MMGNN: A Molecular Merged Graph Neural Network for Explainable Solvation Free Energy Prediction

Wenjie Du[1,2], **Shuai Zhang**[1,2], **Di Wu**[1], **Jun Xia**[3], **Ziyuan Zhao**[4], **Junfeng Fang**[1,*], **Yang Wang**[1,2,*]

[1]University of Science and Technology of China (USTC), Hefei, China
[2]Suzhou Institute for Advanced Research, USTC, Suzhou, China
[3]Zhejiang University, Hangzhou, China
[4]Agency for Science, Technology and Research (A*STAR), Singapore
{duwenjie, shuaizhang, fjf, wdcxy}@mail.ustc.edu.cn, zhaoz@i2r.a-star.edu.sg, junxia@zju.edu.cn, angyan@ustc.edu.cn

Abstract

In this paper, we address the challenge of accurately modeling and predicting Gibbs free energy in solute-solvent interactions, a pivotal yet complex aspect in the field of chemical modeling. Traditional approaches, primarily relying on deep learning models, face limitations in capturing the intricate dynamics of these interactions. To overcome these constraints, we introduce a novel framework, **M**olecular **M**odeling **G**raph **N**eural **N**etwork (MMGNN), which more closely mirrors real-world chemical processes. Specifically, MMGNN exquisitely models atomic interactions such as hydrogen bonds by initially forming indiscriminate connections between intermolecular atoms, which are then refined using an attention-based aggregation method, tailoring to specific solute-solvent pairs. To address the challenges of non-interactive or repulsive atomic interactions, MMGNN incorporates interpreters for nodes and edges in the merged graph, enhancing explainability and reducing redundancy. MMGNN stands as the first framework to exquisitely align with real chemical processes, providing a more accurate and scientifically sound approach to modeling solute-solvent interactions. The infusion of explainability allows for the extraction of key subgraphs, which are pivotal for further research in solute-solvent dynamics. Extensive experimental validation confirms the efficacy and enhanced explainability of MMGNN.

1 Introduction

Understanding solute-solvent interactions in specific solvents is pivotal for areas in physical chemistry, including chemical reactions and electrochemistry [Varghese and Mushrif, 2019; D'Souza *et al.*, 2011]. A comprehensive grasp of these interactions is vital not only for explanatory experimental outcomes but also for guiding the design and control of reactions and properties [Chung *et al.*, 2022; Fang *et al.*, 2024b; Xia *et al.*, 2023b]. In this context, the solvation Gibbs free energy ΔG_{solv} emerges as a critical physicochemical property, dictating a molecule's behavior in solution [Low *et al.*, 2022]. This property is intricately linked to the solute's partition coefficient between gas and solvent phases [Chung *et al.*, 2022]. However, empirically testing solute-solvent free energies for all combinations is impractical due to high costs and extensive time requirements. This challenge necessitates an increased reliance on deep learning models to predict these energies more efficiently [Varghese and Mushrif, 2019].

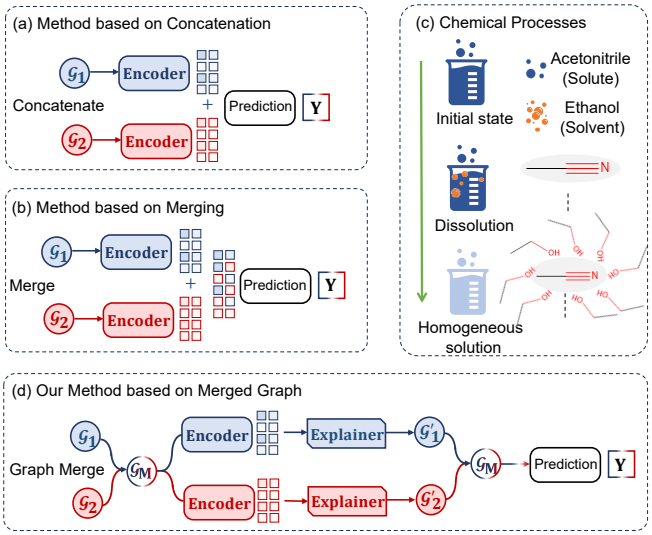

Figure 1: Comparison of different paradigms for Gibbs free energy prediction. (a) Method by concatenation; (b) method by merging; (c) a schematic diagram of the process where acetonitrile (solute) is dissolved in ethanol (solvent); and (d) the illustration of our method. Best viewed in color.

The landscape of solute-solvent interaction modeling is currently dominated by two primary approaches. The first, termed **embedding concatenation**, employs two separate Graph Neural Networks (GNNs) to represent solute and solvent molecules. The individual embeddings from these GNNs are then concatenated for subsequent prediction tasks,

as depicted in Figure 1 (a). While effective, this method may not adequately capture the intricate coupling between solute and solvent. In contrast, the **embedding merging** approach, shown in Figure 1 (b), seeks to address this limitation. It also utilizes two Graph Encoders for initial representation but differs in the subsequent processing of these embeddings. Here, the embeddings are merged through advanced interaction strategies, such as Transformer-based techniques or interactive pruning algorithms, to better reflect the complex interactions before making predictions.

Although the embedding merging approach marks an advancement, it falls short in accurately simulating the real chemical processes involved in dissolution, as depicted in Figure 1 (c). The dissolution process is characterized by the substitution of intramolecular forces within solute and solvent with intermolecular forces between them. This is exemplified by the interaction of acetonitrile and ethanol, where individual acetonitrile molecules are gradually surrounded by ethanol molecules. The formation of hydrogen bonds between atoms (nodes) of these molecules leads to a homogeneous solution. This natural dissolution process suggests a need for an algorithm that explicitly models these intermolecular node interactions, rather than relying solely on embedding-level interactions. Such a model, by more closely mirroring actual chemical processes, is crucial for both scientific accuracy and the effectiveness of the predictions.

In response to the limitations of existing models, we propose the **M**olecular **M**odeling **G**raph **N**eural **N**etwork (MMGNN), a novel framework designed to closely align with the actual chemical processes in solute-solvent interactions. This framework significantly improves the prediction accuracy of Gibbs free energy. Illustrated in Figure 1 (d), MMGNN begins by indiscriminately connecting intermolecular atoms to enhance interactions, such as hydrogen bonds, between molecules. After that, MMGNN assigns variable weights to these connections, reflecting the different constraints of various chemical bonds. An attention-based aggregation method is then employed, enabling the framework to adaptively learn from diverse solute-solvent combinations. The result is a merged graph representation for each pair, highlighting the most significant atomic interactions.

Meanwhile, this framework also addresses potential challenges: (1) the presence of non-existent or repulsive atomic interactions, and (2) increased complexity and convergence difficulties in merged complete graphs for large molecular pairs. Inspired by graph explainability algorithms, MMGNN incorporates interpreters for both nodes and solute-solvent edges in the merged graph. This approach effectively reduces redundancy, focusing only on relevant interactions. The explanatory subgraph is then encoded and utilized in regression models, such as Fully Connected Neural Networks (FCNN), to predict Gibbs free energy with greater precision.

In conclusion, the main contributions of this paper can be summarized as follows:

- **Introduction of MMGNN:** We present the Molecular Modeling Graph Neural Network (MMGNN), a pioneering framework designed to explicitly align with actual chemical processes, enabling more accurate modeling of solute-solvent interactions.
- **Advancement in Explainability:** Our approach integrates explainability into the model, allowing for the extraction of key subgraphs. This feature not only enhances the understanding of MMGNN's predictions but also aids in further research into solute-solvent dynamics.
- **Validation through Extensive Experiments:** We conducted comprehensive experiments to validate MMGNN's performance. The results demonstrate both the accuracy and explainability of our model's predictions, underscoring its effectiveness in the field of physical chemistry.

2 Related Work

2.1 Molecular Relational Learning

Generally, message passing in GNNs or graph convolutional networks (GCNs) refers to the utilization of trainable interaction layers to facilitate the exchange of information among atoms within a local neighborhood [Xia *et al.*, 2023a; Wang *et al.*, 2024; Wang *et al.*, 2023]. Behler and Parrinello are pioneers in modelling interatomic properties [Behler, 2015; Behler, 2016] by summing per-atom contributions in neural network predictions. Tensormol [Yao *et al.*, 2018] employs Behler and Parrinello's approximation to accurately compute the dipole moment in water dimers as one water molecule rotates about the O-H bond, resulting in reasonable outcomes. This method is subsequently refined by integrating the atomic-pairwise into the neural network [Glick *et al.*, 2020] which led to a reduction in the binding energy error by a factor of five. [Lee *et al.*, 2023] provide a molecular relational learning framework that predicts the interaction behavior based on graph information bottleneck theory. In brief, the original method is summing overall the atoms, or fusing information after message passing process but a more rigorous way should consider the atoms or fragments and introducing atom-pair symmetry message passing functions.

2.2 Gibbs Free Energies Prediction Methods

Solvation free energies have been of great interest for many years and have spurred the development of numerous predictive methods [Hildebrand, 1981]. These methods encompass a range of techniques, including molecular dynamics and quantum chemistry methods, as well as empirical or data-driven approaches. While quantum chemistry methods such as the SMx and COSMO(-RS) models [Klamt and Eckert, 2000; Marenich *et al.*, 2007] are based on first-principles calculations of all promising relevant conformers of the solute and solvent molecules, they are often computationally expensive and labor-intensive. In contrast, empirical or data-driven methods could provide a faster alternative for predicting solvation properties. Recent studies have aimed at improving these models for predicting ΔG_{solv}. [Vermeire and Green, 2021] utilized a transfer learning method to achieve a mean absolute error of 0.21 kcal/mol when predicting experimental results from a standard quantum computing dataset. similarly, [Zhang *et al.*, 2022] calculated water solvation Gibbs free energies of more than 100,000 organic compounds, and utilized a graph neural network (GNN) to predict these values achieving an error of 0.4 kcal/mol.

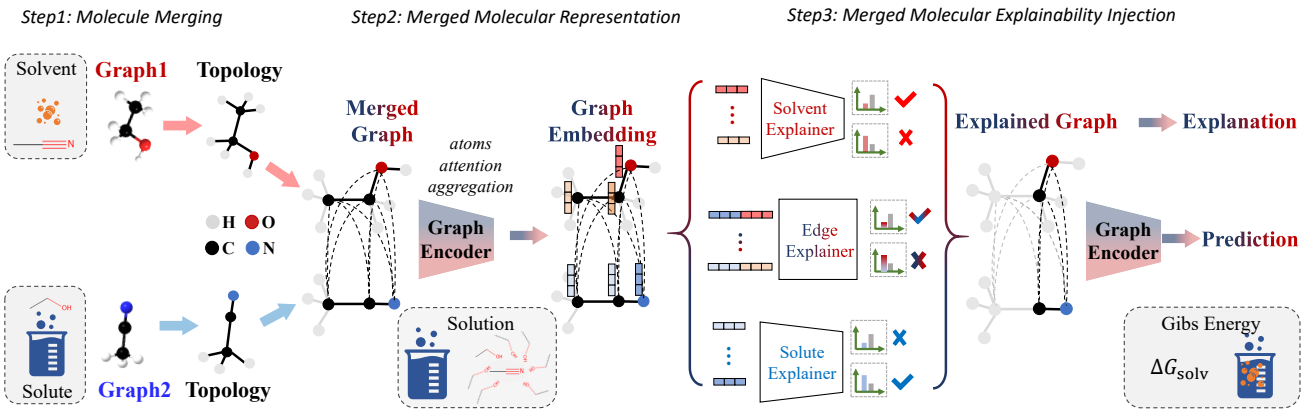

Figure 2: The overall paradigm of our method, which adheres to the realistic chemical processes. Note that the legends in the gray dotted boxes represent the realistic chemical processes.

3 Method

3.1 Overview of the Framework

In this study, we introduce a novel merged molecular graph model, wherein the weighted connection edges adhere to the principles of similarity and compatibility in chemistry. To process an input solute-solvent pair, we initially convert it into a machine-readable topological graph and incorporate relevant weighted edges to enhance molecular information interaction, constructing a merged graph as detailed in Section 3.2 and illustrated in Figure 2. Within this merged graph, we facilitate both inter- and intra-molecular message passing processes. In the initial layer of message passing, each atomic feature of the solute molecule is updated independently and subsequently acquires all corresponding atomic information of the solvent molecule, and vice versa. In the second stage of message passing, each atom of the solute molecule receives information beyond the atomic details in the solvent, including bond and node information with one hop. The final realization of the merged molecular representation is provided in Section 3.3. Additionally, we consider three distinct molecular explainability injection methods, as outlined in Section 3.4, to highlight the most significant atomic interactions. This augmentation aims to enhance the model's generalization capability and align it more closely with biochemical realities.

3.2 Molecule Merging

The molecular structure can be naturally depicted as a topology graph, wherein atoms serve as the nodes and bonds act as the edges [Wen *et al.*, 2021; Zhou *et al.*, 2023b]. The molecular graph can be formally denoted by:

$$\mathcal{G} = \{\mathcal{E}, \mathcal{V}, \mathcal{U}\}, \tag{1}$$

where $\mathcal{E}$ is the set of edges (bonds), and $\mathcal{V}$ is the set of nodes (atoms). $\mathcal{U}$ is the global feature vector which is extracted from each molecule. The detailed features used in molecular representation are recorded in the appendix.

$$\mathcal{E} = (e_k, p_k, q_k)_{k=1}^{N^e}. \tag{2}$$

Specifically, all of bonds are enumerated in a molecule. The composition of $\mathcal{E}$ comprises e_k, p_k and q_k, where e_k represents the edge feature, and p_k and q_k, is the two atoms connected by e_k.

$$\mathcal{V} = \{v_i\}_{i=1}^{N^v}. \tag{3}$$

Similarly, $\mathcal{V}$ enumerates all of atoms in a molecule. v_i is atomic features vector for atom i (e.g., atomic species, valence, *etc.*) and N^v is the total number of atoms in the molecule. So far, the single molecular graph has been fully defined. Next, a merged graph $\widetilde{\mathcal{G}}$ could be generated by establishing a weighted edge between two molecules (solute-solvent) which connects every atom of each molecule to all atoms in the other molecule.

$$\widetilde{\mathcal{G}} = \{\mathcal{R}, \mathcal{E}, \mathcal{V}, \mathcal{U}\}. \tag{4}$$

The set of relation edges are $\mathcal{R}$:

$$\mathcal{R} = \{(r_k, a_i, b_j)\}_{k=1}^{N^a \times N^b}, \tag{5}$$

where $i \in \{1, 2, 3, \ldots, N^a\}$, $j \in \{1, 2, 3, \ldots, N^b\}$. N^a and N^b is the total number of atoms in each of the two molecules, use r_k to represent this new type of edge.

3.3 Merged Molecular Representation

Generally, in this merged molecular graph $\widetilde{\mathcal{G}}$, the message passing process is firstly run in single molecule (i.e. solute and solvent molecules). During the message passing phase, e_k is updated to the new vector e'_{ij} by aggregating the initially bond features, as well as the features of the two atoms, v_i and v_j, and the global features u. In addition, the atomic feature vector, v_i, for each atom i is updated into v'_i:

$$e'_{ij} = e_{ij} + \tau[\mathrm{FC}(v_i + v_j) + \mathrm{FC}(e_{ij}) + \mathrm{FC}(u)], \tag{6}$$

$$\hat{e}_{ij} = \frac{\sigma(e'_{ij})}{\sum_{j' \in N_i} \sigma(e'_{ij'}) + \epsilon}, \tag{7}$$

$$v'_i = v_i + \tau[\mathrm{FC}(v_i + \sum_{j \in N_i} \hat{e}_{ij} \odot \mathrm{FC}(v_j)) + \mathrm{FC}(u)], \tag{8}$$

where v'_i and e'_{ij} are the updated node vectors and edge vectors. τ is the LeakyReLU activation function [Maas *et al.*, 2013]. $\texttt{FC}$ is fully connected layer. $\odot$ denotes the Hadamard product. Then, the intramolecular atomic features could be updated. $\sigma(\cdot)$ is the sigmoid activation function, and ϵ is a fixed constant (0.0001) added for numerical stability. The next atomic update step would achieve intermolecular information passing process. In a molecule a,the atomic feature vector $v_{ai}$for each atom i is further updated as follows:

$$v''_{ai} = (1-\beta)v'_{ai} + \beta \sum_{j \in N_b} \texttt{FC}(\delta(v'_{ai}, v'_{bj}), v'_{bj}), \quad (9)$$

where β are the hyperparameter that control the update rate of information, which makes information differentiated to avoid information decay or overwrite the original information. δ is an additional function that measures the feature similarity between two atoms, which is defined as follows:

$$\delta(v'_{ai}, v'_{bj}) = \frac{1}{|v'_{ai} - v'_{bj}| + \epsilon}. \quad (10)$$

Finally, the global feature vector u and relation bond vector r'_i is updated:

$$u' = u + \tau[\texttt{FC}(\frac{1}{N^v}\sum_{i=1}^{N^v} v'_i + \frac{1}{N^e}\sum_{k=1}^{N^e} e'_k + u)] \quad (11)$$

$$r'_i = \texttt{FC}(\delta(v'_{ai}, v'_{bj})), \quad (12)$$

where N^v and N^e are the number of atoms and bonds in the molecule, respectively. v_{ai} and v_{bi} are two atoms connected by the relation, u is the global feature.

3.4 Merged Molecular Explainability Injection

In this part, we elaborate the process of injecting explainability into the model. The input for the explanation module is the representation of the merged graph. Specifically, for a node, it receives the embedding of the node to output an importance score, and for an edge, it processes the concatenated embeddings of the two nodes connected by the edge to output the importance of the edge. Specifically, we get inspiration from the significant success of the post-hoc explanation methods of the graph neural networks [Ying *et al.*, 2019; Fang *et al.*, 2023b; Fang *et al.*, 2024a] and design our explainers from three perspectives as follows. Note that Figure 2 illustrates the workflow of the explainer designed from the second approach, namely the global mask-based perspective.

- The first one is the **local mask-based perspective**. Specifically, local mask-based methods endeavour to multiply the features v'_{ai} and v'_{bi} with the corresponding masks (all are initialized to 1) p_{ai} and p_{bi} to get $v''_{ai} = v'_{ai}p_{ai}$ and $v''_{bi} = v'_{bi}p_{bi}$. Then, v''_{ai} and v''_{bi} is sent into the model to obtain the updated output ΔG_{solv}. Then they attempt to find the optimal score p_i by minimizing the difference between this processed output ΔG_{solv} and the label Y. To limit the size of explanatory subgraphs, they apply l_1 regularization to the value of mask. In this case, the loss function is:

$$\mathcal{L} = D\left(\Delta G_{\text{solv}}; Y\right) + \sum_i \lambda(p_{ai} + p_{bi}) \cdot \mathbf{1}^{\mathbf{T}}, \quad (13)$$

  where D denotes the distance function; λ is the trade-off parameter.

- The second one is the **global mask-based perspective**, and its only difference from the first approach lies in the method of mask generation. Here, during the process of the adding the mask, the trainable mask p_{ai} and p_{bi} in local mask-based methods is replaced with a trainable MLP_ψ (*i.e.*, $p_{ai} = \text{MLP}_\psi(v'_{ai}), p_{bi} = \text{MLP}_\psi(v'_{bi})$). Meanwhile, following the theory of graph information bottleneck (GIB) [Wu *et al.*, 2020; Yu *et al.*, 2021], these methods instantiates the *information constraint* (ℓ_p) proposed by [Miao *et al.*, 2022], where ℓ_p is defined as:

$$\begin{aligned} \ell_{pa} &= \sum_i p_{ai} \log \frac{p_{ai}}{k} + (1-p_{ai}) \log \frac{1-p_{ai}}{1-k} \\ \ell_{pb} &= \sum_i p_{bi} \log \frac{p_{bi}}{k} + (1-p_{bi}) \log \frac{1-p_{bi}}{1-k}, \end{aligned} \quad (14)$$

  where k is the pre-defined hyperparameter. After applying ℓ_p regularization to the value of mask, the expression of the final loss function is:

$$\mathcal{L} = D\left(\Delta G_{\text{solv}}; Y\right) + \lambda(\ell_{pa} + \ell_{pb}), \quad (15)$$

  where D denotes the distance function; λ is the trade-off parameter.

- The third one is the **gradient-based perspective**. For the v'_{ai} and v'_{bi}, these methods first calculate the absolute values of the features in the derivative of ΔG_{solv} *w.r.t* v'_{ai} and v'_{bi}. After that, their importance score p_{ai} and p_{bi} are defined as the normalized sum of these values. Then, at the end of each training phase, we mask some nodes and edges according to the above derivatives, and input the remain part to the next training phase. More formally:

$$p_{ai} = \left(|\frac{\partial \Delta G_{\text{solv}}}{\partial v'_{ai}}| \cdot \mathbf{1}^{\mathbf{T}}\right), p_{bi} = \left(|\frac{\partial \Delta G_{\text{solv}}}{\partial v'_{bi}}| \cdot \mathbf{1}^{\mathbf{T}}\right). \quad (16)$$

After that, we conduct another merged molecular message passing process as introduced in Section 3.3 to conduct the final interactive prediction. Specifically, the readout phase occurs after the desired number of message-passing layers has taken place (as determined via hyperparameter optimization). To aggregate the set of final feature vectors of v'_{solute} and v'_{solvent} into a single vector suitable for input into a set-to-set (set2set) module, where the solute and solvent molecular vectors are input separately rather than incorporating them as a whole vector that could result in a substantial loss of information during the pooling process. The global feature vector, being a single feature vector, is appended after the set2set steps. This produces a final feature vector which is invariant to permutation of atom, relation and bond indices:

$$\begin{cases} v'_{\text{solute}} = \texttt{F}(\{v''_{ai}\}) \Leftrightarrow v'_{\text{solvent}} = \texttt{F}(\{v''_{bi}\}) \\ g_{\text{interact map}} = v'_{\text{solute}} \parallel v'_{\text{solvent}} \\ \Delta G_{\text{solv}} = \texttt{FC}(g_{\text{interact map}}), \end{cases} \quad (17)$$

where $\texttt{F}$ is a set2set layer and $\parallel$ denotes vector concatenation. $g_{\text{interact map}}$ is then input into a FCNN to obtain the result.

	MAE (↓)				RMSE (↓)			
	FreeSolv	CompSol	Abraham	CompSolv-Exp	FreeSolv	CompSol	Abraham	CompSolv-Exp
D-MPNN	$0.684_{(0.052)}$	$0.179_{(0.013)}$	$0.454_{(0.036)}$	$0.442_{(0.022)}$	$1.164_{(0.055)}$	$0.343_{(0.017)}$	$0.624_{(0.024)}$	$0.672_{(0.051)}$
Explainable GNN	$0.724_{(0.031)}$	$0.184_{(0.012)}$	$0.486_{(0.042)}$	$0.221_{(0.013)}$	$1.276_{(0.045)}$	$0.367_{(0.012)}$	$0.776_{(0.035)}$	$0.404_{(0.054)}$
SolvBERT	$0.588_{(0.034)}$	$0.167_{(0.014)}$	$0.467_{(0.034)}$	$0.382_{(0.023)}$	$1.021_{(0.043)}$	$0.328_{(0.020)}$	$0.652_{(0.022)}$	$0.472_{(0.041)}$
GAT	$0.675_{(0.033)}$	$0.187_{(0.011)}$	$0.457_{(0.043)}$	$0.970_{(0.031)}$	$1.185_{(0.075)}$	$0.390_{(0.012)}$	$0.726_{(0.040)}$	$0.810_{(0.101)}$
GROVER	$0.623_{(0.054)}$	$0.155_{(0.022)}$†	$0.307_{(0.035)}$	$0.382_{(0.023)}$	$1.015_{(0.022)}$	$0.332_{(0.016)}$	$0.475_{(0.044)}$	$0.491_{(0.053)}$
SMD	$0.574_{(0.036)}$	$0.162_{(0.014)}$	$0.374_{(0.024)}$	$0.633_{(0.044)}$	$1.113_{(0.015)}$	$0.317_{(0.011)}$	$0.516_{(0.065)}$	$1.023_{(0.152)}$
Uni-Mol	$0.565_{(0.038)}$	$0.164_{(0.027)}$	$0.322_{(0.071)}$	$0.214_{(0.022)}$	$1.002_{(0.064)}$	$0.303_{(0.020)}$	$0.602_{(0.035)}$	$0.373_{(0.043)}$
Gem	$0.584_{(0.041)}$	$0.174_{(0.011)}$	$0.201_{(0.065)}$	$0.253_{(0.023)}$	$1.131_{(0.059)}$	$0.290_{(0.019)}$	$0.641_{(0.031)}$	$0.551_{(0.023)}$
CIGIN	$0.564_{(0.057)}$	$0.164_{(0.016)}$	$0.254_{(0.010)}$	$0.241_{(0.023)}$	$0.910_{(0.015)}$	$0.318_{(0.020)}$	$0.404_{(0.007)}$	$0.411_{(0.032)}$
CGIB	$0.531_{(0.034)}$†	$0.156_{(0.014)}$	$0.195_{(0.005)}$†	$0.203_{(0.033)}$†	$0.892_{(0.022)}$†	$0.278_{(0.018)}$†	$0.391_{(0.006)}$†	$0.351_{(0.031)}$†
MMGNN	$0.536_{(0.030)}$	$0.146_{(0.010)}$	$0.187_{(0.008)}$	$0.171_{(0.013)}$	$0.902_{(0.026)}$	$0.267_{(0.012)}$	$0.385_{(0.008)}$	$0.303_{(0.033)}$

Table 1: Result of different methods under eight runs on FreeSolv, CompSol and Abraham datasets. (Underlined are the best results while the top-performing baseline is highlighted with a superscript cross. Mean values are reported in the table, with standard deviations in parentheses.)

4 Experimental Results

In this section, we conduct extensive experiments to answer the following questions:

- **RQ1:** Can MMGNN improve the prediction accuracy of ΔG_{solv}?
- **RQ2:** How does the MMGNN perform when explaining the importance of intramolecular and intermolecular features simultaneously?
- **RQ3:** How effective is MMGNN in terms of generalization capabilities?

4.1 Experimental Setup

Datasets. The data set utilized in this study is the *CombiSolv-Exp* database, which is compiled by Vermeire and Green [Vermeire and Green, 2021], and combined the experimental data from multiple sources. And the *FreeSolv* database is published by Mobley and Guthrie [Mobley and Guthrie, 2014], the *CompSol* database from Moine et al. [Moine *et al.*, 2017], and remaining dataset is *Abraham* collected by the Abraham group [Grubbs *et al.*, 2010].

Baselines. We compared MMGNN with some advanced methods, such as D-MPNN [Vermeire and Green, 2021], SolvBert [Yu *et al.*, 2023], SMD [Meng *et al.*, 2023], Explainable GNN [Low *et al.*, 2022], GAT [Velickovic *et al.*, 2017], GROVER [Rong *et al.*, 2020], Uni-Mol [Zhou *et al.*, 2023a], Gem [Fang *et al.*, 2022], these baselines solely concatenate the molecular representations of solute and solvent, but ignore the information interaction process between them. CIGIN [Pathak *et al.*, 2020] and CGIB [Lee *et al.*, 2023] implicitly considered the inter-molecular interactions.

Experimental settings. MMGNN is trained using the Adam optimizer [Kingma and Ba, 2014] ($1 \times 10^{-4} \rightarrow 0.5$) via batch size 50. We employ mean squared error (MSE) as our training loss. The training process was halted if the validation error failed to reduce in 150 epochs, or if the maximum training limit of 1000 epochs was reached. MMGNN is implemented in PyTorch framework via Tesla A100 40GB. Experiment results are presented in following metrics: the mean absolute error (MAE) and root mean squared error (RMSE). The mean and standard deviation are recorded in the table.

	Δ & Backbones	MAE (↓)	RMSE (↓)
Intermolecular	$\beta = 1.0$	$0.190_{(0.022)}$	$0.350_{(0.041)}$
	$\beta = 0.5$	$0.185_{(0.012)}$	$0.343_{(0.034)}$
	$\beta = 0.2$	$\mathbf{0.177}_{(0.015)}$	$\mathbf{0.340}_{(0.022)}$
	$\beta = 0.1$	$0.180_{(0.014)}$	$0.360_{(0.036)}$
	$\beta = 0.0$	$0.200_{(0.027)}$	$0.380_{(0.054)}$
Intramolecular	GAT	$0.197_{(0.011)}$	$0.382_{(0.044)}$
	GIN	$0.212_{(0.018)}$	$0.440_{(0.034)}$
	GCN	$0.265_{(0.025)}$	$0.453_{(0.047)}$
	GAT	$0.203_{(0.024)}$	$0.414_{(0.035)}$
	GCN	$0.177_{(0.015)}$	$0.340_{(0.022)}$
	GIN	$0.181_{(0.017)}$	$0.351_{(0.028)}$

Table 2: Different information interaction update rates and test results of the molecular graph network framework. Interaction coefficient denotes as Δ.

4.2 The Prediction Performance (RQ1)

Similar to previous studies, the random split is a widely used method for evaluating model performance. In each iteration, all methods were trained, validated, and tested on identical datasets in an 8:1:1 ratio. For the three explainability methods: local mask-based, global mask-based, and gradient-based, we conducted separate tests on various datasets and documented the optimal results by MMGNN (this can be found in the appendix). The experimental results are recorded in Table 1. The sensitivity experiment analysis is presented in Table 2. Based on these outcomes, we can delineate three key observations:

Obs.1: MMGNN exhibits the optimal predictive performance. The performance of MMGNN surpasses that of other baselines across three test datasets. In terms of the MAE indicator, MMGNN exhibited an average reduction of 0.025, 0.16, and 0.57 on the CompSol, Abraham, and CompSol-Exp datasets, respectively. Although on the FreeSolv dataset, MMGNN ranks second, slightly behind the CGIB model, this still underscores its notable capability in accurately capturing the characteristics of solute-solvent molecular combinations. Overall, MMGNN consistently demonstrates leading scores across the majority of evaluated datasets, highlighting its effectiveness in precisely characterizing intricate interactions

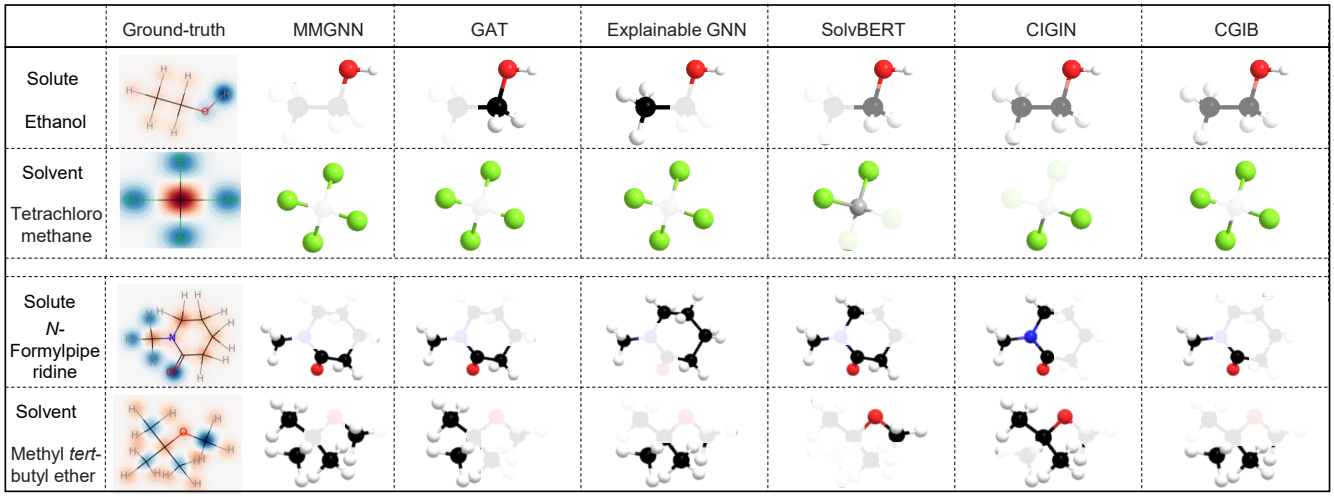

Figure 3: The noteworthy substructures are identified by different methods. Atoms with values above the average score are retained based on the atomic weighted scores. In ground truth, the transition from pink to blue signifies an increase in the significance of the respective atom.

within solute-solvent molecular combination.

Obs.2: $\beta = 0.2$ yielded the best results to control the extent of information exchange. Our experiments revealed that setting $\beta = 0.2$ achieved the best results in controlling the extent of information exchange in Table 2. Various β values were tested to assess the impact on the degree of information exchange. The β value of 0 implies exclusive consideration of intramolecular aspects, resulting in the highest prediction error, as indicated by both MAE and RMSE metrics. Incrementally increasing the interaction level (β ranging from 0 to 0.2) significantly enhanced predictive performance, leading to an 11.5% reduction in MAE. However, further escalation of β resulted in increased MAE, potentially due to information redundancy and repetition causing interference with the model's ability to capture crucial details. Our findings underscore the importance of maintaining a balance between intrinsic and interactive information.

Obs.3: MMGNN is a versatile framework that can be employed with various GNN backbones. We conducted additional tests using GAT, GCN, and GIN to extract intermolecular information. In this scenario, the intramolecular information interaction method utilized the optimal β value of 0.2 from the aforementioned experiments. It is evident that the GCN framework outperformed GAT and GIN, resulting in a reduction of approximately 2.2% and 14.7 % in MAE indicators, respectively. Therefore, we executed the experiments with GCN framework for intramolecular message passing process and hyperparameter $\beta = 0.2$ for intermolecular message passing process.

4.3 Qualitative Evaluation (RQ2)

Due to the current lack of consensus on the key functional groups involved in the energy release process of solution solvents, we have collected some potentially important groups from relevant literature to serve as ground truth for the validation of algorithm explainability [Fang *et al.*, 2023c; Fang *et al.*, 2023a]. The intensity of the color corresponds to the strength of the enhancing (green) or decreasing (pink) effect on ΔG_{solv} prediction.

Obs.4: The significant substructures identified by MMGNN closely align with the ground-truth of their graphs, likely stemming from the class-specific patterns captured by the explanatory subnetwork. However, the baseline explainers could select certain edges that do not belong to the ground-truth. This validates MMGNN's ability to capture substructure interactions aligned with chemistry principles, showcasing its capacity to extract implicit scientific knowledge in solute-solvent molecular interactions.

Obs.5: In contrast to the substructures chosen by baseline explainers, the subgraphs generated by MMGNN exhibit enhanced connectivity. We attribute these distinctions to the alternating and iterative selection processes, a capability not present in baseline explainers.

Obs.6: For specific nodes in molecules (e.g., N, O), certain connected nodes can introduce interference during explanation generation. MMGNN is designed to circumvent these challenges, ensuring explanations are free from interference. In contrast, subgraphs generated by baseline methods may include these disruptive nodes, highlighting the robustness and reliability of MMGNN.

4.4 Generalization Test (RQ3)

To evaluate the model's generalizability on unseen solvents or solute category, a solvent holdout and solute scaffold split cross-validation test are conducted, and the result are present in Figure 4 and Table 3. Here, the four datasets are merged into a single dataset.

Obs.7: MMGNN has superior results on a variety of excluded solvents such as benzene (0.17 kcal mol^{-1}), ethanol (0.23 kcal mol^{-1}), hexane (0.115 kcal mol^{-1}), acetone (0.23 kcal mol^{-1}) (in Figure 4 (a)), *etc*. Nevertheless, it is noteworthy that the water solvent exhibits a higher error compared to the other solvents. This difference is intricately connected to

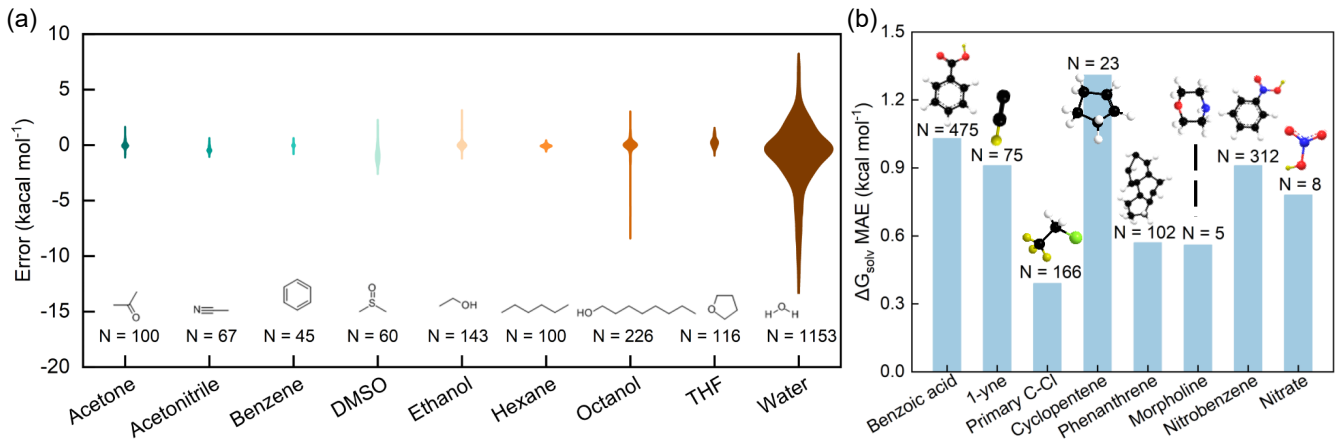

Figure 4: (a) Violin plots of the total error ($\Delta G_{\text{pred}} - \Delta G_{\text{true}}$) for various solvents including acetone, acetonitrile, benzene, DMSO, ethanol, hexane, octanol, THF and water. (b) the scaffold-holdout test.

	MAE (↓)				RMSE (↓)			
	GAT	Explainable GNN	CGIB	**MMGNN**	GAT	Explainable GNN	CGIB	**MMGNN**
Acetone	$0.301_{(0.032)}$	$0.284_{(0.023)}$	$0.252_{(0.024)}$	$0.231_{(0.012)}$	$0.372_{(0.022)}$	$0.355_{(0.021)}$	$0.342_{(0.024)}$	$0.331_{(0.022)}$
Acetonitrile	$0.592_{(0.051)}$	$0.482_{(0.023)}$	$0.454_{(0.023)}$	$0.425_{(0.030)}$	$0.532_{(0.051)}$	$0.484_{(0.041)}$	$0.477_{(0.04)}$	$0.462_{(0.041)}$
Benzene	$0.233_{(0.013)}$	$0.241_{(0.024)}$	$0.202_{(0.014)}$	$0.174_{(0.013)}$	$0.455_{(0.040)}$	$0.531_{(0.104)}$	$0.278_{(0.01)}$	$0.181_{(0.017)}$
DMSO	$1.217_{(0.164)}$	$0.977_{(0.074)}$	$0.955_{(0.483)}$	$0.432_{(0.184)}$	$1.121_{(0.151)}$	$1.050_{(0.071)}$	$1.074_{(0.091)}$	$1.053_{(0.071)}$
Ethanol	$0.291_{(0.024)}$	$0.262_{(0.021)}$	$0.253_{(0.022)}$	$0.231_{(0.018)}$	$0.532_{(0.034)}$	$0.513_{(0.052)}$	$0.435_{(0.051)}$	$0.371_{(0.023)}$
Octanol	$0.452_{(0.023)}$	$0.461_{(0.045)}$	$0.374_{(0.028)}$	$0.321_{(0.024)}$	$0.842_{(0.034)}$	$0.854_{(0.042)}$	$0.827_{(0.031)}$	$0.825_{(0.031)}$
THF	$0.491_{(0.032)}$	$0.453_{(0.024)}$	$0.443_{(0.032)}$	$0.432_{(0.024)}$	$0.512_{(0.041)}$	$0.494_{(0.042)}$	$0.476_{(0.034)}$	$0.462_{(0.032)}$
Water	$2.584_{(0.045)}$	$2.324_{(0.037)}$	$2.117_{(0.043)}$	$2.105_{(0.052)}$	$3.804_{(0.119)}$	$3.204_{(0.12)}$	$3.197_{(0.104)}$	$3.162_{(0.107)}$
Hexane	$0.232_{(0.014)}$	$0.498_{(0.015)}$	$0.143_{(0.015)}$	$0.124_{(0.015)}$	$0.217_{(0.012)}$	$0.472_{(0.031)}$	$0.203_{(0.011)}$	$0.174_{(0.012)}$

Table 3: Cross-test results between MMGNN and other methods on solvents holdout test. (Underlined numbers are the best results)

factors such as the molecule's molar mass and element type. A comprehensive breakdown of results and in-depth analysis can be found in the appendix.

Obs.8: MMGNN model demonstrates effective generalization performance across various types of molecular scaffolds as shown in Figure 4 (b). Heterocyclic and polycyclic molecules exhibit better predictive performance than monocyclic molecules. For instance, phenanthrene and morpholine have predictivity values of 0.57 and 0.56, respectively, while cyclopentene has a predictivity result of 1.31. We hypothesize that monocyclic molecules tend to display nonpolar characteristics, whereas polycyclic and heterocyclic molecules often exhibit polar properties, which makes them easier to differentiate and learn for model. 1-yne and nitrate is significate polar molecular groups which has the relatively low prediction error.

Obs.9: MMGNN outperforms other baselines in generalization test. We selected three representative baseline models for the same generalization testing experiments (in Table 3). MMGNN showed significant superiority over other baseline models. This is mainly attributed to (1) its more fitting modeling of molecular combinations through the use of molecular merge graphs, which aligns better with the underlying physicochemical processes. The explicit enhancement of molecular interaction processes is achieved. (2) The molecular explainability module injection retains core substructures, removes redundant information, and further enhances generalization performance. Constrained by space limitations, additional solute test results are presented in the appendix.

5 Conclusion

In this study, we present MMGNN, a molecular graph neural network designed for predicting solvation free energy. MMGNN exploits the complete connectivity among atoms in solute and solvent molecules to facilitate information transfer and improve model explainability. Demonstrating superior performance on various datasets compared to existing methods, MMGNN highlights the significance of balancing intrinsic and interactive information through our experiments. The model excels in generalization tasks, including solvent holdout and solute scaffold split, superior than other baselines. Its transparent explanations align with well-established chemical insights, which instills confidence among experimental chemists. Furthermore, understanding the interactions between molecules could provide valuable perspectives for understanding drug compound stability and guiding the development of new synthetic routes and catalytic processes.

Ethical Statement

There are no ethical issues.

Acknowledgments

This paper is partially supported by the Project of Stable Support for Youth Team in Basic Research Field, CAS(YSBR-005), Natural Science Foundation of China National Major Research Instrument Development Project (No.12227901), the Innovation Program for Quantum Science and Technology (2021ZD0303303), the National Natural Science Foundation of China (22025304, 22033007, 62072427).

Contribution Statement

Wenjie Du and Shuai Zhang are co-first authors; Yang Wang and Junfeng Fang are corresponding authors.

References

[Behler, 2015] Jörg Behler. Constructing high-dimensional neural network potentials: a tutorial review. *International Journal of Quantum Chemistry*, 115(16):1032–1050, 2015.

[Behler, 2016] Jörg Behler. Perspective: Machine learning potentials for atomistic simulations. *The Journal of chemical physics*, 145(17):170901, 2016.

[Chung *et al.*, 2022] Y. Chung, F. H. Vermeire, H. Y. Wu, P. J. Walker, M. H. Abraham, and W. H. Green. Group contribution and machine learning approaches to predict abraham solute parameters, solvation free energy, and solvation enthalpy. *Journal of Chemical Information and Modeling*, 62(3):433–446, 2022.

[D'Souza *et al.*, 2011] Malcolm J. D'Souza, Anthony M. Darrington, and Dennis N. Kevill. A study of solvent effects in the solvolysis of propargyl chloroformate. *International Scholarly Research Notices*, 2011, 2011.

[Fang *et al.*, 2022] Xiaomin Fang, Lihang Liu, Jieqiong Lei, Donglong He, Shanzhuo Zhang, Jingbo Zhou, Fan Wang, Hua Wu, and Haifeng Wang. Geometry-enhanced molecular representation learning for property prediction. *Nature Machine Intelligence*, 4(2):127–134, 2022.

[Fang *et al.*, 2023a] Junfeng Fang, Wei Liu, Yuan Gao, Zemin Liu, An Zhang, Xiang Wang, and Xiangnan He. Evaluating post-hoc explanations for graph neural networks via robustness analysis. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023.

[Fang *et al.*, 2023b] Junfeng Fang, Wei Liu, An Zhang, Xiang Wang, Xiangnan He, Kun Wang, and Tat-Seng Chua. On regularization for explaining graph neural networks: An information theory perspective, 2023.

[Fang *et al.*, 2023c] Junfeng Fang, Xiang Wang, An Zhang, Zemin Liu, Xiangnan He, and Tat-Seng Chua. Cooperative explanations of graph neural networks. In *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*, pages 616–624, 2023.

[Fang *et al.*, 2024a] Junfeng Fang, Xinglin Li, Yongduo Sui, Yuan Gao, Guibin Zhang, Kun Wang, Xiang Wang, and Xiangnan He. EXGC: bridging efficiency and explainability in graph condensation. *CoRR*, abs/2402.05962, 2024.

[Fang *et al.*, 2024b] 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. *CoRR*, abs/2402.03781, 2024.

[Glick *et al.*, 2020] Zachary L. Glick, Derek P. Metcalf, Alexios Koutsoukas, Steven A. Spronk, Daniel L. Cheney, and C. David Sherrill. Ap-net: An atomic-pairwise neural network for smooth and transferable interaction potentials. *The Journal of Chemical Physics*, 153(4):044112, 2020.

[Grubbs *et al.*, 2010] Laura M. Grubbs, Mariam Saifullah, E. Nohelli, Shulin Ye, Sai S. Achi, William E. Acree Jr, and Michael H. Abraham. Mathematical correlations for describing solute transfer into functionalized alkane solvents containing hydroxyl, ether, ester or ketone solvents. *Fluid phase equilibria*, 298(1):48–53, 2010.

[Hildebrand, 1981] J. H. Hildebrand. A history of solution theory. *Annual Review of Physical Chemistry*, 32:1–23, 1981.

[Kingma and Ba, 2014] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[Klamt and Eckert, 2000] A. Klamt and F. Eckert. Cosmors: a novel and efficient method for the a priori prediction of thermophysical data of liquids. *Fluid Phase Equilibria*, 172(1):43–72, 2000.

[Lee *et al.*, 2023] Namkyeong Lee, Dongmin Hyun, Gyoung S. Na, Sungwon Kim, Junseok Lee, and Chanyoung Park. Conditional graph information bottleneck for molecular relational learning. *arXiv preprint arXiv:2305.01520*, 2023.

[Low *et al.*, 2022] K. Low, M. L. Coote, and E. I. Izgorodina. Explainable solvation free energy prediction combining graph neural networks with chemical intuition. *J Chem Inf Model*, 62(22):5457–5470, 2022.

[Maas *et al.*, 2013] Andrew L Maas, Awni Y Hannun, Andrew Y Ng, et al. Rectifier nonlinearities improve neural network acoustic models. In *Proc. icml*, volume 30, page 3. Atlanta, GA, 2013.

[Marenich *et al.*, 2007] A. V. Marenich, R. M. Olson, C. P. Kelly, C. J. Cramer, and D. G. Truhlar. Self-consistent reaction field model for aqueous and nonaqueous solutions based on accurate polarized partial charges. *Journal of Chemical Theory and Computation*, 3(6):2011–2033, 2007.

[Meng *et al.*, 2023] Fanwang Meng, Hanwen Zhang, Juan Samuel Collins-Ramirez, and Paul W. Ayers. Something for nothing: Improved solvation free energy prediction with -learning. 2023.

[Miao *et al.*, 2022] Siqi Miao, Mia Liu, and Pan Li. Interpretable and generalizable graph learning via stochastic attention mechanism. In *ICML*, pages 15524–15543, 2022.

[Mobley and Guthrie, 2014] David L. Mobley and J. Peter Guthrie. Freesolv: a database of experimental and calculated hydration free energies, with input files. *Journal of computer-aided molecular design*, 28:711–720, 2014.

[Moine *et al.*, 2017] Edouard Moine, Romain Privat, Baptiste Sirjean, and Jean-Noël Jaubert. Estimation of solvation quantities from experimental thermodynamic data: Development of the comprehensive compsol databank for pure and mixed solutes. *Journal of Physical and Chemical Reference Data*, 46(3):033102, 2017.

[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 AAAI Conference on Artificial Intelligence*, volume 34, pages 873–880, 2020.

[Rong *et al.*, 2020] Yu Rong, Yatao Bian, Tingyang Xu, Weiyang Xie, Ying Wei, Wenbing Huang, and Junzhou Huang. Self-supervised graph transformer on large-scale molecular data. *Advances in Neural Information Processing Systems*, 33:12559–12571, 2020.

[Varghese and Mushrif, 2019] Jithin John Varghese and Samir H Mushrif. Origins of complex solvent effects on chemical reactivity and computational tools to investigate them: a review. *Reaction Chemistry & Engineering*, 4(2):165–206, 2019.

[Velickovic *et al.*, 2017] Petar Velickovic, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, Yoshua Bengio, et al. Graph attention networks. *stat*, 1050(20):10–48550, 2017.

[Vermeire and Green, 2021] Florence H Vermeire and William H Green. Transfer learning for solvation free energies: From quantum chemistry to experiments. *Chemical Engineering Journal*, 418:129307, 2021.

[Wang *et al.*, 2023] Binwu Wang, Yudong Zhang, Jiahao Shi, Pengkun Wang, Xu Wang, Lei Bai, and Yang Wang. Knowledge expansion and consolidation for continual traffic prediction with expanding graphs. *IEEE Transactions on Intelligent Transportation Systems*, 2023.

[Wang *et al.*, 2024] Binwu Wang, Pengkun Wang, Yudong Zhang, Xu Wang, Zhengyang Zhou, Lei Bai, and Yang Wang. Towards dynamic spatial-temporal graph learning: A decoupled perspective. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 9089–9097, 2024.

[Wen *et al.*, 2021] Mingjian Wen, Samuel M. Blau, Evan Walter Clark Spotte-Smith, Shyam Dwaraknath, and Kristin A. Persson. Bondnet: a graph neural network for the prediction of bond dissociation energies for charged molecules. *Chemical science*, 12(5):1858–1868, 2021.

[Wu *et al.*, 2020] Tailin Wu, Hongyu Ren, Pan Li, and Jure Leskovec. Graph information bottleneck. In *NeurIPS*, 2020.

[Xia *et al.*, 2023a] Jun Xia, Haitao Lin, Yongjie Xu, Cheng Tan, Lirong Wu, Siyuan Li, and Stan Z Li. Gnn cleaner: Label cleaner for graph structured data. *IEEE Transactions on Knowledge and Data Engineering*, 2023.

[Xia *et al.*, 2023b] Jun Xia, Yanqiao Zhu, Yuanqi Du, Yue Liu, and Stan Z Li. A systematic survey of chemical pretrained models. *IJCAI*, 2023.

[Yao *et al.*, 2018] Kun Yao, John E. Herr, David W. Toth, Ryker McKintyre, and John Parkhill. The tensormol-0.1 model chemistry: a neural network augmented with long-range physics. *Chemical science*, 9(8):2261–2269, 2018.

[Ying *et al.*, 2019] Zhitao Ying, Dylan Bourgeois, Jiaxuan You, Marinka Zitnik, and Jure Leskovec. Gnnexplainer: Generating explanations for graph neural networks. In *NeurIPS*, pages 9240–9251, 2019.

[Yu *et al.*, 2021] Junchi Yu, Tingyang Xu, Yu Rong, Yatao Bian, Junzhou Huang, and Ran He. Graph information bottleneck for subgraph recognition. In *ICLR*. OpenReview.net, 2021.

[Yu *et al.*, 2023] Jiahui Yu, Chengwei Zhang, Yingying Cheng, Yun-Fang Yang, Yuan-Bin She, Fengfan Liu, Weike Su, and An Su. Solvbert for solvation free energy and solubility prediction: a demonstration of an nlp model for predicting the properties of molecular complexes. *Digital Discovery*, 2(2):409–421, 2023.

[Zhang *et al.*, 2022] D. D. Zhang, S. Xia, and Y. K. Zhang. Accurate prediction of aqueous free solvation energies using 3d atomic feature-based graph neural network with transfer learning. *Journal of Chemical Information and Modeling*, 62(8):1840–1848, 2022.

[Zhou *et al.*, 2023a] Gengmo Zhou, Zhifeng Gao, Qiankun Ding, Hang Zheng, Hongteng Xu, Zhewei Wei, Linfeng Zhang, and Guolin Ke. Uni-mol: a universal 3d molecular representation learning framework. 2023.

[Zhou *et al.*, 2023b] Zhengyang Zhou, Qihe Huang, Gengyu Lin, Kuo Yang, Lei Bai, and Yang Wang. Greto: remedying dynamic graph topology-task discordance via target homophily. In *The eleventh international conference on learning representations*, 2023.