Predicting Inhibitors of OATP1B1 via Heterogeneous OATP-Ligand Interaction Graph Neural Network (HOLI-GNN)

Mehrsa Mardikoraem[a,b,1], Joelle N. Eaves[a,b,1], Theodore Belecciu[a,b], Nathaniel Pascual[a,b], Alexander Aljets[a,b], Benjamin J. Orlando[c], Erik M. Shapiro[a,b,d], Daniel R. Woldring[a,b,*]

[a] Institute for Quantitative Health Science and Engineering, Michigan State University, 775 Woodlot Dr, East Lansing, MI 48824, USA

[b] Department of Chemical Engineering and Materials Science, Michigan State University, 428 S Shaw Ln, East Lansing, MI 48824, USA

[c] Department of Biochemistry and Molecular Biology, Michigan State University, 603 Wilson Rd, East Lansing, MI 48823, USA

[d] Department of Radiology, Michigan State University, 846 Service Rd, East Lansing, MI 48824, USA

* Daniel Woldring

Email: woldring@msu.edu

Author Contributions: B.O., E.S., and D.W., conceived the project; M.M., J.E., and D.W. designed research; J.E., M.M., D.W., and T.B., analyzed data and wrote the paper; all authors contributed to discussions about the project and editing the manuscript prior to submission.

[1] M.M. and J.E. contributed equally to this work.

Competing Interests Statement: The authors declare no competing interests.

Classification: Biological Sciences, Biophysics and Computational Biology

Keywords: machine learning, drug-drug interactions, molecular modeling

Abstract

Organic anion transporting polypeptides (OATPs) are crucial for hepatic drug uptake, influencing drug efficacy and toxicity. Predicting OATP-mediated drug-drug interactions (DDIs) is challenging due to limited structural data and inconsistent experimental OATP inhibition data across studies. This study introduces Heterogeneous OATP-Ligand Interaction Graph Neural Network (HOLI-GNN), a novel computational approach that integrates molecular modeling with graph neural networks to enhance the prediction of OATP-mediated drug inhibition. By combining ligand molecular features with protein-ligand interaction data, HOLI-GNN outperforms traditional ligand-based methods. HOLI-GNN achieved median F1 and AUC scores of 0.78 and 0.90, respectively, compared to ECFP- and RDKit-based models built upon XGBoost (F1: 0.68 and 0.78, respectively; AUC: 0.70 and 0.75, respectively). Beyond improving inhibition prediction, we characterize protein residues involved in inhibitory versus non-inhibitory drug interactions, specifically highlighting residues T42, F224, I353, F356, and F386. We speculate that local position shifts in these hydrophobic packing residues, or the inhibition thereof, may be an important aspect of competitive inhibition mechanisms. Our model enhances the performance of OATP inhibitor prediction and, critically, offers interpretable interaction information to inform future mechanistic investigations.

Significance Statement

Concurrent administration of different drugs can cause potentially lethal drug-drug interactions (DDIs), and membrane protein transporters like OATPs can mediate such DDIs. While many current models predict OATP-mediated DDIs, all thus far rely solely on drug features without considering intricate drug-OATP interactions. In this work, we present HOLI-GNN, a graph neural network that leverages both drug and OATP-drug interaction features to predict OATP inhibition. The use of OATP-drug interaction features in the prediction was made possible by the recent publication of cryo-EM structures for OATP1B1 and high-throughput protein-ligand docking. We demonstrate that HOLI-GNN outperforms conventional OATP-mediated DDI predictors which rely solely on drug features, while enabling important mechanistic insights into OATP transport.

Main Text

Introduction

Drug-drug interactions (DDIs) pose a significant challenge in modern pharmacotherapy, as they can lead to the impairment or induction of critical metabolic enzymes and transporter proteins responsible for processing multiple drugs simultaneously. Impairment of certain drug transporters, for instance, can result in toxic serum concentrations of commonly administered drugs like statins, which in turn can cause life-threatening conditions like rhabdomyolysis (1). Such interactions can be further exacerbated by genetic variation (2, 3). Recognizing the importance of predictive DDI models in the drug development process, the US Food and Drug Administration (FDA) has recently issued guidelines advocating for their consideration in regulatory pathways for new drugs (4, 5). While such models have been well-established for metabolic DDIs involving cytochrome P450s (6), there is a notable lack of predictive models for transporter-mediated DDIs (tDDIs). This deficiency is particularly concerning in the context of organic anion transporting polypeptides (OATPs), a family of proteins responsible for transporting a wide range of drugs into the liver for bile-mediated elimination. Numerous alarming tDDIs involving OATPs have been reported, highlighting the urgent need for the development of reliable predictive tDDI models (7, 8). Addressing this knowledge gap is crucial for ensuring the safety and efficacy of medications to ultimately improve patient outcomes.

In addition to hepatic expression for drug transport, OATPs are expressed in other pharmacologically relevant tissues (e.g., kidneys and intestines), where they facilitate the uptake

of a wide variety of endogenous and exogenous small molecules, including hormones and drugs. To accommodate the vast number and variety of substrates processed by these tissues, OATPs exhibit a high degree of promiscuity. This promiscuity gives rise to intricate intermolecular interactions and elaborate regulatory networks that can profoundly influence drug efficacy and toxicity. Understanding the complexities of these interactions is essential for predicting and mitigating the potential risks associated with OATP-mediated drug transport. By elucidating the mechanisms underlying OATP promiscuity and the resulting regulatory networks, researchers can develop more comprehensive and accurate models for predicting tDDIs.

Currently, there is no generalizable and consistent method to accurately predict how a molecule will interact with OATPs, despite valiant efforts to develop machine learning (ML) algorithms to do so (9–12). This is due to two major challenges. First, until very recently (13, 14), no OATP structures had been experimentally solved, substantially limiting options for accurate structure-based analyses. Secondly, there is significant variability in the available in vitro data, resulting in inconsistent data labeling (15, 16).

In the absence of solved OATP structures, past ML models were limited to only physiochemical and/or topological ligand properties, precluding quality protein-ligand interaction analyses. Compounding this limitation is the dynamic nature of OATPs, which are believed to possess a rocker-switch alternating access transport mechanism (13). Recently published cryo-EM structures of OATP1B1 (inward- and outward-facing conformations) and OATP1B3 (inward-facing conformation only) have enabled us to thoroughly explore how protein-ligand interaction features (PLIFs) may be integrated into ML models (13, 14). We present HOLI-GNN, a Heterogeneous OATP-Ligand Interaction Graph Neural Network, as a novel approach that notably outperforms conventional ligand-based models. Taking advantage of both the inward- and outward-facing cryo-EM structures of OATP1B1 (13, 14) and the enhanced spatial sampling afforded by molecular docking simulations, we establish how PLIF-based workflows may expand the breadth of predictive tDDI models.

The lack of consistent in vitro data constrains model applicability to the data on which it was trained. In large part, sources of variability have been attributed to the dependence of inhibitory potency on probe substrate (15–18). For example, inhibition constant (Ki) values for gemfibrozil were measured using 15 different OATP probe substrates. Measured Ki values varied up to 39-fold (9.65 – 381 μM) when measured with estrone-3-sulfate and pravastatin, respectively (15). This variation in values has resulted in ligands being labeled inconsistently in different datasets (i.e., as inhibitors, noninhibitors, active, inactive). Additionally, recent ML-based tDDI prediction models (12, 19) were trained and evaluated on such inconsistent datasets. Others, like Tuerkova et al (20), have made extensive efforts to curate datasets of ligands without conflicting labels, while still considering data from a variety of probe substrates. Drawing from these lessons, we opted to limit our study to a single set of experiments by Karlgren et al. with consistent ligand labeling and a consistent probe substrate (10). The probe substrate in this dataset, estradiol-17β-glucuronide (E2G), has been shown to minimize the risk of false-negative DDI predictions (16).

Furthermore, there is growing evidence of transcriptional, post-translational, and/or degradative regulation involving either the probe substrate or inhibitor ligand (8). While HOLI-GNN was designed under the simplifying assumption of competitive inhibition, we have identified several ligands for which it performs poorly, including dipyridamole, ketoconazole, and telmisartan. These ligands were all classified as OATP1B1 inhibitors based on in vitro assays (10), and each of these ligands are known to be involved in downstream metabolic and/or transport processes (21–24). Thus, we suggest that these ligands be prioritized for future investigations of indirect pathways of OATP transport regulation.

Herein, we present a workflow to extract PLIFs from OATP1B1 conformers docked against a library of experimentally characterized ligands. Collectively, we explore both the powerful utility and

current limitations of PLIF-based tDDI models. Our model, HOLI-GNN, achieves median F1 and AUC scores of 0.78 and 0.90, respectively. Results highlight key differences between OATP1B1 residues involved in interactions with inhibitors and those involved in interactions with noninhibitors. We expand our discussion on the insights of PLIF-centric models and suggest future directions to investigate both competitive and indirect inhibition mechanisms.

Results

To address the challenge of predicting OATP-mediated drug-drug interactions (DDIs), we developed a comprehensive computational approach that integrates molecular modeling, ML classifiers, and graph neural networks (GNNs). This study unfolds in several key stages. Initially, we employed protein-ligand docking simulations followed by a protein-ligand interaction profiler (PLIP) (25) to extract protein-ligand interaction features (PLIFs). These PLIFs enabled comparisons between key residues involved in interactions with inhibitors versus noninhibitors, highlighting specific OATP1B1 residues (T42, A45, P220, F224, I353, F356, F386, and I579) that show notable differences in their interactions. We utilized PLIFs to build and optimize machine learning classifiers, exploring how their predictive performance was impacted by multi-state sampling. Multi-state sampling involved using both inward- and outward-facing OATP conformers and varying the number of docking simulation results (poses) used per protein-ligand pair.

The culmination of this process was the development of HOLI-GNN, designed to capture intricate intra- and inter-molecular interactions between drugs and proteins to enhance inhibition prediction accuracy. For comparison, we also constructed a simplified homogeneous GNN containing only ligand graphs, excluding any protein or protein-ligand interaction information, termed Ligand-GNN. HOLI-GNN achieved median F1 and AUC scores of 0.78 and 0.90, respectively, outperforming Ligand-GNN (F1: 0.69, AUC: 0.86). Both HOLI-GNN and Ligand-GNN models outperformed more traditional, ligand based-methods relying on extended-connectivity fingerprints (ECFPs) and RDkit features (F1: 0.68 and 0.78, respectively; AUC: 0.70 and 0.75, respectively). In addition to improving predictive performance, we hope our approach provides more comprehensive insights to guide future mechanistic studies of OATP-mediated DDIs.

Structure-Based Protein-Drug Interaction Prediction.

We explored the effectiveness of structure-based and drug-based features in predicting ligand-OATP1B1 interactions. Our analyses identified critical residues and interaction types that differentiate inhibitory compounds from noninhibitors, which showed promising agreement with recent findings on OATP transport. For instance, OATP1B1 inhibitors were seen to have significantly more hydrophobic interactions with residues T42, A45, I46, K49, and P220 than noninhibitors (**Figure S4**). This observation aligns with the findings of Shan et al. (13), highlighting the hydrophobic packing and constriction network(s) involved in the binding, transport, and/or substrate-dependent inhibition. Moreover, comparison of interaction profiles between inward and outward conformers yields suggestions for future mechanistic studies (e.g., functional mutagenesis studies to investigate the importance of positions F224, T42, I353 in transport and/or inhibition).

Molecular Docking Simulations Generated Novel OATP-Ligand Interaction Interfaces.

Interaction interfaces between OATP1B1 and ligands were produced via protein-ligand docking simulations. Using high-quality cryo-EM structures of both the inward- and outward-facing conformers of OATP1B1 (13), we conducted thorough docking simulations to generate 1000 docked poses with each of the 222 ligands experimentally characterized by Karlgren et al (10). PLIFs were extracted from each docked pose to capture the intricacies of potential interactions (e.g., interaction distances, angles, atoms involved).

Interacting Residues in Docked Poses are Generally Consistent with Experimental Findings.

To identify key residues that interact more often with inhibitors, the distribution of interactions per OATP1B1 residue was analyzed. **Figure 1** (left) highlights the frequency of any interaction type

occurring with each orthosteric site residue (Y352, F356, and F386) identified by Shan et al. (13) through cryo-EM and functional investigations. The orthosteric site was defined by Shan et al. to be the set of OATP1B1 residues experimentally observed to participate in interactions with many different ligands. For docked poses with the inward-facing conformer (**Figure 1A**, left), inhibitors were observed to have significantly more interactions with each phenylalanine residue in the orthosteric site than noninhibitors (p< 5E-4 for F356 and F386). Conversely, there was no notable difference in the frequency with which inhibitors and noninhibitors interacted with the third orthosteric site residue, Y352, for the inward-facing conformer. Similarly, the distribution for docked poses with the outward-facing conformer (**Figure 1B**, left) showed Y352 did not interact significantly more with inhibitors or noninhibitors, while F356 was more likely to interact with inhibitors.

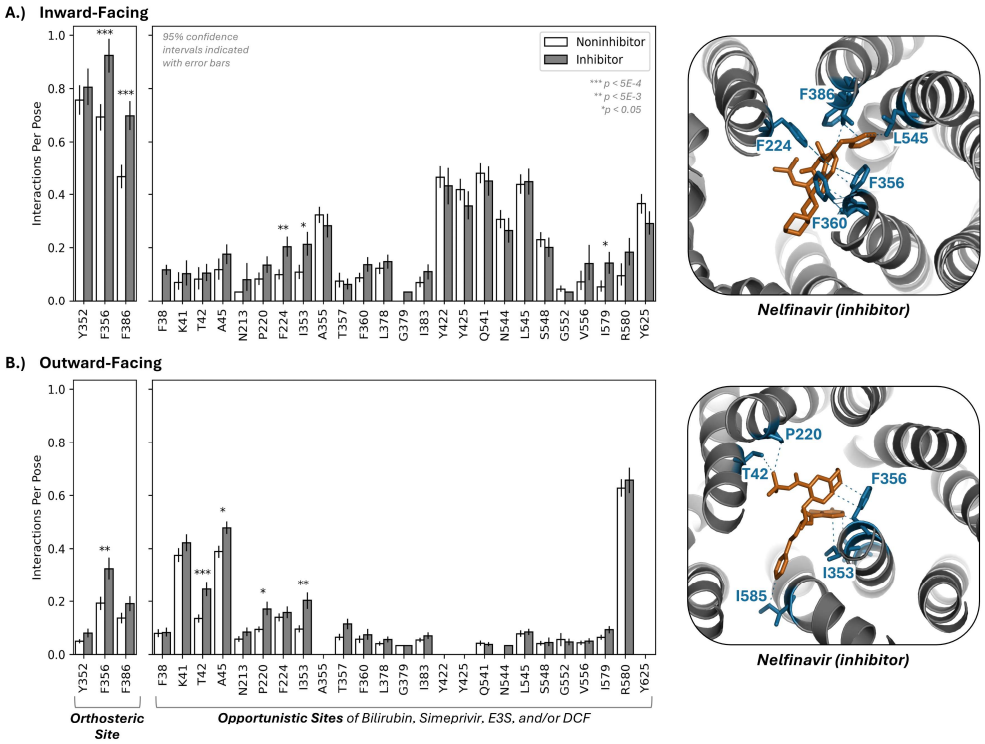

Figure 1. Protein-ligand interaction profiles highlight significant differences in OATP1B1 interaction sites between inhibitors and noninhibitors. Probability of any type of interaction occurring between noninhibitors (white) or inhibitors (grey) and OATP1B1 residues in the previously identified orthosteric site (left) and opportunistic sites (right) for the inward (A) and outward (B) conformers. Probabilities independently normalized for each class. Error bars indicate 95% confidence intervals. Mann-Whitney U p-values with Bonferroni correction indicated by asterisks (*: p< 0.05, **: p< 5e-3, ***: p < 5e-4). Most stable pose of Nelfinavir (orange) docked with the inward (A) and outward (B) OATP1B1. Protein-ligand interactions identified by PLIP (25) shown in blue.

Likewise, interaction distributions were visualized for any residues in the experimentally determined opportunistic sites (13). Opportunistic sites were defined by Shan et al. to be groups of residues important for the binding and/or transport of only certain ligands (i.e., much more ligand-specific than orthosteric sites). It is worth noting that opportunistic site residues participating in interactions indicative of ligand class (i.e., inhibitor, noninhibitor) differ greatly between the inward and outward conformers. For inward-facing poses, residues F224, I353, and I579 participated in significantly more interactions with inhibitors than with noninhibitors. For outward-facing poses, residues T42, P220, and I353 were observed to have significantly more interactions with inhibitors than with

noninhibitors. These findings illustrate the utility of PLIF-based methods for elucidating intricate interaction patterns that can inspire both predictive models and further mechanistic studies.

HOLI-GNN Captures Intricate Protein-Ligand Interactions and Improves Inhibition Prediction. Ultimately, PLIFs were employed to construct our novel HOLI-GNN model. We compared the predictive performance of HOLI-GNN to (1) ligand-centric (ECFP, RDKit) and PLIF-based XGBoost classification models, and (2) Ligand-GNN, a ligand-only GNN that employs comparable state-of-the-art methods to those used in HOLI-GNN.

Our results (**Figure 2**) demonstrate that, with graph representation of the outward-facing OATP1B1 conformer, HOLI-GNN significantly outperforms both ligand-centric and PLIF-based XGBoost classification models in terms of AUC. This is notable, as the same significant improvement is not observed for the Ligand-GNN, indicating that the inclusion of a heterogeneous protein-ligand interaction graph benefits predictive ability. Model performance is further detailed in Table S10.

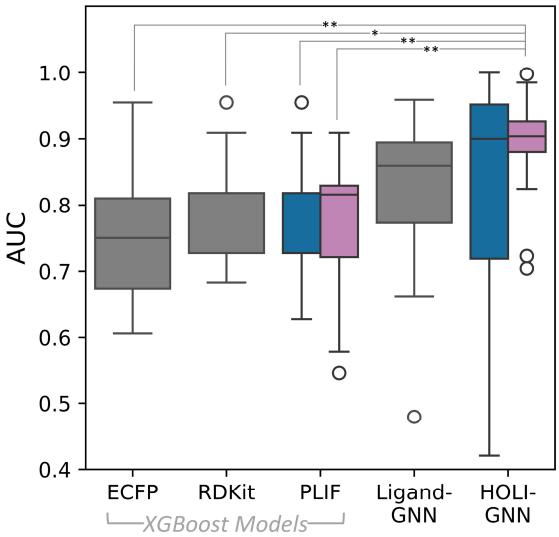

Figure 2. HOLI GNN outperforms ligand-centric and PLIF-based methods using XGBoost when trained on docked poses of outward-facing OATP1B1. Box plots displaying the distribution of scores for ligand-centric XGBoost (ECFP, RDKit), and PLIF XGBoost, Ligand-GNN, and HOLI-GNN (left to right). PLIF-based models for the inward- and outward-facing OATP1B1 conformers were evaluated independently (blue and pink, respectively). Mann-Whitney U p-values with Bonferroni correction indicated by asterisks (*: $p < 0.05$, **: $p < 5e\text{-}3$) show that HOLI-GNN (applied to the outward-facing OATP1B1 conformer) improves AUC scores compared to XGBoost models.

To achieve optimal HOLI-GNN performance, as shown in **Figure 2**, we introduced simple machine learning engineering techniques to maximize the signal obtained from our computational modeling, effectively separating meaningful patterns from noise. This comprehensive approach not only enhanced our model's performance but also contributed to a deeper understanding of OATP inhibition mechanisms.

Consideration of Multiple Docked Poses Expands Binding Site Coverage at the Expense of Signal. Rosetta docking simulations produced 1,000 docked poses of varying quality for each ligand. The Rosetta binding interface energy score was used as a quality metric, with lower energies indicating more stable poses (26). **Figure 3A** illustrates energy distributions of the 100 most stable poses for both inhibitors and noninhibitors docked to the inward and outward conformers (upper and lower panels, respectively). Interestingly, noninhibitors docked to the inward conformer are generally more stable than inhibitors. Conversely, noninhibitors docked to the outward conformer are generally less stable than inhibitors, although to a lesser extent (statistical support in Table S4). This may lend some support to the hypothesis that inhibitors can impair conformational change, whereas transported ligands promote the outward-to-inward conformational change to achieve marginally more stable interactions (13). To investigate the impact of pose sampling on model performance, datasets were created with the best one, two, three, five, ten, 20, 30, 50, 75, and 100 poses per ligand.

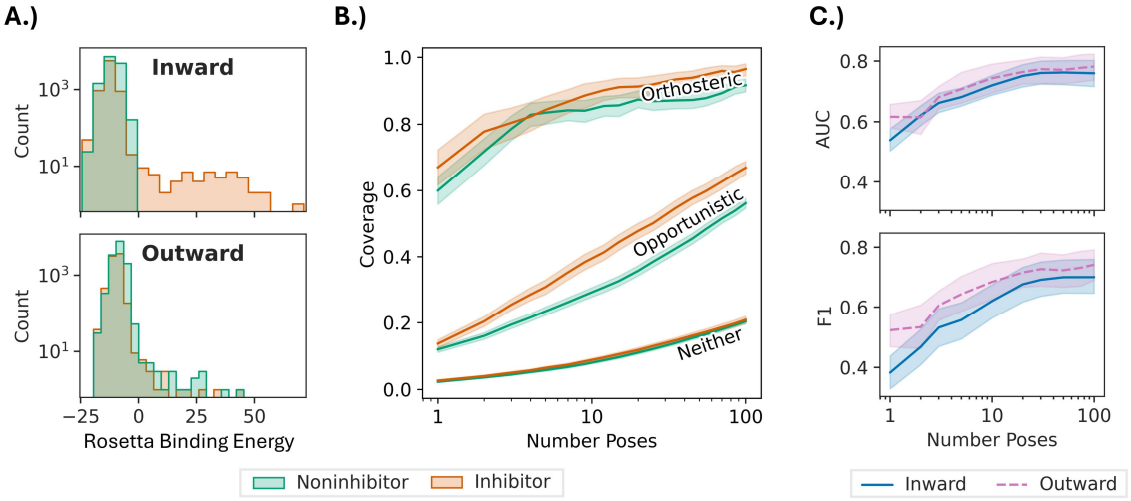

Figure 3. Utilizing multiple docked poses improves model performance, increases coverage of important residues, and reveals differences between the ligand classes; however, it also increases noise. (A) Distribution of Rosetta binding interface energy scores (interface_delta_X, in Rosetta energy units) of the best 100 poses for each ligand docked to the inward- and outward-facing OATP1B1 conformers (top and bottom panels for A, respectively). (B) Fraction of unique residues from the orthosteric site, opportunistic sites, and the remaining protein sequence captured in PLIFs as a function of number of poses. Experimentally characterized noninhibitors shown in green, and inhibitors in red. (C) Performance of the PLIF-based XGBoost models for the inward (blue, solid) and outward (pink, dashed) as a function of the number of docked poses used. Improvement in model performance indicated by AUC (top) and F1 (bottom) plateaus at approximately 30 poses. Shaded regions denote 95% confidence intervals.

The residue-interaction distributions of our docking-derived PLIFs were evaluated for consistency with experimentally identified interaction sites. Our analyses focused on three groups of OATP1B1 residues: (1) orthosteric site residues, (2) opportunistic sites, and (3) residues not part of the orthosteric or opportunistic sites. Specifically, we considered a broadly encompassing set of residues present in opportunistic sites of Simeprevir, estrone-3-sulfate, bilirubin, and 2',7'-dichlorofluorescein (DCF) (13). Herein, we refer to this grouping as the 'opportunistic sites' for brevity. As expected, increasing the number of docking poses included in our datasets resulted in a greater diversity of OATP1B1 residues captured in the PLIFs of each ligand. **Figure 3B** demonstrates how increasing the number of docked poses increases the coverage of the orthosteric, opportunistic, and other OATP1B1 residues in our PLIF datasets.

While expanding the pose sampling space may have increased the binding site coverage in the interaction profiles, it was also responsible for increased noise. Prior to GNN implementation, we assessed simpler classification model performance as a function of number of poses per ligand to optimize our model. The performance of XGBoost models were observed to peak when approximately 30 poses per ligand were included in the PLIF dataset for both the inward and outward conformers (**Figure 3C**). This prompted our use of the 30 best poses for each ligand in HOLI-GNN.

Discussion

ML techniques have been extensively applied to predict interactions between drugs and OATPs, primarily being used in binary classifications (e.g., inhibitor/noninhibitor). Previous models concentrated solely on drug attribute features, likely due to the lack of experimental OATP structural information (9–12, 19, 27). We developed HOLI-GNN to identify intricate patterns in docking-derived protein-ligand interaction features (PLIFs) which support our hypothesis that the OATP transport mechanism reveals important structural determinants of inhibition. The improved predictive performance of HOLI-GNN compared to ligand-centric approaches highlights the benefit of incorporating both ligand and protein structural information.

PLIF-Guided Hypotheses for Competitive Inhibition Mechanism.
The recently solved cryo-EM structures of OATP1B1 provide support for the hypothesized rocker-switch alternating access transport model, wherein conformational change between outward- and inward-facing states involves rigid-body rotations of both the N- and C-terminal domains (NTD, CTD) (13). However, due to limited evidence regarding OATP inhibition mechanisms, it is unclear whether inhibitory ligands are only partially transported and block other molecules from passing through, or if they simply prevent conformational change, allosterically or otherwise (28). In other members of the solute carrier protein family, conformationally prohibitive inhibition has been observed, indicating that this may be the case for OATPs also. For example, dilazep, a drug used to treat renal disorders, is known to block conformational changes in human equilibrative nucleoside transporter 1 (hENT1), another transporter believed to function according to the rocker-switch alternating access transport model (28).

Molecular docking simulations enabled extraction of extensive protein-ligand interaction information. From docking-derived PLIFs, the interesting findings captured in **Figure 1** lead us to two hypotheses for elucidating orthosteric sites in future studies. First, we hypothesize that Y352 is more likely involved in general ligand coordination rather than specific inhibitor binding since it shows no significant interaction difference between inhibitors and noninhibitors for both OATP1B1 conformations. Second, we hypothesize that noninhibitors are less likely to interact with F356 in the outward conformer compared to inhibitors. This may be caused by transported substances being less stabilized than inhibitors.

Multistate Sampling of Docked Poses Extends Perspective Beyond Static Cryo-EM Structures.
Using an ensemble of multiple docked poses per ligand provided a broad view of transporter interactions, both conformationally and energetically. We posit that our enhanced spatial sampling technique (i.e., using an ensemble of 30 docked poses) may be more representative of interactions with transporters relative to the static representations offered by single docked poses or cryo-EM structures. Comparisons between our docking-derived PLIFs to those of experimentally determined, ligand-bound OATP1B1 structures (13) confirmed that our synthetic docked structures not only agreed with the experimental ligand-bound structures but also revealed a broader range of interactions. **Figure S10** presents the distribution of residues involved in PLIFs for the inward and outward cryo-EM structures bound to estrone-3-sulfate and Simeprevir, respectively, in green. For comparison, the PLIF distributions are also displayed for the docked structures, considering only the best pose (pink) and the best 100 poses (orange) for each ligand. The sampling space expansion certainly increases noise (by including interacting residues from unlikely binding poses, **Figure 3B**), however, our simple classifier models allowed us to make great progress in striking a balance between binding site coverage and signal-to-noise ratio.

Considerations for the Competitive Inhibition Assumption.
HOLI-GNN relies on the assumption that inhibitors act via a competitive inhibition mechanism, as our feature representations encompass PLIFs specific to the OATP1B1 transport channel. Competitive inhibition was also assumed in the study that provided the experimental labels for our model (10), as is common practice in OATP1B1 inhibition assays (11, 29). The validity of this

assumption—as applied to both HOLI-GNN and experimental data—has been called into question by emerging evidence of alternative mechanisms by which certain ligands may influence OATP function (8, 30, 31).

Convolution arises from the possible involvement of metabolism, efflux, or non-OATP1B1 uptake of investigated ligands. Indirect inhibition of OATP1B1 may occur as transcriptional modulation via the farnesoid X receptor (FXR), hepatocyte nuclear factor (HNF) 4α and 1α, and liver X receptor (LXR) α (31). In part, this has motivated the standard use of the human embryonic kidney (HEK) 293 cell line for in vitro OATP inhibition assays due to their low basal expression levels of many transporter proteins (30). However, this does not consider the effect of ligand uptake on expression modulation of transporters and/or metabolic enzymes. For example, in developing kidney cells, HNF1α and HNF4α are suspected to regulate expression of several organic anion transporters (OATs) (32–34), which are known to facilitate uptake of multiple ligands in our dataset (31).

Further confounding processes include ligand-induced post-translational and/or degradative regulation. For instance, the protein kinase C activator phorbol 12-myristate 13-acetate (PMA), a compound widely studied in cancer and hematology research (35, 36), results in the phosphorylation and subsequent deactivation of OATP1B3 (37). The extent to which any of these processes may impact in vitro OATP inhibition assays is unclear (15). Timescales of transcriptional regulation in mammalian cells range from tens of minutes to several hours (38), while post-translational events occur in seconds. Thus, some indirect regulatory processes are theoretically plausible within the suggested 30-minute experimental pre-incubation period.

For ligands with potentially indirect means of OATP inhibition, we hypothesize that ligand-centric strategies like our ligand-GNN model may be advantageous for predicting inhibitory activity. This is because ligand-centric models are not constrained by OATP-specific information. However, for ligands whose primary mode of inhibition is competitive (which is generally assumed with OATPs), ligand-centric approaches likely cannot fully capture the intricate protein-ligand interaction features that we hypothesized to be important for determining inhibition.

HOLI-GNN and PLIF-centric analyses enhance our ability to understand OATP inhibition mechanisms, contributing to improved drug safety through better prediction of competitive drug-drug interactions. However, uncertainty in the competitive inhibition assumption compounds the complexities of predictive model development. These challenges necessitate further in vitro assays and highlight the need for advanced computational methods to deconvolute nuanced data patterns. By refining our models with emerging structural and functional data, we can anticipate more accurate predictive tools to advance the drug development process.

Materials and Methods

Data Selection & Labeling

Several studies have been conducted over the past decade that characterize and evaluate OATP interactions with ligands (9–12, 20). Due to notable inter-study variability among the available in-vitro OATP inhibition data (15, 16), we reasoned that sourcing data from a single set of experiments (10) would provide the most consistent framework for our study. Increased caution is warranted in data curation for inhibition modeling, as discrepancies in how inhibitors are defined may result in misguided labeling across ML implementations. To prevent the impacts of such misguided labeling, ML models for OATP inhibition prediction should be trained on in-vitro datasets utilizing one probe molecule, one cell line, and one specific set of experimental conditions, like the dataset we used from Karlgren et al (10).

Experimental inhibition data was obtained from Karlgren et al. for 225 ligands assayed against OATP1B1, OATP1B3, and OATP2B1 (10). We focused only on OATP1B1 since the other OATPs do not yet have cryo-EM structures available in both the inward- and outward-facing conformations.

Of these 225 compounds, 222 were used in our workflow; we did not include 3 ligands due to conformer generation issues and obtainment errors (Table S1). As was done in the original Karlgren study, compounds with reported inhibition percentages greater than 50% were labeled as inhibitors (Class = 1), while those with inhibition percentages less than or equal to 50% were labeled as noninhibitors (Class = 0). For OATP1B1, the distribution of inhibition percent data reported by Karlgren et al. (10) is presented in **Figure S1**.

Data Splitting for Training & Test Sets

A drug-wise splitting approach was adopted to create the holdout test sets to ensure a robust evaluation of model performance. Twenty balanced test sets were created, each containing 10% of the total dataset, with an equal representation of inhibitors and noninhibitors. The remaining 90% of the data was used for training and cross-validation. Table S2 lists the ligands in each test set.

Structural Modeling

Ligand Structure Preparation: Canonical SMILES were compiled from PubChem (39) for the ligands to be used for structural modeling. Ligand representations were converted from SMILES format to three-dimensional structures (SDF format), energetically minimized, and protonated (to pH 7.5) using Open Babel (40). Prior to docking, a total of 250 conformers were generated for each ligand with the BioChemical Library conformation sampling tool (41).

Protein Structure Preparation: The protein sequence of human OATP1B1 was obtained from the UniProt database in FASTA format (42). The cryo-EM structures of both the inward-facing and outward-facing OATP1B1 conformers (PDB IDs 8HND and 8HNB respectively) were obtained from the Protein Data Bank (PDB) (43). These structures were prepared for protein-ligand docking by first removing water molecules, ions, and small molecule ligands manually using PyMOL (44). The structures were then energetically minimized using the Rosetta FastRelax protocol (45) which relieves minor sidechain and backbone clashes that may be present in structures obtained from the PDB.

Ligand-Protein Docking: Molecular docking was performed between the 222 ligands and both the outward and inward conformers of human OATP1B1 using the free Rosetta modeling suite (version 3.12) (46). The RosettaLigand score function was used due to its proven performance in small molecule docking compared to other Rosetta score functions and external docking algorithms (47). A more detailed overview of the Rosetta docking workflow can be found in the Supplemental Information. In brief, the docking workflow involved the following critical steps: (1) concatenating the ligand conformer files with the prepared protein structures; (2) defining the starting coordinates for the ligand within the known transport site of the OATPs (13) in an XML file; (3) specifying the space in which the ligands are allowed to move during the docking simulations in the same XML file; and (4) running the docking simulations for all ligands. To ensure a thorough sampling of the binding space on the OATPs, 1000 docking simulations resulting in different ligand poses (herein referred to as docked poses) were generated for each protein-ligand pair.

Docked Pose Ranking: For each docked pose, Rosetta reports the ligand-protein interface energy in Rosetta Energy Units (REU). This was used as a relative metric to rank the 1000 docked poses for each single ligand-OATP conformer pair. Poses with the lowest energy scores were considered higher-quality docked poses (i.e., those with stronger binding) (26). For all poses, ligand-protein interface scores were recorded in a CSV file for use in subsequent analyses.

Feature Selection & Encoding

Two distinct types of feature sets were generated for each OATP1B1 conformer: traditional ligand-only molecular descriptors (i.e., ligand features) and protein-ligand interaction features (PLIFs).

Ligand Features: Two types of vectorized molecular representations, extended-connectivity fingerprints (ECFP6, henceforth referred to as ECFP) and RDKit physicochemical descriptors, were

generated for each of the 222 docked ligands using the pinky and RDKit Python libraries, respectively. The resulting data was organized in comma-separated values (CSV) files, with each row containing the vectorized representation for a single ligand.

Protein-Ligand Interaction Features (PLIFs): The Protein-Ligand Interaction Profiler (PLIP) (25) was used to generate interaction profiles (XML file format) for each docked pose. PLIP produces intricate features to describe each interaction present in a docked pose, including interaction type, distance, angle, donor and acceptor atom IDs, etc. (Table S4). These comprehensive PLIFs were loaded into a CSV file, with each row containing vectorized PLIFs for a single docked pose.

Simple Classification Models

Data Preprocessing: Prior to training the machine learning models, several preprocessing steps were applied to the ligand features and PLIFs. One-hot encoding was performed on categorical features to convert them into a binary representation suitable for the models. Variance and multicollinearity correlation thresholding were employed to remove features with low variance (<0.001) and high correlation (>0.99), respectively, reducing the dimensionality of the feature space and mitigating the risk of overfitting. Feature normalization using StandardScaler (48) was applied to ensure that all features have zero mean and unit variance, improving the convergence and stability of the machine learning algorithms.

Classifier Training & Evaluation: Several ML algorithms, including logistic regression, support vector machines (SVM), random forests, and XGBoost, were trained and evaluated using the preprocessed data. The training process involved k-fold cross-validation, where the training data was split into k subsets, and the models were trained and validated k times, using a different subset for validation each time. This approach helps to assess the models' performance and stability across different subsets of the data.The trained models were then evaluated on the holdout test sets to assess their generalization performance. Metrics such as accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC) were calculated to provide a comprehensive assessment of the model's performance in predicting OATP inhibition.

Heterogeneous GNN for Inhibition Prediction of OATPs (HOLI-GNN)

A heterogeneous graph neural network (HeteroGNN, **Figure 4**) was constructed to capture the intricate interactions between the proteins and ligands. The graph contains two types of nodes: amino acid residues and ligand atoms. Each amino acid node is represented by 28 features (Table S5), including one-hot encoded residue type, molecular weight, aromaticity, isoelectric point, hydrophobicity, flexibility, and secondary structure propensities. These features capture the essential physicochemical properties of the protein residues involved in protein-ligand interactions. Ligand atom nodes are comprised of 79 atomic features (Table S6), including one-hot encoded atom type, number of heavy neighbors, formal charge, hybridization state, presence in rings, aromaticity, atomic mass, van der Waals radius, covalent radius, chirality, and number of implicit hydrogens. This set of atomic features provides a detailed representation of the ligand's chemical structure and properties.

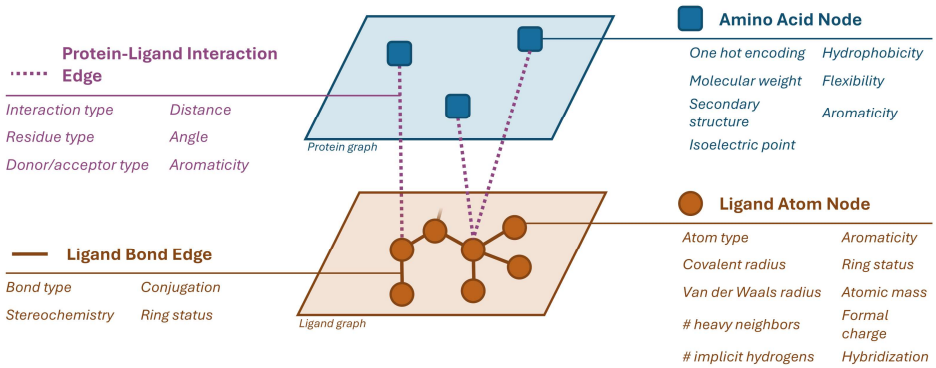

Figure 4. HOLI-GNN (Heterogeneous OATP-Ligand Interaction Graph Neural Network) encompasses a ligand graph (orange, nodes: atoms, edges: intra-ligand bonds) and a protein graph (blue, nodes: amino acids) connected by protein-ligand interaction edges (magenta). Below each node or edge type in the simple visualization, lists of features stored within them are provided, with expanded details in supplemental Table S5-Table S9.

Edges in the constructed graph represent the intermolecular interactions between amino acids and ligand atoms, as well as the intramolecular interactions among atoms within the ligand. Edges which connect heterogenous nodes encode protein-ligand interactions, represented by 332 edge PLIFs (Table S9), such as interaction type, distance, angle, and other geometric properties. To represent intra-ligand bonds, 10 atom-to-atom edge features (Table S8) include bond type, conjugation, presence in rings, and stereochemistry. This graph construction process captures the local and global structural information of the protein-ligand complexes, providing a comprehensive representation for subsequent analysis. Lastly, Table S10 presents the edge formulation indices and their corresponding sizes, detailing the interactions between amino acids (labeled 'aa') and atoms, as well as the bonds between atoms.

HOLI-GNN Architecture: The HOLI-GNN architecture is designed to model protein-drug interactions and perform a graph classification task. It uses separate convolution layers for amino acid-to-atom, atom-to-amino acid, and atom-to-atom interactions. Edge features are first processed through multi-layer perceptrons (MLPs) to capture interaction characteristics. Convolutional layers (i.e., NNConv) then update node features based on neighboring nodes and processed edge features, propagating information across the protein-ligand interface. A HeteroConv layer aggregates outputs from these convolutions, combining information from different interaction types (i.e., intermolecular and intramolecular). Linear projection layers then map atom and amino acid features to a common space. Global mean pooling is applied to create fixed-size representations for the ligand and protein. These pooled features are concatenated and passed through a classification network consisting of linear layers with ReLU activation and dropout. The network outputs a single value predicting the compound's inhibitor status. This architecture integrates heterogeneous information from protein-ligand complexes, considering both local interactions and global structure to predict inhibitory activity.

HOLI-GNN Training and Evaluation: HOLI-GNN was trained using the AdamW optimizer with weight decay. The OneCycleLR learning rate scheduler was used for adaptive learning rates. Mixed precision training improved computational efficiency. Binary cross-entropy with logits served as the loss function. Early stopping, based on validation loss with a patience of 5 epochs, prevented overfitting. Each model trained for up to 100 epochs or until early stopping occurred. For each of the 20 data splits, a separate model was trained and evaluated on the corresponding test set. Multiple metrics were calculated: accuracy, precision, recall, F1 score, balanced accuracy, and AUC-ROC. This thorough evaluation provided a robust understanding of the model's predictive capabilities across various classification aspects.

Acknowledgments

This work was financially supported by the Michigan State University Chemical Engineering and Materials Science Department Startup Funds and the National Science Foundation Graduate Research Fellowship (Award No. 2235783). ML model development was partially supported by USDA-NIFA (2023-67013-39901). Much of the computational work in this project was performed on resources managed by the Michigan State University Institute for Cyber-Enabled Research. The authors would also like to express their gratitude to Dr. Bruno Hagenbuch of the University of Kansas Medical Center for very informative discussions at the start of the project.

References

1. S. Ezad, H. Cheema, N. Collins, Statin-induced rhabdomyolysis: A complication of a commonly overlooked drug interaction. *Oxf Med Case Reports* 2018, 86–88 (2018).
2. A. Kalliokoski, M. Niemi, Impact of OATP transporters on pharmacokinetics. *Br J Pharmacol* [Preprint] (2009).
3. I. G. Asiimwe, M. Pirmohamed, Drug–Drug–Gene Interactions in Cardiovascular Medicine. *Pharmgenomics Pers Med* [Preprint] (2022).
4. United States Food and Drug Administration, “Clinical Drug Interaction Studies-Cytochrome P450 Enzyme-and Transporter-Mediated Drug Interactions Guidance for Industry” (2020).
5. United States Food and Drug Administration, “In Vitro Drug Interaction Studies-Cytochrome P450 Enzyme-and Transporter-Mediated Drug Interactions Guidance for Industry” (2020).
6. F. Cheng, Z. Zhao, Machine learning-based prediction of drug-drug interactions by integrating drug phenotypic, therapeutic, chemical, and genomic properties. *J Am Med Inform Assoc* 21 (2014).
7. K. Yoshida, C. Guo, R. Sane, Quantitative Prediction of OATP-Mediated Drug-Drug Interactions With Model-Based Analysis of Endogenous Biomarker Kinetics. *CPT Pharmacometrics Syst Pharmacol* 7, 517–524 (2018).
8. K. Alam, et al., Regulation of organic anion transporting polypeptides (OATP) 1B1- and OATP1B3-mediated transport: An updated review in the context of OATP-mediated drug-drug interactions. *Int J Mol Sci* [Preprint] (2018).
9. T. R. Lane, et al., Machine Learning Models Identify New Inhibitors for Human OATP1B1. *Mol Pharm* 19, 4320–4332 (2022).
10. M. Karlgren, et al., Classification of inhibitors of hepatic organic anion transporting polypeptides (OATPs): Influence of protein expression on drug-drug interactions. *J Med Chem* 55, 4740–4763 (2012).
11. T. De Bruyn, et al., Structure-based identification of oatp1b1/3 inhibitorss. *Mol Pharmacol* 83, 1257–1267 (2013).
12. H. Duan, et al., In Silico prediction of inhibitors for multiple transporters via machine learning methods. *Mol Inform* (2024). https://doi.org/10.1002/minf.202300270.
13. Z. Shan, et al., Cryo-EM structures of human organic anion transporting polypeptide OATP1B1. *Cell Res* 33, 940–951 (2023).
14. A. D. Ciută, et al., Structure of human drug transporters OATP1B1 and OATP1B3. *Nat Commun* 14 (2023).
15. S. J. McFeely, T. K. Ritchie, I. Ragueneau-Majlessi, Variability in In Vitro OATP1B1/1B3 Inhibition Data: Impact of Incubation Conditions on Variability and Subsequent Drug Interaction Predictions. *Clin Transl Sci* 13, 47–52 (2020).
16. S. Izumi, et al., Investigation of the impact of substrate selection on in vitro organic anion transporting polypeptide 1B1 inhibition profiles for the prediction of drug-drug interactions. *Drug Metabolism and Disposition* 43, 235–247 (2015).

17. S. J. McFeely, et al., Identification and Evaluation of Clinical Substrates of Organic Anion Transporting Polypeptides 1B1 and 1B3. *Clin Transl Sci* 12, 379–387 (2019).
18. K. Alam, T. Farasyn, A. Crowe, K. Ding, W. Yue, Treatment with proteasome inhibitor bortezomib decreases organic anion transporting polypeptide (OATP) 1B3-mediated transport in a substrate-dependent manner. *PLoS One* 12 (2017).
19. N. Shaikh, M. Sharma, P. Garg, Selective Fusion of Heterogeneous Classifiers for Predicting Substrates of Membrane Transporters. J Chem Inf Model 57, 594–607 (2017).
20. A. Tuerkova, et al., Identifying Novel Inhibitors for Hepatic Organic Anion Transporting Polypeptides by Machine Learning-Based Virtual Screening. *J Chem Inf Model* [Preprint] (2022).
21. R. M. Esquejo, et al., Dipyridamole Inhibits Lipogenic Gene Expression by Retaining SCAP-SREBP in the Endoplasmic Reticulum. *Cell Chem Biol* 28, 169-179.e7 (2021).
22. L. M. M. Vermeer, C. D. Isringhausen, B. W. Ogilvie, D. B. Buckley, Evaluation of ketoconazole and its alternative clinical CYP3A4/5 inhibitors as inhibitors of drug transporters: The in vitro effects of ketoconazole, ritonavir, clarithromycin, and itraconazole on 13 clinically-relevant drug transporters. *Drug Metabolism and Disposition* 44, 453–459 (2016).
23. H. Huang, et al., Inhibition of drug metabolism by blocking the activation of nuclear receptors by ketoconazole. *Oncogene* 26, 258–268 (2007).
24. J. G. Park, et al., Connectivity mapping of angiotensin-PPAR interactions involved in the amelioration of non-alcoholic steatohepatitis by Telmisartan. *Sci Rep* 9 (2019).
25. M. F. Adasme, et al., PLIP 2021: Expanding the scope of the protein-ligand interaction profiler to DNA and RNA. *Nucleic Acids Res* 49, W530–W534 (2021).
26. K. W. Kaufmann, J. Meiler, Using RosettaLigand for Small Molecule Docking into Comparative Models. *PLoS One* 7 (2012).
27. H. You, et al., Computational classification models for predicting the interaction of compounds with hepatic organic ion importers. *Drug Metab Pharmacokinet* 30, 347–351 (2015).
28. N. J. Wright, S. Y. Lee, Recent advances on the inhibition of human solute carriers: Therapeutic implications and mechanistic insights. *Curr Opin Struct Biol* [Preprint] (2022).
29. M. G. Soars, P. Barton, M. Ismair, R. Jupp, R. J. Riley, The development, characterization, and application of an OATP1B1 inhibition assay in drug discovery. *Drug Metabolism and Disposition* 40, 1641–1648 (2012).
30. G. Ahlin, et al., Endogenous gene and protein expression of drug-transporting proteins in cell lines routinely used in drug discovery programs. *Drug Metabolism and Disposition* 37, 2275–2283 (2009).
31. S. Zhou, Y. Shu, Special Section on New Era of Transporter Science: Unraveling the Functional Role of Orphan Transporters-Minireview Transcriptional Regulation of Solute Carrier Drug Transporters. *Drug Metabolism and Disposition* [Preprint] (2022).
32. T. F. Gallegos, G. Martovetsky, V. Kouznetsova, K. T. Bush, S. K. Nigam, Organic anion and cation slc22 “drug” transporter (oat1, oat3, and oct1) regulation during development and maturation of the kidney proximal tubule. *PLoS One* 7 (2012).
33. G. Martovetsky, J. B. Tee, S. K. Nigam, Hepatocyte nuclear factors 4α and 1α regulate kidney developmental expression of drug-metabolizing enzymes and drug transporters. *Mol Pharmacol* 84, 808–823 (2013).
34. S. K. Nigam, et al., Handling of drugs, metabolites, and uremic toxins by kidney proximal tubule drug transporters. *Clinical Journal of the American Society of Nephrology* 10, 2039–2049 (2015).
35. J. Veh, et al., Phorbol-12-myristate-13-acetate is a potent enhancer of B cells with a granzyme B+ regulatory phenotype. *Front Immunol* 14 (2023).
36. E. Tahara, H. Kadara, L. Lacroix, D. Lotan, R. Lotan, Activation of protein kinase C by phorbol 12-myristate 13-acetate suppresses the growth of lung cancer cells through KLF6 induction. *Cancer Biol Ther* 8, 801–807 (2009).

37. J. Powell, et al., Novel mechanism of impaired function of organic anion- transporting polypeptide 1B3 in human hepatocytes: Post-translational regulation of OATP1B3 by protein kinase C activation. *Drug Metabolism and Disposition* 42, 1964–1970 (2014).
38. N. C. Lammers, Y. J. Kim, J. Zhao, H. G. Garcia, A matter of time: Using dynamics and theory to uncover mechanisms of transcriptional bursting. *Curr Opin Cell Biol* [Preprint] (2020).
39. S. Kim, et al., PubChem 2023 update. Nucleic Acids Res 51, D1373–D1380 (2023).
40. N. M. O'Boyle, et al., Open Babel: An Open chemical toolbox. *J Cheminform* 3 (2011).
41. J. Mendenhall, B. P. Brown, S. Kothiwale, J. Meiler, BCL::Conf: Improved Open-Source Knowledge-Based Conformation Sampling Using the Crystallography Open Database. *J Chem Inf Model* 61, 189–201 (2021).
42. R. Apweiler, The Universal Protein resource (UniProt). *Nucleic Acids Res* 36 (2008).
43. H. M. Berman, et al., "The Protein Data Bank" (2000).
44. S. Rosignoli, A. Paiardini, Boosting the Full Potential of PyMOL with Structural Biology Plugins. *Biomolecules* [Preprint] (2022).
45. P. Conway, M. D. Tyka, F. DiMaio, D. E. Konerding, D. Baker, Relaxation of backbone bond geometry improves protein energy landscape modeling. *Protein Science* 23, 47–55 (2014).
46. J. Meiler, D. Baker, ROSETTALIGAND: Protein-small molecule docking with full side-chain flexibility. *Proteins: Structure, Function and Genetics* 65, 538–548 (2006).
47. S. T. Smith, J. Meiler, Assessing multiple score functions in Rosetta for drug discovery. *PLoS One* 15 (2020).
48. K. Balabaeva, S. Kovalchuk, Comparison of Temporal and Non-Temporal Features Effect on Machine Learning Models Quality and Interpretability for Chronic Heart Failure Patients in Procedia Computer Science, (Elsevier B.V., 2019), pp. 87–96.