Journal Title Here, 2022, pp. 1–11

doi: DOI HERE
Advance Access Publication Date: Day Month Year
Paper

PAPER

FD-CGNN

First Author,[1,*] Second Author,[2] Third Author,[3] Fourth Author[3] and Fifth Author[4]

[1]Department, Organization, Street, Postcode, State, Country, [2]Department, Organization, Street, Postcode, State, Country, [3]Department, Organization, Street, Postcode, State, Country and [4]Department, Organization, Street, Postcode, State, Country

[*]Corresponding author. email-id.com

FOR PUBLISHER ONLY Received on Date Month Year; revised on Date Month Year; accepted on Date Month Year

Abstract

Motivation: Accurately predicting the metabolic stability of molecules is crucial in modern drug discovery, as it directly affects pharmacokinetics, safety, and efficacy. Although current models primarily focus on atom-level and bond-level structural information, they often overlook the incorporation of global molecular characteristics. This overemphasis on local features limits the robustness of learned molecular embeddings and hinders predictive performance in tasks such as metabolic stability estimation. Addressing this issue requires a more integrated approach to molecular representation learning—one that combines both local structure and global chemical context to improve generalization and predictive accuracy. **Result:** We introduce FD-CGNN, a novel molecular representation learning framework featuring a dual-channel architecture that integrates molecular fingerprints with a cross-dependent graph neural network. The two channels are designed to capture complementary global and local molecular features, and joint prediction further boosts model performance. Moreover, multiple classes of molecular fingerprints are incorporated to enrich global semantic information and enhance the representational capacity of the model. We evaluated the proposed approach across multiple experimental settings for human and rat liver microsomal stability. The results demonstrate that our FD-CGNN model consistently and significantly outperforms existing mainstream methods across all performance metrics, exhibiting superior generalization capacity and predictive effectiveness.

Key words: graph neural networks; molecular fingerprint; metabolic stability; molecular structure

Introduction

Hepatic microsomal metabolic stability is a critical pharmacokinetic parameter in the early stages of drug development, as it directly affects a compound's clearance, half-life, and oral bioavailability, thereby determining its absorption, distribution, metabolism, and excretion (ADME) properties [1, 2, 3, 4]. By analyzing metabolic stability data, researchers can more effectively guide lead optimization and structural modification during the initial phases of drug discovery [5, 6].

Currently, the most commonly used in vitro models for metabolic stability studies include hepatic microsomes, hepatocytes,and hepatic S9 fractions, among which hepatic microsomes and hepatocytes are the most widely applied [7]. Hepatic microsomes are vesicles derived from the endoplasmic reticulum membrane of hepatocytes and are enriched with cytochrome P450 (CYP450) enzymes, making them well suited for investigating phase I metabolic reactions such as oxidation and reduction [8].Hepatocyte models retain the full metabolic machinery of intact cells, supporting both phase I and phase II reactions and thereby offering a closer approximation of the in vivo metabolic environment.The hepatic S9 fraction, comprising both microsomal and cytosolic components, encompasses a complete panel of metabolic enzymes and is therefore suitable for more comprehensive metabolism studies. Studies have demonstrated a strong correlation between in vitro microsomal metabolic stability and in vivo drug clearance, which has reinforced the widespread use of these models for predicting metabolic behavior in vivo [9].

Despite their essential role in elucidating metabolic mechanisms, these in vitro assays often require expensive instrumentation, complex experimental protocols, and substantial time and labor investments [10]. As a result, computational approaches based on machine learning and deep learning have drawn increasing attention in recent years, particularly for predicting the metabolic stability of human (HLM) and rat (RLM) liver microsomes, where they have shown promising application potential [11, 12, 13].

Early studies on molecular metabolic stability prediction mainly used traditional machine learning with handcrafted descriptors and fingerprints [14, 15, 16]. Perryman et al. applied a Laplace-corrected Naïve Bayes model with ECFP_6 fingerprints to classify compounds as "stable" or "unstable" [17]. Podlewska et al. proposed MetStabOn, integrating kNN, NB, and RF with PaDEL descriptors covering physicochemical properties and topological fingerprints [18, 19, 20]. Ryu et al. curated a high-quality HLM dataset and trained RF models for clearance-related prediction [21], while Li et al. evaluated tree-based models (e.g., GBDT, XGBoost) using ECFP fingerprints and achieved strong performance [22]. However, these approaches depend heavily

on manually designed features, limiting their ability to capture complex substructure interactions and reducing generalization.

To better model molecular topology, deep learning methods—especially GNNs—have become widely adopted [23, 24]. Renn et al. used GCNs to learn molecular representations from SMILES-derived graphs via neighborhood aggregation [25], but standard GCNs struggle with bond-type differentiation and may suffer from redundant message propagation. D-MPNN addresses these issues by directing message passing along bonds, mitigating over-smoothing [22]. However, current GNN-based methods remain limited in structural expressiveness: message passing is typically confined to local neighborhoods, making it difficult to incorporate global semantics; moreover, many models overlook bond-level topology and struggle to capture long-range dependencies.

To address the limitations of single-perspective GNNs, recent studies have explored multi-level structural modeling to enhance molecular graph representations. Ma et al. proposed CD-MVGNN, which constructs two complementary graph encoders—one atom-centered and the other bond-centered—and strengthens information exchange across these perspectives through a cross-dependent message passing mechanism, enabling the model to obtain a more complete structural characterization of molecules [27].Along this direction, Wang et al. introduced the MS-BACL model, which constructs a bond graph based on the "atom–bond–atom" structure, allowing the model to incorporate both atomic and bond information during message passing. Moreover, by applying contrastive learning between the molecular graph and the bond graph, MS-BACL learns more robust molecular representations [28].

Another line of research has explored combining graph neural networks with additional molecular modalities, such as SMILES sequences or molecular fingerprints, to compensate for the limitations of a single structural perspective. FinGAT strengthens the coordination between fingerprint-based and structural information by introducing a fingerprint-enhanced graph attention network. In this framework, sequence-derived 2D fingerprints are integrated with graph-based structural representations to support molecular feature learning for antibiotic activity prediction [30]. Du et al. later proposed CMMS-GCL for metabolic stability prediction, adopting a cross-modal learning paradigm that models molecular graphs using a GNN and encodes SMILES sequences via a bidirectional GRU (BiGRU), while contrastive learning is employed to enhance semantic consistency between the two modalities [31].

Although graph-based molecular modeling approaches have achieved notable progress, the interaction across different structural perspectives remains limited, and the absence of global chemical priors restricts their capacity to capture the multi-scale determinants of metabolic stability. Multimodal fusion methods likewise present inherent limitations: feature integration across modalities is often shallow, resulting in insufficient semantic alignment; moreover, most existing approaches do not incorporate multi-view graph information—such as atom-level and bond-level structural cues—thereby hindering the construction of complementary molecular representations. As a result, neither graph-structured models nor multimodal frameworks are able to simultaneously leverage local structural patterns and global chemical semantics in a comprehensive manner. To overcome these challenges, we introduce FD-CGNN (Fingerprint-augmented Cross-dependent Graph Neural Network), a dual-channel architecture designed to separately model local structural dependencies and global fingerprint-based priors. The structural channel employs atom-centered and bond-centered graph encoders to derive fine-grained molecular representations, whereas the semantic channel utilizes molecular fingerprints to capture high-level chemical knowledge. The three channels produce independent predictions for metabolic stability and are jointly optimized under a multi-task learning scheme, enabling the model to learn more robust, complementary, and task-relevant molecular representations through parallel supervision. Overall, the main contributions of this work can be summarized as follows:

• A deep learning (DL) architecture named FD-CGNN was presented for predicting molecular metabolic stability, in which global fingerprint semantics and atom–bond graph features are processed in parallel through independent prediction channels rather than embedded into a unified structure.

• To better integrate global and local features, we design a multi-channel architecture where atom, bond, and fingerprint encoders make independent predictions, promoting more effective collaborative learning.

• Experiments on multiple datasets (HLM and RLM) show that FD-CGNN outperforms existing methods in metabolic stability prediction and exhibits superior generalization across species.

METHOD

Problem Definition

We formulate metabolic stability prediction as a supervised learning task. A molecule can be represented as a graph:

$$G = (V, E), \tag{1}$$

where, V is the set of atoms and E is the set of chemical bonds. Each atom $v \in V$ is described by a feature vector x_v (e.g., atom type, degree, valence, aromaticity), and each bond $(u, v) \in E$ is described by a feature vector e_{uv} (e.g., bond type, conjugation, ring membership).

In addition, each molecule is also represented by a fingerprint vector F, which provides global chemical information. The goal of metabolic stability prediction is to learn a mapping:

$$f(G, F) \to y, \tag{2}$$

where y is the metabolic stability label.

Overview of FD-CGNN

The framework is composed of two main components:

• Cross-dependent GNN Encoders: These encoders model atom-centered and bond-centered structural information in parallel, and exchange features through a cross-dependent message passing mechanism.

• Fingerprint Encoder: This module learns embedded representations from multiple molecular fingerprints, allowing the model to incorporate global chemical information.

During training, FD-CGNN produces predictions from the atom channel, bond channel, and fingerprint channel, and jointly optimizes them through a weighted multi-task loss function.To provide a clearer illustration of the overall architecture, we present the framework of FD-CGNN in Figure 1.

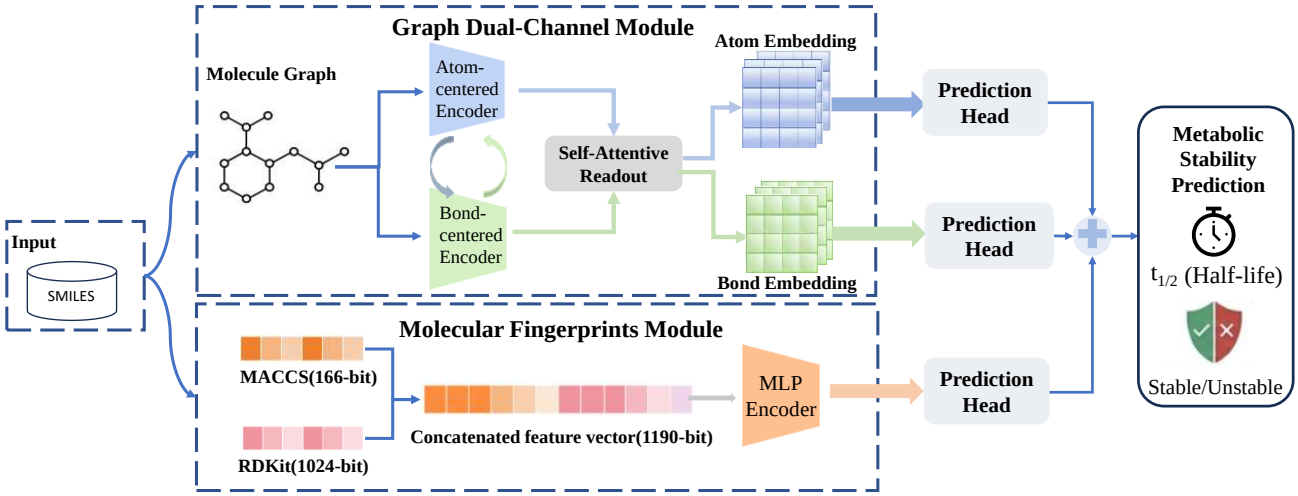

Fig. 1. Overall framework of the proposed FD-CGNN.

Cross-dependent GNN Encoders

In molecular structure modeling, FD-CGNN employs a dual-channel message passing mechanism to jointly capture atomic-level and bond-level structural information. The model consists of two parallel encoders:

Atom-centered encoder: updates atom embeddings by aggregating messages from neighboring atoms and associated bonds, thereby capturing inter-atomic dependencies.

Bond-centered encoder: constructs a line graph representation in which bonds are treated as nodes and learns bond-to-bond relationships to propagate bond-level information.

To enhance structural interactions between these two perspectives, FD-CGNN introduces a cross-dependent mechanism that enables bidirectional information flow during message propagation.

Let $h_v^{(t)}$ and $h_{uv}^{(t)}$ denote the atom and bond embeddings at the t-th iteration, respectively. The update rules are formulated as:

$$h_v^{(t+1)} = \sigma\left(W_a \cdot h_v^{(t)} + \sum_{u \in N(v)} \phi_a\left(h_u^{(t)}, h_{uv}^{(t)}, e_{uv}\right)\right), \tag{3}$$

$$h_{uv}^{(t+1)} = \sigma\left(W_b \cdot h_{uv}^{(t)} + \phi_b\left(h_u^{(t)}, h_v^{(t)}, e_{uv}\right)\right). \tag{4}$$

where $\phi_a(\cdot)$ and $\phi_b(\cdot)$ denote the message aggregation functions, $\sigma(\cdot)$ is a nonlinear activation function, and W_a, W_b are learnable transformation matrices.

In this design, the atom representations are updated based on the latest bond embeddings, while the bond representations incorporate contextual information from the connected atoms. This bidirectional dependency allows mutual reinforcement between atom- and bond-level representations, enabling a more expressive and robust molecular embedding.

After multiple propagation layers, molecular-level embeddings are obtained through attention-based readout functions:

$$h_{\text{atom}} = \text{Readout}_a\left(\{h_v^{(T)} \mid v \in V\}\right), \tag{5}$$

$$h_{\text{bond}} = \text{Readout}_b\left(\{h_{uv}^{(T)} \mid (u,v) \in E\}\right). \tag{6}$$

where Readout$(\cdot)$ aggregates node or bond embeddings via a self-attention mechanism. The resulting embeddings h_{atom} and h_{bond} are then fed into independent feed-forward networks (FFNs) for downstream prediction.

This cross-dependent message passing architecture enables FD-CGNN to jointly encode local atomic environments and bond-level topological dependencies, resulting in richer and more chemically meaningful molecular representations.

Fingerprint Encoder

Although graph neural networks (GNNs) can effectively capture local structural features through atom–bond message passing, their ability to model global molecular context remains limited due to the bounded receptive field. To compensate for the lack of global chemical knowledge in structural representations, FD-CGNN incorporates molecular fingerprints as global structural priors, complementing the locality inherent in graph-based modeling.

In this work, we employ two widely used and conceptually complementary classes of fingerprints:

- **MACCS Keys** [32]: a rule-based 166-bit binary fingerprint designed to encode common functional groups, ring systems, and heteroatom environments. It offers strong interpretability and reflects expert-defined structural patterns.
- **RDKit Topological Fingerprint (RDKFingerprint)**: a 1024-bit path-based fingerprint that enumerates atom paths to capture long-range dependencies and non-local topological features, which are particularly valuable for metabolic stability modeling.

These two descriptors are concatenated to form a unified 1190-bit binary feature vector:

$$F \in \{0,1\}^{1190}. \tag{7}$$

To incorporate such heterogeneous and high-dimensional fingerprint information into the neural architecture, FD-CGNN employs a fingerprint encoder that transforms the fingerprint

vector via a multi-layer perceptron (MLP):

$$h_{\text{fp}} = \text{MLP}(F). \tag{8}$$

Through nonlinear transformations and dropout regularization, the MLP maps the sparse binary space into a continuous latent representation, enabling the model to simultaneously leverage the expert-defined chemical semantics provided by MACCS Keys and the topological patterns captured by RDKFingerprint.

By integrating these two complementary sources of molecular knowledge—one emphasizing chemically meaningful substructures and the other capturing broader topological relationships—FD-CGNN gains access to a richer chemical feature space that extends beyond local graph structure, thereby substantially enhancing its ability to model the structural determinants of metabolic stability.

Metabolic stability predictor

FD-CGNN integrates the three molecular representations—atom-level, bond-level, and fingerprint-derived—through a multi-branch prediction module. Each representation is processed by an independent feed-forward network, producing three parallel prediction outputs. The atom-level branch captures local atomic environments, the bond-level branch reflects connectivity-driven structural patterns, and the fingerprint branch incorporates global chemical descriptors that complement the graph-based embeddings.

During training, all three branches are jointly optimized using a weighted multi-task loss:

$$L_{\text{total}} = L_{\text{atom}} + \lambda_{\text{bond}} L_{\text{bond}} + \lambda_{\text{fp}} L_{\text{fp}}, \tag{9}$$

where task-specific weights regulate the relative contributions of the bond and fingerprint branches. Binary cross-entropy is applied as the loss function for the binary classification task. This multi-branch supervision allows each modality to learn informative and complementary features, thereby improving the robustness and generalization of the metabolic stability predictor.

EXPERIMENT RESULTS

Dataset

To rigorously evaluate the performance of the proposed FD-CGNN model, we employed two benchmark datasets related to molecular metabolic stability. The first dataset, referred to as HLM, contains metabolic stability measurements obtained from human liver microsomes. Because no universally accepted definition of metabolic stability currently exists, we followed the classification criteria adopted in previous studies using the same dataset. Molecules with a half-life ($t_{1/2}$) greater than 30 minutes were labeled as metabolically stable (Label 1), whereas those with shorter half-lives were labeled as unstable (Label 0). Under this definition, the HLM dataset comprises 5,878 molecules, including 3,784 stable compounds (64.4%) and 2,094 unstable compounds (35.6%).

The second dataset, RLM, is a cross-species benchmark involving metabolic stability measurements from rat liver microsomes [33]. Consistent with the classification threshold used for the HLM dataset, the RLM dataset includes 5,590 molecules, among which 2,278 compounds (40.8%) are categorized as stable and 3,312 compounds (59.2%) as unstable. Together, these two datasets provide complementary perspectives on human and rat metabolic stability, enabling a comprehensive assessment of the generalization performance of FD-CGNN.

Table 1. Statistics of metabolic stability datasets.

Dataset	Total	Positive (stable)	Negative (unstable)
HLM dataset	5878	3784	2094
RLM dataset	5590	2278	3312

To assess the structural diversity of the datasets, we calculated the pairwise Tanimoto similarity based on ECFP fingerprints. The average Tanimoto coefficients for the HLM and RLM datasets were 0.118 and 0.123, respectively, indicating relatively low similarity between molecular pairs. This suggests that both datasets exhibit substantial structural diversity, which is crucial for evaluating the robustness and generalization capability of predictive models.

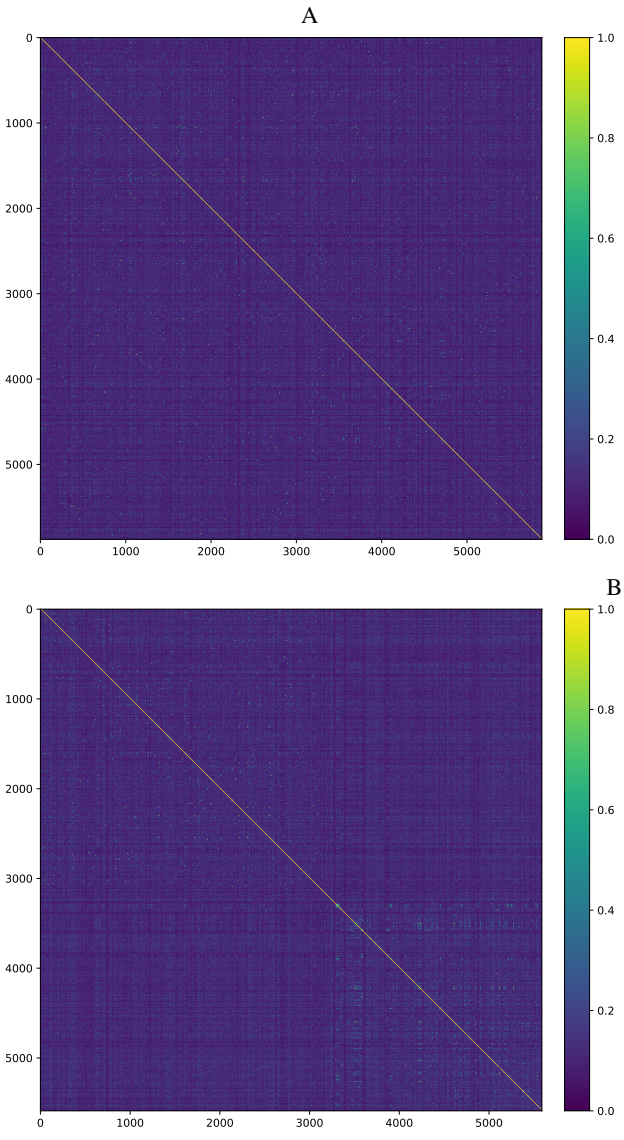

Fig. 2. Tanimoto similarity heatmaps of molecular datasets. (a) HLM; (b) RLM.

Table 2. Performance evaluation of FD-CGNN and baseline methods on HLM metabolic stability prediction.

Models	AUC	ACC	F1	MCC
D-MPNN	0.845±0.016	0.785±0.012	0.836±0.011	0.526±0.025
GBDT	0.802±0.012	0.750±0.011	0.827±0.021	0.425±0.021
XGBoost	0.850±0.014	0.796±0.013	0.849±0.015	0.543±0.019
FinGAT	0.831±0.018	0.776±0.019	0.775±0.020	0.523±0.041
FP-GNN	0.840±0.015	0.784±0.016	0.833±0.011	0.530±0.053
AttentiveFP	0.865±0.016	0.771±0.017	0.810±0.013	0.546±0.024
CMMS-GCL	0.865±0.019	0.812±0.018	0.853±0.011	0.570±0.034
MS-BACL	0.870±0.012	0.817±0.013	0.850±0.013	0.595±0.031
FD-CGNN	**0.913±0.001**	**0.834±0.003**	**0.870±0.005**	**0.643±0.013**

Table 3. Performance evaluation of FD-CGNN and baseline methods on RLM metabolic stability prediction.

Models	AUC	ACC	F1	MCC
D-MPNN	0.837±0.011	0.754±0.012	0.775±0.009	0.476±0.023
GBDT	0.791±0.012	0.727±0.016	0.793±0.015	0.422±0.019
XGBoost	0.839±0.017	0.765±0.011	0.810±0.013	0.508±0.025
FinGAT	0.818±0.019	0.739±0.014	0.791±0.013	0.482±0.028
FP-GNN	0.826±0.011	0.769±0.021	0.804±0.011	0.512±0.024
AttentiveFP	0.790±0.013	0.704±0.012	0.726±0.015	0.422±0.017
CMMS-GCL	0.843±0.014	0.783±0.010	0.757±0.012	0.520±0.013
MS-BACL	0.823±0.013	0.771±0.011	0.799±0.010	0.513±0.018
FD-CGNN	**0.872±0.005**	**0.805±0.005**	**0.833±0.004**	**0.575±0.011**

Experimental setup

All models were implemented using the PyTorch and PyTorch Geometric (PyG) libraries. Our proposed FD-CGNN framework employs a core graph encoder component, the cross-dependent GNN encoder, which features a 2-layer cross-dependent message passing architecture (`depth = 2`) with a hidden dimension of 256. For training, we utilized the Adam optimizer with a batch size of 128 for 300 epochs. A learning rate warm-up strategy was adopted, starting from 1×10^{-4} and rising to a maximum learning rate of 1×10^{-3}. To mitigate overfitting, a dropout rate of 0.3 was applied across the layers.

As detailed in Section 2.5, the total loss L_{total} is a weighted multi-task objective, composed of three components: (1) the main task loss (weight $\lambda_{\text{atom}} = 1.0$), (2) the auxiliary fingerprint channel loss (weight $\lambda_{\text{fp}} = 0.5$), and (3) the auxiliary bond-level loss (weight $\lambda_{\text{bond}} = 0.1$).

Performance comparison with other models

To comprehensively evaluate the effectiveness of FD-CGNN, we compared it against a diverse set of baseline models, including both traditional machine learning algorithms and advanced deep learning approaches. A brief introduction to these methods is provided below:

- GBDT and XGBoost [22]: Gradient Boosted Decision Trees and eXtreme Gradient Boosting models trained on ECFP fingerprints for molecular property prediction, representing a traditional cheminformatics approach to capture global structural features.
- D-MPNN [22]: A message passing neural network that propagates information along directed chemical bonds to capture the directionality of atomic interactions.
- AttentiveFP [26]: A graph neural network combining gated recurrent units (GRUs) and attention-based aggregation to refine atom-level embeddings.
- FP-GNN [29]: A hybrid framework that combines GNN-based structural encodings with multiple expert-defined fingerprints (e.g., MACCS, PubChem).
- FinGAT [30]: A fingerprint-augmented graph attention network that integrates sequence-derived fingerprints with graph structural features.
- CMMS-GCL [31]: A framework that encodes molecular graphs with GNNs and SMILES sequences with BiGRUs, applying contrastive learning to enforce semantic alignment between modalities.
- MS-BACL [28]: A structure-enhanced model that constructs an auxiliary bond graph based on atom–bond–atom triplets and aligns it with the main graph using contrastive learning.

Together, these models provide a broad and rigorous benchmark for evaluating the robustness and generalization capabilities of FD-CGNN across diverse representation paradigms. All baseline methods were trained and evaluated under consistent 10-fold cross-validation settings using AUC, ACC, F1-score, and MCC as evaluation metrics.

FD-CGNN achieves the best performance on the HLM dataset, with an average AUC of 0.913, Accuracy of 0.834, F1-score of 0.870, and MCC of 0.643. Compared to the second-best method (MS-BACL), FD-CGNN improves AUC by 4.3%, Accuracy by 1.7%, F1-score by 2.4%, and MCC by 4.8%, highlighting its superior capacity in capturing global and local structural features for metabolic stability prediction.

On the RLM dataset, FD-CGNN again outperforms all baselines, achieving an AUC of 0.872, Accuracy of 0.805, F1-score of 0.833, and MCC of 0.544. These correspond to improvements of 4.9% in AUC, 2.2% in Accuracy, 3.4% in

F1-score, and 6.1% in MCC over the next best model (CMMS-GCL), reinforcing the model's robustness and generalization in cross-species metabolic stability prediction.

Taken together, these results show that FD-CGNN not only outperforms traditional machine learning and state-of-the-art graph neural network models but also exhibits clear advantages over multimodal baselines such as CMMS-GCL and MS-master in modeling global chemical semantics. Although these models enhance structural representation—either through graph–SMILES cross-modal contrastive learning or through the construction of auxiliary bond graphs—they do not incorporate expert-defined global descriptors such as molecular fingerprints. By integrating MACCS, RDKit, and other fingerprint-based priors with detailed local structural modeling, FD-CGNN achieves a more comprehensive understanding of structure–property relationships, resulting in improved predictive robustness.

Moreover, while FP-GNN combines graph structures with fingerprint features, its graph encoder remains atom-centered and does not explicitly model bond-level dependencies, which limits its ability to capture more intricate structural contexts. In contrast, the dual-channel architecture of FD-CGNN jointly encodes atom- and bond-level information and fuses these signals through attention-based aggregation, enabling a richer and more chemically coherent molecular representation. This unified design contributes directly to the superior performance of FD-CGNN across metabolic stability prediction tasks.

Ablation experiment

To assess more precisely the contribution of fingerprint-based descriptors within FD-CGNN, we performed a series of ablation studies on both the HLM and RLM datasets. The investigated model variants included a fingerprint-free baseline (w/o FP), models incorporating a single fingerprint type (Morgan, MACCS, or RDKit), and several two-fingerprint combinations, such as Morgan+MACCS and MACCS+RDKit. A consolidated overview of these results is provided in Table 4.

For the HLM dataset, the w/o FP variant yielded the lowest predictive performance, indicating that structural information derived solely from the molecular graph is insufficient and that global chemical descriptors contribute substantially to model robustness. Incorporation of any single fingerprint improved performance, with RDKit providing the most pronounced enhancement. When multiple fingerprints were combined, predictive accuracy increased further; notably, the MACCS+RDKit pairing afforded the highest performance, suggesting complementary strengths between the chemically interpretable substructure patterns captured by MACCS and the topological features encoded by RDKit.

A comparable trend was observed in the RLM dataset. Once again, the fingerprint-free model performed the poorest, reinforcing the importance of global chemical semantics in stability prediction. Among the single-fingerprint configurations, RDKit showed the most favorable behavior, while the MACCS+RDKit combination surpassed all other variants. Relative to the w/o FP baseline, this combination delivered marked improvements across both datasets, demonstrating that integrating multiple sources of chemical prior information enhances both predictive stability and generalization.

Overall, these ablation results underscore the significant role of molecular fingerprints in strengthening FD-CGNN. The consistently superior performance of the MACCS+RDKit combination highlights the value of integrating chemically curated descriptors with graph-based structural representations when modeling the determinants of metabolic stability.

Parameter Analysis of Fingerprint Weight λ_{fp}

To investigate the influence of global fingerprint information on model performance, we conduct a parameter sensitivity analysis on the hyperparameter λ_{fp}, which controls the contribution of the fingerprint-based channel. As shown in Figure **??**, when evaluated on the HLM dataset, performance across all metrics (AUC, Accuracy, F1-Score, MCC) gradually improves as λ_{fp} increases from 0.1 to 0.7. Notably, the model achieves peak performance at $\lambda_{\text{fp}} = 0.7$, suggesting that the fingerprint features offer complementary global chemical context that enhances the expressiveness of the structural representations. However, further increasing λ_{fp} to 0.9 slightly reduces performance, indicating that overly amplifying global semantics may overshadow local graph features, thus hurting representation balance.

We further validate this trend on the RLM dataset, where a similar pattern emerges. As λ_{fp} increases, the model's AUC, Accuracy, F1, and MCC scores consistently rise up to $\lambda_{\text{fp}} = 0.5$ or 0.7, then plateau or slightly decline at $\lambda_{\text{fp}} = 0.9$. These results confirm the robustness of fingerprint integration across species and underscore the importance of balancing global and local feature contributions in molecular representation learning.

Investigation of novel substructure

To more rigorously assess the generalization ability of FD-CGNN to previously unseen chemotypes, we employ a leave-one-cluster-out evaluation protocol on both the HLM and RLM datasets. Molecular similarities are computed using ECFP fingerprints with the Tanimoto coefficient, and the resulting structures are partitioned into five clusters using K-means (k = 5). PCA visualizations (Figure 6 for HLM; Figure 7 for RLM) show clear inter-cluster separation and strong intra-cluster cohesion, indicating that the clustering captures the dominant structural modes of each dataset and provides an appropriate basis for evaluating out-of-distribution performance.

Under the leave-one-cluster-out protocol, an entire structural cluster is withheld during training, and the model is evaluated exclusively on this held-out cluster. This design more closely reflects real-world scenarios in molecular discovery, where predictive models must often make inferences on scaffolds that diverge substantially from those seen during training. Consequently, the protocol constitutes a more stringent test of structural extrapolation than random splits.

As shown in Figure 3, FD-CGNN exhibits consistently strong performance across all held-out clusters for both datasets, while CMMS-GCL and MS-master display pronounced variability when confronted with clusters exhibiting greater structural divergence. The aggregated results across AUC, ACC, F1, and MCC further reinforce this observation, demonstrating that FD-CGNN learns molecular representations that are robust to shifts in structural distribution and transfer effectively across clusters.

Overall, the leave-one-cluster-out experiments indicate that FD-CGNN captures structural determinants relevant to metabolic stability and generalizes them to novel scaffolds and rare

Table 4. Ablation study of fingerprint configurations on HLM and RLM datasets.

Dataset	Model/Variant	Configuration	AUC	ACC	F1-score	MCC
HLM	w/o FP	—	0.880±0.003	0.811±0.005	0.858±0.006	0.579±0.009
	Single Fingerprint	Morgan	0.884±0.003	0.814±0.004	0.853±0.003	0.600±0.009
		MACCS	0.887±0.002	0.808±0.004	0.850±0.003	0.588±0.004
		RDKit	0.900±0.002	0.816±0.004	0.856±0.002	0.605±0.008
	Dual Fingerprint	Morgan+MACCS	0.883±0.003	0.811±0.004	0.851±0.003	0.594±0.008
		Morgan+RDKit	0.890±0.002	0.817±0.002	0.855±0.005	0.607±0.003
		MACCS+RDKit	**0.913±0.001**	**0.834±0.003**	**0.870±0.005**	**0.643±0.013**
RLM	w/o FP	—	0.849±0.005	0.758±0.004	0.801±0.005	0.494±0.009
	Single Fingerprint	Morgan	0.853±0.003	0.784±0.003	0.822±0.006	0.550±0.004
		MACCS	0.862±0.002	0.789±0.003	0.825±0.003	0.560±0.006
		RDKit	0.856±0.003	0.790±0.004	0.827±0.003	0.562±0.007
	Dual Fingerprint	Morgan+MACCS	0.858±0.003	0.789±0.004	0.826±0.003	0.560±0.008
		Morgan+RDKit	0.849±0.002	0.786±0.002	0.823±0.005	0.553±0.003
		MACCS+RDKit	**0.872±0.005**	**0.805±0.005**	**0.833±0.004**	**0.575±0.011**

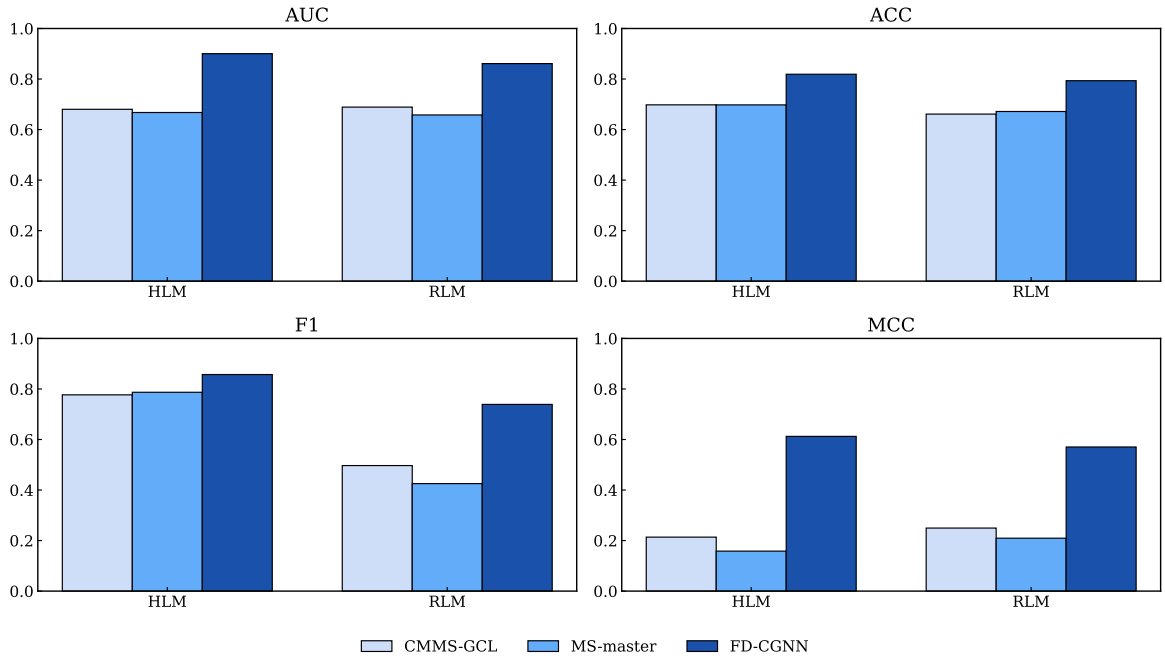

Fig. 3. Performance comparison of FD-CGNN and baseline models on HLM and RLM datasets under the leave-one-cluster-out evaluation.

substructures. Given the sensitivity of metabolic stability to local structural environments—and the frequent need in early-stage drug discovery to evaluate molecules outside the existing chemical domain—such structural generalization is particularly valuable.

Attention Interpretation of FD-CPNN

After establishing the overall performance of the model, we further conducted an interpretability analysis to examine the basis of FD-CGNN's predictions using its self-attention mechanism. The attention module enables the model to focus on structural regions that are highly correlated with metabolic stability, thereby offering an intuitive explanation of its decision process. Specifically, the attention heatmaps show that the model assigns higher weights to fragments known to have substantial influence on stability: on the one hand, aromatic polycyclic frameworks and electron-withdrawing groups such as trifluoromethyl substituents, which generally enhance metabolic stability; on the other hand, benzylic C–H bonds, O-alkyl linkages in ethers, and amine functionalities, which are commonly recognized metabolic soft spots targeted by metabolic enzymes. To examine this behavior, we selected eight representative molecules from each of the human liver microsome (HLM) and rat liver microsome (RLM) datasets (four stable and four unstable compounds) and visualized their attention distributions.

The results indicate that, for stable compounds, the model tends to focus on rigid structural features such as fused aromatic rings, cyano groups, and trifluoromethyl substituents. These fragments are typically resistant to enzymatic attack due to their strong electron-withdrawing effects or steric hindrance, thereby contributing to enhanced metabolic stability. For example, in one

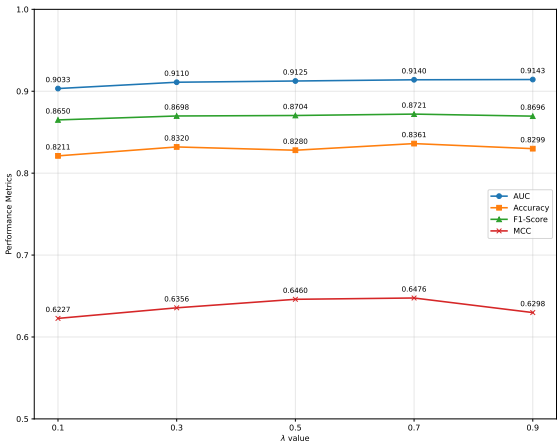

Fig. 4. Parameter Analysis of Fingerprint Weight λ_{fp}

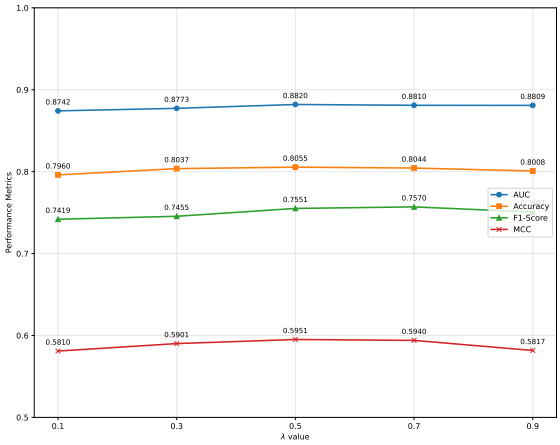

Fig. 5. Parameter Analysis of Fingerprint Weight λ_{fp}

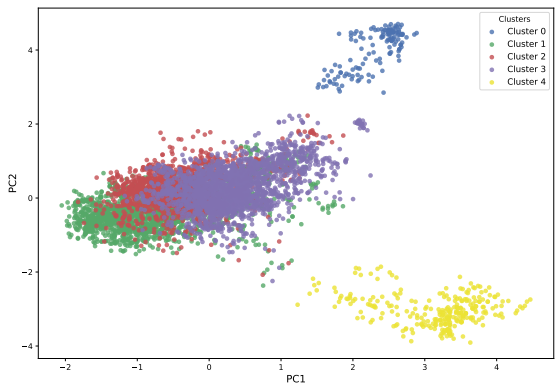

Fig. 6. PCA visualization of K-means clustering on the HLM dataset.

stable HLM molecule, the model concentrates its attention on aromatic rings and trifluoromethyl groups, which form a rigid conjugated backbone and suppress the reactivity of potential metabolic sites. Similarly, in a stable RLM example, the highlighted cyano group represents a metabolically inert moiety whose strong electron-withdrawing effect reduces the reactivity of adjacent carbons.

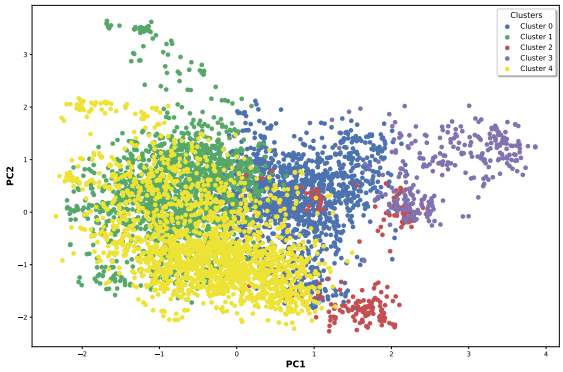

Fig. 7. PCA visualization of K-means clustering on the RLM dataset.

In contrast, for unstable compounds, the model distinctly focuses on structurally flexible, oxidation-prone metabolic soft spots. In an unstable HLM molecule, attention is concentrated on benzylic C–H bonds, which are classical oxidation sites readily attacked by cytochrome P450 enzymes, leading to hydroxylated products and rapid clearance. In an unstable RLM compound, the model prominently highlights tertiary amines, which frequently undergo microsomal N-dealkylation to form more polar and easily excreted metabolites. These attention patterns are highly consistent with well-established drug metabolism mechanisms, demonstrating the model's ability to accurately identify metabolically vulnerable structural motifs.

In summary, the attention visualizations reveal that FD-CGNN possesses a strong capability to recognize structural determinants of metabolic stability: stable fragments in stable compounds receive high attention weights, whereas metabolically sensitive moieties in unstable compounds are precisely captured. This not only strengthens confidence in the model's predictions but also provides structurally interpretable insights that can inform downstream molecular optimization and mechanistic analysis.

Conclusion

In this study, we reviewed existing approaches for predicting the metabolic stability of small molecules and identified several limitations inherent in these methods. Traditional machine-learning models rely mainly on handcrafted descriptors or molecular fingerprints, but lack the capacity to capture fine-grained structural relationships within molecules. Deep-learning methods, particularly GNN-based approaches, utilize molecular graphs to learn structural features; however, many of them focus exclusively on atom-level message passing and fail to incorporate global chemical information or the complementary role of chemical bonds. To address these issues, we proposed FD-CGNN, a fingerprint-augmented cross-dependent graph neural network that integrates both graph-based structural features and fingerprint-derived global descriptors into a unified learning framework.

FD-CGNN employs a dual-channel message passing mechanism that simultaneously models atom-centered and bond-centered structural information, allowing the network to capture more comprehensive chemical relationships. Moreover, the incorporation of MACCS and RDKit fingerprints enables the

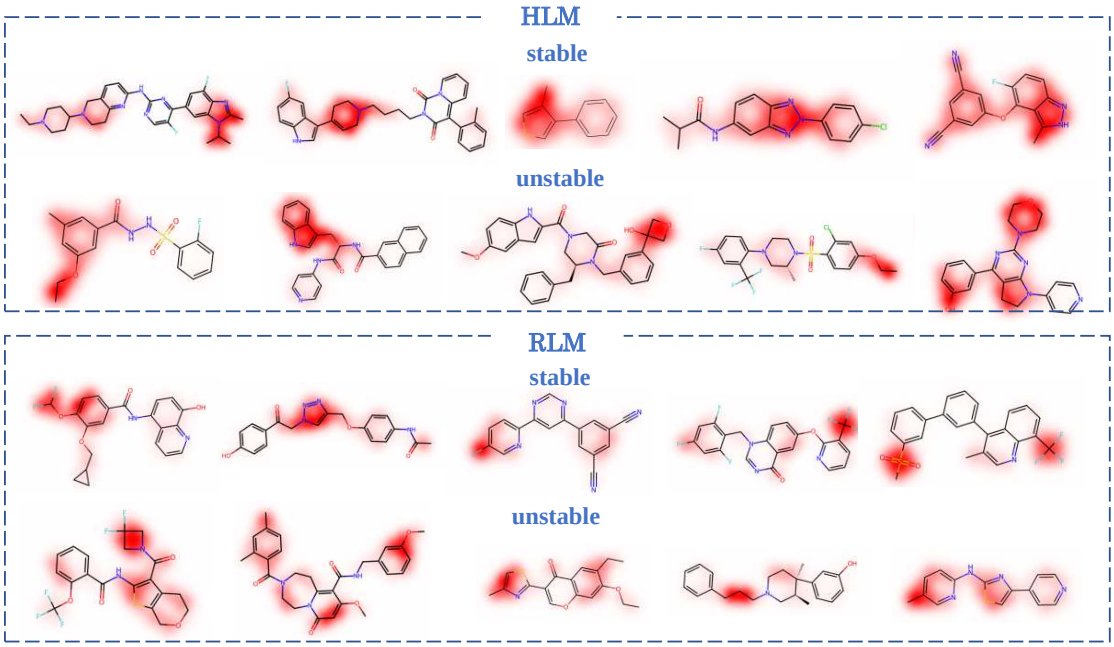

Fig. 8. HLM and RLM Attention.

model to assimilate higher-level molecular semantics beyond what graph topology alone can provide. To validate the effectiveness of the proposed architecture, we conducted extensive comparison and ablation studies on HLM, RLM, and external datasets. Experiments across human and rat microsomal stability datasets further reveal cross-species similarities and differences in metabolic behavior. Additional analyses were performed to examine the contribution of key functional groups and substructures known to influence metabolic stability, confirming that FD-CGNN is capable of identifying metabolically stable motifs as well as vulnerable soft spots within molecular structures.

The strong performance of FD-CGNN relies particularly on its cross-dependent encoding module, which enables reciprocal information flow between atom-level and bond-level representations, and its fingerprint encoder, which injects global chemical knowledge into the learned embeddings. These modules are flexible and can be integrated into other GNN-based molecular modeling frameworks to enhance their structural representation capabilities. Nevertheless, FD-CGNN still exhibits certain limitations. First, the model primarily relies on structural information and does not incorporate additional data modalities such as biological assay readouts, textual annotations, or molecular dynamics features. Second, training solely on species-specific datasets may restrict its generalization to broader chemical spaces or unseen metabolic environments.

In future work, we plan to extend FD-CGNN by incorporating multi-source molecular information to improve representation completeness, and by exploring pre-training strategies or large chemical language models to capture universal chemical knowledge and further enhance model generalization. We also intend to apply the FD-CGNN framework to other ADME-related endpoints to evaluate its applicability across a broader range of pharmacokinetic properties. These efforts may pave the way toward building more versatile and generalizable tools for computational drug design and molecular property prediction.

Key Points

- FD-CGNN achieves state-of-the-art performance in molecular metabolic stability prediction, significantly outperforming baseline models.
- The model integrates local structural features from atom–bond–atom graphs and global semantic features from molecular fingerprints, enabling comprehensive and chemically meaningful molecular representation.
- FD-CGNN demonstrates strong cross-species generalization across both human (HLM) and rat (RLM) metabolic stability datasets.

Author contributions statement

Must include all authors, identified by initials, for example: S.R. and D.A. conceived the experiment(s), S.R. conducted the experiment(s), S.R. and D.A. analysed the results. S.R. and D.A. wrote and reviewed the manuscript.

Acknowledgments

The authors thank the anonymous reviewers for their valuable suggestions. This work is supported in part by funds from the National Science Foundation (NSF: # 1636933 and # 1920920).

References

1. Robert S. Obach. Prediction of human clearance of twenty-nine drugs from hepatic microsomal intrinsic clearance data: an examination of in vitro half-life approach and nonspecific

binding to microsomes. *Drug Metabolism and Disposition*, 27(11):1350–1359, 1999.

2. Neil J. Hewitt, Maria J. Lechón, J. Brian Houston, et al. Primary hepatocytes: current understanding of the regulation of metabolic enzymes and transporter proteins, and pharmaceutical practice for the use of hepatocytes in metabolism, enzyme induction, transporter, clearance, and hepatotoxicity studies. *Drug Metabolism Reviews*, 39(1):159–234, 2007.
3. Wen N. Wu and Linda A. McKown. In Vitro Drug Metabolite Profiling Using Hepatic S9 and Human Liver Microsomes. In Z. Yan and G. W. Caldwell (eds), *Optimization in Drug Discovery*. Methods in Pharmacology and Toxicology. Humana Press, pages 163–177, 2004.
4. Słoczynska K., Gunia-Krzyżak A., Koczurkiewicz P., et al. Metabolic stability and its role in the discovery of new chemical entities. *Acta Pharmaceutica*, 69:345–361, 2019.
5. Paweł Baranczewski, et al. Introduction to in vitro estimation of metabolic stability and drug interactions of new chemical entities in drug discovery and development. *Pharmacological Reports*, 58(4):453–472, 2006.
6. Gajula S. N. R., Nadimpalli N., Sonti R., et al. Drug metabolic stability in early drug discovery to develop potential lead compounds. *Drug Metabolism Reviews*, 53:459–477, 2021.
7. Robert S. Obach. The prediction of human clearance from hepatic microsomal metabolism data. *Current Opinion in Drug Discovery & Development*, 4(1):36–44, 2001.
8. Yuki Naritomi, Satoshi Terashita, Akihiro Kagayama, and Yuichi Sugiyama. Utility of hepatocytes in predicting drug metabolism: comparison of hepatic intrinsic clearance in rats and humans in vivo and in vitro. *Drug Metabolism and Disposition*, 31(5):580–588, 2003.
9. Giorgia Ghibellini, L.S. Vasist, E.M. Leslie, W.D. Heizer, R.J. Kowalsky, B.F. Calvo, and K.L.R. Brouwer. In vitro–in vivo correlation of hepatobiliary drug clearance in humans. *Clinical Pharmacology & Therapeutics*, 81(3):406–413, 2007.
10. Kirchmair J., Göller A. H., Lang D., et al. Predicting drug metabolism: experiment and/or computation? *Nature Reviews Drug Discovery*, 14:387–404, 2015.
11. Albert P. Li. Screening for human ADME/Tox drug properties in drug discovery. *Drug Discovery Today*, 6(7):357–366, 2001.
12. E. E. Litsa, P. Das, and L. E. Kavraki. Machine learning models in the prediction of drug metabolism: challenges and future perspectives. *Expert Opinion on Drug Metabolism & Toxicology*, 17(11):1245–1247, 2021.
13. X. Yang, Y. Wang, R. Byrne, G. Schneider, and S. Yang. Concepts of artificial intelligence for computer-assisted drug discovery. *Chemical Reviews*, 119(18):10520–10594, 2019. doi: 10.1021/acs.chemrev.8b00728.
14. Liang Tao, Peng Zhang, Cheng Qin, Shao-Yong Chen, Chao Zhang, Zhiwei Chen, Feng Zhu, Shao-Yong Yang, Yong-Qiang Wei, and Yu-Zong Chen. Recent progresses in the exploration of machine learning methods as in-silico ADME prediction tools. *Advanced Drug Delivery Reviews*, 86:83–100, 2015.
15. Y. Hu, R. Unwalla, R. A. Denny, J. Bikker, L. Di, and C. Humblet. Development of QSAR models for microsomal stability: identification of good and bad structural features for rat, human and mouse microsomal stability. *Journal of Computer-Aided Molecular Design*, 24:23–35, 2010.
16. C. W. Yap. PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. *Journal of Computational Chemistry*, 32:1466–1474, 2011.
17. Perryman A. L., Stratton T. P., Ekins S., Freundlich J. S. Predicting mouse liver microsomal stability with "pruned" machine learning models and public data. *Pharmaceutical Research*, 33:433–449, 2016.
18. Podlewska S., Kafel R. MetStabOn – Online platform for metabolic stability predictions. *International Journal of Molecular Sciences*, 19(4):1040, 2018.
19. M. Shen, Y. Xiao, A. Golbraikh, V. K. Gombar, and A. Tropsha. Development and validation of k-nearest-neighbor QSPR models of metabolic stability of drug candidates. *Journal of Medicinal Chemistry*, 46:3013–3020, 2003.
20. P. Zhao, Y. Peng, X. Xu, Z. Wang, Z. Wu, W. Li, Y. Tang, and G. Liu. In silico prediction of mitochondrial toxicity of chemicals using machine learning methods. *Journal of Applied Toxicology*, 41:1518–1526, 2021.
21. Ryu J. Y., Lee J. H., Lee B. H., et al. PredMS: a random forest model for predicting metabolic stability of drug candidates in human liver microsomes. *Bioinformatics*, 38(2):364–368, 2022.
22. Li L., Zhou L., Liu G., et al. In silico prediction of human and rat liver microsomal stability via machine learning methods. *Chemical Research in Toxicology*, 35(9):1614–1624, 2022.
23. Wieder O., Kohlbacher S., Kuenemann M., et al. A compact review of molecular property prediction with graph neural networks. *Drug Discovery Today: Technologies*, 37:1–12, 2020.
24. D. Jiang, Z. Wu, C.-Y. Hsieh, G. Chen, B. Liao, Z. Wang, C. Shen, D. Cao, J. Wu, and T. Hou. Could graph neural networks learn better molecular representation for drug discovery? A comparison study of descriptor-based and graph-based models. *Journal of Cheminformatics*, 13(1):12, 2021. doi: 10.1186/s13321-020-00479-8.
25. Renn A., Su B.-H., Liu H., et al. Advances in the prediction of mouse liver microsomal studies: from machine learning to deep learning. *WIREs Computational Molecular Science*, 11:e1479, 2021.
26. Xiong Z., Wang D., Liu X., et al. Pushing the boundaries of molecular representation for drug discovery with the graph attention mechanism. *Journal of Medicinal Chemistry*, 63(16):8749–8760, 2019.
27. Ma H., Bian Y., Rong Y., et al. Cross-dependent graph neural networks for molecular property prediction. *Bioinformatics*, 38:2003–2009, 2022.
28. Wang T., Li Z., Zhuo L., et al. MS-BACL: enhancing metabolic stability prediction through bond graph augmentation and contrastive learning. *Briefings in Bioinformatics*, 25:bbae127, 2024.
29. Cai H., Zhang H., Zhao D., et al. FP-GNN: a versatile deep learning architecture for enhanced molecular property prediction. *Briefings in Bioinformatics*, 23:bbac408, 2022.
30. Hou Yee Choo, JunJie Wee, Cong Shen, and Kelin Xia. Fingerprint-Enhanced Graph Attention Network (FinGAT) Model for Antibiotic Discovery. *Journal of Chemical Information and Modeling*, 63(10):2928–2935, 2023. DOI: 10.1021/acs.jcim.3c00045.

31. Bing-Xue D., Long Y., Li X., et al. CMMS-GCL: cross-modality metabolic stability prediction with graph contrastive learning. *Bioinformatics*, 39(8):btad503, 2023.
32. J. L. Durant, B. A. Leland, D. R. Henry, et al. Reoptimization of MDL keys for use in drug discovery. *Journal of Chemical Information and Computer Sciences*, 42:1273–1280, 2002.
33. Yun S., Nam G., Koo J. HiMolformer: Integrating graph and sequence representations for predicting liver microsome stability with SMILES. *Computational Biology and Chemistry*, 113:108263, 2024.

Author Name. This is sample author biography text. The values provided in the optional argument are meant for sample purposes. There is no need to include the width and height of an image in the optional argument for live articles. This is sample author biography text this is sample author biography text this is sample author biography text this is sample author biography text this is sample author biography text this is sample author biography text this is sample author biography text this is sample author biography text.

Author Name. This is sample author biography text this is sample author biography text this is sample author biography text this is sample author biography text this is sample author biography text this is sample author biography text this is sample author biography text this is sample author biography text this is sample author biography text this is sample author biography text.