Subgraph Information Bottleneck with Causal Dependency for Stable Molecular Relational Learning

Peiliang Zhang[1,2], **Jingling Yuan**[1,3*], **Chao Che**[4], **Yongjun Zhu**[2], **Lin Li**[1]
[1]School of Computer Science and Artificial Intelligence,
Wuhan University of Technology, Wuhan, 430070, China
[2]Department of Library and Information Science, Yonsei University, Seoul, 03722, Korea
[3]Hubei Key Laboratory of Transportation Internet of Things,
Wuhan University of Technology, Wuhan, 430070, China
[4]Key Laboratory of Advanced Design and Intelligent Computing (Dalian University),
Ministry of Education, Dalian, 116622, China
cheungbl@ieee.org, yjl@whut.edu.cn, chechao@gmail.com, zhu@yonsei.ac.kr, cathylilin@whut.edu.cn

Abstract

Molecular Relational Learning (MRL) is widely applied in molecular sciences. Recent studies attempt to retain molecular core information (e.g., substructures) by Graph Information Bottleneck but primarily focus on information compression without considering the causal dependencies of chemical reactions among substructures. This oversight neglects the core factors that determine molecular relationships, making maintaining stable MRL in distribution-shifted data challenging. To bridge this gap, we propose the Causal Subgraph Information Bottleneck (CausalGIB) for stable MRL. CausalGIB leverages causal dependency to guide substructure representation and integrates subgraph information bottleneck to optimize the core substructure representation, generating stable representations. Specifically, we distinguish causal and confounding substructures by noise injection and substructure interaction based on causal analysis. Furthermore, by minimizing the discrepancy between causal and confounding information within subgraph information bottleneck, CausalGIB captures core substructures composed of causal substructures and aggregates them into molecular representations to improve their stability. Experimental results on nine datasets demonstrate that CausalGIB outperforms state-of-the-art models in two tasks and significantly enhances model's stability in distribution-shifted data.

1 Introduction

Molecular Relational Learning (MRL) predicts the interaction between molecular pairs by mining structures and properties [Fang *et al.*, 2024; Du *et al.*, 2024]. MRL has garnered significant attention in natural science research due to its widespread applications in new material design and drug discovery [Pei *et al.*, 2024; Zhang *et al.*, 2024a]. The molecular structural information mining-based methods have advantages in MRL [Wang *et al.*, 2024; Job *et al.*, 2023; Zhang *et al.*, 2025b]. However, these methods exhibit varying levels of performance fluctuations in real-world scenarios. Consequently, enhancing model stability is essential for the effective implementation of MRL.

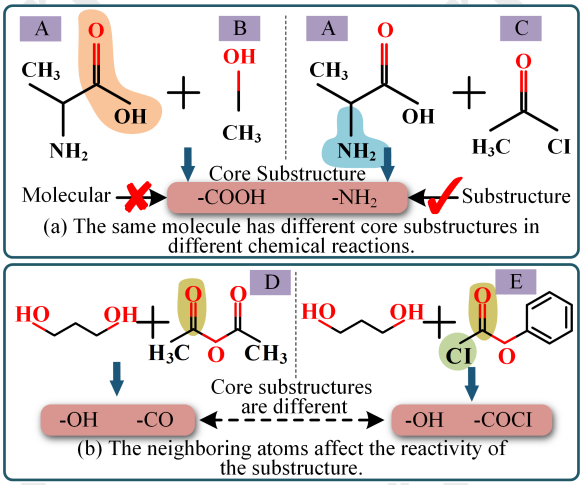


Figure 1: The motivating example.

The Graph Information Bottleneck (GIB) [Wu *et al.*, 2020] has recently been applied to extracting and identifying core subgraphs. GIB generates representation that represent the minimal sufficient information of the input graph by information compression [Yuan *et al.*, 2024; Fan *et al.*, 2022; Seo *et al.*, 2024]. The theory provides a novel solution for MRL, with the key problem being to efficiently capture the molecular representation that are most beneficial for MRL [Li *et al.*, 2024; Seo *et al.*, 2023]. Related research addresses the above problem by maximizing the mutual information between molecules and prediction targets [Yang *et al.*, 2023; Li *et al.*, 2024]. Specifically, these methods control the information flow between molecules and the targets to gener-

* Corresponding author.

ate representation that contain minimal sufficient information about molecules [Lee *et al.*, 2023a]. The generalizability of GIB-based MRL are improved by integrating the interaction information into molecular representations.

The core challenge of MRL is the instability of model predictions caused by distribution differences between training data and real-world data [Lee *et al.*, 2023b]. An effective way to deal with this challenge is to capture the core substructures that determine molecular reactions [Fang *et al.*, 2024]. Nevertheless, GIB-based MRL have the following limitations: (1) **Missing Substructure Causal Information in Chemical Reactions:** The nature of chemical reactions is the dynamic rearrangement between substructures [Hu *et al.*, 2017]. The substructures of the same paired molecule causally related to prediction targets in different chemical reactions may be different [Iwasaki and Nozaki, 2024]. As shown in Fig. 1(a), when molecule A reacts with B, the reactive substructure in A is -COOH, while with C, it is -NH_2. GIB typically focuses on information compression and fails to model causal interactions between substructures, making it challenging to identify the substructures that truly determine chemical reactions. (2) **Challenges in Capturing Nuanced Differences Between Substructures**: Substructures are susceptibly impacted by active atoms within the molecule, altering their properties in chemical reactions [Zhang *et al.*, 2025a]. As shown in Fig. 1(b), molecules D and E contain a -CO, but the substructures involved in reactions with the same molecule differ. This is because the -CI in E has the stronger electron adsorption, enhancing the reactivity of -CO, which leads to the -COCI substructure involved in chemical reaction. Considering MRL solely at the molecular level makes capturing these nuanced differences between substructures difficult. These limitations restrict the ability of GIB-based relevant methods to maintain stable relational learning.

In this paper, we first establish the causal associations between substructures and molecular relationships in MRL and analyze the key factors influencing the backdoor paths. Subsequently, we propose the Causal Subgraph Information Bottleneck (CausalGIB), which incorporates the Causally-driven Substructure Representation Learning (CauSRL) and the Synergistic Optimization of Subgraph Information Bottleneck (SOSIB). CauSRL integrates causal dependencies into the optimization of substructure representation through noise injection and interactions, effectively distinguishing causal and confounding substructures. SOSIB encourages the model to learn the maximum causal mutual information between causal substructures and prediction targets while minimizing the information among confounding substructures, thereby generating stable molecular representations. The main contributions are as follows.

- We propose CausalGIB that introduces subgraph causal information into GIB to improve the stability of MRL. To the best of our knowledge, it is the first work to address the problem of causal substructure selection in different chemical reactions.
- We provide a formal validation for the design of the loss function, grounded in theoretical reasoning, to effectively capture subtle distinctions between substructures by minimizing the disparity between confounding and causal information.
- Experiments on two tasks in nine real-world datasets and three different data distributions demonstrate that the predictive performance and stability of CausalGIB outperforms seven state-of-the-art models.

2 Related Work

Molecular Relational Learning. MRL is an essential task in molecular representation, encompassing various applications [Pei *et al.*, 2024; Zhang *et al.*, 2024b; Job *et al.*, 2024]. We focus on Molecular Interaction (MI) and Drug-Drug Interaction (DDI) predictions. In MI prediction, CIGIN [Pathak *et al.*, 2020] employed message-passing network and collaborative attention to encode molecular atoms and predict solvation energy. CMRL [Lee *et al.*, 2023b] combined molecular representation with causal relationships, identifying substructures related to chemical reactions. In DDI prediction, SA-DDI [Yang *et al.*, 2022], SSI-DDI [Nyamabo *et al.*, 2021], and DSN-DDI [Li *et al.*, 2023] extracted substructure information from drugs in various ways to represent drug molecules.

Graph Information Bottleneck. GIB is an extension and development of the Information Bottleneck theory within graph theory, which has been widely applied in molecular relational learning [Sun *et al.*, 2022], multi-agent communication [Ding *et al.*, 2024b; Ding *et al.*, 2024a], and graph learning interpretability [Yuan *et al.*, 2024]. Specifically, PGIB [Seo *et al.*, 2023] introduced prototype information into the GIB framework and used it as a compression bottleneck to improve the interpretability of graph representation learning. [Ding *et al.*, 2024b; Ding *et al.*, 2024a] optimized robust communication in multi-agent systems by setting mutual information between message and action choices as the GIB optimization objectives. CGIB [Lee *et al.*, 2023a] combined molecular relational learning with the GIB framework to model the nature of chemical reactions.

Despite the advantages of the related methods in MRL, they fail to consider the causal dependencies between substructure pairs. Therefore, we aim to improve the model's stability in distribution-shifted data by introducing causal information to identify core substructures of molecules.

3 Preliminaries

In this section, we define stable MRL and construct the Causal Substructure Model (CSM). Based on the analysis results of CSM, we identify the key factors influencing the backdoor paths in MRL.

3.1 Stable MRL

For any molecule G, it can be represented as $G = (V, E, X, A)$. Here, $V = \{v_1, v_2, \cdots, v_N\}$ denotes the set of nodes. $E \in N \times N$ represents the connections between atoms within the molecule, which is closely related to the adjacency matrix A. If $(v_i, v_j) \in E$, then $A_{ij} = 1$; otherwise, $A_{ij} = 0$. X is the feature matrix, consisting of the atom feature representations. For the given molecular pair (G_x, Y_{xy}, G_y), the

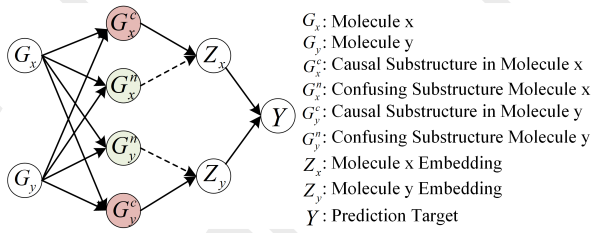


Figure 2: The CSM for stable MRL.

objective of Stable MRL is to generate a stable molecular embedding representation $\mathcal{Z}(G^{st})$ and to use it for molecule pair interaction prediction in different data distributions. Formally, $Y_{xy} = F_s(\mathcal{Z}(G_x^{st}), \mathcal{Z}(G_y^{st}))$. Therefore, CausalGIB aims to learn stable representations of molecules and improve the model's predictive performance.

3.2 CSM for Stable MRL

We define and construct the CSM for stable MRL, as shown in Fig. 2. The causal relationships illustrated in Fig. 2 are described as follows:

- $G_x \rightarrow G_x^c \leftarrow G_y$: G_x^c is the causal core substructure of molecule G_x, jointly determined by G_x and G_y, and varies with changes in G_y.
- $G_y \rightarrow G_y^c \leftarrow G_x$: G_y^c is the causal core substructure of molecule G_y, jointly determined by G_y and G_x, and varies with changes in G_x.
- $G_x \rightarrow G_x^n \leftarrow G_y$: G_x^n is the confounding substructure of molecule Gx, representing the substructure involved of G_x other than G_x^c.
- $G_y \rightarrow G_y^n \leftarrow G_x$: G_y^n is the confounding substructure of molecule G_y, representing the substructure involved of G_x other than G_y^c.
- $G_x^c \rightarrow Z_x$: Z_x is the representation of the molecule G_x, obtained by encoding the causal core substructure G_x^c.
- $G_y^c \rightarrow Z_y$: Z_y is the representation of the molecule G_y, obtained by encoding the causal core substructure G_y^c.
- $Z_x \rightarrow Y \leftarrow Z_y$: Y represents the interaction between molecular pairs, which depends on both G_x and G_y in molecular relationship learning.

Through backdoor path analysis, we identify the core factors influencing molecular relationships Y. The backdoor paths are $G_x^n \leftarrow G_y \rightarrow G_y^c \leftarrow G_x \rightarrow G_x^c \rightarrow Z_x \rightarrow Y$, $G_y \rightarrow G_x^n \leftarrow G_x \rightarrow G_x^c \rightarrow Z_x \rightarrow Y$, $G_x \rightarrow G_y^n \leftarrow G_y \rightarrow G_y^c \rightarrow Z_y \rightarrow Y$, and $G_x^n \leftarrow G_x \rightarrow G_x^c \leftarrow G_y \rightarrow G_y^c \rightarrow Z_y \rightarrow Y$. By analyzing the backdoor paths in the CSM, we observe that G_x^c and G_y^c serve as the backdoor criteria constraining molecular relationships Y. Considering the specificity of molecular relationship learning, i.e., the substructure is either core substructure G^c or confounding substructure G^n. Therefore, our objective is to establish stable interactions between G^c and Y by eliminating the influence of G^n on G^c.

4 Methodology

Motivated by CSM, we propose the CausalGIB with joint optimization of predictive loss and confusing information for stable MRL, as shown in Fig. 3. CausalGIB models substructure dynamic binding in CSM by noise injection and substructure interactions, utilizing it to optimize substructure embeddings (Sec. 4.1). Furthermore, we propose the optimization objective for Subgraph Information Bottleneck (SIB) to perturb confounding information while preserving the association between causal substructures and the target (Sec. 4.2).

4.1 Causally-Driven Substructure Representation Learning (CauSRL)

CauSRL utilizes the GNN encoder and an adjacency aggregation strategy to generate representations for irregular substructures. Subsequently, we intervene with causal dependencies between substructures via noise injection to identify substructures with strong causal relationships.

Irregular Substructure Representation Learning

We utilize substructure coefficient and neighbor feature weighted aggregation to extract molecular substructures. For the molecule G, the substructure G^s consists of the central atom v_c and its K-hopping neighboring nodes v_c^k. Formally, $\mathcal{Z}(G^s) = \sum_{k=1}^{K} \sum_{v_s \in v_c^k} \mathcal{C}^k \text{GNN}^k(v_s)$. GNN$(*)$ can be selected from GIN [Xu *et al.*, 2019], MPNN [Gilmer *et al.*, 2017], GAT [Veličković *et al.*, 2018], or GCN [Kipf and Welling, 2017]. The calculation of $\mathcal{C}^k$ is defined as follows:

$$\mathcal{C}^k = \frac{\sigma(\mathcal{Z}(v_n)) \cdot \log \sigma(\mathcal{Z}(v_n))}{\sum_{v_u \in v_c^k} \sigma(\mathcal{Z}(v_u)) \cdot \log \sigma(\mathcal{Z}(v_u))} \tag{1}$$

Here, σ is the activation function. $\mathcal{Z}(*)$ is the embedding representation of node. We use $\mathcal{Z}(G^s)$ to obtain the molecular representation $\mathcal{Z}(G) = \mathcal{Z}(G^{s_1}) || \cdots || \mathcal{Z}(G^{s_J})$. J is the number of molecular substructures. Next, we compute the importance R_i of different substructures G^{s_i} within the molecule G, which is used for subsequent noise injection.

$$R_i = \sigma(\mathcal{Z}(G^{s_i}), \mathcal{Z}(G)) + \sum_{j \neq i}^{J} \sigma(\mathcal{Z}(G^{s_j}), \mathcal{Z}(G)) \tag{2}$$

Causal Substructure Recognition with Dependency Intervention (CSDI)

To better identify causal relationships between substructures, we introduce causal dependence to model the interaction behaviour. Causal dependence refers to a scenario where a change in one variable, due to an external intervention, directly induces a change in another variable, with this effect being unexplained by other variables. In other words, if a change in the topology of a substructure increases its interaction probability with the paired substructure, the two substructures exhibit a higher causal dependency.

Consequently, we interfere with the topological information of substructures through noise injection and quantify causal dependency by causal contribution scores to identify and select causal substructures. Specifically, given the adjacency matrix A of molecule G, our objective is to inject noise into the adjacent atoms of the substructure G^s and reconstruct the associations between them via the molecule's adjacency matrix A to obtain an intervening topologically informed substructure G^{ns}. Formally, $G^{ns} = G^s + \triangle\ v_{ns}$, where $\triangle\ v_{ns}$

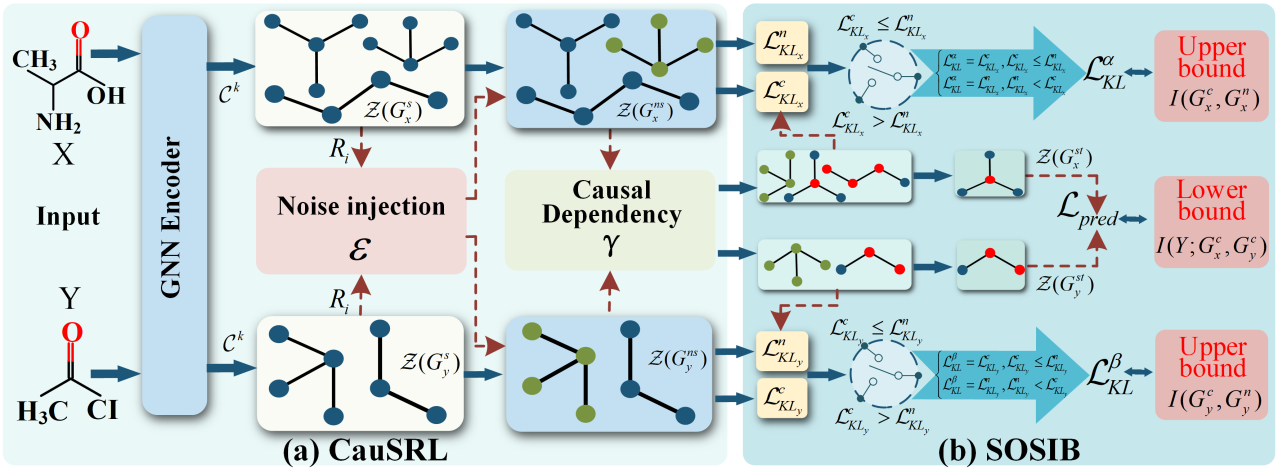


Figure 3: The model structure of CausalGIB. (a) CauSRL incorporates noise injection to compute causal dependencies between substructures. (b) SOSIB perturbs confounding information and preserves the association between causal substructures and target.

represents the added atomic nodes. The embedding representation of the substructure after noise injection is defined as:

$$\mathcal{Z}(G^{ns}) = \lambda_i \mathcal{Z}(G^{s_i}) + \varepsilon\Big(\sum_{v_{ns}\in\mathcal{N}} \mathcal{Z}(v_{ns})(1-\log\sigma(\mathcal{Z}(v_{ns})))\Big)$$
$$\lambda_i \sim \text{Bernoulli}\left(\text{Sigmoid}\left(R_i\right)\right), \varepsilon \sim (\mu_{\mathcal{Z}(v_{ns})}, \sigma^2_{\mathcal{Z}(v_{ns})}) \tag{3}$$

where $\mathcal{N}$ represents the set of neighboring atoms for substructure G^s, and ε is the noise function. μ and σ^2 are mean and variance of $\mathcal{Z}(v_{ns})$, respectively.

We compute the causal contribution difference γ between substructures before and after noise injection by incorporating stochastic interaction behaviours. γ is represented by the ratio of causal dependencies between substructures.

$$\gamma = \frac{\mathbb{E}_{\sim G_x^{ns}, G_y^{s_j}}[G_x^{ns}|G_y^{s_j}] - \mathbb{E}_{\sim G_x^{s_i}, G_y^{s_j}}[G_x^{s_i}|G_y^{s_j}]}{\mathbb{E}_{\sim G_x^{s_i}, G_y^{s_j}}[G_x^{s_i}|G_y^{s_j}]} \tag{4}$$

where G_x^{ns} represents the substructure generated from $G_x^{s_i}$ after noise injection. $G_y^{s_j}$ is the paired substructure of $G_x^{s_i}$, belonging to molecules G_y and G_x, respectively.

The numerator in Eq (4) denotes the average causal effect of the perturbed substructure G_x^{ns} on the paired substructure $G_y^{s_j}$, while the denominator represents the causal dependency of $G_x^{s_i}$ on $G_y^{s_j}$. The value of γ indicates whether the causal dependency between substructures changes before and after noise injection. Based on γ, we identify the causal substructures G^c between paired molecules. Specifically, if $\gamma > 0$, G_x^{ns} is classified as causal substructure; otherwise, $G_x^{s_i}$ is considered causal substructure. The remaining substructures are designated as confounding substructures G^n.

After CSDI screening, CausalGIB leverages SIB further to optimize the representations of G^c and G^n and guide the model learning. We identify the SIB-optimized causal substructures as core substructures G^{co} and use them to generate stable molecular representations $\mathcal{Z}(G^{st})$.

4.2 Synergistic Optimization of Subgraph Information Bottleneck (SOSIB)

Considering that the nature of MRL is dynamic interaction between substructures [Lee *et al.*, 2023b], we extend GIB to the subgraph level. We derive the optimization objective for the SIB and exhibit its transformational solution process.

Subgraph Information Bottleneck

Based on the analysis of CSM, we need to generate a stable MRL that meets the following conditions: perturbing the confounding information and maintaining the causal substructure information. Therefore, we propose the optimization objective of SIB from the subgraph level.

Definition 1. (***SIB***) *Given a pair of graphs and their interaction* (G_x, Y, G_y), (G_x^c, G_y^c) *and* (G_x^n, G_y^n) *are the causal and confounding subgraph pairs of* (G_x, G_y) *and* Y*, respectively. Under the SIB principle, the optimal causal subgraphs and the optimization objective are defined as:*

$$G^o = \arg\min\left(-I(Y; G_x^c, G_y^c) + I(G^c, G^n)\right) \tag{5}$$

Optimizing Eq (5) maximizes the preservation of causal information while minimizing the perturbation of confounding information on MRL. To simplify the computation of mutual information, we reformulate and solve the upper bounds for $-I(Y; G_x^c, G_y^c)$ and $I(G^c, G^n)$ to optimize Eq (5).

Minimizing $I(G^c, G^n)$ in Eq (5)

The goal of minimizing $I(G^c, G^n)$ is to extract subgraphs that contain less confounding information. Inspired by [Lee *et al.*, 2023a], we minimize the upper bound of I_{si} as the conditionally independent distribution between G^c, G^n, and R in Eq (2). Since G^c, G^n are subgraphs of G, according to principles of information theory, we have $I_{si} = I(G^c, G^n) \leq \min(I(G^c, G), I(G^n, G))$. The minimization of the upper bound of I_{si} can be expressed as:

$$I_{si} \leq \min\Big(\iint \mathcal{F}_{cn}(G^c; G) dG^c dG, \iint \mathcal{F}_{cn}(G^n; G) dG^n dG\Big)$$
$$= \min(\mathcal{L}_{KL}^c, \mathcal{L}_{KL}^n) \tag{6}$$

where p is a posterior distribution function.

Eq (6) indicates that minimizing $(\mathcal{L}^c_{KL}, \mathcal{L}^n_{KL})$ provides an upper bound for $I(G^c, G^n)$. In the context of MRL, for a paired molecule (G_x, G_y), the $I(G^c, G^n)$ is refined as $I(G^c_x, G^n_x) + I(G^c_y, G^n_y)$.

Minimizing $-I(Y; G^c_x, G^c_y)$) in Eq (5)

The goal of minimizing $-I(Y; G^c_x, G^c_y))$ is to capture the stable representation of causal substructures that determine molecular relationship learning. We maximize $I_{ca} = I(Y; G^c_x, G^c_y)$ through γ in Eq (4), the causal substructures G^c_x, G^c_y, and the prediction target Y:

$$\begin{aligned} I_{ca} \geq & \int p(Y)(\iint \mathcal{F}_{ca}(G^c_x, G^c_y|Y) dG^c_x dG^c_y) dY \\ & = -\mathcal{L}_{pred}(q(\gamma(G^c_{x_n}, G^c_{y_m})|Y)|\hat{Y}) \end{aligned} \tag{7}$$

where $q(G^c_x, G^c_y|Y)$ is the variational approximation distribution used to approximate the posterior distribution $p(G^c_x, G^c_y|Y)$.

The Eq (7) indicates that minimizing the prediction loss $\mathcal{L}_{pred}$ achieves the minimization of $-I(Y; G^c_x, G^c_y)$. Maximizing the mutual information between γ and Y helps the model identify the causal relationship between G^c and Y in the CSM backdoor paths.

4.3 Model Loss and Analysis

We define the model training objective as the sum of $\mathcal{L}_{pred}$ and $\mathcal{L}_{KL}$ as follows:

$$\mathcal{L} = \mathcal{L}_{pred} + \alpha \mathcal{L}^{\alpha}_{KL} + \beta \mathcal{L}^{\beta}_{KL} \tag{8}$$

where α, β are hyperparameters. $\mathcal{L}_{KL} = \min(\mathcal{L}^c_{KL}, \mathcal{L}^n_{KL})$ controls the impact of confounding substructures on the causal substructure representation. $\mathcal{L}_{pred}$ is computed by Mean Absolute Error Loss for regression task and Cross-Entropy Loss for classification task.

5 Experiments

In this section, we analyze the experiments to answer the following questions: **RQ1**: Can CausalGIB improve the stability of MRL in distribution-shifted data? **RQ2**: How does the performance of CausalGIB in MRL, and whether CausalGIB is susceptible to backbone? **RQ3**: Can the results of ReAlign-Fit be visually supported? **RQ4**: What is the key to CausalGIB's performance improvement?

Following the related research [Lee *et al.*, 2023a; Lee *et al.*, 2023b; Boulougouri *et al.*, 2024], we conduct extensive MI prediction and DDI prediction experiments on nine datasets. To better simulate the real world data distribution, we set up three different data distributions (Original Partitioning (Original), Rule-based Partitioning (P1) and Graph-based Partitioning (P2)) in DDI prediction and use it for CausalGIB learning.

5.1 Stability Analysis of CausalGIB (RQ1)

We compared the stability of CausalGIB in distribution-shifted data with four backbone models (GCN [Kipf and Welling, 2017], GAT [Veličković *et al.*, 2018], MPNN [Gilmer *et al.*, 2017], and GIN [Xu *et al.*, 2019]) and seven domain models (MIRACLE [Wang *et al.*, 2021], CIGIN [Pathak *et al.*, 2020], SSI-DDI [Nyamabo *et al.*, 2021], DSN-DDI [Li *et al.*, 2023], CMRL [Lee *et al.*, 2023b], CGIB [Lee *et al.*, 2023a] and MMGNN [Du *et al.*, 2024]). We quantify model stability by calculating the performance retention (PR) in distribution shifts, inter-distribution performance variation (PV) for AUROC, ACC, F1, Pre, and AUPR.

$$PR_M = \frac{Eva^{P_i}_M}{Eva^{Ori}_M} \times 100\% \tag{9}$$

$$PV_M = \frac{\sum_{i,j} |Eva^{P_i}_M - Eva^{Ori_j}_M|}{\sum_i P_i} \times 100\% \tag{10}$$

where Eva represents the evaluation metrics. Eva^P_M and Eva^{Ori}_M denote the prediction performance in the original and drift distributions, respectively. $M \in \{AUROC, ACC, F1, Pre, AUPR\}$ and $P_i \in \{P1, P2\}$. It is important to clarify that we conduct model training and testing separately in three distinct data distribution scenarios (Original, P1, and P2), rather than directly applying the model trained on distribution Original for inference in distributions P1 and P2.

The experimental results are shown in Table 1 and Table 2. From Tables 1 and 2, the stability of the MRL models is influenced by data distribution.

Overall Analysis of Model Stability: As data distribution differences increase, MRL models exhibit varying degrees of performance degradation. However, it is noteworthy that CausalGIB consistently maintains the highest in different distributions. Compared to the second-best model, CausalGIB achieves approximately 3% to 6% improvements in RPD, PR, and PV. These experimental results indicate that CausalGIB learns more stable molecular representations under distribution shifts, which we consider a key factor to its success.

Analysis of the Importance of Causal Dependency: As shown in Table 2, methods that do not adequately account for causal relationships (e.g., CIGIN, CGIB) exhibit at least a 1.5% lower PV_{Pre} in DDI prediction compared to methods that fully consider causal dependencies (e.g., CausalGIB). Similar trends are observed in the experimental results shown in Table 1. This underscores the critical importance of causal dependencies between prediction targets and substructures for the stability of MRL.

Adaptability to Distribution-Shifted Data: In Table 1, backbone models demonstrate poor adaptability to data shifts, showing significant performance drops and fluctuations. In contrast, CausalGIB achieves the best results in multiple evaluation metrics on different datasets compared to other MRL domain models. In P1 scenario, the performance fluctuations of CausalGIB range from 15% to 20%, which is significantly smaller than those of MMGNN and CMRL. These findings suggest that CausalGIB exhibits acceptable adaptability to data distribution shifts. We attribute this result to two main reasons: the molecular representations generated by CausalGIB contain robust causal information between substructures and prediction targets. On the other hand, CausalGIB mitigates the influence of confounding substructures on molecular representations by weakening the associations between confounding and causal substructures.

Model	ZhangDDI				HetionteDDI				DrugBankDDI			
	Original → P1		Original → P2		Original → P1		Original → P2		Original → P1		Original → P2	
	PR_{ACC}↑	PR_{AUR}↑	PR_{ACC}↑	PR_{AUR}↑	PR_{ACC}↑	PR_{AUR}↑	PR_{ACC}↑	PR_{AUR}↑	PR_{ACC}↑	PR_{AUR}↑	PR_{ACC}↑	PR_{AUR}↑
GCN (ICLR'17)	79.03	77.82	56.46	48.41	76.41	73.33	66.45	57.93	72.78	79.79	51.36	45.27
GAT (ICLR'18)	81.27	76.86	59.35	51.49	77.19	73.11	69.84	57.11	73.73	80.04	51.51	49.98
MPNN (ICML'17)	80.82	73.25	55.36	48.90	76.67	73.13	71.08	58.08	76.19	73.44	53.26	47.01
GIN (ICLR'19)	81.96	76.62	61.05	47.61	79.01	74.06	72.26	58.95	78.55	76.94	52.96	45.20
MIRACLE (WWW'21)	82.35	76.64	53.56	51.95	79.94	75.69	73.72	58.20	80.69	77.18	50.09	51.35
CIGIN (AAAI'20)	84.21	80.52	57.79	53.60	77.66	74.44	70.09	60.33	76.60	82.97	54.36	60.04
SSI-DDI (BIB'21)	83.81	**87.60**	56.30	51.76	76.62	71.45	67.64	57.82	85.34	78.37	50.92	59.74
DSN-DDI (BIB'23)	84.91	81.69	60.58	55.95	76.00	74.50	68.24	61.09	77.81	80.85	52.07	59.92
CMRL (KDD'23)	84.44	81.96	62.21	60.59	76.34	72.11	69.95	60.28	78.78	77.09	57.02	60.14
CGIB (ICML'23)	85.02	85.08	63.24	61.67	80.02	77.36	70.17	64.10	81.35	79.25	58.54	60.21
MMGNN (IJCAI'24)	85.40	80.39	60.96	60.94	76.67	77.53	70.67	64.21	76.68	80.17	58.00	59.65
CausalGIB	**88.14**	86.90	**68.04**	**65.31**	**82.86**	**81.49**	**75.50**	**70.64**	**86.27**	**87.36**	**65.51**	**67.34**
	PV_{ACC}↓	PV_{AUR}↓	PV_{F1}↓	PV_{AUP}↓	PV_{ACC}↓	PV_{AUR}↓	PV_{F1}↓	PV_{AUP}↓	PV_{ACC}↓	PV_{AUR}↓	PV_{F1}↓	PV_{AUP}↓
GCN (ICLR'17)	26.87	33.81	24.39	29.14	25.04	31.94	28.50	27.20	31.87	34.08	33.46	38.34
GAT (ICLR'18)	24.98	33.00	21.88	29.81	23.29	32.92	25.14	28.99	31.52	31.94	30.69	39.20
MPNN (ICML'17)	26.99	35.95	22.46	33.76	23.18	32.72	23.25	30.10	30.80	38.15	35.19	36.44
GIN (ICLR'19)	24.39	35.30	19.90	33.50	21.29	31.16	22.59	27.85	29.18	36.04	31.07	32.63
MIRACLE (WWW'21)	25.94	33.23	18.44	34.59	20.16	30.70	**18.04**	28.05	29.41	32.94	31.21	31.52
CIGIN (AAAI'20)	24.76	30.68	16.94	27.42	22.98	30.60	23.26	26.94	29.05	31.00	30.80	35.33
SSI-DDI (BIB'21)	25.95	27.63	21.61	21.93	25.43	33.71	27.88	29.01	28.84	29.43	37.47	32.47
DSN-DDI (BIB'23)	22.89	29.09	18.53	23.71	25.34	29.88	26.60	28.25	31.92	27.73	32.97	35.79
CMRL (KDD'23)	22.60	26.93	17.46	21.87	24.51	31.50	28.81	30.78	29.62	29.84	28.80	31.59
CGIB (ICML'23)	22.35	24.97	16.93	22.72	22.69	27.30	26.83	31.12	27.77	28.75	**24.65**	28.20
MMGNN (IJCAI'24)	22.89	27.35	**16.56**	23.79	23.87	27.54	25.72	32.71	30.58	28.62	29.89	30.23
CausalGIB	**19.60**	**22.86**	17.01	**21.79**	**19.24**	**23.26**	20.32	**25.11**	**23.04**	**22.27**	27.67	**23.54**

Table 1: The PR_M and PV_M of CausalGIB and comparison methods on DDI data in different distributions, with the best results highlighted in **bold** and the second results highlighted in underline. → reflects the performance disparity between two data distributions, without implying an inferential relationship. Unit: %.

Model	ZhangDDI	HetionteDDI	DrugBankDDI
MIRACLE	19.21	24.14	25.34
CIGIN	18.70	25.01	26.31
SSI-DDI	20.74	25.21	**18.33**
DSN-DDI	18.30	23.62	19.00
CMRL	19.41	28.90	22.36
CGIB	20.96	25.75	22.29
MMGNN	19.44	27.14	23.64
CausalGIB	**18.17**	**23.45**	19.99

Table 2: The PV_{Pre} of CausalGIB and comparison methods on DDI data in different distributions, with the best results highlighted in **bold** and the second results highlighted in underline. Unit: %.

5.2 Predictive Performance Analysis (RQ2)

We further compared the predictive performance of CausalGIB and baseline methods on regular data to evaluate whether CausalGIB remains competitive in independent and identically distributed data. The experimental results are reported in Table 3. CausalGIB exhibits pronounced advantages even in independent and identically distributed data.

Comprehensive Performance Analysis: CausalGIB demonstrates the best overall performance in MI and DDI predictions. In MI prediction, CausalGIB improves RMSE by 4% to 6% in multiple datasets. CausalGIB's comprehensive performance outperforms other comparative methods in DDI prediction. This results highlights the importance of mitigating confounding information and strengthening causal substructure-target relationships to enhance predictive performance.

Substructure Importance Analysis: Methods integrating substructure information demonstrate superior predictive performance compared to those that disregard it, highlighting the pivotal role of substructure interactions in MRL. Encouraging models to capture substructure-level features significantly enhances the prediction of molecular interactions, consistent with the principle that substructures are fundamental determinants of chemical reactivity.

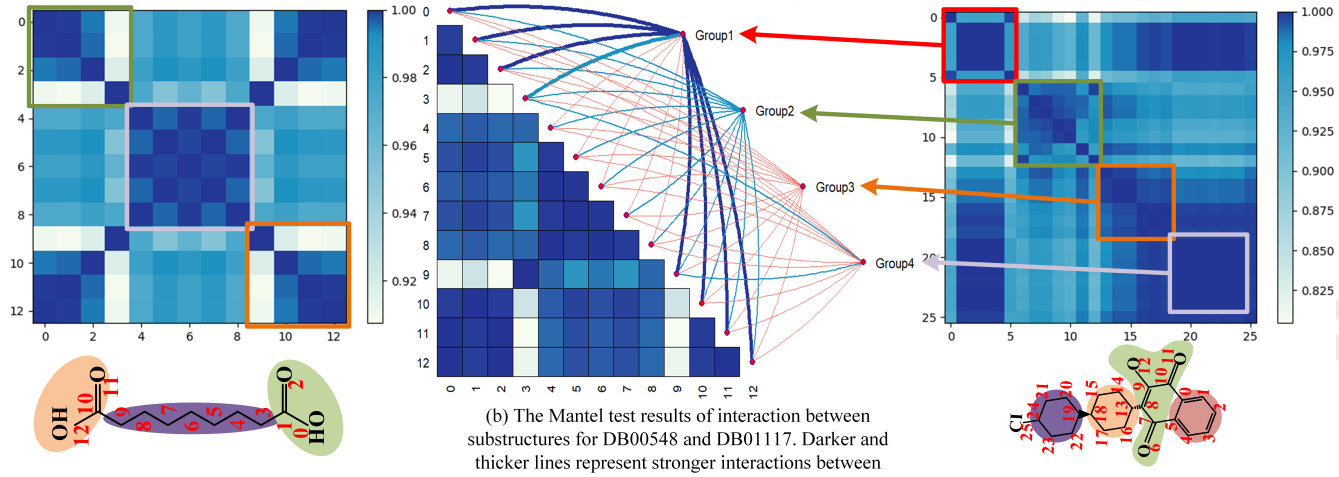
5.3 Visualization Analysis (RQ3)

To validate CausalGIB's ability to capture causal substructures, we visualize the node features and the interactions between substructures, as shown in Fig. 4.

CausalGIB effectively captures substructure information in drug molecules, aligning closely with actual molecular substructures and highlighting intra-molecular atomic interactions. As shown in Fig. 4, it demonstrates strong recognition of interactions between functional groups, such as the symmetrical -COOH groups (green and orange areas) in DB00548 and Group1 (benzene ring marked by red box) in DB01117. This aligns with domain knowledge that chemical reactions primarily occur between core substructures. CausalGIB identifies critical substructures and underscores their pivotal role in driving chemical reactions.

Model	Chromophore			MNSol	FreeSolv	CompSol	Abraham	CombiSolv
	Absorption	Emission	Lifetime					
Substructure ✗								
GCN (ICLR'17)	25.75±1.48	31.87±1.70	0.866±0.015	0.675±0.021	1.192±0.042	0.389±0.009	0.738±0.041	0.672±0.022
GAT (ICLR'18)	26.19±1.44	30.90±1.01	0.859±0.016	0.731±0.007	1.280±0.049	0.387±0.010	0.798±0.038	0.662±0.021
MPNN (ICML'17)	24.43±1.55	30.17±0.99	0.802±0.024	0.682±0.017	1.159±0.032	0.359±0.011	0.601±0.035	0.568±0.005
GIN (ICLR'19)	24.92±1.67	32.31±0.26	0.829±0.027	0.669±0.017	1.015±0.041	0.331±0.016	0.648±0.024	0.595±0.014
CIGIN (AAAI'20)	19.32±0.35	25.09±0.32	0.804±0.010	0.607±0.024	0.905±0.014	0.308±0.018	0.411±0.008	0.451±0.009
Substructure ✓								
CMRL (KDD'23)	17.93±0.31	24.30±0.22	0.776±0.007	0.551±0.017	0.815±0.046	**0.255±0.011**	0.374±0.011	0.421±0.008
CGIB (ICML'23)	18.11±0.20	23.90±0.35	0.771±0.005	**0.538±0.007**	0.852±0.022	0.276±0.017	0.390±0.006	0.422±0.005
MMGNN (IJCAI'24)	18.65±0.34	25.33±0.43	0.801±0/007	0.546±0.011	0.902±0.026	0.267±0.012	0.385±0.008	**0.303±0.033**
CausalGIB	**16.82±0.25**	**22.95±0.33**	**0.769±0.005**	0.541±0.010	**0.799±0.034**	0.261±0.013	**0.371±0.008**	0.419±0.008

Table 3: The performance of CausalGIB and comparative methods in MI prediction, with the best results highlighted in **bold** and the second results highlighted in underline.



(a) The node visualization results for DB00548.

(b) The Mantel test results of interaction between substructures for DB00548 and DB01117. Darker and thicker lines represent stronger interactions between substructures.

(c) The node visualization results for DB01117.

Figure 4: The visualization of node features and interaction strengths between DB01117 and DB00548.

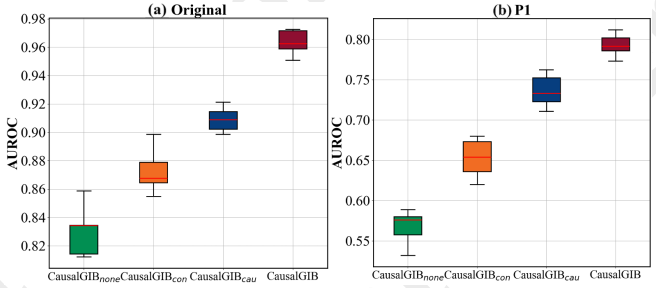


Figure 5: The experimental results of model analysis in HetionteDDI dataset.

5.4 Ablation Experiment (RQ4)

To analyze the impact of confounding information and causal information on the overall performance of the model, we compared the following model variants: CausalGIB$_{\text{none}}$ (no optimizations applied in the model architecture), CausalGIB$_{\text{con}}$ (only minimizing $I(G^c, G^n)$), and CausalGIB$_{\text{cau}}$ (only maximizing $I(Y;G_x^c,G_y^c)$). The experimental results are shown in Fig. 5. As shown in Fig. 5, minimizing $I(G^c, G^n)$ and maximizing $I(Y;G_x^c,G_y^c)$ improve AUROC by 4%, 8% and 8%, 16%, respectively. Notably, the performance gain from maximizing $I(Y;G_x^c,G_y^c)$ significantly surpasses that of minimizing $I(G^c, G^n)$. While minimizing $I(G^c, G^n)$ has limited impact on stability, maximizing $I(Y; G_x^c, G_y^c)$ enhances predictive stability, though its contribution alone remains insufficient. The best performance and stability are achieved when both objectives are optimized concurrently, highlighting SIB's role in refining substructure differentiation and enhancing predictive stability.

6 Conclusion

In this paper, we propose the Causal Subgraph Information Bottleneck to enhance the stability of MRL. CausalGIB leverages causal dependencies to guide substructure representation and integrates SIB to optimize the core substructure representation, generating stable representations. Experimental results confirm that CausalGIB significantly enhances stability in distribution shifts. Future work will explore its validation in diverse datasets and tasks to establish broader applicability.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No.62472332, No.62276196, No.62076045), the China Scholarship Council, the Open Bidding for Selecting the Best Candidates Project of Wuhan East Lake High-Tech Development Zone (No.2024KJB322), the Hubei Provincial International Science and Technology Cooperation Project (No.2024EHA031), the 111 Project (No.D23006) and the National Foreign Expert Project of China (No.D20240244).

References

[Bolougouri *et al.*, 2024] Maria Boulougouri, Pierre Vandergheynst, and Daniel Probst. Molecular Set Representation Learning. *Nature Machine Intelligence*, 6:754–763, 2024.

[Ding *et al.*, 2024a] Shifei Ding, Wei Du, Ling Ding, Lili Guo, and Jian Zhang. Learning Efficient and Robust Multi-Agent Communication via Graph Information Bottleneck. In *Proceedings of the 38th AAAI Conference on Artificial Intelligence*, volume 38, pages 17346–17353, 2024.

[Ding *et al.*, 2024b] Shifei Ding, Wei Du, Ling Ding, Jian Zhang, Lili Guo, and Bo An. Robust Multi-Agent Communication With Graph Information Bottleneck Optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(5):3096–3107, 2024.

[Du *et al.*, 2024] Wenjie Du, Shuai Zhang, Jun Xia Di Wu, Ziyuan Zhao, Junfeng Fang, and Yang Wang. MMGNN: A Molecular Merged Graph Neural Network for Explainable Solvation Free Energy Prediction. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, pages 5808–5816, 2024.

[Fan *et al.*, 2022] Shaohua Fan, Xiao Wang, Yanhu Mo, Chuan Shi, and Jian Tang. Debiasing Graph Neural Networks via Learning Disentangled Causal Substructure. In *Proceedings of the 36th International Conference on Neural Information Processing Systems*, volume 35, pages 24934–24946, 2022.

[Fang *et al.*, 2024] Junfeng Fang, Shuai Zhang, Chang Wu, Zhengyi Yang, Zhiyuan Liu, Sihang Li, Kun Wang, Wenjie Du, and Xiang Wang. MolTC: Towards Molecular Relational Modeling In Language Models. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics*, pages 1943–1958, 2024.

[Gilmer *et al.*, 2017] Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. Neural Message Passing for Quantum Chemistry. In *Proceedings of the 34th International Conference on Machine Learning*, page 1263–1272, 2017.

[Hu *et al.*, 2017] Ye Hu, Dagmar Stumpfe, and Jurgen Bajorath. Recent Advances in Scaffold Hopping: Miniperspective. *Journal of Medicinal Chemistry*, 60(4):1238–1246, 2017.

[Iwasaki and Nozaki, 2024] Takanori Iwasaki and Kyoko Nozaki. Counterintuitive Chemoselectivity in the Reduction of Carbonyl Compounds. *Nature Reviews Chemistry*, 8:518–534, 2024.

[Job *et al.*, 2023] Simi Job, Xiaohui Tao, Yuefeng Li, Lin Li, and Jianming Yong. Topic integrated opinion-based drug recommendation with transformers. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 7(6):1676–1686, 2023.

[Job *et al.*, 2024] Simi Job, Xiaohui Tao, Lin Li, Haoran Xie, Taotao Cai, Jianming Yong, and Qing Li. Optimal treatment strategies for critical patients with deep reinforcement learning. *ACM Transactions on Intelligent Systems and Technology*, 15(2):1–22, 2024.

[Kipf and Welling, 2017] Thomas N Kipf and Max Welling. Semi-Supervised Classification with Graph Convolutional Networks. In *Proceedings of the 5th International Conference on Learning Representations*, 2017.

[Lee *et al.*, 2023a] Namkyeong Lee, Dongmin Hyun, Gyoung S. Na, Sungwon Kim, Junseok Lee, and Chanyoung Park. Conditional Graph Information Bottleneck for Molecular Relational Learning. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202, pages 18852–18871, 2023.

[Lee *et al.*, 2023b] Namkyeong Lee, Kanghoon Yoon, Gyoung S Na, Sein Kim, and Chanyoung Park. Shift-Robust Molecular Relational Learning with Causal Substructure. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 1200–1212, 2023.

[Li *et al.*, 2023] Zimeng Li, Shichao Zhu, Bin Shao, Xiangxiang Zeng, Tong Wang, and Tie-Yan Liu. DSN-DDI: An Accurate and Generalized Framework for Drug–Drug Interaction Prediction by Dual-View Representation Learning. *Briefings in Bioinformatics*, 24(1):bbac597, 2023.

[Li *et al.*, 2024] Changsheng Li, Kaihang Mao, Shiye Wang, Ye Yuan, and Guoren Wang. Self-Supervised Graph Information Bottleneck for Multiview Molecular Embedding Learning. *IEEE Transactions on Artificial Intelligence*, 5(4):1554–1562, 2024.

[Nyamabo *et al.*, 2021] Arnold K Nyamabo, Hui Yu, and Jian-Yu Shi. SSI–DDI: Substructure–Substructure Interactions for Drug–Drug Interaction Prediction. *Briefings in Bioinformatics*, 22(6):bbab133, 2021.

[Pathak *et al.*, 2020] Yashaswi Pathak, Siddhartha Laghuvarapu, Sarvesh Mehta, and U Deva Priyakumar. Chemically Interpretable Graph Interaction Network for Prediction of Pharmacokinetic Properties of Drug-Like Molecules. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence*, volume 34, pages 873–880, 2020.

[Pei *et al.*, 2024] Hongbin Pei, Taile Chen, A Chen, Huiqi Deng, Jing Tao, Pinghui Wang, and Xiaohong Guan. Hago-net: Hierarchical Geometric Massage Passing for Molecular Representation Learning. In *Proceedings of*

the 38th AAAI Conference on Artificial Intelligence, volume 38, pages 14572–14580, 2024.

[Seo *et al.*, 2023] Sangwoo Seo, Sungwon Kim, and Chanyoung Park. Interpretable Prototype-based Graph Information Bottleneck. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*, volume 36, pages 76737–76748, 2023.

[Seo *et al.*, 2024] Sangwoo Seo, Sungwon Kim, Jihyeong Jung, Yoonho Lee, and Chanyoung Park. Self-Explainable Temporal Graph Networks based on Graph Information Bottleneck. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2024.

[Sun *et al.*, 2022] Qingyun Sun, Jianxin Li, Hao Peng, Jia Wu, Xingcheng Fu, Cheng Ji, and S Yu Philip. Graph Structure Learning with Variational Information Bottleneck. In *Proceedings of the 36th AAAI Conference on Artificial Intelligence*, volume 36, pages 4165–4174, 2022.

[Veličković *et al.*, 2018] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph Attention Networks. In *Proceedings of the 6th International Conference on Learning Representations*, 2018.

[Wang *et al.*, 2021] Yingheng Wang, Yaosen Min, Xin Chen, and Ji Wu. Multi-View Graph Contrastive Representation Learning for Drug-Drug Interaction Prediction. In *Proceedings of the 30th ACM on Web Conference*, pages 2921–2933, 2021.

[Wang *et al.*, 2024] Ruijia Wang, Haoran Dai, Cheng Yang, Le Song, and Chuan Shi. Advancing Molecule Invariant Representation via Privileged Substructure Identification. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 3188–3199, 2024.

[Wu *et al.*, 2020] Tailin Wu, Hongyu Ren, Pan Li, and Jure Leskovec. Graph Information Bottleneck. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, volume 33, pages 20437–20448, 2020.

[Xu *et al.*, 2019] Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. How Powerful are Graph Neural Networks? In *Proceedings of the 7th International Conference on Learning Representations*, 2019.

[Yang *et al.*, 2022] Ziduo Yang, Weihe Zhong, Qiujie Lv, and Calvin Yu-Chian Chen. Learning Size-Adaptive Molecular Substructures for Explainable Drug–Drug Interaction Prediction by Substructure-Aware Graph Neural Network. *Chemical Science*, 13(29):8693–8703, 2022.

[Yang *et al.*, 2023] Nianzu Yang, Kaipeng Zeng, Qitian Wu, and Junchi Yan. Molerec: Combinatorial Drug Recommendation with Substructure-Aware Molecular Representation Learning. In *Proceedings of the 32nd ACM on Web Conference*, pages 4075–4085, 2023.

[Yuan *et al.*, 2024] Haonan Yuan, Qingyun Sun, Xingcheng Fu, Cheng Ji, and Jianxin Li. Dynamic Graph Information Bottleneck. In *Proceedings of the 33rd ACM on Web Conference*, page 469–480, 2024.

[Zhang *et al.*, 2024a] Peiliang Zhang, Chao Che, Bo Jin, Jingling Yuan, Ruixin Li, and Yongjun Zhu. Nch-dda: Neighborhood contrastive learning heterogeneous network for drug–disease association prediction. *Expert Systems with Applications*, 238:121855, 2024.

[Zhang *et al.*, 2024b] Peiliang Zhang, Jingling Yuan, Lin Li, Wen Luo, Jiwei Hu, and Xin Li. Key substructure learning with chemical intuition for material property prediction. In *International Conference on Database Systems for Advanced Applications*, pages 87–103. Springer, 2024.

[Zhang *et al.*, 2025a] Peiliang Zhang, Jingling Yuan, Qing Xie, Yongjun Zhu, and Lin Li. Representational alignment with chemical induced fit for molecular relational learning. *arXiv preprint arXiv:2502.07027*, 2025.

[Zhang *et al.*, 2025b] Xin Zhang, Peiliang Zhang, Jingling Yuan, and Lin Li. Zero-shot learning for materials science texts: Leveraging duck typing principles. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 1129–1137, 2025.