WORLD JOURNAL OF PHARMACY AND PHARMACEUTICAL RESEARCH

Review Article

Coden USA: WJPPCL

2026
Volume: 03
Issue: 06
Page: 110-133

"GRAPH NEURAL NETWORKS (GNNS) IN DRUG DISCOVERY AND DRUG TARGET AFFINITY (DTA) PREDICTION"

Komal Jain[1*], Dr. Shailendra Chawda[2]

Research Scholar[1], Professor[2]

Mandsaur Institute of Pharmacy, Mandsaur University, Mandsaur.

Received: 21 April 2026 **Revised: 11 May 2026** **Accepted: 31 May 2026**

Corresponding Author: Komal Jain

Address: Research Scholar, Mandsaur Institute of Pharmacy, Mandsaur University, Mandsaur.

DOI: https://doi.org/10.5281/zenodo.20483779,

ABSTRACT:

Graph Neural Networks (GNNs) are deep learning models that process graph-structured data, which consists of nodes and edges. They capture complex relationships and interaction, making them ideal for tasks like molecular drug discovery, Drug Target Interactions (DTI), and Drug Target (Binding) Affinity (DTA). GNNs revolutionize drug discovery by treating molecules as graphs, with atoms as nodes and chemical bonds as edges. This allows models to understand the complex three-dimensional structures and patterns relevant to biological activity. In poly-pharmacology, GNNs effectively model drug-drug interactions (DDIs) and predict multi-target activities. They improve predictive accuracy, lower development costs, and reduce late-stage failures. This review examines the application of GNNs in various phases of drug discovery, including lead discovery, optimization, synthetic route design, drug-target interaction prediction, and molecular property profiling, while addressing challenges in translational medicine. Computational drug-target affinity prediction has the potential to accelerate drug discovery. Currently, pre-training models have achieved significant success in various fields due to their ability to train the model using vast amounts of unlabelled data. However, given the scarcity of drug-target interaction data, pre-training models can only be trained separately on drug and target data, resulting in features that are insufficient for drug-target affinity prediction. To address this issue, a graph neural pre-training-based drug-target affinity prediction method (GNPDTA). This approach comprises three stages. In the first stage, two pre-training models are utilized to extract low-level

features from drug atom graphs and target residue graphs, leveraging a large number of unlabelled training samples. In the second stage, two 2D convolutional neural networks are employed to combine the extracted drug atom features and target residue features into high-level representations of drugs and targets. Finally, in the third stage, a predictor is used to predict the drug-target affinity. This approach fully utilizes both unlabelled and labelled training samples, enhancing the effectiveness of pre-training models for drug-target affinity prediction. In our experiments, GNPDTA outperforms other deep learning methods, validating the efficacy of our approaches.

KEY WORDS: Graph neural networks; Lead discovery; Lead optimization; Synthetic route; Drug—target interaction; Property prediction; Virtual screening; De novo drug design, drug-target affinity, pre-training model, graph isomorphism network, deep neural network, feature extraction.

INTRODUCTION:

To facilitate precision medicine in complex diseases like cancer, researchers are increasingly employing computational approaches to explore the interactions between drugs and cancer cells. Recently, many machine learning and deep learning methods have been effectively applied to predict drug response levels with high accuracy. However, most of these methods focus on phenotypic screening and lack reasonable interpretability, which obscures our understanding of the mechanisms behind drug reactions. To advance precision medicine, it is essential to clarify how drugs work and to promote the discovery of new drugs. A proper representation of a drug molecule is crucial for any prediction method related to drug response. Recent reviews of molecular representations identify three main categories: linear notations, molecular fingerprints (FPs), and graph notations. Linear notations represent a molecule using a vector of strings, with two commonly used examples being the IUPAC International Chemical Identifier (InChI) and the Simplified Molecular-Input Line-Entry System (SMILES). SMILES strings are more prevalent because they encode the chemical structure as a string of ASCII characters.

Molecular fingerprints, such as the Molecular Access System (MACCS) and Chemically Advanced Template Search, identify key structures within a molecule and represent them using a binary vector, where each bit indicates the presence of a specific structure. A notable drawback of this type of representation is that it can only recognize pre-defined structures, which may hinder the discovery of novel compounds. To address this limitation, circular

fingerprints, like Extended Connectivity Fingerprints (ECFPs) based on the Morgan algorithm, have been developed. These fingerprints iteratively search for substructures within molecules rather than defining them in advance. While this approach preserves critical structural information, it does lose positional information, making it difficult to determine where these substructures occur within the molecule. DeepDSC integrates Morgan fingerprints of drugs into the latent features of cancer cell lines, which are learned by an auto encoder. S2DV applies word2vec to tokenize ECFP features or SIMLES to create drug representations. Ma et al. utilized Atom Pairs (AP), MACCS, and circular fingerprints as drug descriptors and conducted a quantitative structure-activity relationship (QSAR) study using a Deep Neural Network. Recently, graph notations have gained attention in the field of drug representation. Previous methods, which prioritized computational simplicity and the limited capabilities of graph learning, often compromised the detail that can be captured in molecular structures. However, with the emergence of Graph Neural Networks (GNN) in deep learning, it has become possible to store and analyze molecular information using graph-based representations. Various GNN models have been applied in the pharmaceutical domain, effectively learning the latent representations of molecular graphs while balancing descriptive power and complexity. A Graph Convolution Network (GCN) model has been proposed to predict the chemical properties of molecules and discover porous materials. The traditional message-passing mechanism of GNNs tends to diminish the influence of distant nodes, which can be misleading in real molecules where atoms that are far apart can still interact, such as through intra-molecular hydrogen bonds. To address this issue, the Attentive FP model leverages a graph attention mechanism to recognize the impact of one node on another. This model improves node updates by balancing topological distance with potential interactions through the attention mechanism. The GraphDRP model enhanced prediction precision by replacing the drug-CNN module with GNN to better capture drug features. Additionally, models like DeepCDR, TGSA, and DualGCN further explored the integration of multi-omics profiles for improved representation of cancer cell lines. Beyond modeling drugs with GNNs, SWNet introduced a self-attention mechanism to account for drug similarity when learning cell features. An algebraic graph-assisted bidirectional transformer (AGBT) model was developed to encode the 3D structures of molecules into algebraic graphs. Additionally, a Molecular Topographic Map (MTM) was generated from atom features using Generative Topographic Mapping (GTM) to represent drugs in a graph format.

Predicting drug-target affinity (DTAP) is a crucial research topic in drug development, which can be used for discovering drug on-target and off-target effects. However, due to the increasing number of targets, it is difficult to fully validate the drug-target affinity (DTA) of drugs using biochemical experiments. In recent years, with the development of artificial intelligence technology, the use of computational methods for preliminary prediction of DTA has become an economically effective method. Graph neural networks (GNN) can extract features from graph-structured data and have been widely used in DTA, as drugs and targets are typically graph-structured data. Different graph can enhance understanding of atomic connectivity and residue interactions. The strengths of attention mechanisms lie in their ability to focus on relevant parts of the graph, enhancing the model's representational power. Various deep learning frameworks offer diverse and innovative approaches for DTAP. Combining GNNs with other network architectures provides a comprehensive approach for DTAP, leveraging the strengths of different network types.

1. Limitations of Existing CADD Approaches:

Accurate prediction of molecular properties is a fundamental challenge in drug design and discovery. Drug molecules inherently possess distinct physicochemical as well as physiologic—toxicological properties. Nevertheless, the process of acquiring their relevant property parameters is frequently time consuming and labour-intensive. Over recent decades, computer aided drug design (CADD) has emerged as a transformative approach to address these challenges. CADD uses high performance computers that can rapidly simulate many steps in drug design, including generative molecular modeling, molecular property prediction, and virtual screening. However, due to algorithmic constraints, molecules generated by such CADD methods often turn out to be chemically unfeasible or synthetically inaccessible, and predictions are not entirely accurate due to current limitations of existing CADD methods.

1.1. Limitations of Virtual Molecule Generation:

Commonly used CADD methods encounter difficulties in simultaneously generating a large number of completely novel molecules that possess high synthesizability and bioactivity. Common methods for obtaining a large number of molecules involve screening large databases (such as ZINC, ChEMBL, etc.) and computer-based generation. Screening standard libraries is limited to 10 million known compounds. Comparing with the chemical space that containing as many as 1060 drug-like molecules that adhere to Lipinski's rule of five, there is

a large blank. Molecule generation without the help of neural networks or deep learning can explore new chemical space, yet it fails to strike a balance among novelty, synthesizability and bioactivity.

1.2. Limitations of synthetic route design:

The synthetic accessibility score (SAscore) represents a traditional and specific approach of CADD for evaluating the synthesizability of a molecule. By establishing a threshold for the SAscore, the synthesizability of molecules can be categorized. However, this approach of setting thresholds based on SAscore tends to be inflexible. Programs such as RECAP, BRICS, and eMolFrag are employed to decompose molecules with high synthesizability into blocks. However, these programs differ in their focus when breaking down the molecules and cannot handle complex rings effectively. As a result, a substantial amount of manual labour is still needed to devise a comprehensive synthetic route.

1.3. Limitations of virtual screening:

Virtual screening plays a key role in the drug discovery process, and its core methods include structure-based virtual screening (SBVS) and ligand-based virtual screening (LBVS). Drug—target interaction (DTI) studies serve as the foundation of SBVS. Molecular docking, a typical research instrument for DTI studies is applicable to protein binding sites of known or predictable structure, providing highly accurate results by modeling ligandreceptor interactions. However, it is complicated by the time complexity [O(n2)], which requires a deeper understanding of the protein. These problems mainly restrict its applicability to approximately 35% of proteins with explicit structure information. Although AlphaFold prediction covers 98.2% of proteins, the accuracy of the predictions still awaits verification. Especially in folding prediction of complex domain of protein, AlphaFold showed an unsatisfactory Performance. Moreover, commonly used molecular docking (e.g., AutoDock, Discovery Studio, and Schro¨dinger's Glide) often produce different results when dealing with the same problem due to different underlying architectures. DTA represents an alternative approach for implementing SBVS. DTA is used to assess the binding potential between small molecules and proteins and to provide information on binding strength. In the early stages, computational models for predicting DTA predominantly relied on machine learning methods. Although traditional machine learning methods can accurately predict DTA, they involve complex and time-consuming feature engineering. This typically necessitates manual data processing and annotation, followed by extraction of valuable

features. In addition, the performance of these models tends to degrade as the size of the dataset increases. As a result, traditional machine learning methods have significant limitations in the field of DTA predictions. LBVS is a drug-screening method that relies on molecular properties such as quantitative structure-activity relationship (QSAR), pharmacophores and structural similarity. To realize LBVS, the prerequisite is an explicit QSAR model or pharmacophore model. Developing such models demands a substantial number of experimentally validated drugs, a task that is improbable to be accomplished in the short term. In addition, the screening process needs a large compound library, making the LBVS faces similar challenges as the virtual molecule generation.

1.4. GNNs redefine CADD:

GNNs upgrade drug design from step-by-step rule-based optimization in traditional CADD to autonomous discovery through end-to-end molecular topology learning. In traditional CADD, we often need to abstract a molecular descriptor for defining certain properties of a drug molecule through feature engineering, but a finite feature set cannot capture complex topological relationships. In addition, traditional CADD adopts a step-by-step pipeline of virtual screening, efficacy optimization, and ADMET prediction. These fragmented tasks may lead to error accumulation and ultimately produce unreliable results. In contrast, drug discovery leveraging GNNs eliminates the need for manual feature specification; instead, it directly takes the entire molecular graph as input. Whether GNNs are used for only one part of drug discovery or the entire process is handed over to AI, directly utilizing graphs as input minimizes information loss and, in theory, reduces errors. Nevertheless, when faced with extremely large graph datasets, or small datasets featuring complex topologies, the characteristic of GNNs to use molecular graphs as inputs may shift from an advantage to a disadvantage. This gives rise to issues like the dimensionality catastrophe, which further leads to difficulties in fitting effective results.

2. GNNs pioneer the future of CADD

2.1. GNNs revolutionize drug discovery:

Graph neural networks (GNNs) were first proposed by Franco Scarselli as new neural network models in 2004. They represented a generalization of two prevailing approaches, the 6164 Rui Wang, Chunlin Zhuang graph-centric approach, and the node-centric approach. A graph is defined as a collection of node sets, edge sets, and global information with certain mapping relationships. Based on this concept, a molecule can be embedded into a graph.

Atoms can be regarded as nodes; chemical bonds can be regarded as edges and the entire molecule can be regarded as global information. By converting this information into matrices, a molecule is successfully transformed into a mathematical language that is more readily learned by computers. Besides, graphs can simplify knowledge containing complex relationships. Terms like "Selegiline", "MAO-B inhibitors" and "drugs for Parkinson disease" can be treated as nodes, while edges represent their containing relationships. When there are enough terms, a large network can be formed, known as a knowledge graph. The core concept of GNNs is the message passing mechanism. That means taking a node as the centre, integrating the information of the neighbouring nodes through the edges, gradually expanding the learning scope, and finally uncovering the hidden relationships between nodes and nodes. Unlike strings, images, or feature matrixes (tables where each row represents a sample and each column represents a feature, being used as input for machine learning models), graphs can accurately reflect the nodes and edges, making them suitable for the training of GNNs. Through learning the information between nodes, GNNs can capture hidden information and generate outputs for common classification or regression tasks in drug design. Graph neural networks (GNNs) have introduced ground breaking solutions to these challenges.

2.2. Comparative advantages and limitations of GNNs versus alternative neural networks:

The optimal selection of neural network architecture for drug discovery tasks depends critically on the data structure and problem constraints. Below, we draw a comparison between GNNs and other widely-used alternatives: convolutional neural networks (CNNs), recurrent neural networks (RNNs), and transformers.

Table 1 Comparative advantages and limitations of GNNs *versus* alternative neural network.

Feature	Graph neural network (GNNs)	Convolutional neural network (CNNs)	Recurrent neural network (RNNs)	Transformer
Data structure	Graphs (Nodes, edges)	Grids (Images, voxels)	Sequences (Ordered tokens)	Sequences (Ordered tokens + relations)
Core mechanism	Message passing: aggregates neighbor features to update node states.	Convolution: Fixed kernels on local receptive fields.	Recurrence: Hidden state tracks temporal dependencies.	Self-attention: Weights relationships between all pairs of tokens globally.
Key strengths	• Explicit relational modeling. • Topological invariance. • SAR modeling. • Natural fit for molecular graphs.	• Strong spatial feature extraction. • Shift-invariance. • Proven on images/3D grids.	• Handles variable-length sequences. • Captures sequential dependencies. • Low computational cost per step.	• Captures long-range dependencies. • Parallelizable training. • Context-aware representations. • Strong on sequence-to-sequence tasks.
Critical limitations	• Over-smoothing: deep layers lose node distinctions. • Scalability: high-degree graphs are resource intensive. • Dynamic struggles: poor with conformational kinetics. • Long-range: global context capture can be weak.	• Lack topology: cannot natively encode graphs. • Fixed size: needs padding/cropping. • Spatial distortion: 3D rasterization may create artifacts. • Global context: diluted in deep networks.	• Vanishing gradients: struggles with long sequences. • Sequential bottleneck: slow training. • Temporal bias: later tokens dominate. • Local focus: weak explicit global reasoning.	• Computational cost: $O(N^2)$ memory for sequence length. • Data hunger: requires large datasets. • Structural agnosticism: ignores explicit topology/bond info. • Interpretability: attention can be opaque.
Best-suited drug tasks	• Molecule property prediction. • *De novo* molecular graph generation. • DTI prediction (structure-based). • Reaction prediction.	• Binding affinity from 3D structures. • Protein structure classification (from density). • High-content imaging analysis.	• SMILES generation. • Protein sequence feature extraction.	• Protein language modeling. • AI-generated molecules *via* language prompts. • Multi-modal knowledge integration. • Graph transformers: combines graph & sequence benefits.

3. GNNs for lead discovery and optimization

GNNs can capture the hidden relationship between atoms and bonds of molecules, to help with lead discovery and optimization. Combining GNNs with more information, like additional inter-atom relationships, fragment information, SMILES or molecular fingerprint, and prior knowledge, GNNs can fasten the process.

3.1. GNNs for de novo drug design:

De novo drug design works based on an in-depth understanding of the structure and functions of the targets including enzymes, receptors, or other biomolecules. This approach aims to explore a wide range of chemical space to discover new molecular entities. The involvement of GNNs in the process renders de novo drug design feasible for targets with limited research and accelerates traditional de novo drug design. To discover novel molecules, the most typical models are inference-based generators, which rely on Markov chains (a random process of transition of something from one state to another) or Monte Carlo trees (a numerical simulation method that takes probabilistic phenomena as the object of study). AstraZeneca, collaborated with Mercado et al and developed a model based on hierarchical GNN, GraphINVENT. This model achieved the probabilistic generation of new molecules, adding one chemical bond at a time. It can quickly learn to construct molecules like those in the training set without explicit programming of chemical rules. GraphINVENT, as an effective program, has been serving as the starting point for the design of future molecule generators, such as state-of-the-art reinforcement molecule generators, de novo PROTAC design models, polymer design models, GRELinker for lead optimization, etc. Fang et al integrated a multi-targeted deep reinforcement learning pipeline, called QADD, for enabling an accurate quality assessment of the drug potential of the generated molecules and intermediate molecules. Through the newly designed pipeline, they discovered a series of compounds that exhibited strong binding affinities to DRD2. These compounds have better quantitative estimates of drug-likeness (QED) and SAscore values than the positive drug risperidone. Zhu et al developed a pharmacophore-guided deep learning framework for molecule generation, named PGMG, which enabled intelligent design of bioactive molecules through integration of pharmacophore features. The method established a flexible and controllable generation framework by implementing pharmacophore-driven guidance, generating over 1000 novel molecular structures across distinct therapeutic targets. Experimental validation demonstrated that these generated molecules successfully recapitulate key pharmacophore features of known active compounds, thereby confirming its

effectiveness in structure-based drug design (SBDD). Novel inhibitors targeting TGFßR1, designed by PGMG, exhibited better QED, biological affinity, synthetic accessibility (SAscore), as well as lower cardio-toxicity—all superior to those of the leads. Atz et al used the idea of "graph-to-sequence-to-graph" to achieve "zero shot" (a deep learning method that allows the model to build up with insufficient data in the initial phase) construction of compound libraries. They developed a method named DRAGONFLY, which employed a framework that encoded input molecule graphs using a transformer (neural network architecture for processing sequence data) and a GNN model. Based on DRAGONFLY, a series of PPAR inhibitors were generated. The compound was further validated as a novel and active compound by a co-crystallography study.

3.2. GNNs for lead optimization:

Fragment-based drug discovery (FBDD) combines the central idea of combinatorial chemistry to design novel drug molecules by rational combinations of organic structural fragments. Xie et al proposed a multi-target drug molecule discovery method, called MARS. This method employed the Markov chain Monte Carlo trees and simulated an annealing algorithm to iteratively modify molecular structure fragments for generating efficient candidates. MARS integrated an adaptive molecular analysis strategy with dynamically trained GNNs for molecular characterization and screening to successfully design novel molecules targeting GSK3ß or JNK3. The generated compounds not only exhibited favourable QED (0.76 ± 0.02) and SAscore (0.90 ± 0.03), but also showed promising dual-targeting inhibitory activities. GRELinker , a generative fragment linker model proposed by Zhang et al, combined a gated GNN with reinforcement learning to generate molecules with proposed properties. Novel A2AAR inhibitors designed by GRELinker were predicted to have superior docking scores compared to the leads Scaffold hopping is a strategy for compound optimization starting with a known active compound. Novel chemical structures are obtained by changing the core structure of the molecule. This technique can help avoid patent-protected chemical space or improve the pharmacokinetic properties, selectivity, potency, and other relevant characteristics. Zheng et al proposed a new multimodal model, one that combined multiple inputs, named DeepHop, for scaffold hopping. The trained DeepHop model was shown to generate approximately 70% of molecules with enhanced bioactivity, high 3D similarity, and low 2D scaffold similarity compared to the leads. In a case study using DeepHop, the design of inhibitors for Src-family kinases yielded the most promising results. The generated molecules were demonstrated to have not only a high degree

of spatial structural similarity, but also enhanced predicted activities in different folds. Hu et al. constructed a variational auto encoder based on multi-view GNNs to enable molecule generation and scaffold hopping. Its effectiveness was demonstrated through the design of a series of LRRK2 inhibitors for Parkinson's disease.

3.3. GNNs for drug repurposing:

Drug repurposing aims to discover new applications for clinical drugs, as well as for drug candidates in preclinical or clinical studies. This line of research can utilize existing toxicological and pharmacokinetic evaluations, which significantly cuts down the development time and research costs by 40%. GNNs can rapidly learn drug—disease relationships to fasten drug repurposing process. Cui et al proposed a GNN model, GraphRepur, for repurposing drugs for breast cancer. This model combined two main computational methods—drug network-based and drug signature-based approaches—to predict potential repurposed drugs. Subsequent reports have confirmed that the topranked drugs identified by the model show a promising therapeutic potential for breast cancer. In addition to disease-specific models, more drug repositioning studies have focused on pervasive modeling for multiple diseases. Network information, like drug—disease association networks, drug—drug similarity networks, and disease—disease similarity networks, can be characterized by graph-based data as the training data. Meng et al constructed knowledge graph about drug—disease association, drug—drug similarity, and disease—disease similarity individually. Their resulting model, named DRWBNCF, successfully identified potential drugs for breast cancer and small cell lung cancer. Huang et al built a zero-shot drug repurposing GNN model, TxGNN, realized drug repurposing for multiple rare diseases. Clinical experts who participated in the case studies expressed a high degree of confidence in the predictive results of TxGNN, demonstrating that the predictive results of TxGNN were consistent with medical reasoning. However, setting up a knowledge graph for model training involves a massive amount of work. Tayebi et al found a way to solve the problem by combining drug—drug interactions (DDIs), drug categorical information, and drug structural information. Based on these, they built an end-to-end model for drug repurposing, called EKGDR, which directly utilizes input data without additional processing. It is characterized by its ability to automatically capture complex associations in the knowledge graph (e.g., DDIs, classification, structure) without mankind's intervention (e.g., labeling drugs and diseases). That could greatly reduce the dependence on humans in the model design and training process. EKGDR enabled the prediction of 20 potential drug-

repositioning candidates for treating Alzheimer's disease and Parkinson's disease. Shao et al constructed a graph convolutional neural network (GCN; a convolutional neural network on graphs) on a heterogeneous graph and a graph attention network (GAT; a graph neural network introduced attention mechanism) to realize the DTI prediction. This work resulted in the development of the DTI-HETA model that was also an end-to-end model, successfully applying for some targets, such as ADRA2C, ERBB4, and JAK2. Similarly, to minimize the amount of human work involved in constructing the knowledge graph, Du et al introduced heuristic search (searching algorithm incorporates a heuristic functions like genetic algorithms, simulated annealing that mimic the behaviour of nature) to the knowledge graph learning in GNNs. They built the GNN model named convolutional KGCNH. With the help of heuristic search, KGCNH can learn like a human automatically, facilitating prediction on drug—disease associations, thereby realizing reposition of the drug targets. The current mainstream idea for future research on drug repurposing prediction models is to design multimodal models using GNNs in combination with other models. Wang et al integrated pre-training with GNNs and came up with PT-KGNN. Using pre-training techniques on biomedical knowledge graphs, abundant semantic and structural information can be learned, thereby improving the performance of downstream processes.

4. GNNs for synthetic route design

Chemical reactions are essentially the processes of breaking and forming bonds, adding and removing atoms from chemical molecules. This nature fits the advantages of GNNs for processing information built with nodes and edges. Synthetic feasibility assessment treats the synthesizability of a molecule as a regression or categorization task. GNNs can refine the properties that should be considered in these tasks, giving a more precise and detailed synthesizability score. Yu et al applied the GAT mechanism to predict synthesizability and construct a novel synthesizability scoring model GASA. GASA can not only predicts synthesizability, but also provide an atomic-level explanation of the predictions based on its GAT. The analysis of retrosynthesis using GNNs is also a hot research topic. Liu et al built a synthetic prediction model called RetroGNN via GNNs, which realized 105 times faster retrosynthetic analysis than using retrosynthetic software (e.g., BRICS in RDkit). Lin et al combined GNNs with transformers and proposed a retrosynthesis prediction model G2GT. In case studies, reactions predicted by G2GT were highly overlapped with the reaction design by synthesis experts. Similarly, Zhong et al had built an end-to-end GNN model for multi-

step retrosynthesis planning, called Graph2Edits. In the case study of nirmatrelvir, all prediction results from Graph2Edits almost exactly copied the routes from Pfizer.

5. GNNs for virtual screening

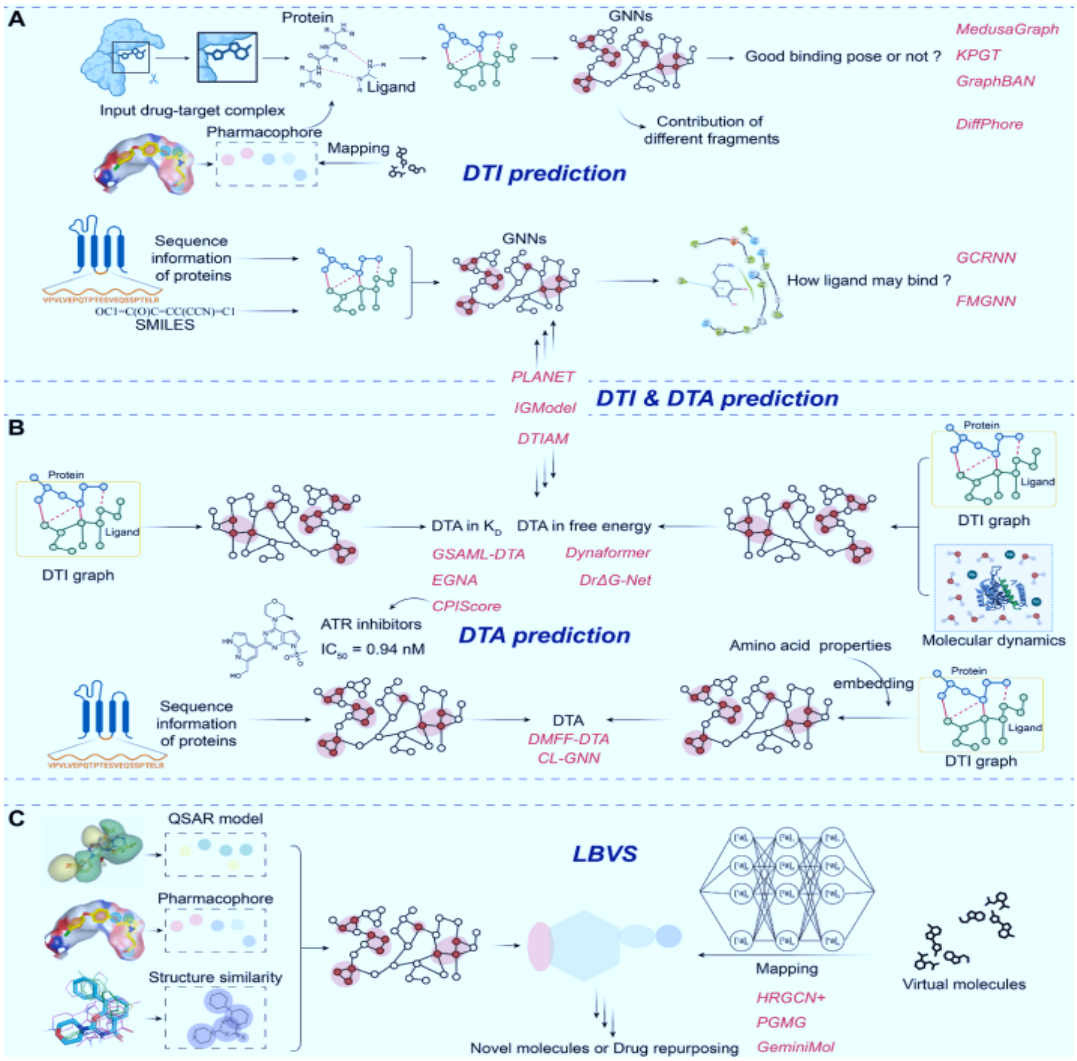

Figure: GNN models for DTI and DTA prediction and LBVS. (A) Schematic illustration of different models that realized DTI prediction. (B) Schematic illustration of different models realized DTA prediction. (C) Schematic illustration of different models realized LBVS.

5.1. GNNs for DTIs:

Clear DTIs can explain "how a drug molecule binds with target", and it is the basis of drug molecular mechanism and conformational studies. GNNs can help us obtain accurate DTIs closer to reality faster and give basic predictions of DTIs in the absence of protein structure information. In general, existing computational DTI study protocols include two main steps: sampling and scoring. Sampling means searching the space where the drug binds to the

target, finding as many poses as possible. Scoring means accessing the binding affinity of the drug and target of each pose .GNNs can revise both sampling and scoring to improve the accuracy of DTI studies. The schematic flowchart of DTI prediction is presented in Jiang et al constructed a prediction model, named MedusaGraph, for recognizing good poses of protein—ligand complexes. Their model utilized GNNs for two different functions, one for prediction and another for scoring, on predicting the poses and evaluating the poses. These two GNNs complements could find the best pose for DTI with high efficiency. Similarly, KPGT proposed by Li et al used self-supervised learning (SSL; an unsupervised learning technique that is often used to explore the hidden information inherent in the learning object automatically) in their framework, and found two inhibitors of HPK1 and FGFR1 for cancer therapy. GraphBAN proposed by Hadipour et al introduced teacher-student learning model to GNN, which made GraphBAN have ability to summary from training dataset. With the help of GraphBAN, they successfully got Pin1 inhibitors. Feeding GNNs with additional knowledge can further improve DTI prediction accuracy. Pharmacophores are important information for DTI prediction, distilling and summarizing essential chemical interaction patterns. DiffPhore proposed by Li et al learned from 3D ligand—pharmacophore graph data to realize accurate DTI predictions. Using DiffPhore, they successfully identified structurally distinct inhibitors of human glutamine cyclase for Alzheimer's disease therapy and further validated their binding modes by cocrystal analysis. GNNs can predict DTI without protein structures. Elbasani et al.65 proposed a model combining 3 different networks, containing a GNN for molecule representation, a bipartite RNN, and a CNN with long short-term memory (LSTM; an algorithm can make model perform well on long sequence) for protein sequence vectorization for model to learn. A highly accurate DTI prediction model called GCRNN was developed. With the help of GCRNN, predicting DTI of structure information was realized by using only protein primary structure information instead of 3D structure information. Tang et al proposed a novel hybrid model, FMGNN, which combined a factorization machine and a GNN to extract both low-order and high-order features. FMGNN enhanced the prediction accuracy by incorporating the pharmacophore features of drug molecules and the physicochemical properties of amino acid residues and the results from biological experiments.

5.2. GNNs for DTA:

DTA can explain "how strong a drug molecule binds with target", and it is the most basic indicator to assess the binding ability of small molecule with a target. The schematic

flowchart of DTA prediction is presented in the figure above. The typical measurements of DTA include dissociation constant (KD), inhibition constant (Ki), or 50% inhibitory concentration (IC50). Liao et al. constructed GSAML-DTA integrating the self-attention mechanism and GNN, which outperformed the stateof-the-art of DTA prediction methods on two test datasets. GSAML-DTA can predict KD values and the contributions of different motifs to the KD. Similarly, Xia et al. developed an empirical GNN, named EGNA, containing different parts for proteins, ligands and their interactions. They treated DTA prediction as a regression task and realized an accurate function for GTA prediction. Liang et al utilized the functionality of transformers and GCNs to enhance the prediction of DTA by constructing a GNN model, called CPIScore. This model was designed based on the transformer architecture to capture comprehensive global contextual information of protein and ligand sequences. Meanwhile, the GCN component was used to extract local features from the small-molecule graph. When applied to the generated compound library, CPIScore successfully identified potent small-molecule inhibitors of ATRs, notably one of which was confirmed experimentally with inhibitory activity below 1 nmol/L. For proteins without an accurate structure, using its amino acids sequence to perform its structure information is a good idea. He et al proposed DMFF-DTA, a bimodal neural network model that integrated sequence and graph information from drugs and proteins for DTA prediction70. Moreover, embedding the amino acid properties into the model improved the convergence speed and boosted accuracy of protein feature learning. Zhang et al introduced SSL into the GNN training process, being utilized to learn from a large dataset of 371458 different unlabelled protein—ligand complexes. The resulted DTA prediction model, called CL-GNN, showed the ability to effectively predict DTA and quantify the contribution of different residues to DTI. Besides, binding free energy can also characterize DTA. However, popular methods (e.g., FEP, MM-GBSA and MMPBSA) are often too expensive to use in large-scale molecular screening. GNNs can both learn from existing cases and predict different forms of DTA with less computational costs. Metcalf et al proposed a variogram neural network, named DrΔG-Net, which combined DTI graphs and molecular dynamics (MD) databases. Significantly, it circumvented the use of expensive MD and FEP methods, thus allowing for the prediction of DTA for specific proteins. Min et al constructed a MD dataset of 3218 protein—ligand complexes, and developed a graph-based deep learning model, Dynaformer. The model predicted binding affinities by learning dynamic geometric features of protein—ligand interactions from MD trajectories files, results of MD that are used to reflect the movement of the complex. Utilizing the ability of GNNs to capture implicit information from

graph data can predict "how a drug molecule binds with target" and "how strong a drug molecule binds with target" by applying a same model. Zhang et al. constructed a multi-targeted GNN called PLANET, learning from the target protein in combination with the 3D representation of its binding pocket and the 2D chemical structure of the ligand. As a result, it can simultaneously predict the DTA and DTI. PLANET can achieve accuracy comparable to a traditional docking program, Glide, yet consuming less than 1% of Glide's computational time. Wang et al. developed a GNN model named IGModel, by incorporating geometric information of protein—ligand interactions. Through their studies, they were able to show that the integration of spatial data significantly enhanced the model's performance across multiple dimensions. IGModel can provide DTI and DTA at the same time. In some case studies, the root mean square errors (RMSEs; measurements for error between predictions and realities) between their predicted DTI and the true situation approached Lu et al constructed a complete working framework, DTIAM, by combining the models of small-molecule substructure learning GNN, protein sequence learning Transformer, and a DTI or DTA prediction module. DTIAM was not only able to predict DTI and DTA together, but also can help to elucidate drugs' mechanisms of action. With the help of DTIAM, Lu et al successfully obtained several potent compounds, including TMEM16A inhibitors, EGFR inhibitors and CDK4/6 inhibitors.

Data sets	Drugs	Targets	Used drug-targets pairs
Kiba	2,111	229	118,254
Davis	68	442	30,056
DTC	5,983	118	67,894
Metz	1,471	170	35,307
Tox-Cast	7,657	328	342,869

TABLE: Simple statistics for the sample information of five DTA dataset.

5.3. GNNs for LBVS:

Based on the QSAR, pharmacophores, and structural similarity of leads, GNNs can capture the hidden information of leads and map virtual molecules to realize LBVS. QSAR models

can predict the biological activity of new molecules without the necessity of experimental testing, thus greatly speeding up the drug discovery process and reducing costs. Wu et al developed a new generation of GNN models, called HRGCN+, by combining molecular graphs and molecular descriptors as inputs to a modified GNN. The model demonstrated excellent performance in 11 drug discovery datasets and was a novel computationally inexpensive, predictive, and noise-resistant QSAR modeling method. GNNs can also be used in LBVS based on pharmacophores or molecular structural similarity. PGMG, mentioned above, can generate molecules to meet the requirements of LBVS in some specific situations. Wang et al developed a GNN model named GeminiMol, it incorporated conformational space profiles into molecular representation learning, which could enhance the ability to capture complicated interplay between molecular structure and conformational space. With this ability, GeminiMol can realize matching tasks between molecules and pharmacophores or QSAR models.

6. GNNs for molecular property prediction

6.1. GNNs for physical property prediction

Aqueous solubility is a key physicochemical property for small molecules in drug discovery. Many studies have focused on solubility prediction over the past decades. Much software can predict water solubility, such as the ADMET calculator from Discovery Studio and QikProp or CIQikProp from Schro¨dingee. However, most of them have high RMSEs (logRMSE>1.0) on commonly used datasets. To address this problem, Deng et al constructed a new aqueous solubility dataset, containing a small dataset of 2609 compounds with only solubility records in the absence of salt at 25 ◦C. Based on the dataset, a GCN model named Multi-channel GCN was proposed. The model developed by a solubility regression integrated GCNs and two machine learning models. Using the modified GNN, the solubility prediction problem can be well solved. Ahmad et al developed SolPredictor by introducing residual information to GNN, enabling the model to capture implicit relationships between distant atoms, thus reducing the RMSE. Accurate and rapid pKa estimation is vital in drug discovery. Xiong et al developed the GAT model Graph-pKa by combining multi-example learning into GNN, performing well in the prediction of macroscopic pKa values. Similarly, An et al also used the GAT, but extended the prediction range by learning from the iBonD database and successfully realized the prediction of pKa of molecules in different solvents, including pure water, pure non-aqueous, and mixed solvents. Adding some QM properties to the graph data can contribute to the construction of a more accurate and superior model. Miao et structured

the GR-pKa model by using the five QM properties related to molecular thermodynamics and kinetics as the molecular characterization. By integrating these with molecular graph information, they achieved the accurate prediction of macroscopic pKa values. There are already some well-established computational platforms or methodologies for molecular property prediction studies. Pan et al developed MolGpKa based on GCN is a user-friendly server online (https://xundrug.cn/molgpka), for predicting pKa. Chemprop, a deep learning model developed by Heid et al, is highlighted as a user-friendly package to predict physicochemical properties, like aqueous solubility, pKa, etc. It brought the benefits of deep learning to GNNs, enabling the prediction of important drug properties such as octanol-water partition. Wong et al systematized and theorized the process of drug discovery using Chemprop, having built an explainable deep learning platform for molecular discovery. It enabled novel small molecule discovery of anticancer, antiviral and senolytic drugs in 1—2 weeks. The above models and their characteristics are summarized in Table.

6.2. GNNs for ADMET prediction

Pharmacokinetic properties, especially absorption, distribution, metabolism, excretion and toxicity (ADMET) properties, are crucial in drug design. To address these challenges, the introduction of neural networks for predicting pharmacokinetic parameters and pharmacological or toxicological properties has emerged as an advanced solution. Obrezanova et al developed a GCN model for the prediction of pharmacokinetic parameters in rats. For drug metabolism prediction, Du et al constructed a multimodal GNN model, named CMMS-GCL, to realize the prediction of drug hepatic clearance. Pu et al constructed a multimodal GNN model for the prediction of clearance, highlighting the application in chiral drugs in liver microsome. For drug toxicity prediction, Wang et al proposed MMGIN, a multimodal model for toxicity. Wang et al built a GCN model and Vinh et al built a GAT for the cardiotoxicity prediction. Xuan et al reported GCRS for side-effect and adverse reactions prediction. Cui et al developed Bontox by integrating models such as Transformer, SVM, XGBoost, and molecular maps with GNN, ViT, and pre-trained KPGT models.This is the first online predictive model for drug osteotoxicity. In DDI prediction field, Feng et al built a GNN, named SGRL-GNN, focusing on the pharmacological changes caused by DDIs. Unlike SGRL-GNN, Meta3D-DDI proposed by Lv et al predicted DDIs of novel drugs by learning molecular 3D graph features. MM-GANN-DDI proposed by Feng et al integrated six types of inputs of drug data and combined them with GAT to predict DDIs for novel drugs without prior knowledge. Similarly, Gao et al built a GNN model, named AutoDDI, by introducing

the reinforcement learning search algorithm. AutoDDI can build special networks for each molecule automatically and access drug substructures for DDI prediction. ADMET models on the basis of GNN frameworks have been generated for general prediction on drug molecules. FP-GNN developed by Cai et al and HiGNN developed by Zhu et al combined graphs with structure information to predict multiple properties. HSL-RG developed by Ju et al made predictions of the scarcity of molecules with desired properties. HimGNN developed by Han et al and TransFoxMol developed by Gao et al combined GNNs with transformers to increase the accuracy of prediction. Gu et al developed admetSAR3.0 and provided a user-friendly platform for searching, predicting, and optimizing compounds in one web.

7. DTA PREDICTIONS AND OTHER FACTORS

Predicting drug-target affinity (DTAP) is a crucial research topic in drug development, which can be used for discovering drug on-target and off-target effects. However, due to the increasing number of targets, it is difficult to fully validate the drug-target affinity (DTA) of drugs using biochemical experiments. In recent years, with the development of artificial intelligence technology, the use of computational methods for preliminary prediction of DTA has become an economically effective method. Graph neural networks (GNN) can extract features from graph-structured data and have been widely used in DTA, as drugs and targets are typically graph-structured data.

Different graph can enhance understanding of atomic connectivity and residue interactions. The strengths of attention mechanisms lie in their ability to focus on relevant parts of the graph, enhancing the model's representational power. Various deep learning frameworks offer diverse and innovative approaches for DTAP. Combining GNNs with other network architectures provides a comprehensive approach for DTAP, leveraging the strengths of different network types. Although the above deep neural networks based on GNNs can improve the performance of DTAP in various ways, most of the methods cannot address the issue of the scarcity of labelled training samples for DTA. There have been some methods that have noticed this problem. A pre-trained language model based on bidirectional encoder representations from transformers is designed to extract semantic features of SMILES molecules (Qiu et al., 2024). Multiple Transformer-Encoder blocks were designed to capture and learn the proteomics, chemical, and pharmacological contexts (Monteiro et al., 2024). Transformer-based architecture was utilized to learn representation for drugs (Rafiei et al., 2024). GCN-BERT utilized two RoBERTa models to extract features for the drug and target

(Lennox et al., 2021). CPCProt divided protein sequences into fxedsize segments and trained an autoregressor to distinguish subsequent segments of the same protein from random protein segments (Lu et al., 2020). SubMDTA proposed a self-supervised pretraining model based on substructure extraction and multi-scale features (Pan et al., 2023). ProtBert was utilized to extract the feature for the target (Zhang Xianfeng et al., 2023). Two modalities ProtBERT-BFD from ProtTrans2 and PSSM based descriptors are used to represent the target (Liyaqat et al., 2023). Four contrastive loss functions are considered to learn a more powerful model, such as Max-margin contrastive loss function, Triplet loss function, Multi-class N-pair Loss Objective, and NT-Xent loss function (Dehghan et al., 2024). These methods use pre-training to extract better features, but they fail to notice the significant difference between the pre-training objectives and samples, and the training objectives and samples for DTAP. Specifically, pre-training uses samples of drugs or targets individually, while DTAP utilizes samples of drug-target pairs for model training. To overcome the before-mentioned issues and further improve the DTAP performance of GNNs, this paper proposes a graph neural pre-training-based drug-target affinity (GNPDTA) prediction method. This approach divides the feature extraction for DTAP into two stages. In the first stage, a graph neural pre-training model is employed to extract low-level features of drugs and targets separately. During the process of drug-target affinity generation, we observe that drug-target affinity is generally related to their local fragments. Since target sequences tend to be longer, the pre-training model primarily extracts features of target fragments. Drug SMILES usually consist of 50 atoms, so the pre-training model focuses on extracting features of drug graph nodes. In the second stage, a convolutional neural network is utilized to combine the features of adjacent target fragments and drug graph nodes, resulting in features for predicting drug-target affinity. GNPDTA has the following main contributions. One is that GNPDTA can Minimize discrepancy between pre-training and DTAP objectives. By devoting the initial stage exclusively to feature extraction, the proposed method effectively bridges the gap between pre-training objectives (which focus on individual drug or target samples) and DTAP objectives (which consider drug-target pairs). This ensures that the extracted features are highly pertinent to the DTAP task. Another is that The GNPDTA method introduces a two-stage approach tailored specifically for drug-target affinity prediction (DTAP). This strategy intelligently leverages distinct models and training methodologies at each stage, maximizing the utilization of both unlabelled and labelled data to enhance feature extraction effectiveness.

LIMITATIONS:

GNNs, as an emerging deep learning paradigm, have gained widespread application across various domains, particularly in drug discovery and target prediction. Owing to their inherent advantage in handling non-Euclidean data such as molecular structures and protein interaction networks, GNNs have become indispensable tools in DTI prediction. Nevertheless, despite their promise, GNN-based methods face several significant limitations and challenges that hinder their practical deployment.

CONCLUSION

GNNs have emerged as a core technology in drug target prediction, owing to their unique ability to process non-Euclidean data, such as molecular graphs and biological networks. Due to the difficulties in accurately identifying and characterizing drug targets, the current drug discovery pipelines remain prohibitively costly, time-consuming and inefficient. Presently, we systematically analyze GNN architectures and their applications, providing a timely and comprehensive perspective that addresses this critical bottleneck and offering practical guidance for accelerating AI-driven drug discovery.

The importance of this study lies in filling key gaps left by prior works. Instead of treating GNNs as a monolithic tool, we establish a principled classification paradigm that dissects the design principles, mechanisms and applicable conditions of representative models such as GCNs, GATs and GAEs. Beyond summarization, we highlight how multimodal fusion, high-order graph reasoning and dynamic learning strategies can overcome real-world challenges such as data scarcity, limited interpretability and the complexity of polypharmacology.

For medicinal chemists and pharmaceutical industries intending to use GNNs in drug discovery, greater attention should be paid to the latest improved network architectures, such as signed attention, to explore their potential role in drug discovery. In the realm of industry application, greater emphasis may be placed on integrating different models together to develop mature working pipelines.

REFERENCES:

1. Deng JY, Yang ZB, Wang HH, Ojima I, Samaras D, Wang F. A systematic study of key elements underlying molecular property prediction. Nat Commun., 2023; 14: 6395.
2. Liyaqat T, Ahmad T, Saxena C. Advancements in molecular property prediction: a survey of single and multimodal approaches. Arch Comput Methods Eng 2024. Available from: https://doi.org/10.1007/ s11831-025-10317-5.
3. Stanley M, Segler M. Fake it until you make it? Generative de novo design and virtual screening of synthesizable molecules. Curr Opin Struct Biol., 2023; 82: 102658.
4. Walters WP, Murcko M. Assessing the impact of generative AI on medicinal chemistry. Nat Biotechnol., 2020; 38: 143—5.
5. Reymond JL. The chemical space project. Acc Chem Res., 2015; 48: 722—30.
6. Sadybekov AA, Sadybekov AV, Liu Y, Iliopoulos-Tsoutsouvas C, Huang XP, Pickett J, et al. Synthon-based ligand discovery in virtual libraries of over 11 billion compounds. Nature., 2022; 601: 452—9.
7. Schneider G, Fechner U. Computer-based de novo design of druglike molecules. Nat Rev Drug Discov., 2005; 4: 649—63.
8. Bickerton GR, Paolini GV, Besnard J, Muresan S, Hopkins AL. Quantifying the chemical beauty of drugs. Nat Chem., 2012; 4: 90—8.
9. Hajduk PJ, Greer J. A decade of fragment-based drug design: strategic advances and lessons learned. Nat Rev Drug Discov., 2007; 6: 211—9.
10. Paggi JM, Pandit A, Dror RO. The art and science of molecular docking. Annu Rev Biochem., 2024; 93: 389—410.
11. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with alphafold. Nature., 2021; 596: 583—9.
12. He XH, Li JR, Shen SY, Xu HE. Alphafold3 versus experimental structures: assessment of the accuracy in ligand-bound G proteincoupled receptors. Acta Pharmacol Sin., 2025; 46: 1111—22.
13. Wang HW. Prediction of protein—ligand binding affinity via deep learning models. Brief Bioinform., 2024; 25: bbae081.
14. Li ZP, Zeng YN, Jiang MF, Wei B. Deep drug—target binding affinity prediction base on multiple feature extraction and fusion. ACS Omega, 2025; 10: 2020—32.
15. Sliwoski G, Kothiwale S, Meiler J, Lowe Jr EW. Computational methods in drug discovery. Pharmacol Rev., 2014; 66: 334—95.

16. Dai JY, Zhou ZY, Zhao YR, Kong FJ, Zhai ZW, Zhu ZS, et al. Combined usage of ligand- and structure-based virtual screening in the artificial intelligence era. Eur J Med Chem., 2025; 283: 117162.
17. Valverde JR. Molecular modelling: principles and applications. Brief Bioinform., 2001; 2: 199—200.
18. Rong Y, Bian YT, Xu TY, Xie WY, Wei Y, Huang WB, et al. Self-supervised graph transformer on large-scale molecular data. NeurIPS., 2020: 12559—71.
19. Scarselli F, Tsoi AC, Gori M, Hagenbuchner M. Graphical-based learning environments for pattern recognition. In: Structural, syntactic, and statistical pattern recognition. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004.
20. Simon G, Aliferis C. Data design in biomedical AI/ML. In: Simon GJ, Ceditors Aliferis, editors. Artificial intelligence and machine learning in health care and medical sciences: best practices and pitfalls. Cham (CH): Springer, 2024; 341—75.
21. Zhou J, Cui GQ, Hu SD, Zhang ZY, Yang C, Liu ZY, et al. Graph neural networks: a review of methods and applications. AI Open, 2020; 1: 57—81. 6174 Rui Wang, Chunlin Zhuang
22. Wu Z, Pan S, Chen F, Long G, Zhang C, Yu PS. A comprehensive survey on graph neural networks. IEEE Trans Neural Netw Learn Syst., 2021; 32: 4—24.
23. Garg V. Generative AI for graph-based drug design: recent advances and the way forward. Curr Opin Struct Biol., 2024; 84: 102769.
24. Xie XY, Gui L, Qiao BX, Wang GH, Huang S, Zhao YM, et al. Deep learning in template-free de novo biosynthetic pathway design of natural products. Brief Bioinform., 2024; 25: bbae495—501.
25. Lu HR, Wang L, Ma XL, Cheng J, Zhou MC. A survey of graph neural networks and their industrial applications. Neurocomputing, 2025; 614: 128761.
26. Rui Wang, Chunlin Zhuang. Graph neural networks driven acceleration in drug discovery
27. Acta Pharmaceutica Sinica B,Volume 15, Issue 12,2025, Pages 6163-6177, ISSN 2211-3835, https://doi.org/10.1016/j.apsb.2025.10.011.
28. Popova M, Isayev O, Tropsha A. Deep reinforcement learning for de novo drug design, 2018; 4: 7885.
29. Dwivedi VP, Joshi CK, Luu AT, Laurent T, Bengio Y, Bresson XJAEP. Benchmarking graph neural networks. J Mach Learn Res., 2022; 24: 1730—77.
30. Xiong JC, Xiong ZP, Chen KX, Jiang HL, Zheng MY. Graph neural networks for automated de novo drug design. Drug Discov Today 021; 26: 1382—93.

31. PangC, Qiao JB, Zeng XX, Zou Q,Wei LY. Deep generativemodelsin de novo drug molecule generation. J Chem Inf Model, 2024; 64: 2174—94.
32. Mercado R, Rastemo T, Lindelo¨f E, Klambauer G, Engkvist O, Chen H, et al. Graph networks for molecular design. Mach Learn Sci Technol., 2021; 2: 025023.
33. Atance SR, Diez JV, Engkvist O, Olsson S, Mercado R. De novo drug design using reinforcement learning with graph-based deep generative models. J Chem Inf Model, 2022; 62: 4863—72.
34. Nori D, Coley CW, Mercado RJae-p. De novo protac design using graph-based deep generative models. NeurIPS 2022. Available from: https://doi.org/10.48550/arXiv.2211.02660.
35. Yue TL, Tao L, Varshney V, Li Y. Benchmarking study of deep generative models for inverse polymer design. Digit Discov., 2025; 4: 910—26.
36. Zhang H, Huang JC, Xie JJ, Huang WF, Yang YD, Xu MY, et al.Grelinker: a graph-based generative model for molecular linker design with reinforcement and curriculum learning. J Chem Inf Model, 2024; 64: 666—76.
37. Fang Y, Pan X, Shen HB. De novo drug design by iterative multiobjective deep reinforcement learning with graph-based molecular quality assessment. Bioinformatics, 2023; 39: btad157.
38. Zhu HM, Zhou RY, Cao DS, Tang J, Li M. A pharmacophore-guided deep learning approach for bioactive molecular generation. Nat Commun., 2023; 14: 6234.
39. Atz K, Cotos L, Isert C, Ha˚kansson M, Focht D, Hilleke M, et al. Prospective de novo drug design with deep interactome learning. Nat Commun., 2024; 15: 3408.
40. Seneci P. Comprehensive medicinal chemistry ii. Oxford: Elsevier, 2007.
41. Xie YT, Shi CC, Zhou H, Yang YW, Zhang WN, Yu Y, et al. Mars: markov molecular sampling for multi-objective drug discovery. ICLR 2021. Available from: https://doi.org/10.48550/arXiv.2103.10432.
42. Zheng SJ, Lei ZR, Ai HT, Chen HM, Deng DG, Yang YD. Deep scaffold hopping with multimodal transformer neural networks. J Cheminf., 2021; 13: 87.
43. Hu C, Li S, Yang CX, Chen J, Xiong Y, Fan GS, et al. Scaffoldgvae: scaffold generation and hopping of drug molecules via a variational autoencoder based on multi-view graph neural networks. J Cheminf., 2023; 15: 91.
44. Chen JH, Lin AQ, Luo P. Advancing pharmaceutical research: a comprehensive eview of cutting-edge tools and technologies. Cur Pharma Anal., 2024; 21: 1—19.

45. Cui C, Ding XY, Wang DY, Chen LF, Xiao F, Xu TY, et al. Drug repurposing against breast cancer by integrating drug-exposure expression profiles and drug—drug links based on graph neural network. Bioinformatics, 2021; 37: 2930—7.
46. Meng YJ, Lu CC, Jin M, Xu JL, Zeng XX, Yang JL. A weighted bilinear neural collaborative filtering approach for drug repositioning. Brief Bioinform., 2022; 23: bbab581.
47. Xiong ZK, Huang F, Wang ZY, Liu SC, Zhang W. A multimodal framework for improving in silico drug repositioning with the prior knowledge from knowledge graphs. IEEE/ACM Trans Comput Biol Bioinform., 2022; 19: 2623—31.
48. Huang KX, Chandak P, Wang QW, Havaldar S, Vaid A, Leskovec J, et al. A foundation model for clinician-centered drug repurposing. Nat Med., 2024; 30: 3601—13.
49. Tayebi J, BabaAli B. EKGDR: an end-to-end knowledge graph-based method for computational drug repurposing. J Chem Inf Model, 2024; 64: 1868—81.
50. Shao KH, Zhang YH, Wen YQ, Zhang ZN, He S, Bo XC. Dti-HETA: prediction of drug—target interactions based on gcn and gat on heterogeneous graph. Brief Bioinform., 2022; 23: bbac109.