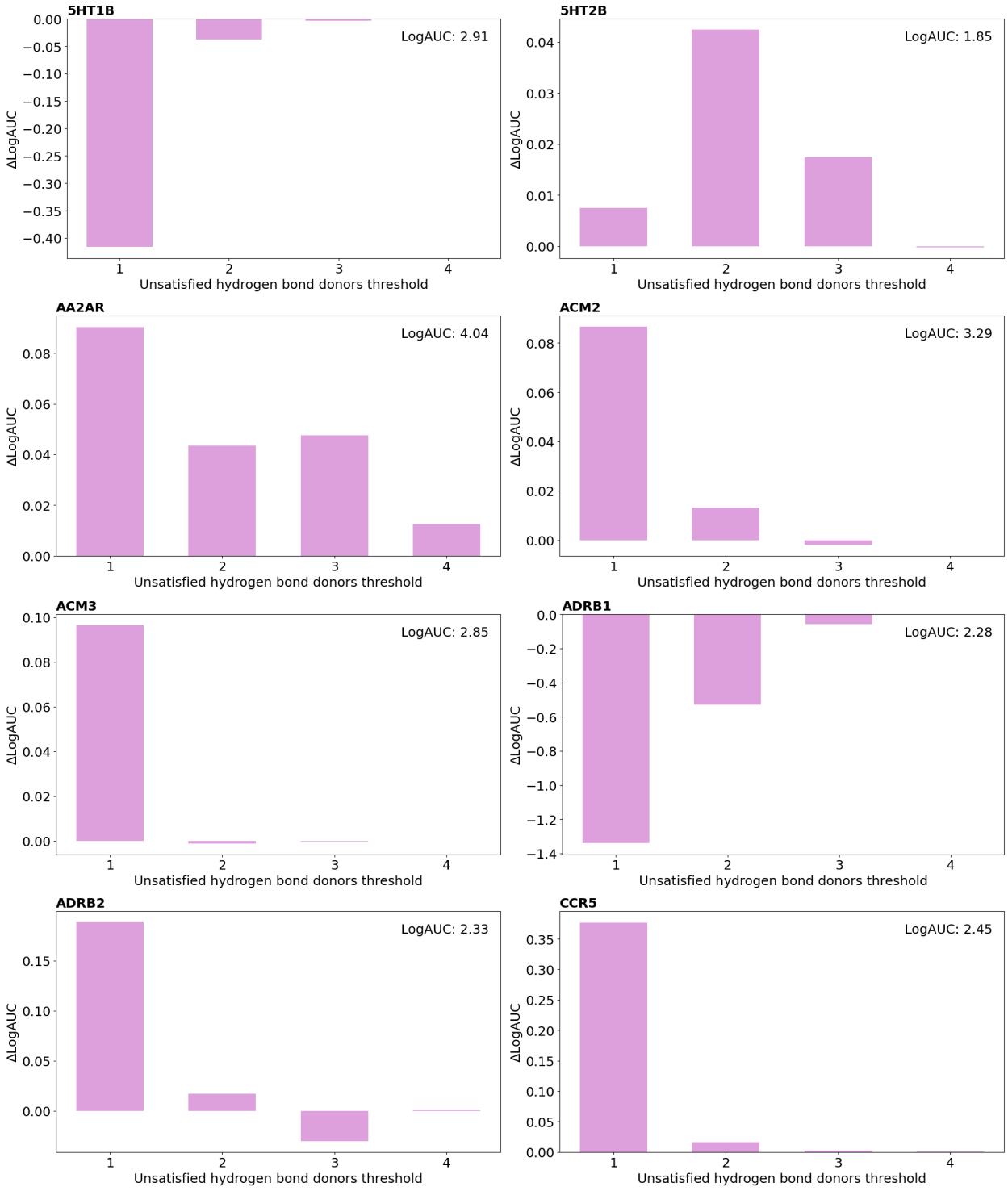
Supplementary figure series 2.5.10. Retrospective virtual screening results: benchmarking enrichment with strain filtering on the GPCR-Bench database docked with GNINA.

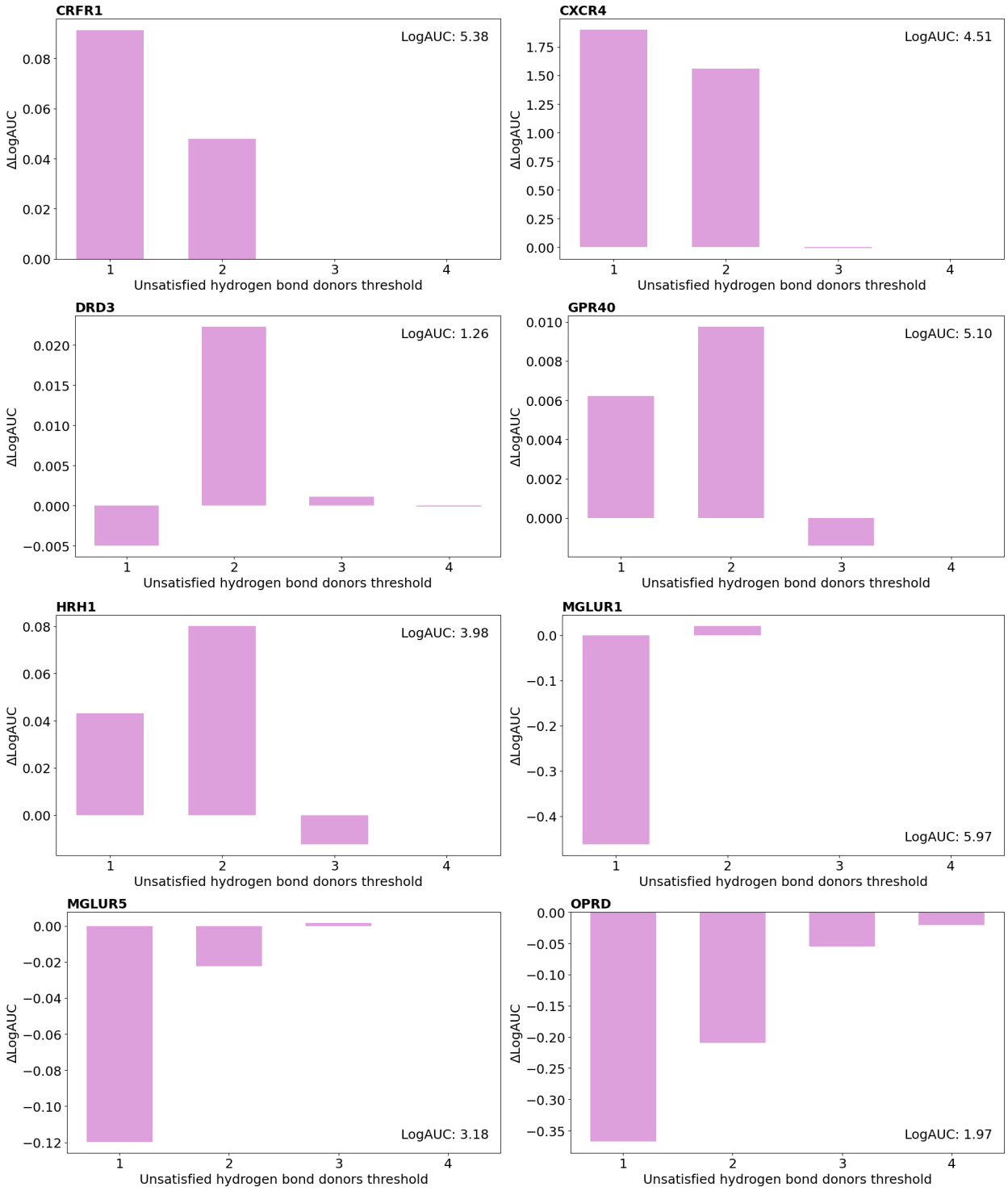


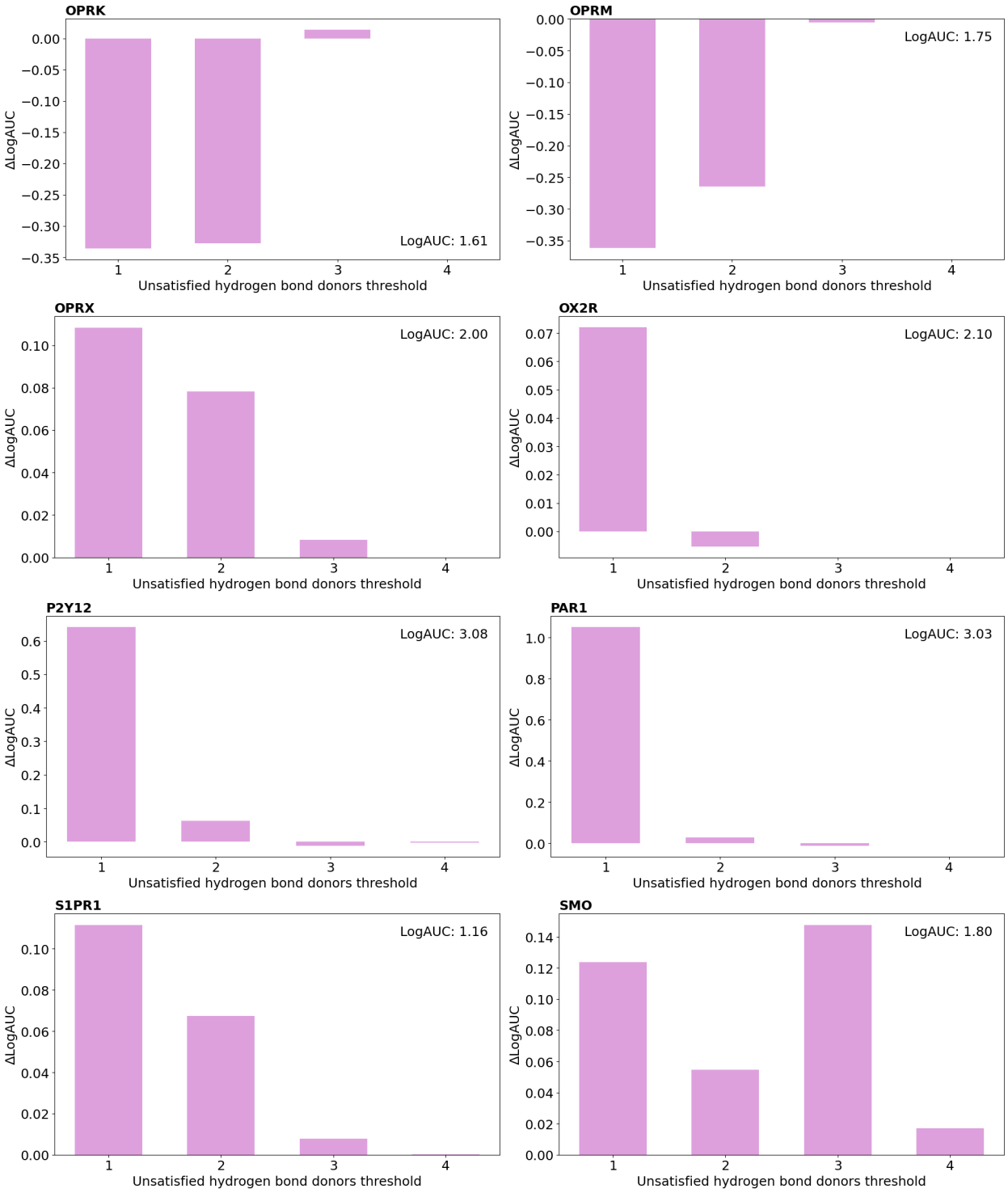

CRFR1
LogAUC: 5.40
CXCR4
LogAUC: 4.51
DRD3
LogAUC: 1.26
GPR40
LogAUC: 5.24
HRH1
LogAUC: 3.97
MGLUR1
LogAUC: 5.98
MGLUR5
LogAUC: 3.29
OPRD
LogAUC: 1.94
ΔLogAUC
Strain threshold


OPRK
LogAUC: 1.63
OPRM
LogAUC: 1.75
OPRX
LogAUC: 2.00
OX2R
LogAUC: 2.10
P2Y12
LogAUC: 3.13
PAR1
LogAUC: 2.85
S1PR1
LogAUC: 1.15
SMO
LogAUC: 1.88
ΔLogAUC
Strain threshold

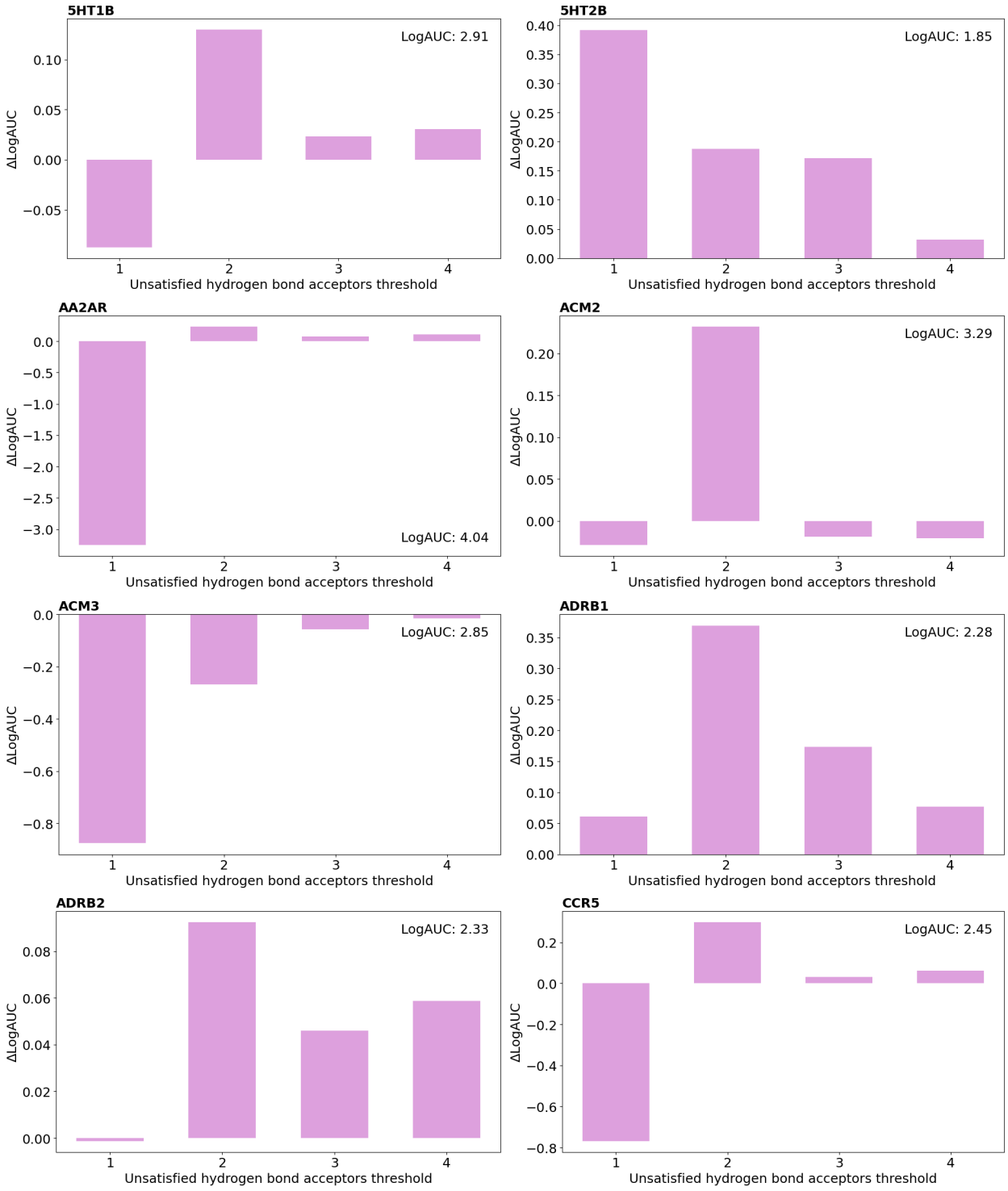
Supplementary figure series 2.5.11. Retrospective virtual screening results: benchmarking enrichment with unsatisfied hydrogen bond donors filtering on the GPCR-Bench database docked with GNINA.

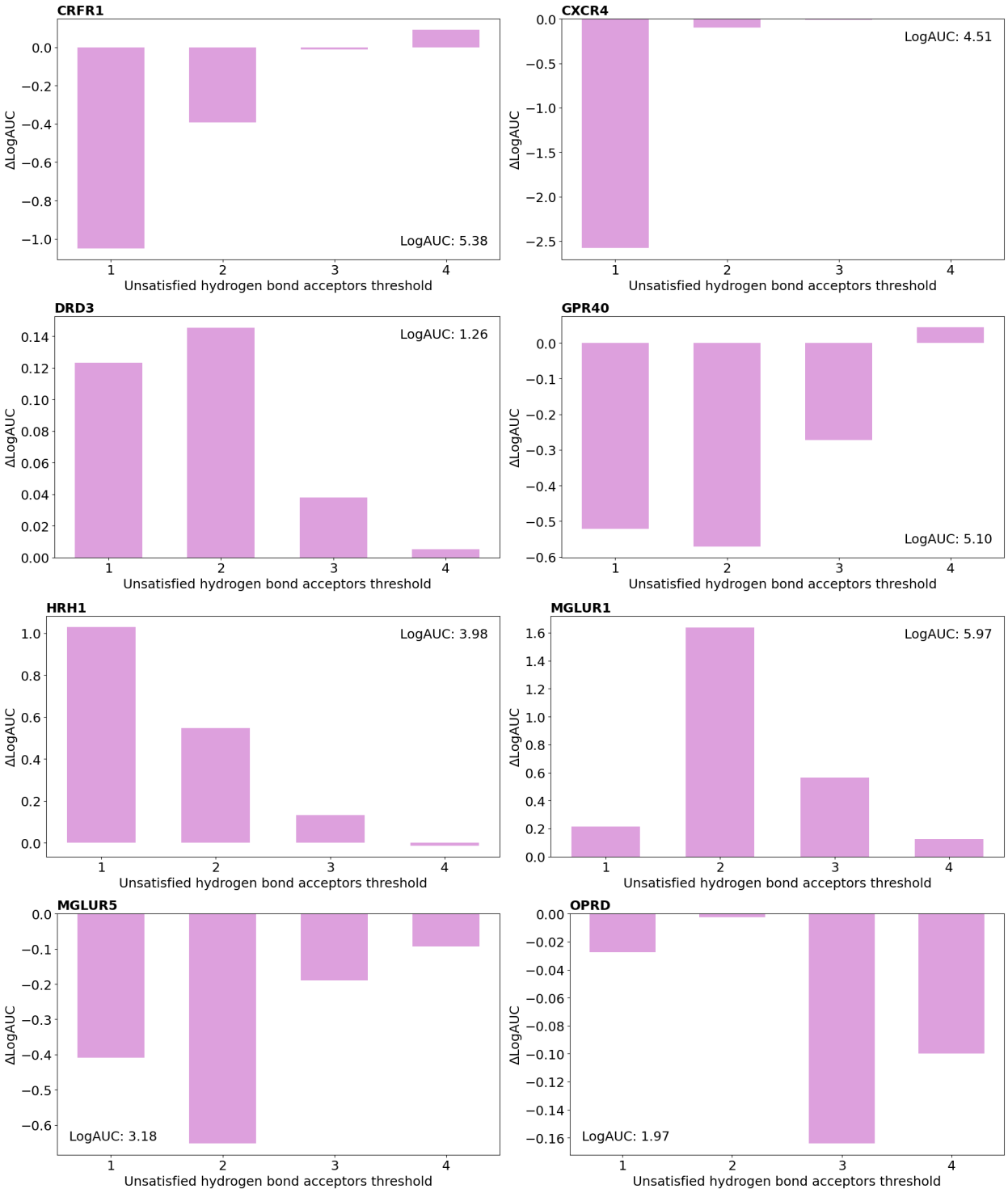


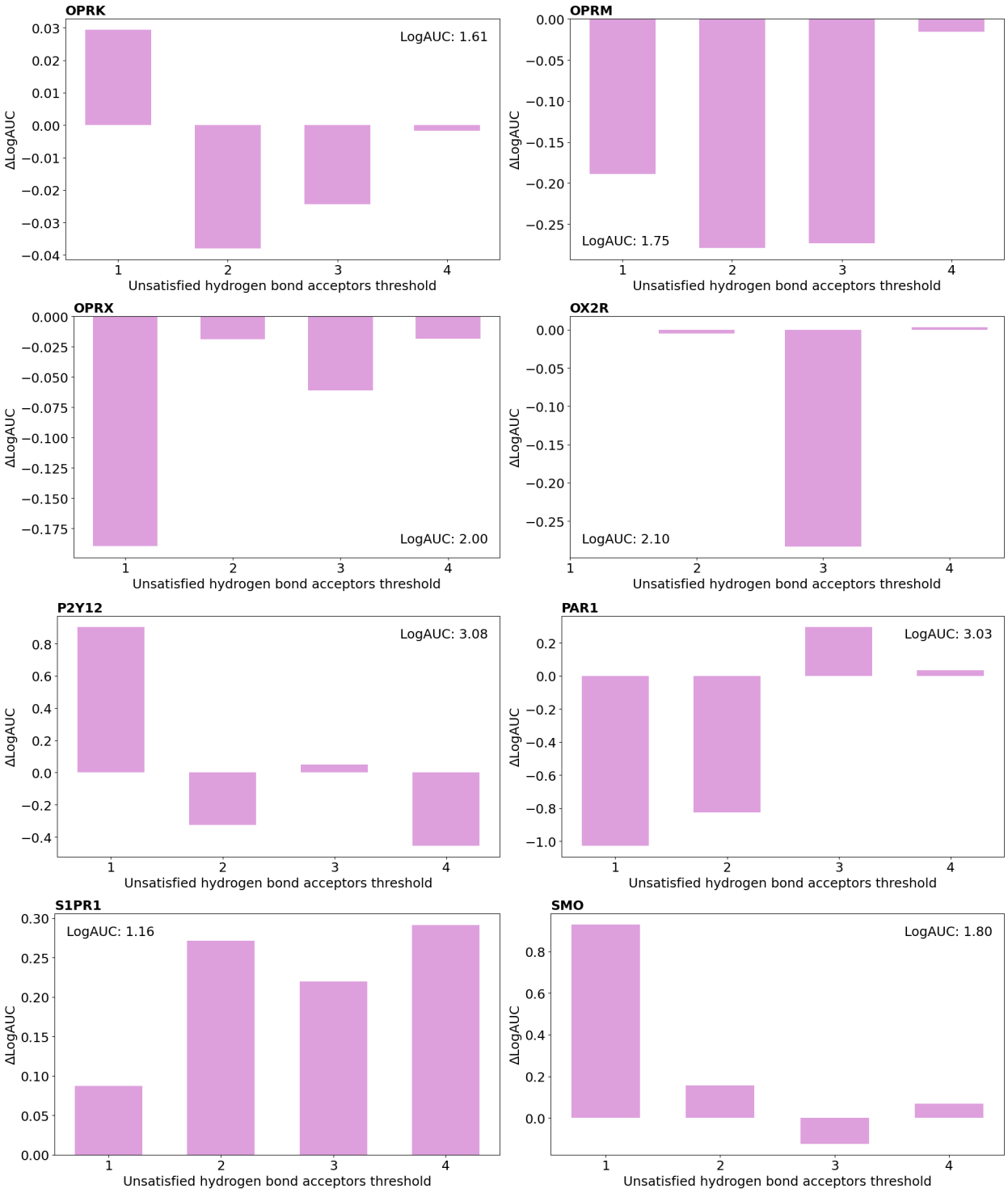

CRFR1
LogAUC: 5.38
CXCR4
LogAUC: 4.51
DRD3
LogAUC: 1.26
GPR40
LogAUC: 5.10
HRH1
LogAUC: 3.98
MGLUR1
LogAUC: 5.97
MGLUR5
LogAUC: 3.18
OPRD
LogAUC: 1.97
ΔLogAUC
Unsatisfied hydrogen bond donors threshold


OPRK
LogAUC: 1.61
OPRM
LogAUC: 1.75
OPRX
LogAUC: 2.00
OX2R
LogAUC: 2.10
P2Y12
LogAUC: 3.08
PAR1
LogAUC: 3.03
S1PR1
LogAUC: 1.16
SMO
LogAUC: 1.80
ΔLogAUC
Unsatisfied hydrogen bond donors threshold

Supplementary figure series 2.5.12. Retrospective virtual screening results: benchmarking enrichment with unsatisfied hydrogen bond acceptors filtering on the GPCR-Bench database docked with GNINA.




CRFR1
LogAUC: 5.38
CXCR4
LogAUC: 4.51
DRD3
LogAUC: 1.26
GPR40
LogAUC: 5.10
HRH1
LogAUC: 3.98
MGLUR1
LogAUC: 5.97
MGLUR5
LogAUC: 3.18
OPRD
LogAUC: 1.97
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold


OPRK
LogAUC: 1.61
OPRM
LogAUC: 1.75
OPRX
LogAUC: 2.00
OX2R
LogAUC: 2.10
P2Y12
LogAUC: 3.08
PAR1
LogAUC: 3.03
S1PR1
LogAUC: 1.16
SMO
LogAUC: 1.80
ΔLogAUC
Unsatisfied hydrogen bond acceptors threshold

Chapter 3: Simulated ML-Accelerated Virtual Screening Campaigns

As discussed in the previous chapter, it is important to conduct retrospective VS before prospective VS; this is critical to the selection of medicinal chemistry filters (strain and unsatisfied hydrogen bonds) used in the post-processing of molecular dockings – these filters showed system-dependent performance in the enrichment of docking-ranked compound lists. Moreover, some proteins showed significant improvement in enrichment with both strain and unsatisfied hydrogen bond filters considered at certain thresholds. Herein, the interest is to test these medicinal chemistry filters for their potential to improve enrichment in large-scale ML-accelerated VS. MOO could be used to handle multiple molecule design criteria in parallel, in the search for molecules with optimal molecular docking scores, strains, and unsatisfied hydrogen bonds. Scalarization is a common MOO method which aggregates multiple objectives into a single objective function – this requires assigning the relative importance of each objective. Several MOO AI-augmented molecular docking models were discussed in Section 1.3. MTL presents an attractive method to predict multiple tasks in a joint objective; the shared architecture increases computational efficiency and has the potential to improve predictions through shared information. Common MTL implementations aggregate separate objectives with different loss functions into a single scalar objective through weighted sum or adaptive methods. The consideration of multiple objectives in the shared learning representation can lead to either an increase in inductive knowledge transfer or destructive interference, where optimization requires the balance of multiple objectives with potentially conflicting demands. No current model aims to improve AI-accelerated molecular docking predictions through the incorporation of multiple medicinal chemistry properties used in the classification of drug candidates or explores MTL to this end. This chapter tests selected proteins from our benchmark (Section 2.2) in simulated large-scale VS to assess whether considering strain and unsatisfied hydrogen bonds together with docking scores within a MTL framework improves enrichment and model performance.

3.1 Methods

3.1.1 Data

3.1.1.1 Morgan Fingerprints

Molecular fingerprints are bit string or vector representations that indicate the presence or absence of specific molecular substructures.[1] The most common molecular fingerprints are Morgan fingerprints[2] (or extended-connectivity fingerprints[3], ECFPs).[4] Morgan fingerprints are generated through a circular fingerprint algorithm that centres on each non-hydrogen atom or fragment; predefined parameters (bit length, radius etc.) are then used to map circular atom neighbourhoods at the specified radius in an iterative manner.[5] The selected parameters used were 1024 bit length, radius 2, and chiral information incorporated.[6] Morgan fingerprints are fast to compute and include the relevant local, chemical information (functional groups, sterics, bond connectivity, chirality, etc.) – these have shown success in ML-based VS.[7]

3.1.1.2 Molecular Docking Scores

Molecular Docking programs AutoDock (Section 2.1.2.1.1) and GNINA (Section 2.1.2.1.2) were used as outlined in chapter 2 Section 2.1.2. The training, validation, and testing (TVT) data was preprocessed to remove Not a Number (NaN) values.

3.1.1.3 Strain

The strain filter was used as outlined in chapter 2 Section 2.1.3. TVT data was preprocessed to remove NaN and negative values.

3.1.1.4 Hydrogen Bonds

The IChem program was used as outlined in chapter 2 Section 2.1.4 to count molecular unsatisfied hydrogen bond donors and acceptors.

3.1.1.5 Protein-Ligand Datasets

Selected (4) LIT-PCBA datasets were explored in simulated large-scale VS based on their benchmark performance (Section 2.2). Proteins were selected based on their high enrichment,

reasonable number of actives, and diverse protein families (Table. 3.1). Both AutoDock and GNINA were considered.

Table 3.1 Simulated virtual screening campaigns (benchmark): selected protein datasets.

Docking Program	Protein	Threshold (strain; u-hbd[a]; u-hba[b])	Actives[c] (docked)	Actives (filtered)	LogAUC[d] (docked)	LogAUC[d] (filtered)	ΔLogAUC
AD	MTORC1	4;2;2	99	26	1.26	2.31	1.05
AD	MAPK1	4;0;3	287	17	1.71	2.60	0.89
GNINA	TP53	4;1;1	51	4	1.34	3.24	1.89
GNINA	VDR	3;2;0	606	4	1.41	2.49	1.06

[a]unsatisfied hydrogen bond donors. [b]unsatisfied hydrogen bond acceptors. [c]Actives include stereo-enumerated (considered active). [d]LogAUC values are scaled (x10); logAUC random classifier is 1.45.

MTORC1, TP53, and VDR (Table 3.1) showed to be poor molecular docking systems in benchmark: their LogAUC values were below that of a random classifier (1.45). Additional strain and unsatisfied hydrogen bond filtering increased LogAUC values – this incentivizes that systems with poor molecular docking performance can be recovered with strain and unsatisfied hydrogen bond filtering.

3.1.1.6 Chemical Database

The selected LIT-PCBA protein datasets were each combined with the Enamine[8] REAL diversity set containing 80 million compounds. Uniform Manifold Approximation and Projection[9] (UMAP) is a non-linear dimensionality reduction technique that projects high-dimensional data into low-dimensional space (2D). The chemical space was mapped with the UMAP algorithm to check that actives, inactives, and Enamine REAL diversity set compounds share the same chemical space (Figure 3.1). UMAP was fit on a random 1 million compound sample of the Enamine dataset using the Jaccard distance metric, 25 nearest neighbours, and a minimum distance of 0.001. The fitted model was used to transform protein active and inactive datasets into the learnt embedding space. Herein, all Enamine REAL diversity set compounds are considered inactive in the simulated large-scale VS.

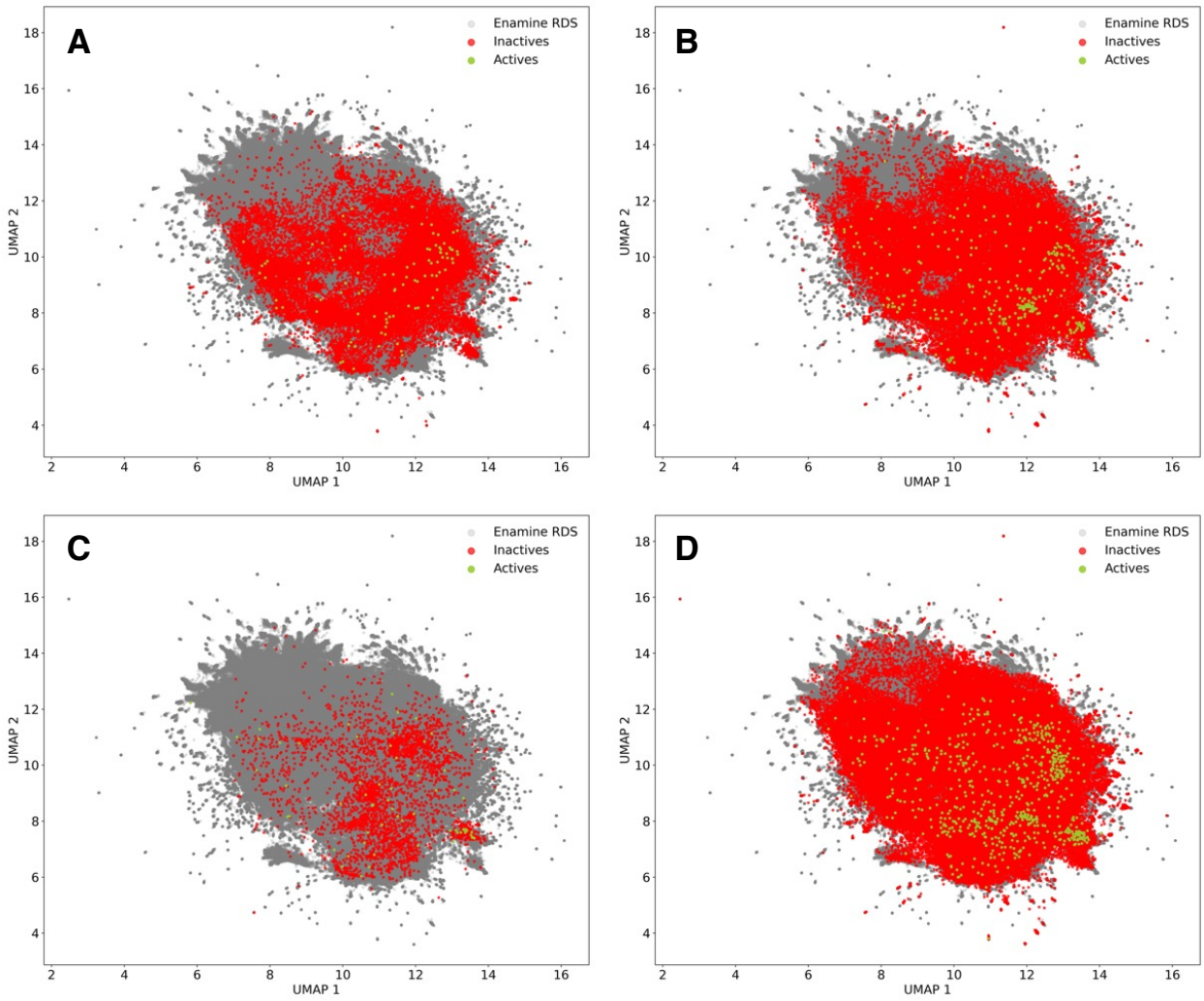


Figure 3.1. UMAP chemical space visualization. (A) MTORC1 and enamine datasets. (B) MAPK1 and enamine datasets. (C) TP53 and enamine datasets. (D) VDR and enamine datasets. Enamine (RDS) is represented by a random sample of 1 million compounds. All protein dataset actives and inactives are plotted.

3.1.1.7 Data Splitting

Each large protein dataset was randomly sampled (999k molecules) and split into training (333k), validation (333k), and testing (333k) splits.[6]

3.1.2 Model Architecture

3.1.2.1 Multi-Layer Perceptron

The perceptron is a simple, supervised learning algorithm and single-layer artificial neuron designed for binary classification, Figure 3.2.

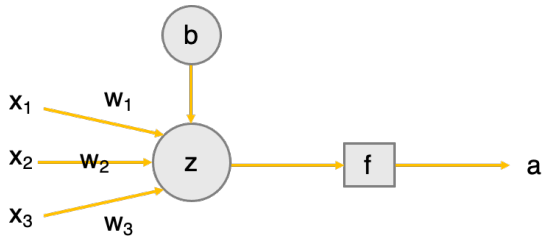


Figure 3.2. Perceptron. The input features (x) are multiplied by their weights (w) and summed with the neuron's bias (b). The pre-activation, linear combination is computed at the neuron and passed through a non-linear activation function (f) to receive the activation output (a) of the perceptron.

(Equation 3.1)

$$z = w \cdot x + b$$

where, z is the pre-activation, linear combination computed at the neuron, w is the weight vector, x is the input feature vector, and b is the bias.

(Equation 3.2)

$$a = f(z)$$

where, a is the activation output, f is the non-linear activation function, and z is the pre-activation output in Equation 3.1.

Multi-Layer Perceptron (MLP) is a supervised learning algorithm, referred to as a feed-forward artificial neural network, Figure 3.3.

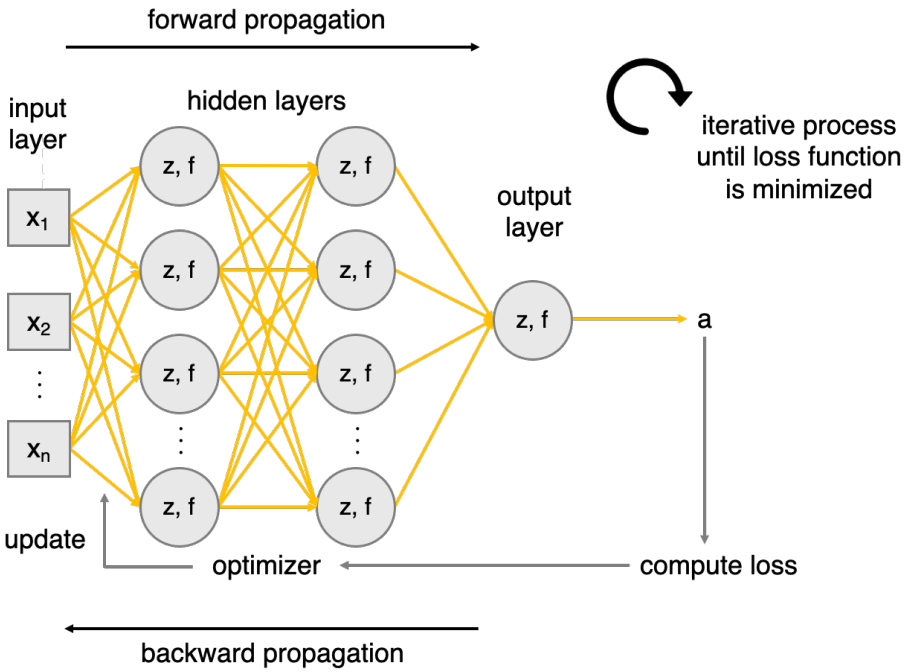


Figure 3.3. Multi-layer perceptron. Forward propagation: input features are passed through the fully connected hidden layers, where each neuron computes a weighted sum that is then followed with non-linear activation. The final layer outputs the prediction, which is used to compute the loss with a loss function. Back propagation: gradients of the loss with respect to each parameter (weight and bias) are computed and the optimizer (gradient descent algorithm) updates the parameters to minimize loss in an iterative training process.

The MLP architecture includes an input layer, hidden layer(s), and output layer, with trainable weights and biases at each node that are adjusted during learning to model complex, non-linear relationships. The input layer holds the feature vector. The hidden layer(s) consists of neuron(s) located between the input and output layer; these layer(s) are considered hidden because their activations are not observed and represent intermediate feature transformations. Neurons in the hidden layer(s) calculate the weighted sum of inputs in the previous layer multiplied by their weight and run the weighted sum to an activation function, Equation 3.1 The activation function is a mathematical function applied to the output of a neuron, Equation 3.2; it transforms the weighted sum (weighted input and bias) into an activation value that is passed onto the next layer – this enables the network to learn complex, non-linear relationships. The output layer is the final layer in a neural network that contains the prediction. All MLP architectures were implemented in PyTorch.

3.1.2.2 Model Training and Hyperparameter Optimization

All the MLP architectures herein contain 1024-bit Morgan fingerprint input feature vectors, multiple hidden layers (4 in STL and 3 in MTL models), and task-specific output layers. Batch normalization, Rectified Linear Unit activation, and dropout regularization was applied at each hidden layer. Model training was performed with mini-batch gradient descent. The Adam[10] optimizer was used to update model parameters. Early stopping was applied to prevent overfitting in training, with patience 5. Optuna[11] was used for automatic hyperparameter optimization (30 trials, 3 parallel jobs, and trial pruning with patience 5). Optuna hyperparameters include hidden shape (512-1536), learning rate (1e-4-5e-5), dropout rate (0.1-0.6), batch size (256, 512, 1024); epochs were fixed at 50. Models were optimized on validation loss, used both in training and Optuna hyperparameter optimization.

3.1.2.2.1 Multi-Layer Perceptron – Single-Task Learning Model

Regression models were used in the prediction of molecular docking scores and strain. Multi-class classification models (5 classes) were used in the prediction of unsatisfied hydrogen bond donors and acceptors.

3.1.2.2.2 Multi-Layer Perceptron – Multi-Task Learning Model

The MTL MLP models have 4 task-specific output layers: 2 regression and 2 multi-class classification heads. The shared hidden layers learn a common molecular representation across related prediction tasks – parameter sharing can introduce an inductive bias that promotes knowledge transfer (regularization effect: learns common structures, improves generalization, and reduces overfitting). Hard parameter sharing was implemented, where the parameters are jointly optimized across all tasks. The MTL model was used in the joint prediction of molecular docking scores, strain, and unsatisfied hydrogen bond donors and acceptors.

Table 3.2. Model-specific components.

Model	Task	Loss	Metrics
MLP-R	Regression	MSE	Pearson Correlation Coefficient, RMSE, R^2
MLP-C	Classification	Cross Entropy	Precision, Recall
MLP-MTL	Multi-Task Learning: regression + classification	Uncertainty-weighted: MSE + Cross Entropy	Pearson Correlation Coefficient, RMSE, R^2, Precision, Recall

3.1.3 Loss Functions

3.1.3.1 Loss Functions – Single-Task Learning

3.1.3.1.1 Mean Squared Error – Regression

Mean Squared Error (MSE) is a common regression, loss function that measures the average of the squared difference between the predicted and actual values. MSE was implemented as the regression loss function.

(Equation 3.3)

$$L_{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$$

where, N is the number of samples, i is the sample, y_i is the actual value, and $\hat{y}_i$ is the predicted value.

3.1.3.1.2 Cross Entropy – Multi-Class Classification

Cross Entropy (CE) is a common multi-class classification loss function that measures the difference between the predicted probability distribution and actual distribution. The CE version implemented operates on the network logits.

(Equation 3.4)

$$L_{CE} = \frac{1}{N}\sum_{i=1}^{N} -\log\left(\frac{e^{z_{i,y_i}}}{\sum_{c=1}^{k} e^{z_{i,c}}}\right)$$

where, N is number of samples, i is the sample, k is the number of classes, e^{z_i} is the logit, and y_i is the actual class label.

3.1.3.2 Loss Function – Multi-Task Learning

MSE (Section 3.1.3.1.1) and Cross Entropy (Section 3.1.3.1.2) were used as the regression and classification loss functions, respectively, in the MTL model.

The conventional method to combine multiple objective losses is a weighted linear sum of the losses for each individual task:

(Equation 3.5)

$$L_T = \sum_i w_i L_i$$

where, L_T is total loss, i is the task, w_i is the task-specific weight, and L_i is task-specific loss.

Concerns with this method include appropriate weight selection and resource demands.

Homoscedastic uncertainty weighing proposed in Kendell et al.[12] work was used to implement MTL with multiple regression and classification objectives – this method handles task-specific loss weighing through a trainable parameter (σ) to balance the contributions of each task during model training:

(Equation 3.6)

$$L_T = \sum_i \frac{1}{2\sigma_i^2} L_i + \log\sigma_i$$

where, L_T is the total loss, i is the task, σ_i is the task-specific trainable parameter, and L_i is the task-specific loss.

Homoscedastic uncertainty weighing balances task-specific gradient magnitudes and normalizes different loss scales. Tasks with higher uncertainty (σ) contribute less to the total loss, whereas tasks with lower uncertainty (σ) contribute more. The trainable parameter (σ) was initialized at a small positive number (0.1) to attribute moderate importance to each task before dynamic, task-specific weighing.

3.1.4 Machine Learning Metrics

The metric selected to assess the classification of active versus inactive molecules is LogAUC as used in the VS benchmark and described in Section 2.5.

3.1.4.1 Regression Metrics

3.1.4.1.1 Root Mean Squared Error

Root Mean Squared Error (RMSE) is the square root of the average of the squared differences between the predicted and actual values. The RMSE scale ranges from 0 to positive infinity, where 0 indicates perfect performance, and higher numbers indicate worse performance.

(Equation 3.7)

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}}(y_i - \hat{y}_i)^2}$$

where, N is the number of samples, i is the sample, y_i is the actual value, and $\hat{y}_i$ is the predicted value.

3.1.4.1.2 Pearson Correlation Coefficient

Pearson Correlation Coefficient (r) is a statistical measure of the strength and direction of the linear association between two continuous variables. The correlation scale ranges from -1 to 1, where 0 indicates no correlation, 1 indicates total positive correlation, and -1 indicates total negative correlation.

(Equation 3.8)

$$r = \frac{\sum_{i=1}^{N}(\hat{y}_i - \bar{\hat{y}})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(\hat{y}_i - \bar{\hat{y}})^2 (y_i - \bar{y})^2}}$$

where, N is number of samples, i is the sample, y_I is the actual value, $\hat{y}_I$ is the predicted value, $\bar{y}$ is the mean of the actual values, and $\bar{\hat{y}}$ is the mean of the predicted values.

3.1.4.1.3 R-squared

R-squared (R^2) or coefficient of determination is a statistical measure of the variation in the dependent variable that is predictable from the independent variable(s). The R^2 scale ranges from 0 to 1, where 0 indicates none of the dependent variable's variance is explained by the independent variable, 1 indicates all the dependent variable's variance is explained by the independent variable, and between 0 and 1 indicates the extent to which the dependent variable's variance can be predicted from the independent variable.

(Equation 3.9)

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}$$

where, N is the number of samples, i is the sample, y_I is the actual value, $\hat{y}_I$ is the predicted value, and $\bar{y}$ is the mean of the actual values.

3.1.4.2 Multi-Class Classification Metrics

3.1.4.2.1 Precision

Precision is the proportion of the model's positive classifications that are true positives. The precision scale ranges from 0 to 1, where 0 indicates none of the predicted positives are correct, and 1 indicates that all the predicted positives are correct.

(Equation 3.10)

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives+False Positives}}$$

The weighted-average precision was used to handle class imbalance: computes an independent precision for each class and averages each class-specific precision with the weights proportional to the number of true samples in the class.

(Equation 3.11)

$$\text{Precision}_{\text{weighted}} = \sum_{c=1}^{c} \left(\frac{N_c}{N}\right) \text{Precision}_c$$

where, c is the class, N_c is the number of true samples in the class, and N is the total number of samples.

3.1.4.2.2 Recall

Recall is the proportion of all true positive samples that were classified positive. The recall scale ranges from 0 to 1, where 0 indicates none of the true positive samples were predicted positive, and 1 indicates that all true positive samples were predicted positive.

(Equation 3.12)

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives+False Negatives}}$$

The weighted-average recall was used to handle class imbalance. It computes an independent recall for each class and averages each class-specific recall with the weights proportional to the number of true samples in the class.

(Equation 3.13)

$$\text{Recall}_{\text{weighted}} = \sum_{c=1}^{c} \left(\frac{N_c}{N}\right) \text{Recall}_C$$

where, c is the class, N_c is the number of true samples in the class, and N is the total number of samples.

3.2 Results and Discussion

3.2.1 Model Performance

MTL was compared to STL in the prediction of molecular docking scores, strains, and unsatisfied hydrogen bond donors and acceptors (Figure 3.4-3.7).

Figure 3.4 summarizes model performances for molecular docking score predictions on the testing dataset. STL showed better performance in Pearson correlation coefficient, RMSE, and R^2 compared to MTL. The MTL architecture did not contribute to improved model performance for the prediction of molecular docking scores, suggesting that the additional tasks do not increase inductive transfer in the shared representation environment. Model performance was very system-dependent, as substantial differences in model performances were observed (ranges around Pearson correlation coefficient 0.60–0.90, RMSE, 0.50–0.90, and R^2 0.30–0.80).

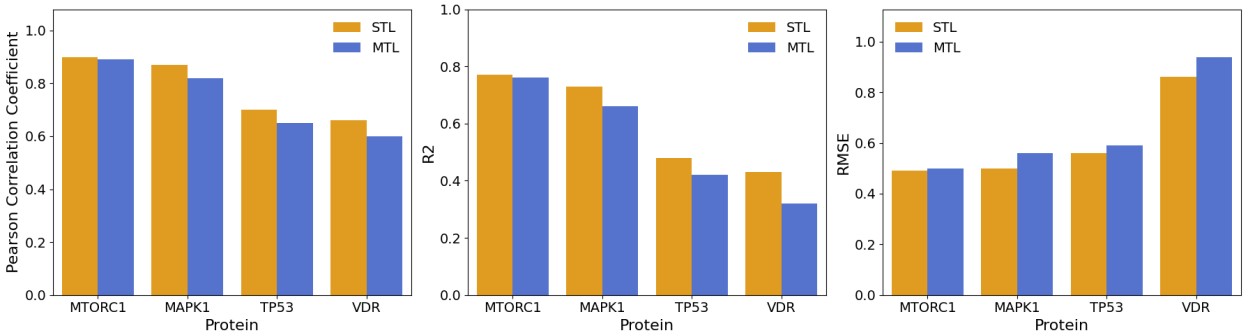


Figure 3.4. STL and MTL model performances for molecular docking score predictions. Model performances are reported on the testing dataset. Pearson Correlation Coefficient, R^2, and RMSE metrics (left to right).

Figure 3.5. summarizes model performances for strain prediction on the testing dataset. STL and MTL model performances (Pearson correlation coefficient, RMSE, and R^2) were near identical across all datasets. The MTL architecture did not contribute to improved model performance, but the incorporated other tasks did not interfere with strain prediction. This suggests that strain is a robust task in the joint learning environment. All datasets and molecular docking programs considered, STL and MTL models showed strong predictive power (Pearson correlation coefficients > 0.85 and R^2 > 0.70).

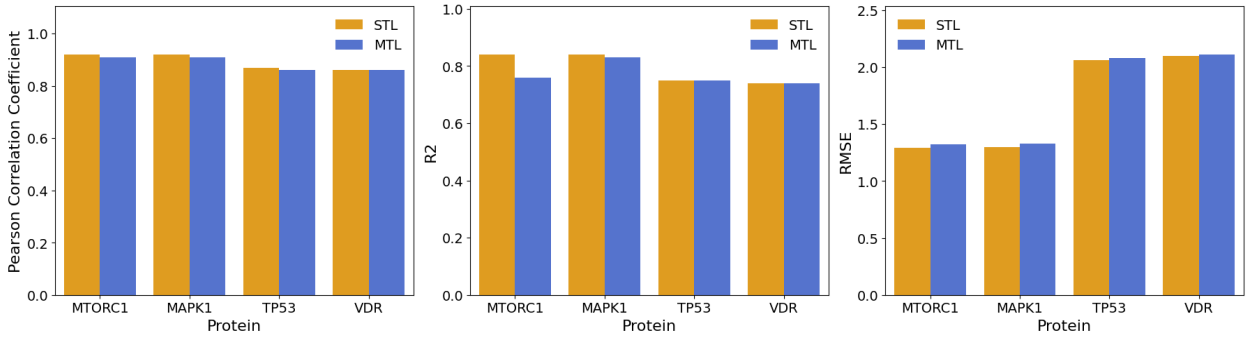


Figure 3.5. STL and MTL model performances for strain predictions. Model performances are reported on the testing dataset. Pearson Correlation Coefficient, R^2, and RMSE metrics (left to right).

Figure 3.6. summarizes model performances for unsatisfied hydrogen bond donor predictions on the testing dataset. STL and MTL model performances (class-weighted precision and recall) were near identical across all datasets – this similar pattern was also observed in the prediction of strain. The STL and MTL models showed strong performance (precision > 0.95 and recall > 0.95). Class-weighted precision and recall were near identical.

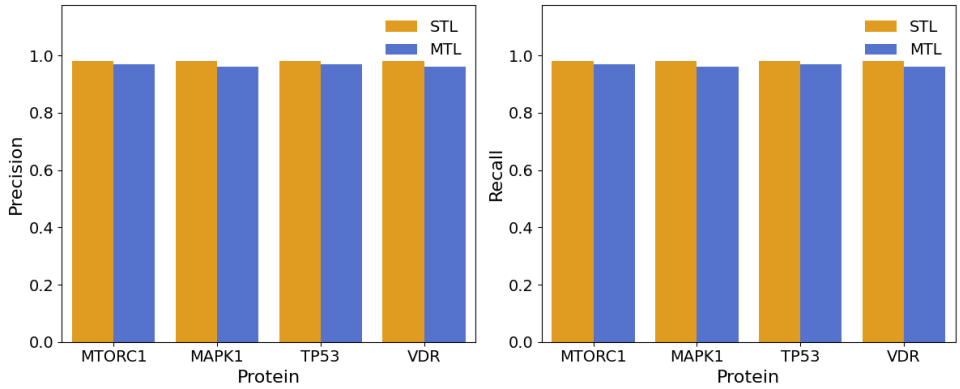


Figure 3.6. STL and MTL model performances for unsatisfied hydrogen bond donor predictions. Model performances are reported on the testing dataset. Precision (left) and recall (right) metrics.

Figure 3.7. summarizes model performances for unsatisfied hydrogen bond acceptor predictions on the testing dataset. STL showed better performances (class-weighted precision and recall) than MTL; the additional tasks in the MTL architecture did not improve model performance for unsatisfied hydrogen bond acceptor predictions. The STL models showed very strong performance (precision > 0.97 and recall > 0.97); MTL models showed a modest reduction compared to STL, but still strong performance (precision > 0.91 and recall > 0.91). Class-

weighted precision and recall were near identical – as seen in unsatisfied hydrogen bond donor predictions.

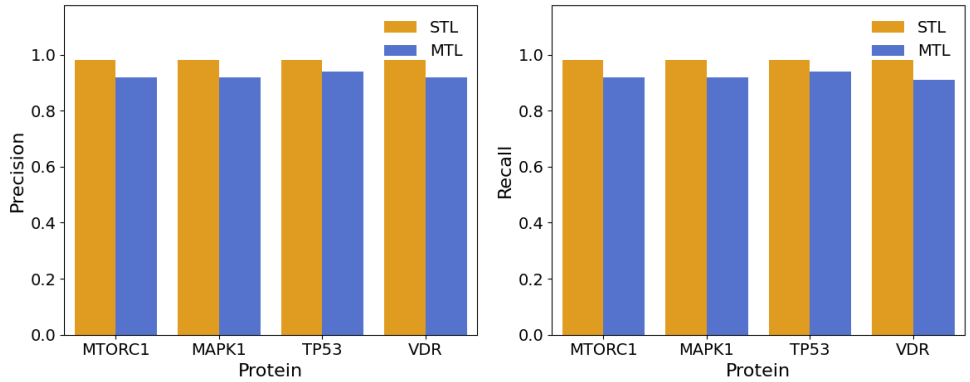


Figure 3.7. STL and MTL model performances for unsatisfied hydrogen bond acceptor predictions. Model performances are reported on the testing dataset. Precision (left) and recall (right) metrics.

Moreover, the MTL architecture reduced computational runtime (3-4x), Figure 3.8, at comparable model performances to STL. Although STL showed very slightly better performance overall, MTL provides significant runtime speed and resource advantages – this is of significant importance in ultra large-scale VS.

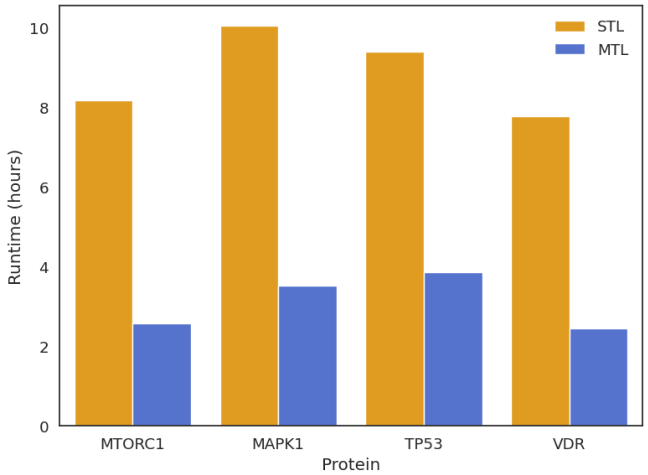


Figure 3.8. Model runtime comparisons between STL and MTL. Runtimes are the cumulative hours to run training, validation (HPO), and testing on molecular docking scores, strain, unsatisfied hydrogen bond donors, and unsatisfied hydrogen bond acceptors. Separate models (4) were run with STL; the cumulative runtime is shown.

3.2.2 Early Enrichment

Early enrichment in simulated large-scale VS was assessed with LogAUC (Section 2.2). All the selected proteins (MTORC1, MAPK1, TP53, and VDR) showed an improvement in LogAUC in benchmark (Table 3.1) when both strain and unsatisfied hydrogen bond filtering was considered. This suggested that these filters can be used to substantially improve enrichment in large-scale VS through the removal of artifactual, high scoring molecules.

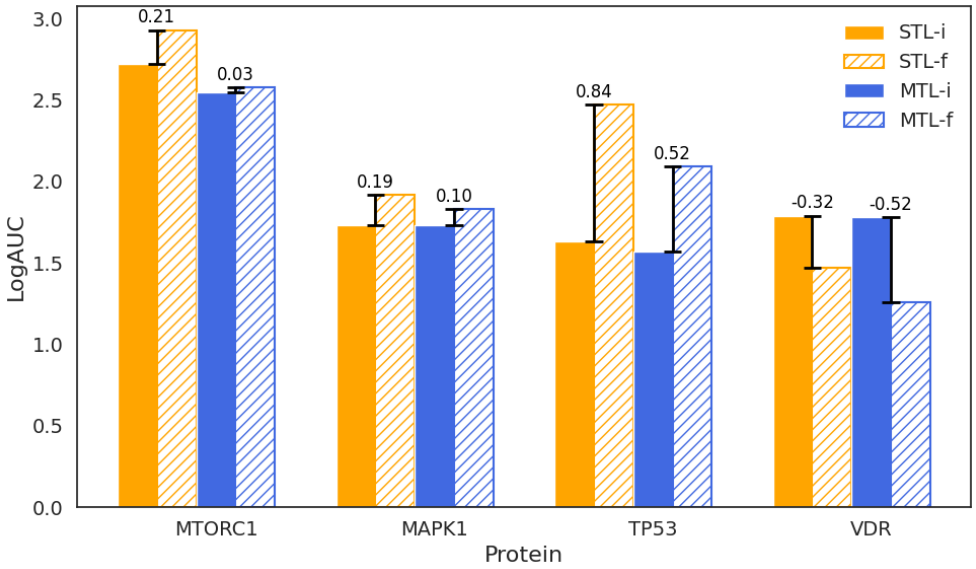


Figure 3.9. STL and MTL simulated large-scale VS results. LogAUC before (i) and after (f) filtering with strain and unsatisfied hydrogen bonds, and ΔLogAUC (vertical difference lines and values) are shown for each protein dataset at the thresholds reported in Table 3.1.

The simulated large-scale VS with STL and MTL models are shown, Figure 3.9. MTORC1 in simulated large-scale VS with the STL model showed a better improvement in ΔLogAUC (0.21) compared to the MTL model (0.03), after filtering with strain and unsatisfied hydrogen bonds. The same pattern is observed with MAPK1, where the STL model showed better improvement in ΔLogAUC (0.19) compared to the MTL model (0.10). TP53 showed the highest improvement in ΔLogAUC out of all the proteins explored in this simulated large-scale VS. The enrichment (LogAUC) with molecular docking scores alone was comparable to random classification (close to 1.45); strain and unsatisfied hydrogen bond filtering improved enrichment with ΔLogAUC values of 0.84 and 0.52 with the STL and MTL models, respectively. For VDR, both STL and MTL models did not improve enrichment with strain and unsatisfied hydrogen bonds filtering, ΔLogAUC is negative in both cases. These simulated large-scale VS demonstrated that strain and

unsatisfied hydrogen bond filtering can increase early enrichment and MTL can be used in this setting, but the performance is system-dependent.

The reduction in database size was observed as a major advantage to strain and unsatisfied hydrogen bond filtering, Figure 3.10, where enrichment (ΔLogAUC) is positive.

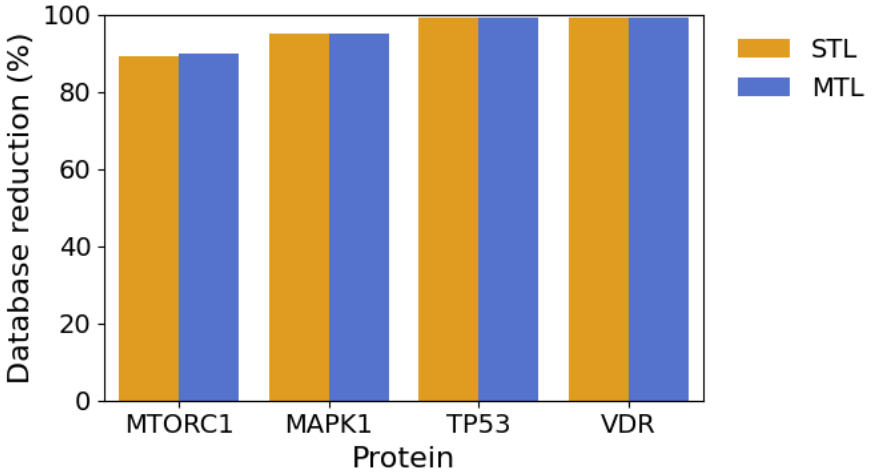


Figure 3.10. Database reduction with filtering in STL and MTL simulated large-scale VS. The filters applied are strain, unsatisfied hydrogen bond donors, and unsatisfied hydrogen bond acceptors.

MTORC1, MAPK1, and TP53 showed positive enrichment (ΔLogAUC) and significant database reductions (>90%) with strain and unsatisfied hydrogen bond filtering. This demonstrates that these selected filters can provide significant advantages in large-scale VS.

3.3 Conclusions

In this chapter, we explored how medicinal chemistry filters (strain and unsatisfied hydrogen bonds) can improve early enrichment in large-scale VS, and how MTL can be used to this end – however the performance and effectiveness is system-dependent. In MTORC1, MAPK1, and TP53 VS, these filters improved enrichment with both STL and MTL models. There are some examples in which the consideration of these filters is not productive, i.e. both STL and MTL models in the VS of VDR. STL models performed slightly better than MTL models across all VS in the predictions of all medicinal chemistry properties explored (molecular docking scores, strain, and unsatisfied hydrogen bonds), leading to slightly improved enrichments (LogAUC). This suggests that shared representation learning does not confer to better model performance in

this setting, but MTL offers substantial runtime reductions (3-4x) compared to STL. Moreover, this method showed significant database size reductions (>90%) across all VS. These resource advantages are particularly relevant in the VS of ultra-large chemical libraries in the ever-expanding chemical space.

3.4 References

(1) Liu, S.; Chen, M.; Yao, X.; Liu, H. Fingerprint-Enhanced Hierarchical Molecular Graph Neural Networks for Property Prediction. *J. Pharm. Anal.* **2025**, *15* (6), 101242. https://doi.org/10.1016/j.jpha.2025.101242.

(2) Morgan, H. L. The Generation of a Unique Machine Description for Chemical Structures-A Technique Developed at Chemical Abstracts Service. *J. Chem. Doc.* **1965**, *5* (2), 107–113. https://doi.org/10.1021/c160017a018.

(3) Rogers, D.; Hahn, M. Extended-Connectivity Fingerprints. *J. Chem. Inf. Model.* **2010**, *50* (5), 742–754. https://doi.org/10.1021/ci100050t.

(4) Capecchi, A.; Probst, D.; Reymond, J.-L. One Molecular Fingerprint to Rule Them All: Drugs, Biomolecules, and the Metabolome. *J. Cheminform.* **2020**, *12* (1), 43. https://doi.org/10.1186/s13321-020-00445-4.

(5) Yang, X.; Wang, Y.; Byrne, R.; Schneider, G.; Yang, S. Concepts of Artificial Intelligence for Computer-Assisted Drug Discovery. *Chem. Rev.* **2019**, *119* (18), 10520–10594. https://doi.org/10.1021/acs.chemrev.8b00728.

(6) Gentile, F.; Yaacoub, J. C.; Gleave, J.; Fernandez, M.; Ton, A.-T.; Ban, F.; Stern, A.; Cherkasov, A. Artificial Intelligence–Enabled Virtual Screening of Ultra-Large Chemical Libraries with Deep Docking. *Nat. Protoc.* **2022**, *17* (3), 672–697. https://doi.org/10.1038/s41596-021-00659-2.

(7) Gentile, F.; Agrawal, V.; Hsing, M.; Ton, A.-T.; Ban, F.; Norinder, U.; Gleave, M. E.; Cherkasov, A. Deep Docking: A Deep Learning Platform for Augmentation of Structure Based Drug Discovery. *ACS Cent. Sci.* **2020**, *6* (6), 939–949. https://doi.org/10.1021/acscentsci.0c00229.

(8) *Enamine REAL database*. Enamine.net. https://enamine.net/compound-collections/real-compounds/real-database.

(9) McInnes, L.; Healy, J.; Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv*. 2018. https://doi.org/10.48550/ARXIV.1802.03426.

(10) Kingma, D. P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv*. 2014. https://doi.org/10.48550/ARXIV.1412.6980.

(11) Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; Koyama, M. Optuna: A Next-Generation Hyperparameter Optimization Framework. *arXiv*. 2019. https://doi.org/10.48550/ARXIV.1907.10902.

(12) Kendall, A.; Gal, Y.; Cipolla, R. Multi-Task Learning Using Uncertainty to Weigh Losses for Scene Geometry and Semantics. *arXiv*. 2017. https://doi.org/10.48550/ARXIV.1705.07115.

3.5 Appendix

Supplementary Table 1. MTL model (best) training: learnable parameter (σ) in uncertainty-weighted method.

Docking Program	Protein	σ (u-hbd[a])	σ (u-hba[b])	σ (docking score)	σ (strain)
AD	MTORC1	0.29	0.49	0.54	0.95
	MAPK1	0.41	0.53	0.51	0.78
GNINA	TP53	0.24	0.41	0.59	1.78
	VDR	0.43	0.57	0.90	0.95

[a]unsatisfied hydrogen bond donors. [b]unsatisfied hydrogen bond acceptors.

Chapter 4: Conclusions and Future Directions

The work presented in this thesis outlines the research and design involved in the development of a new MOO ML-accelerated molecular docking model that includes the prediction of multiple, 3D medicinal chemistry properties (strain and unsatisfied hydrogen bonds) to aid in the selection of potential bioactive molecules, towards the development of next-generation drug discovery tools and therapeutics.

The second chapter in this work tested strain and unsatisfied hydrogen bond filters in the selection of hit candidates in simulated VS campaigns: GPCR-Bench and LIT-PCBA databases were docked with molecular docking programs AutoDock and GNINA and subjected to strain and unsatisfied hydrogen bond filters at different threshold combinations. These simulated campaigns showed system-dependence in the early enrichment of docking-ranked compound lists. In contrast to the common strain[1–5] and unsatisfied hydrogen bond[1,4,6–10] thresholds used in literature, no single threshold or combination improved enrichment across all protein-ligand datasets explored. However, several proteins showed significant enrichment with the tested filters at specific thresholds. Therefore, these filters can be effective post-processing tools in VS when retrospectively confirmed to aid the selection of candidates for the protein target of interest. This work contributed a strain and unsatisfied hydrogen bonds benchmark with an exhaustive search of all filter threshold combinations.

The third chapter in this work developed and tested a new MOO ML-accelerated molecular docking model that incorporates selected 3D medicinal chemistry filters to assess the impact on early enrichment in simulated large-scale VS. This method showed system-dependent performance – most systems studied showed an improvement in enrichment (ΔLogAUC) with values between 0.03-0.84. Moreover, there were substantial reductions in chemical database sizes (90-99%) across all simulated VS. MTL models showed the potential to improve early enrichment in large-scale VS, while accelerating runtime (3-4x) compared to STL, without significant performance loss. Existing MOO ML-accelerated molecular docking models[11,12] focus on the prediction of docking scores and other experimental, drug-like characteristics. This work is the first which aims to reduce artifact contamination in molecular docking through the incorporation of 3D medicinal chemistry properties (strain and unsatisfied hydrogen bonds) to

improve hit identification in large-scale VS.

To summarize, a new MOO ML-accelerated molecular docking model was developed to support artifact filtering and selection of promising candidates in the early stages of drug discovery. These models offer a method to explore and narrow down ultra-large chemical spaces in the search of novel drug candidates, but as with all current, available methods, its success is not universal – modelling protein-ligand binding is complex.

These models are inherently dependent on the quality of available data (protein structures and ligand activities) and the methods and tools used to acquire their labels. To address the observed system-dependence in future large-scale VS campaigns, the combination of objectives can be selected based on specific, available information about the pharmaceutical target of interest – including other tasks such as multiple docking scores, pharmacophore matching, and specific protein-ligand interactions could be explored. Testing other MTL partial/hybrid or soft parameter sharing architectures[13] could improve model performance through learning task-specialized representations. Moreover, active learning[14] could be incorporated in the model's architecture to label selected, informative data in an iterative manner for training, where the most informative unlabeled data is identified based on the likelihood of improving model performance.

This work provides an open-source, medicinal chemistry-informed ML-accelerated molecular docking model in contribution to the development of new drug discovery tools. To validate the effectiveness of these models in early enrichment, prospective, ultra large-scale VS on a pharmaceutical target of interest should be conducted, proceeded with experimental testing on the top virtual hit candidates.

4.1 References

(1) Liu, F.; Wu, C.-G.; Tu, C.-L.; Glenn, I.; Meyerowitz, J.; Kaplan, A. L.; Lyu, J.; Cheng, Z.; Tarkhanova, O. O.; Moroz, Y. S.; Irwin, J. J.; Chang, W.; Shoichet, B. K.; Skiniotis, G. Large Library Docking Identifies Positive Allosteric Modulators of the Calcium-Sensing Receptor. *Science* **2024**, *385* (6715), eado1868. https://doi.org/10.1126/science.ado1868.

(2) Lyu, J.; Kapolka, N.; Gumpper, R.; Alon, A.; Wang, L.; Jain, M. K.; Barros-Álvarez, X.; Sakamoto, K.; Kim, Y.; DiBerto, J.; Kim, K.; Glenn, I. S.; Tummino, T. A.; Huang, S.; Irwin, J. J.; Tarkhanova, O. O.; Moroz, Y.; Skiniotis, G.; Kruse, A. C.; Shoichet, B. K.; Roth, B. L. AlphaFold2 Structures Guide Prospective Ligand Discovery. *Science* **2024**, *384* (6702), eadn6354. https://doi.org/10.1126/science.adn6354.

(3) Fink, E. A.; Xu, J.; Hübner, H.; Braz, J. M.; Seemann, P.; Avet, C.; Craik, V.; Weikert, D.; Schmidt, M. F.; Webb, C. M.; Tolmachova, N. A.; Moroz, Y. S.; Huang, X.-P.; Kalyanaraman, C.; Gahbauer, S.; Chen, G.; Liu, Z.; Jacobson, M. P.; Irwin, J. J.; Bouvier, M.; Du, Y.; Shoichet, B. K.; Basbaum, A. I.; Gmeiner, P. Structure-Based Discovery of Nonopioid Analgesics Acting through the α_{2A} -Adrenergic Receptor. *Science* **2022**, *377* (6614), eabn7065. https://doi.org/10.1126/science.abn7065.

(4) Alon, A.; Lyu, J.; Braz, J. M.; Tummino, T. A.; Craik, V.; O'Meara, M. J.; Webb, C. M.; Radchenko, D. S.; Moroz, Y. S.; Huang, X.-P.; Liu, Y.; Roth, B. L.; Irwin, J. J.; Basbaum, A. I.; Shoichet, B. K.; Kruse, A. C. Structures of the Σ2 Receptor Enable Docking for Bioactive Ligand Discovery. *Nature* **2021**, *600* (7890), 759–764. https://doi.org/10.1038/s41586-021-04175-x.

(5) Gu, S.; Smith, M. S.; Yang, Y.; Irwin, J. J.; Shoichet, B. K. Ligand Strain Energy in Large Library Docking. *J. Chem. Inf. Model.* **2021**, *61* (9), 4331–4341. https://doi.org/10.1021/acs.jcim.1c00368.

(6) Sindt, F.; Seyller, A.; Eguida, M.; Rognan, D. Protein Structure-Based Organic Chemistry-Driven Ligand Design from Ultralarge Chemical Spaces. *ACS Cent. Sci.* **2024**, *10* (3), 615–627. https://doi.org/10.1021/acscentsci.3c01521.

(7) Sindt, F.; Bret, G.; Rognan, D. On the Difficulty to Rescore Hits from Ultralarge Docking Screens. *J. Chem. Inf. Model.* **2025**, *65* (11), 5553–5566. https://doi.org/10.1021/acs.jcim.5c00730.

(8) Zhou, G.; Rusnac, D.-V.; Park, H.; Canzani, D.; Nguyen, H. M.; Stewart, L.; Bush, M. F.; Nguyen, P. T.; Wulff, H.; Yarov-Yarovoy, V.; Zheng, N.; DiMaio, F. An Artificial Intelligence Accelerated Virtual Screening Platform for Drug Discovery. *Nat. Commun.* **2024**, *15* (1), 7761. https://doi.org/10.1038/s41467-024-52061-7.

(9) Chakrabarti, M.; Tan, Y. S.; Balius, T. E. Considerations Around Structure-Based Drug Discovery for KRAS Using DOCK. In *KRAS*; Stephen, A. G., Esposito, D., Eds.; Methods in Molecular Biology; Springer US: New York, NY, 2024; Vol. 2797, pp 67–90. https://doi.org/10.1007/978-1-0716-3822-4_6.

(10) Singh, I.; Li, F.; Fink, E. A.; Chau, I.; Li, A.; Rodriguez-Hernández, A.; Glenn, I.; Zapatero-Belinchón, F. J.; Rodriguez, M. L.; Devkota, K.; Deng, Z.; White, K.; Wan, X.; Tolmachova, N. A.; Moroz, Y. S.; Kaniskan, H. Ü.; Ott, M.; García-Sastre, A.; Jin, J.; Fujimori, D. G.; Irwin, J. J.; Vedadi, M.; Shoichet, B. K. Structure-Based Discovery of Inhibitors of the

SARS-CoV-2 Nsp14 N7-Methyltransferase. *J. Med. Chem.* **2023**, *66* (12), 7785–7803. https://doi.org/10.1021/acs.jmedchem.2c02120.

(11) Graff, D. E.; Shakhnovich, E. I.; Coley, C. W. Accelerating High-Throughput Virtual Screening through Molecular Pool-Based Active Learning. *Chem. Sci.* **2021**, *12* (22), 7866–7881. https://doi.org/10.1039/D0SC06805E.

(12) Mehta, S.; Goel, M.; Priyakumar, U. D. MO-MEMES: A Method for Accelerating Virtual Screening Using Multi-Objective Bayesian Optimization. *Front. Med.* **2022**, *9*, 916481. https://doi.org/10.3389/fmed.2022.916481.

(13) Ruder, S. An Overview of Multi-Task Learning in Deep Neural Networks. arXiv 2017. https://doi.org/10.48550/ARXIV.1706.05098.

(14) Wang, L.; Zhou, Z.; Yang, X.; Shi, S.; Zeng, X.; Cao, D. The Present State and Challenges of Active Learning in Drug Discovery. *Drug Discov. Today* **2024**, *29* (6), 103985. https://doi.org/10.1016/j.drudis.2024.103985.