DrugReason: Dynamic Multi-View Reasoning over Knowledge Graph and Language Evidence for Drug Repurposing

Zijie Liu[1,*] **Hongxuan Li**[1,*] **Zhen Tan**[2] **Jinhao Duan**[1]
Baixiang Huang[3] **Zunpeng Liu**[4] **Kai Shu**[3] **Tianlong Chen**[1,†]
[1]University of North Carolina at Chapel Hill
[2]Stevens Institute of Technology
[3]Emory University
[4]Massachusetts Institute of Technology
[*]Equal contribution.
[†]Corresponding author: `tianlong@cs.unc.edu`

Abstract

Drug repurposing aims to identify new therapeutic uses for existing compounds and, compared with de novo drug discovery, offers a faster and more cost-effective path to clinical translation. However, the space of candidate drug–disease pairs is enormous and their underlying relationships often depend on complex multi-hop biological mechanisms, making it difficult to reliably predict which pairs represent true therapeutic relationships. Existing approaches tackle this from two directions: knowledge graph-based methods organize curated biomedical evidence into structured relational networks for grounded multi-hop reasoning, while LLM-based methods leverage pretrained knowledge to generate flexible mechanistic rationales. Yet neither is sufficient alone — KGs are confined to observed graph structure while LLMs lack factual grounding and risk hallucination. To address this gap, we propose DrugReason, a multi-view reasoning framework that integrates grounded KG reasoning with LLM-generated mechanistic inference for drug repurposing. DrugReason adaptively routes diverse reasoning paths to specialized experts conditioned on the query context, while a cross-expert distillation objective enables knowledge sharing without sacrificing expert specialization. Experiments on PharmaDB, DDInter, and DrugBank show that DRUGREASON improves average performance over strong single-view reasoning baselines and achieves competitive or superior results compared with graph-based alternatives, while providing interpretable routing-based predictions.

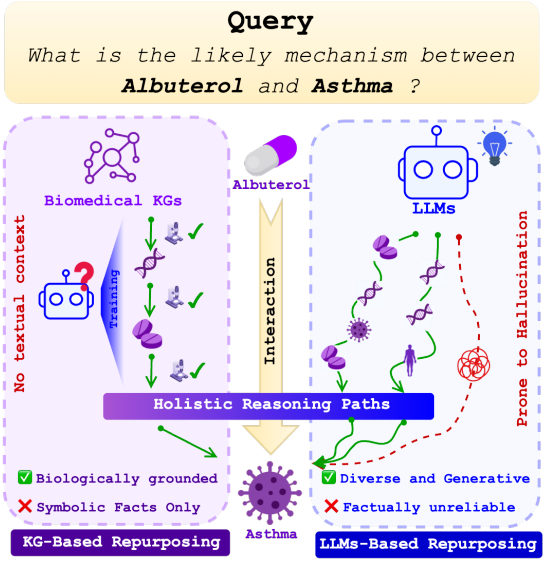


Figure 1: **Motivation for evidence-adaptive multi-view reasoning.** KG-based methods provide grounded biomedical evidence through curated relational structure, while LLM-based methods generate flexible mechanistic hypotheses but may hallucinate without factual grounding. DrugReason combines KG-derived and LLM-generated reasoning views and adaptively routes them according to the query context for drug repurposing.

1 Introduction

Drug repurposing aims to identify new therapeutic uses for existing compounds, offering a faster and more cost-effective alternative to de novo drug discovery (Pushpakom et al., 2019; Hopkins, 2008). Despite its promise, computational drug repurposing remains challenging because candidate drug–disease relationships often depend on complex biological mechanisms involving drugs, genes, proteins, pathways, and diseases (Hopkins, 2008; Himmelstein et al., 2017a; Chandak et al., 2023). These mechanisms are rarely captured by a single direct association; instead, they often require reasoning over diverse biomedical evidence and multi-hop biological connections. In this paper, we study biomedical relation prediction over drug-involving entity pairs, encompassing both drug–disease therapeutic associations and drug–drug interactions.

Existing approaches address this problem from two directions. Knowledge graph (KG)-based methods organize curated biomedical entities and relations into structured networks, enabling grounded reasoning over multi-hop paths between drugs, diseases, genes, and pathways (Himmel-

Proceedings of the 2026 Conference on Empirical Methods in Natural Language Processing, pages 17254–17266
October 24-29, 2026 ©2026 Association for Computational Linguistics

stein et al., 2017a; Chandak et al., 2023; Perdomo-Quinteiro and Belmonte-Hernández, 2024a). Such graph structure provides interpretable evidence for repurposing relations, especially when relevant biological routes are explicitly represented. In parallel, large language model (LLM)-based methods leverage pretrained knowledge to generate natural-language mechanistic rationales, such as causal chains, pathway-level explanations and alternative biological hypotheses (Singhal et al., 2023a; Chen et al., 2025a; Liu et al., 2025a).

However, neither direction is sufficient alone. KG-based methods are constrained by observed graph structure and relation coverage, while LLM-based methods may generate unsupported or hallucinated mechanisms without factual grounding (Ji et al., 2023). Recent work has therefore explored grounding LLMs with biomedical knowledge, retrieval, or graph-derived context (Soman et al., 2024a; Cabello et al., 2025; Li et al., 2024; Sohn et al., 2025; Abdullahi et al., 2025a; Wei et al., 2024a). Yet combining KG and LLM evidence does not resolve the core challenge. Different reasoning paths encode different types of information: KG paths expose structured graph evidence, while CoT- and ToT-style rationales elicit sequential or hypothesis-level reasoning (Abdullahi et al., 2025a; Wei et al., 2022; Yao et al., 2023). Static fusion strategies such as concatenation, voting, or fixed aggregation may not fully use such heterogeneous path-specific signals, a limitation also observed in multi-view and multimodal fusion settings (Sahu and Vechtomova, 2021; Han et al., 2022). This motivates an adaptive routing mechanism over diverse reasoning views.

To this end, we propose DRUGREASON, a multi-view reasoning framework for drug repurposing. For each candidate pair, DRUGREASON combines KG-derived evidence with LLM-generated mechanistic inference to construct multiple reasoning views. Rather than concatenating or averaging these views, DRUGREASON adaptively routes each view to a specialized expert based on the query and evidence content. Cross-expert distillation further shares transferable biomedical signals across routed experts while preserving their specialization. Experiments on PharmaDB, DDInter, and DrugBank show that DrugReason improves average performance over strong single-view reasoning baselines and achieves competitive or superior results compared with graph-based and static multi-view alternatives. Ablation studies show that adaptive routing and cross-expert mutual distillation both contribute to the overall gains, while routing distributions provide evidence-level transparency into how different reasoning views support each prediction. Our main contributions are as follows:

- We introduce a multi-view reasoning formulation for drug repurposing that integrates KG-derived evidence and LLM-generated mechanistic rationales across drug–disease and drug–drug relation types.
- We propose DRUGREASON, which adaptively routes heterogeneous reasoning views to two specialized generative experts with cross-expert mutual distillation for knowledge sharing without collapsing specialization.
- Experiments on PharmaDB, DDInter, and DrugBank show strong improvements over single-view and static-fusion baselines, with ablations confirming the contribution of both routing and distillation.

2 Related Work

Knowledge-Grounded Reasoning for Drug Discovery. Knowledge graph (KG)-based methods support drug repurposing by organizing curated biomedical entities and relations into structured networks (Himmelstein et al., 2017b,c; Perdomo-Quinteiro and Belmonte-Hernández, 2024b). Path-based methods provide interpretable drug–disease mechanisms (Abdullahi et al., 2025b; Guney et al., 2016), while GNN-based methods learn expressive entity representations but are often less transparent and remain bounded by observed graph structure (Gharizadeh et al., 2024; Liu et al., 2024a; Huang et al., 2024). Recent KG-to-LLM approaches, including K-Paths, DrugReAlign, and graph2prompt, verbalize KG paths as prompts to improve factual grounding (Abdullahi et al., 2025b; Wei et al., 2024b; Soman et al., 2024b). However, most of them use KG evidence as static context rather than adaptively deciding how different evidence views should be interpreted for each query.

LLM Biomedical Reasoning. Large language models and biomedical foundation models have shown strong potential on medical and biomedical reasoning tasks (Singhal et al., 2023b; Chen et al., 2025b; Luo et al., 2022; Mackenzie et al., 2025; Liu et al., 2025b). In drug discovery, systems such as DrugAgent, MedReason, Tx-LLM,

DrugMCTS, and PharmAgents combine domain knowledge, external tools, or agent collaboration to support interpretable therapeutic prediction (Liu et al., 2024b; Wu et al., 2025; Chaves et al., 2024; Yang et al., 2025; Gao et al., 2025). These methods broaden mechanistic reasoning, but generated rationales can still be unreliable without structured grounding and evidence-specific control.

Diverse LLM Reasoning Strategies. Reasoning prompts such as Chain-of-Thought and Tree-of-Thought encourage LLMs to decompose complex questions and explore alternative hypotheses (Wei et al., 2022; Yao et al., 2023). Such strategies increase reasoning diversity, including in biomedical hypothesis generation, but they typically treat generated paths uniformly. In contrast, DRUGREASON treats KG-derived and LLM-generated rationales as heterogeneous reasoning views and routes each view to a specialized expert.

3 Method

3.1 Problem Formulation and Overview

Given a candidate pair (u, v), where u is a drug and v is either a disease or another drug depending on the benchmark, the goal is to predict a relation label $y \in \mathcal{Y}$. For drug–disease pairs, y denotes a therapeutic association label; for drug–drug pairs, y denotes a drug–drug relation label. Since our experts are generative LLMs, each label y is represented by a canonical label verbalizer $\nu(y)$, i.e., the text sequence that the model should generate for that relation label.

For each pair, DRUGREASON constructs a set of reasoning views

$$\mathcal{R}(u, v) = \{r_1, \ldots, r_m\}, \tag{1}$$

where each r_j is a serialized textual evidence instance generated from one reasoning strategy. For KG-derived views, r_j is an entity definition or a verbalized graph path; for LLM-generated views, r_j is a mechanistic rationale generated under a specific prompt. Thus, all evidence sources are represented as text while preserving their different reasoning roles.

DRUGREASON proceeds in three stages. First, it constructs $\mathcal{R}(u, v)$ by combining KG-derived evidence with LLM-generated mechanistic inference. Second, each reasoning view is serialized with the candidate query and routed to one of two generative LLM experts. The selected expert predicts a relation label for that view. Third, routed view-level predictions are aggregated into the final pair-level prediction. During training, we further apply symmetric mutual distillation between the two LLM experts so that they can exchange softened predictive signals without removing routed specialization.

3.2 Reasoning View Construction

For each candidate pair (u, v), DRUGREASON constructs reasoning views that describe the same query from different evidence sources. We use the term *reasoning view* to denote a serialized textual evidence instance generated from one reasoning strategy. KG-derived views provide structured grounding from curated biomedical resources, while LLM-generated views provide mechanistic inference in natural language. All views are serialized into a unified text format so that they can be routed and processed under the same framework.

KG-derived views. The definition view provides semantic grounding for the entities in the query. We retrieve biomedical definitions for the drug and disease/drug entities from PrimeKG (Chandak et al., 2023). These definitions help identify entity type, biological role, and pharmacological context before relation prediction.

The KG-path view provides graph-grounded multi-hop evidence between the two entities. We extract relational paths from the biomedical KG using K-Paths (Abdullahi et al., 2025a), which retrieves concise and diverse loopless paths between query entities. A path is represented as

$$P_j = (u, \rho_1, v_1, \rho_2, \ldots, \rho_l, v), \quad l \leq l_{\max}, \tag{2}$$

where ρ_i denotes a KG relation and v_i denotes an intermediate biomedical entity. Each extracted path is verbalized into text and used as one KG-derived reasoning view. We denote the KG-derived view set as $\mathcal{R}^{\mathrm{KG}}(u, v)$.

LLM-generated views. KG-derived views are grounded in curated graph structure, but they may not fully express the mechanism behind a candidate relation. During training, we use a teacher LLM to construct additional mechanistic reasoning views under four prompting strategies. For LLM-generated views, we use four prompting strategies that capture different forms of mechanistic inference: chain-of-thought reasoning describes a sequential causal explanation connecting the candidate pair; tree-of-thought reasoning explores and compares multiple biological hypotheses; forward–backward validation reasons from u

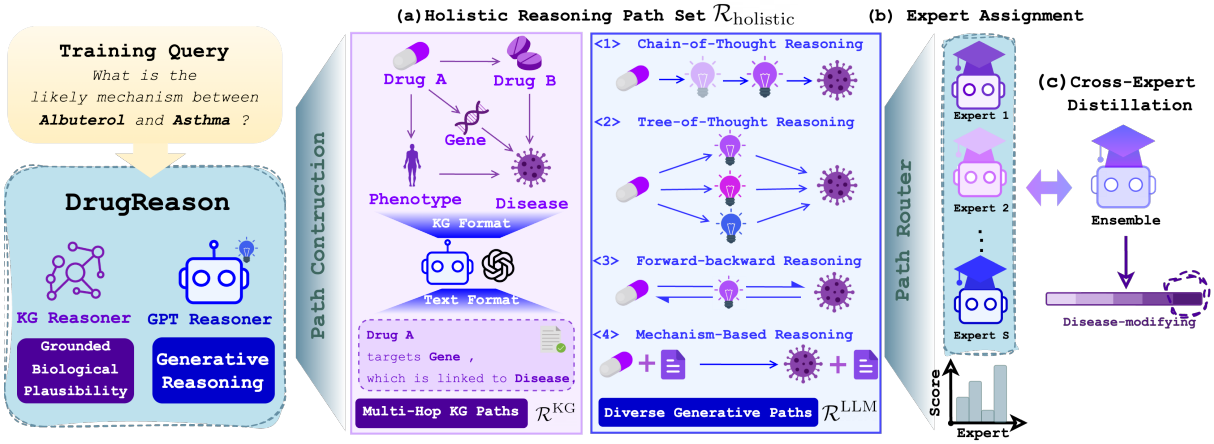


Figure 2: **Overview of DRUGREASON.** Given a candidate pair, DRUGREASON constructs multiple reasoning views from KG-derived evidence and LLM-generated mechanistic inference. Each view is serialized with the query context and encoded for adaptive routing. A conditional router assigns each view to one of two generative LLM experts, Llama or Qwen, which predicts a relation label from the routed view. During training, cross-expert mutual distillation exchanges softened label distributions between the two routed experts while preserving their specialization. At inference, routed view-level predictions are aggregated to produce the final relation label.

to v and from v back to u to check directional consistency; and mechanism-focused reasoning emphasizes molecular targets, pathways, pharmacological processes, and therapeutic mechanisms. These teacher-generated views are used only as training-time evidence for routing and expert learning. DRUGREASON does not query the teacher LLM at inference. The prompt templates and filtering procedure are provided in Appendix A.

The LLM-generated views form $\mathcal{R}^{\text{LLM}}(u, v)$. The full reasoning-view set is

$$\mathcal{R}(u, v) = \mathcal{R}^{\text{KG}}(u, v) \cup \mathcal{R}^{\text{LLM}}(u, v). \quad (3)$$

View serialization. Each reasoning view is serialized with the candidate query into a unified input:

$$x_j = [q(u, v); r_j], \quad (4)$$

where $q(u, v)$ is a short textual description of the candidate pair and r_j is the corresponding reasoning view. We apply only deterministic formatting checks to remove empty outputs, malformed responses, and template artifacts such as explicit label fields. This step does not use human annotation and does not rewrite retained views.

3.3 Adaptive Routing to Generative LLM Experts

The reasoning views in $\mathcal{R}(u, v)$ encode different types of information and have different textual formats. A KG path is concise and relation-centered, while a mechanism-focused rationale is longer and narrative. Instead of concatenating all views or sending every view to the same expert, DRUGREASON first routes each reasoning view to a specialized generative LLM expert and then aggregates the routed label distributions. In our implementation, the expert set is

$$\mathcal{E} = \{E^{(1)}, E^{(2)}\}, \quad (5)$$

where $E^{(1)}$ and $E^{(2)}$ correspond to the Llama and Qwen experts, respectively. The two experts remain separate throughout training and inference; routing determines which expert scores each view, while mutual distillation later exchanges only soft predictive distributions. Routing induces specialization by allowing each expert to learn from the views assigned to it.

For each serialized reasoning view $x_j = [q(u, v); r_j]$, we compute a routing representation

$$h_j = \text{Enc}(x_j) \in \mathbb{R}^d, \quad (6)$$

where $\text{Enc}(\cdot)$ is a lightweight text encoder used only for routing. The router maps h_j to expert logits:

$$z_j = \text{MLP}(h_j) \in \mathbb{R}^S, \quad (7)$$

where $S = 2$ in our implementation. We use a straight-through Gumbel–Softmax router:

$$\begin{aligned} \tilde{\boldsymbol{\alpha}}_j &= \text{softmax}\left((z_j + \mathbf{g})/\tau\right), \\ g^{(s)} &\sim \text{Gumbel}(0, 1), \end{aligned} \quad (8)$$

and take the hard assignment

$$\boldsymbol{\alpha}_j = \mathbf{e}_{\arg\max_{s \in [S]} \tilde{\alpha}_j^{(s)}}. \quad (9)$$

Here, $\boldsymbol{\alpha}_j$ is one-hot in the forward pass, while gradients are propagated through the soft routing probabilities $\tilde{\boldsymbol{\alpha}}_j$.

Because the experts are generative LLMs, each expert predicts a relation label by scoring canonical label verbalizers. Let $\nu(y) = (w_1, \ldots, w_{|\nu(y)|})$ denote the token sequence for label y. For expert $E^{(s)}$, we compute the generative score of label y under input x_j as

$$a_{j,y}^{(s)} = \frac{1}{|\nu(y)|} \sum_{t=1}^{|\nu(y)|} \log P_{E^{(s)}}(w_t \mid x_j, w_{<t}), \tag{10}$$

where length normalization avoids favoring shorter label strings. The view-level label distribution of expert s is

$$p_j^{(s)}(y) = \frac{\exp(a_{j,y}^{(s)})}{\sum_{y' \in \mathcal{Y}} \exp(a_{j,y'}^{(s)})}. \tag{11}$$

The routed view-level prediction is

$$p_j = \sum_{s=1}^{S} \alpha_j^{(s)} p_j^{(s)}. \tag{12}$$

The pair-level prediction aggregates the routed predictions across all reasoning views:

$$\hat{p} = \frac{1}{|\mathcal{R}(u,v)|} \sum_{j=1}^{|\mathcal{R}(u,v)|} p_j, \quad \hat{y} = \arg\max_{y \in \mathcal{Y}} \hat{p}_y. \tag{13}$$

3.4 Cross-Expert Mutual Distillation

Adaptive routing encourages the two LLM experts to specialize on different reasoning views. In the first training stage, the router assigns reasoning views to experts, allowing each model to learn from the subset of views routed to it. However, this routed specialization can also separate useful information across experts. For example, one expert may learn stronger patterns from graph-grounded views, while the other may better capture signals from generated mechanistic rationales.

To reduce this separation without removing specialization, we apply symmetric mutual distillation in the second training stage. During this stage, both experts are evaluated on each serialized view to compute the mutual distillation loss, while the supervised prediction loss remains routed according to $\boldsymbol{\alpha}_j$. For each expert, we first compute temperature-scaled label distributions from the generative label scores:

$$p_{j,T}^{(s)}(y) = \frac{\exp(a_{j,y}^{(s)}/T)}{\sum_{y' \in \mathcal{Y}} \exp(a_{j,y'}^{(s)}/T)}, \tag{14}$$

where T is the distillation temperature.

We use a symmetric KL objective so that the two experts teach each other:

$$\begin{aligned} \mathcal{L}_{\text{mutual}} = \frac{T^2}{2|\mathcal{R}|} \sum_{j=1}^{|\mathcal{R}(u,v)|} \Big[& \\ & \mathrm{KL}\Big(\mathrm{stopgrad}(p_{j,T}^{(1)}) \,\|\, p_{j,T}^{(2)}\Big) \\ & + \mathrm{KL}\Big(\mathrm{stopgrad}(p_{j,T}^{(2)}) \,\|\, p_{j,T}^{(1)}\Big)\Big]. \end{aligned} \tag{15}$$

The stop-gradient operator prevents the target distribution in each KL term from being updated by the same term. Since the objective is bidirectional, each expert serves as both a soft teacher and a student through the two KL terms. This mutual distillation exchanges softened predictive information between the Llama and Qwen experts without sharing parameters. Because routed supervised learning is still maintained, the experts preserve their specialization while benefiting from each other's predictive distributions.

3.5 Training Objective and Inference

We train DRUGREASON in two stages. In Stage 1, the router assigns reasoning views to the two LLM experts, and each expert learns from the views routed to it. For a serialized reasoning view x_j, let $\alpha_j^{(s)} \in \{0, 1\}$ denote the hard routing assignment to expert s. The routed supervised generation loss is

$$\begin{aligned} \mathcal{L}_{\text{pred}} = -\frac{1}{|\mathcal{R}|} \sum_{j=1}^{|\mathcal{R}|} \sum_{s=1}^{2} \alpha_j^{(s)} \frac{1}{|\nu(y)|} \\ \sum_{t=1}^{|\nu(y)|} \log P_{E^{(s)}}(w_t \mid x_j, w_{<t}). \end{aligned} \tag{16}$$

Equivalently, this loss trains only the selected expert to generate the canonical label verbalizer for each routed reasoning view.

To avoid expert collapse while allowing confident view-level assignments, we regularize the

average routing distribution:

$$\bar{\alpha}^{(s)} = \frac{1}{|\mathcal{R}|} \sum_{j=1}^{|\mathcal{R}|} \tilde{\alpha}_j^{(s)}, \qquad \mathcal{H}_{\text{load}} = -\sum_{s=1}^{2} \bar{\alpha}^{(s)} \log \bar{\alpha}^{(s)}. \tag{17}$$

The Stage 1 objective is

$$\mathcal{L}^{(1)} = \mathcal{L}_{\text{pred}} - \lambda_1 \mathcal{H}_{\text{load}}. \tag{18}$$

In Stage 2, we keep the routed supervised generation objective and add symmetric mutual distillation between the two LLM experts:

$$\mathcal{L}^{(2)} = \mathcal{L}_{\text{pred}} - \lambda_1 \mathcal{H}_{\text{load}} + \lambda_2 \mathcal{L}_{\text{mutual}}. \tag{19}$$

where λ_1 controls routing entropy regularization and λ_2 controls mutual distillation. The routed supervised loss preserves expert specialization, while the mutual distillation term allows the two experts to exchange soft predictive information.

At inference time, each reasoning view is encoded for routing, assigned to one expert, and then evaluated by the selected generative LLM expert. We use the selected expert's canonical label-verbalizer scores to obtain p_j, aggregate routed view-level distributions across all reasoning views, and output the final label:

$$\hat{p} = \frac{1}{|\mathcal{R}(u,v)|} \sum_{j=1}^{|\mathcal{R}(u,v)|} \sum_{s=1}^{2} \alpha_j^{(s)} p_j^{(s)}, \qquad \hat{y} = \arg\max_{y \in \mathcal{Y}} \hat{p}_y. \tag{20}$$

The mutual distillation loss and routing entropy regularizer are used only during training.

4 Experiments

Experimental Setup. We evaluate DRUGREASON on three drug-repurposing-oriented biomedical relation prediction benchmarks: DDInter (Xiong et al., 2022) and DrugBank (Wishart et al., 2018) for drug–drug interaction prediction, and PharmaDB (Himmelstein et al., 2017a) for drug–disease therapeutic association prediction. All experiments follow an inductive protocol over Hetionet (Himmelstein et al., 2017b), where test triples contain entities unseen during task-specific training. We report accuracy as the primary metric and Macro-F1 for additional control analyses.

We use `GPT-4o` (OpenAI, 2024) only as a training-time teacher to generate mechanistic rationales. The two generative experts are `Llama3.2-3B-Instruct` (Dubey et al., 2024) and `Qwen2.5-3B-Instruct` (Yang et al., 2024), with a lightweight `RoBERTa-base` encoder (Liu et al., 2019) used only for routing serialized reasoning views. We report expert-specific results after routed training and mutual distillation, while both experts remain part of the same DRUGREASON framework.

We compare with no-reasoning prompting, single-view reasoning baselines including K-Path (Abdullahi et al., 2025b), Chain-of-Thought (Wei et al., 2022), and Tree-of-Thought (Yao et al., 2023), static fusion controls, KGE baselines including TransE (Bordes et al., 2013) and DistMult (Yang et al., 2014), and graph/transformer baselines including GCN (Kipf and Welling, 2017), GAT (Veličković et al., 2018), GIN (Xu et al., 2019), GraphSAGE (Hamilton et al., 2017), Graph Transformer (Dwivedi and Bresson, 2020), and Transformer (Vaswani et al., 2017). Additional implementation details and hyperparameters are provided in Appendix B.

RQ1: Why do we need adaptive multi-view reasoning? Table 1 compares DRUGREASON with no-reasoning prompting and individual KG- or LLM-based reasoning views. Entity names alone are insufficient: on DDInter, no-reasoning prompting achieves only 0.20% with `Llama3.2` and 3.34% with `Qwen2.5`, while DRUGREASON reaches 60.00% and 72.00%. On PharmaDB, DRUGREASON also improves from 43.25% to 68.25% with `Llama3.2` and from 45.63% to 65.48% with `Qwen2.5`.

Single-view reasoning helps but is uneven. KG-derived views can be strong when graph evidence aligns with the target relation, e.g., Definition reaches 63.40% on DDInter with `Qwen2.5`, and K-Path reaches 41.86% on DrugBank with `Llama3.2`. However, no single view is reliable across all datasets and backbones, while LLM-only views are often unstable without structured grounding. By adaptively routing and aggregating heterogeneous views, DRUGREASON obtains the best result in five of six dataset–expert settings. The exception is DrugBank with `Llama3.2`, where K-Path remains stronger, suggesting that fine-grained DDI labels can sometimes favor direct graph-path evidence.

RQ2: Is naive fusion enough? To rule out the possibility that DRUGREASON improves sim-

Category	Prompt Type	Llama3.2			Qwen2.5		
		DDInter	DrugBank	PharmaDB	DDInter	DrugBank	PharmaDB
No Reasoning	-	0.20	17.30	43.25	3.34	19.80	45.63
KG-only Reasoning	k-Path	33.30	**41.86**	49.21	18.20	3.94	9.92
	Definition	35.40	36.01	56.75	63.40	16.55	58.33
LLM-only Reasoning	Mechanism	1.50	0.81	44.44	20.10	6.18	13.10
	Chain of Thought	20.40	14.42	45.24	11.80	1.55	0.79
	Forward-Backward	–	15.91	53.17	20.30	0.85	1.59
	Tree of Thought	26.00	13.48	45.63	24.30	14.77	–
Our Method	-	**60.00**	25.99	**68.25**	**72.00**	**33.98**	**65.48**

Table 1: **Main comparison across reasoning strategies.** KG-only and LLM-only views provide useful but uneven gains, while DRUGREASON achieves the best average performance across datasets and backbones by adaptively combining heterogeneous reasoning views.

Dataset	Best Single	Best Static	Best KGE
DDInter	34.00 / 23.36	20.40 / 8.61	43.00 / 30.58
PharmaDB	49.60 / 37.63	46.43 / 35.44	37.70 / 36.43

Table 2: **Control baselines for static fusion and graph embeddings.** Each cell reports Acc./Macro-F1 (%). Best Single denotes the best individual prompt; Best Static is selected by accuracy among concat-all, majority vote, and mean vote; Best KGE is selected by accuracy between TranSE and DistMult.

Type	Model	DDInter	Phar.
GNN	GAT	43.20	67.11
	GCN	41.60	65.35
	GIN	25.40	47.81
	GraphSAGE	60.70	62.72
	GraphTrans.	71.00	61.00
General	Transformer	66.00	65.79
Ours	Llama3.2	60.00	**68.25**
	Qwen2.5	**72.00**	65.48

Table 3: **Comparison with graph-based and transformer baselines.** Results are reported as accuracy (%). The strongest DRUGREASON configuration is competitive with or better than graph and transformer baselines while producing language-level reasoning evidence.

ply by using more evidence, we compare against prompt-based static fusion (concat-all, majority/mean voting) and KGE baselines (TransE, DistMult). As shown in Table 2, static fusion is unreliable: on DDInter, the best single prompt achieves 34.00/23.36 while the best static fusion reaches only 20.40/8.61, and on PharmaDB static fusion still falls below the best single prompt. The KGE controls reveal a complementary limitation—TransE and DistMult are competitive on DDInter but weaker on PharmaDB, and on DrugBank the best KGE result is only 11.80/4.59, confirming that graph embeddings alone struggle with fine-grained 86-way prediction. Together, these controls support our core motivation: heterogeneous evidence must be interpreted adaptively rather than simply concatenated, voted, or embedded. Full static-fusion and KGE results are reported in Appendix C.1 and Appendix C.2.

RQ3: How does DRUGREASON compare with graph-based baselines? We compare DRUGREASON with graph-based and transformer baselines to evaluate whether routed language reasoning can compete with models that directly operate on KG structure. As shown in Table 3, DRUGREASON is competitive with or better than graph-based baselines on both tasks. On DDInter, Qwen2.5 reaches 72.00%, slightly exceeding GraphTrans. (71.00%) and outperforming GAT, GCN, GIN, GraphSAGE, and the general Transformer. On PharmaDB, Llama3.2 reaches 68.25%, exceeding GAT (67.11%). Performance is backbone-dependent—Llama3.2 trails GraphTrans. on DDInter and Qwen2.5 falls slightly below GAT on PharmaDB—so we do not claim universal dominance. The key result is that verbalized KG evidence and mechanistic language reasoning can match or exceed strong graph baselines while retaining interpretable language-level evidence for each prediction.

RQ4: Do routing and mutual distillation matter? Table 4 isolates the effects of conditional routing and cross-expert mutual distillation. Removing the router causes large performance drops, especially on DDInter with Qwen2.5, where accuracy falls from 72.00% to 12.80%. This confirms that heterogeneous reasoning views should not be assigned uniformly. Mutual distillation also improves average robustness by exchanging softened predictive signals between the two experts. Its benefit is not monotonic in every cell: on PharmaDB with Llama3.2, removing distillation slightly improves accuracy from 68.25% to 68.65%. Overall, however, the full model achieves the best average performance for both experts, showing that routing

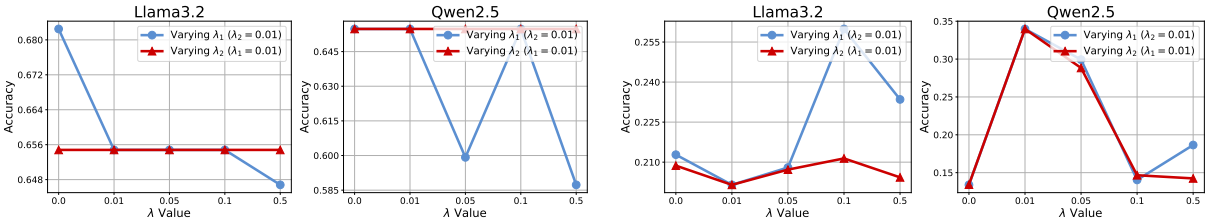
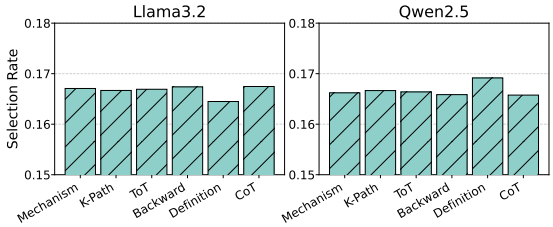


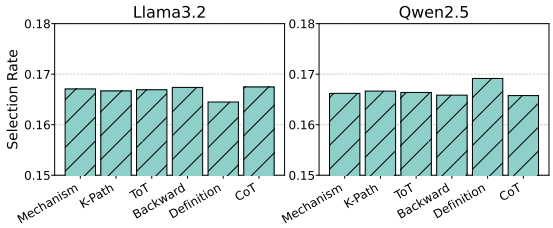
Figure 3: **Sensitivity of DRUGREASON to routing and distillation weights.** Varying λ_1 (routing regularization) has a larger impact than λ_2 (distillation weight) for both backbones, with performance generally degrading at larger values. `Llama3.2` remains relatively stable across most of the range, while `Qwen2.5` shows a sharper drop at $\lambda_1 = 0.05$ before partially recovering, indicating greater sensitivity to routing regularization strength.

`Llama3.2`			
	No Mut.	No Rtr.	Ours
PharmaDB	**68.65**	29.37	68.25
DDInter	23.10	20.70	**60.00**
DrugBank	16.60	24.35	**25.99**
Avg.	36.12	24.81	**51.41**
`Qwen2.5`			
	No Mut.	No Rtr.	Ours
PharmaDB	50.40	53.57	**65.48**
DDInter	59.90	12.80	**72.00**
DrugBank	4.63	12.80	**33.98**
Avg.	38.31	26.39	**57.15**

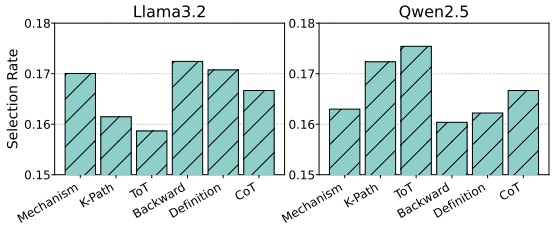
Table 4: **Ablation study on mutual distillation (Mut.) and conditional routing (Rtr.).** Results are reported as accuracy (%). Removing routing causes large drops, and removing mutual distillation generally weakens average performance. The full model achieves the best average results, showing that routing and cross-expert knowledge sharing are complementary.



(a) DDInter.



(b) DrugBank.



(c) PharmaDB.

Figure 4: **Distribution of routed reasoning views across strategies.** The router maintains non-collapsed utilization of KG-derived and LLM-generated views, with dataset- and backbone-specific variation.

and cross-expert knowledge sharing are complementary.

Router Analysis. Figure 4 analyzes how reasoning views are assigned across experts. The router does not collapse to a single reasoning strategy: KG-derived and LLM-generated views all receive non-trivial routing mass, with variations across datasets and backbones. This suggests that DRUGREASON keeps access to heterogeneous evidence sources rather than relying on one dominant prompt type. We view this analysis as a diagnostic of non-collapsed routing behavior; the ablation results in Table 4 provide direct evidence that routing-based evidence assignment affects prediction.

Hyper-parameter Sensitivity. Figure 3 studies the effect of λ_1 and λ_2, which control routing entropy regularization and mutual distillation. `Llama3.2` remains relatively stable across settings, while `Qwen2.5` is more sensitive, especially to λ_1. This is consistent with the role of λ_1 in controlling expert assignment: overly weak or overly strong routing regularization can affect the balance between specialization and evidence sharing. The final training hyperparameters are listed in Appendix B.

5 Conclusion

We presented DRUGREASON, a dynamic multi-view reasoning framework for drug-repurposing-oriented biomedical relation prediction. It integrates KG-derived factual grounding with training-time LLM-generated mechanistic rationales, adaptively routes heterogeneous reasoning views to two specialized generative experts, and applies

cross-expert mutual distillation to share predictive signals while preserving specialization. Experiments on DDInter, DrugBank, and PharmaDB show that explicit reasoning evidence is critical, single-view and static-fusion strategies are often unstable, and DRUGREASON achieves strong overall performance against reasoning, graph-based, and transformer baselines. These results highlight heterogeneous, query-dependent evidence interpretation as a promising direction for reliable and interpretable drug repurposing.

Limitations

DRUGREASON is intended as a computational reasoning framework for prioritizing candidate relations, not as experimental validation of therapeutic efficacy or safety. Its performance depends on the coverage and quality of the underlying biomedical KG, as missing or noisy graph relations can affect the constructed reasoning views. The LLM-generated views may also contain unsupported mechanistic claims, although routing and mutual distillation help reduce reliance on any single view. In addition, using two generative LLM experts increases computational cost compared with single-model prompting or graph-only baselines. Future work will explore more efficient routing, stronger automatic rationale verification, and broader evaluation on prospective drug repurposing settings.

Acknowledgments

This research was partially supported by the National Science Foundation (NSF) under Award No. IIS-2551752

References

Tassallah Abdullahi, Ioanna Gemou, Nihal V Nayak, Ghulam Murtaza, Stephen H Bach, Carsten Eickhoff, and Ritambhara Singh. 2025a. K-paths: Reasoning over graph paths for drug repurposing and drug interaction prediction. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2*, pages 5–16.

Tassallah Abdullahi, Ioanna Gemou, Nihal V Nayak, Ghulam Murtaza, Stephen H Bach, Carsten Eickhoff, and Ritambhara Singh. 2025b. K-paths: Reasoning over graph paths for drug repurposing and drug interaction prediction. *arXiv preprint arXiv:2502.13344*.

Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko. 2013. Translating embeddings for modeling multi-relational data. *Advances in neural information processing systems*, 26.

Laura Cabello, Carmen Martin-Turrero, Uchenna Akujuobi, Anders Søgaard, and Carlos Bobed. 2025. Meg: Medical knowledge-augmented large language models for question answering. *Preprint*, arXiv:2411.03883.

Payal Chandak, Kexin Huang, and Marinka Zitnik. 2023. Building a knowledge graph to enable precision medicine. *Scientific Data*, 10(1):67.

Juan Manuel Zambrano Chaves, Eric Wang, Tao Tu, Eeshit Dhaval Vaishnav, Byron Lee, S Sara Mahdavi, Christopher Semturs, David Fleet, Vivek Natarajan, and Shekoofeh Azizi. 2024. Tx-llm: A large language model for therapeutics. *arXiv preprint arXiv:2406.06316*.

Qingyu Chen, Yan Hu, Xueqing Peng, Qianqian Xie, Qiao Jin, Aidan Gilson, Maxwell B Singer, Xuguang Ai, Po-Ting Lai, Zhizheng Wang, and 1 others. 2025a. Benchmarking large language models for biomedical natural language processing applications and recommendations. *Nature communications*, 16(1):3280.

Qingyu Chen, Yan Hu, Xueqing Peng, Qianqian Xie, Qiao Jin, Aidan Gilson, Maxwell B Singer, Xuguang Ai, Po-Ting Lai, Zhizheng Wang, and 1 others. 2025b. Benchmarking large language models for biomedical natural language processing applications and recommendations. *Nature communications*, 16(1):3280.

Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, and 1 others. 2024. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*.

Vijay Prakash Dwivedi and Xavier Bresson. 2020. A generalization of transformer networks to graphs. *arXiv preprint arXiv:2012.09699*.

Bowen Gao, Yanwen Huang, Yiqiao Liu, Wenxuan Xie, Wei-Ying Ma, Ya-Qin Zhang, and Yanyan Lan. 2025. Pharmagents: Building a virtual pharma with large language model agents. *arXiv preprint arXiv:2503.22164*.

Ali Gharizadeh, Karim Abbasi, Amin Ghareyazi, Mohammad RK Mofrad, and Hamid R Rabiee. 2024. Hgtdr: Advancing drug repurposing with heterogeneous graph transformers. *Bioinformatics*, 40(7):btae349.

Emre Guney, Jörg Menche, Marc Vidal, and Albert-László Barábasi. 2016. Network-based in silico drug efficacy screening. *Nature communications*, 7(1):10331.

Will Hamilton, Zhitao Ying, and Jure Leskovec. 2017. Inductive representation learning on large graphs. *Advances in neural information processing systems*, 30.

Zongbo Han, Changqing Zhang, Huazhu Fu, and Joey Tianyi Zhou. 2022. Trusted multi-view classification with dynamic evidential fusion. *IEEE transactions on pattern analysis and machine intelligence*, 45(2):2551–2566.

Daniel Scott Himmelstein, Antoine Lizee, Christine Hessler, Leo Brueggeman, Sabrina L Chen, Dexter Hadley, Ari Green, Pouya Khankhanian, and Sergio E Baranzini. 2017a. Systematic integration of biomedical knowledge prioritizes drugs for repurposing. *elife*, 6:e26726.

Daniel Scott Himmelstein, Antoine Lizee, Christine Hessler, Leo Brueggeman, Sabrina L Chen, Dexter Hadley, Ari Green, Pouya Khankhanian, and Sergio E Baranzini. 2017b. Systematic integration of biomedical knowledge prioritizes drugs for repurposing. *elife*, 6:e26726.

Daniel Scott Himmelstein, Antoine Lizee, Christine Hessler, Leo Brueggeman, Sabrina L Chen, Dexter Hadley, Ari Green, Pouya Khankhanian, and Sergio E Baranzini. 2017c. Systematic integration of biomedical knowledge prioritizes drugs for repurposing. *elife*, 6:e26726.

Andrew L Hopkins. 2008. Network pharmacology: the next paradigm in drug discovery. *Nature chemical biology*, 4(11):682–690.

Kexin Huang, Payal Chandak, Qianwen Wang, Shreyas Havaldar, Akhil Vaid, Jure Leskovec, Girish N Nadkarni, Benjamin S Glicksberg, Nils Gehlenborg, and Marinka Zitnik. 2024. A foundation model for clinician-centered drug repurposing. *Nature Medicine*, 30(12):3601–3613.

Ziwei Ji, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung. 2023. Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12):1–38.

Thomas N Kipf and Max Welling. 2017. Semi-supervised classification with graph convolutional networks. *International Conference on Learning Representations*.

Mingchen Li, Halil Kilicoglu, Hua Xu, and Rui Zhang. 2024. Biomedrag: A retrieval augmented large language model for biomedicine. *Preprint*, arXiv:2405.00465.

Junkai Liu, Fuyuan Hu, Quan Zou, Prayag Tiwari, Hongjie Wu, and Yijie Ding. 2024a. Drug repositioning by multi-aspect heterogeneous graph contrastive learning and positive-fusion negative sampling strategy. *Information Fusion*, 112:102563.

Sizhe Liu, Yizhou Lu, Siyu Chen, Xiyang Hu, Jieyu Zhao, Yingzhou Lu, and Yue Zhao. 2024b. Drugagent: Automating ai-aided drug discovery programming through llm multi-agent collaboration. *arXiv preprint arXiv:2411.15692*.

Xiaohong Liu, Hao Liu, Guoxing Yang, Zeyu Jiang, Shuguang Cui, Zhaoze Zhang, Huan Wang, Liyuan Tao, Yongchang Sun, Zhu Song, and 1 others. 2025a. A generalist medical language model for disease diagnosis assistance. *Nature medicine*, 31(3):932–942.

Xiaohong Liu, Hao Liu, Guoxing Yang, Zeyu Jiang, Shuguang Cui, Zhaoze Zhang, Huan Wang, Liyuan Tao, Yongchang Sun, Zhu Song, and 1 others. 2025b. A generalist medical language model for disease diagnosis assistance. *Nature medicine*, 31(3):932–942.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.

Renqian Luo, Liai Sun, Yingce Xia, Tao Qin, Sheng Zhang, Hoifung Poon, and Tie-Yan Liu. 2022. Biogpt: generative pre-trained transformer for biomedical text generation and mining. *Briefings in bioinformatics*, 23(6):bbac409.

Elyse Mackenzie, Roger Cheng, and Pengfei Zhang. 2025. Gpt meets pubmed: a novel approach to literature review using a large language model to crowdsource migraine medication.

OpenAI. 2024. Gpt-4o system card. *arXiv preprint arXiv:2410.21276*.

Pablo Perdomo-Quinteiro and Alberto Belmonte-Hernández. 2024a. Knowledge graphs for drug repurposing: a review of databases and methods. *Briefings in Bioinformatics*, 25(6):bbae461.

Pablo Perdomo-Quinteiro and Alberto Belmonte-Hernández. 2024b. Knowledge graphs for drug repurposing: a review of databases and methods. *Briefings in Bioinformatics*, 25(6):bbae461.

Sudeep Pushpakom, Francesco Iorio, Patrick A Eyers, K Jane Escott, Shirley Hopper, Andrew Wells, Andrew Doig, Tim Guilliams, Joanna Latimer, Christine McNamee, and 1 others. 2019. Drug repurposing: progress, challenges and recommendations. *Nature reviews Drug discovery*, 18(1):41–58.

Gaurav Sahu and Olga Vechtomova. 2021. Adaptive fusion techniques for multimodal data. In *Proceedings of the 16th conference of the European chapter of the Association for Computational Linguistics: Main Volume*, pages 3156–3166.

Karan Singhal, Shekoofeh Azizi, Tao Tu, S Sara Mahdavi, Jason Wei, Hyung Won Chung, Nathan Scales, Ajay Tanwani, Heather Cole-Lewis, Stephen Pfohl, and 1 others. 2023a. Large language models encode clinical knowledge. *Nature*, 620(7972):172–180.

Karan Singhal, Shekoofeh Azizi, Tao Tu, S Sara Mahdavi, Jason Wei, Hyung Won Chung, Nathan Scales, Ajay Tanwani, Heather Cole-Lewis, Stephen Pfohl, and 1 others. 2023b. Large language models encode clinical knowledge. *Nature*, 620(7972):172–180.

Jiwoong Sohn, Yein Park, Chanwoong Yoon, Sihyeon Park, Hyeon Hwang, Mujeen Sung, Hyunjae Kim, and Jaewoo Kang. 2025. Rationale-guided retrieval augmented generation for medical question answering. *Preprint*, arXiv:2411.00300.

Karthik Soman, Peter W Rose, John H Morris, Rabia E Akbas, Brett Smith, Braian Peetoom, Catalina Villouta-Reyes, Gabriel Cerono, Yongmei Shi, Angela Rizk-Jackson, Sharat Israni, Charlotte A Nelson, Sui Huang, and Sergio E Baranzini. 2024a. Biomedical knowledge graph-optimized prompt generation for large language models. *Preprint*, arXiv:2311.17330.

Karthik Soman, Peter W Rose, John H Morris, Rabia E Akbas, Brett Smith, Braian Peetoom, Catalina Villouta-Reyes, Gabriel Cerono, Yongmei Shi, Angela Rizk-Jackson, and 1 others. 2024b. Biomedical knowledge graph-optimized prompt generation for large language models. *Bioinformatics*, 40(9):btae560.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems*, 30.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. 2018. Graph attention networks. *International Conference on Learning Representations*.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, and 1 others. 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35:24824–24837.

Jinhang Wei, Linlin Zhuo, Xiangzheng Fu, XiangXiang Zeng, Li Wang, Quan Zou, and Dongsheng Cao. 2024a. Drugrealign: a multisource prompt framework for drug repurposing based on large language models. *BMC biology*, 22(1):226.

Jinhang Wei, Linlin Zhuo, Xiangzheng Fu, XiangXiang Zeng, Li Wang, Quan Zou, and Dongsheng Cao. 2024b. Drugrealign: a multisource prompt framework for drug repurposing based on large language models. *BMC biology*, 22(1):226.

David S Wishart, Yannick D Feunang, An C Guo, Elvis J Lo, Ana Marcu, Jason R Grant, Tanvir Sajed, Daniel Johnson, Carin Li, Zinat Sayeeda, and 1 others. 2018. Drugbank 5.0: a major update to the drugbank database for 2018. *Nucleic acids research*, 46(D1):D1074–D1082.

Juncheng Wu, Wenlong Deng, Xingxuan Li, Sheng Liu, Taomian Mi, Yifan Peng, Ziyang Xu, Yi Liu, Hyunjin Cho, Chang-In Choi, and 1 others. 2025. Medreason: Eliciting factual medical reasoning steps in llms via knowledge graphs. *arXiv preprint arXiv:2504.00993*.

Guoli Xiong, Zhijiang Yang, Jiacai Yi, Ningning Wang, Lei Wang, Huimin Zhu, Chengkun Wu, Aiping Lu, Xiang Chen, Shao Liu, and 1 others. 2022. Ddinter: an online drug–drug interaction database towards improving clinical decision-making and patient safety. *Nucleic acids research*, 50(D1):D1200–D1207.

Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. 2019. How powerful are graph neural networks? *International Conference on Learning Representations*.

An Yang, Binyuan Hui, Bubing Zhang, and 1 others. 2024. Qwen2.5 technical report. *arXiv preprint arXiv:2412.15115*.

Bishan Yang, Wen-tau Yih, Xiaodong He, Jianfeng Gao, and Li Deng. 2014. Embedding entities and relations for learning and inference in knowledge bases. *arXiv preprint arXiv:1412.6575*.

Zerui Yang, Yuwei Wan, Yinqiao Li, Yudai Matsuda, Tong Xie, and Linqi Song. 2025. Drugmcts: a drug repurposing framework combining multi-agent, rag and monte carlo tree search. *arXiv preprint arXiv:2507.07426*.

Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Tom Griffiths, Yuan Cao, and Karthik Narasimhan. 2023. Tree of thoughts: Deliberate problem solving with large language models. *Advances in neural information processing systems*, 36:11809–11822.

A Prompt Templates and Input Filtering

This appendix provides the prompt templates used to construct reasoning views. These templates correspond to the reasoning strategies discussed in Section 3.2. The gold label is never included in model inputs. For teacher-generated rationales, we use the teacher LLM only during training to construct mechanistic evidence views; the teacher is not queried during inference. We apply deterministic filtering to remove empty outputs, malformed responses, and template artifacts such as explicit label fields. We do not rewrite or manually improve retained rationales.

Filtering. After generation, we retain only non-empty, well-formed rationales. We remove explicit label fields, answer-only artifacts, malformed outputs, and responses that fail to follow the requested format. This filtering is deterministic and does not involve rewriting the content of retained rationales. This design ensures that LLM-generated views provide mechanistic evidence rather than direct label leakage.

View Type	Template
No reasoning	`Given the candidate pair [ENTITY_1] and [ENTITY_2], predict the relation label for the target benchmark. Return only one valid label.`
Definition	`Definition of [ENTITY_1]: [DEF_1]. Definition of [ENTITY_2]: [DEF_2]. Based on these biomedical definitions, predict the relation label for the candidate pair. Return only one valid label.`
K-Path	`Candidate pair: [ENTITY_1] and [ENTITY_2]. Extracted biomedical paths: [PATH_1]; [PATH_2]; ... Based on these graph paths, predict the relation label. Return only one valid label.`
Mechanism	`Focus on the biomedical mechanism connecting [ENTITY_1] and [ENTITY_2]. Describe relevant molecular targets, pathways, pharmacological processes, or therapeutic mechanisms. Do not restate the answer directly; justify the relationship mechanistically.`
Chain-of-Thought	`Reason step by step about the possible biomedical relationship between [ENTITY_1] and [ENTITY_2]. Consider pharmacological effects, disease mechanisms, pathway connections, and interaction evidence.`
Tree-of-Thought	`Generate multiple plausible biomedical hypotheses for the relationship between [ENTITY_1] and [ENTITY_2]. Compare these hypotheses and identify which one is most biologically plausible.`
Forward–Backward	`First reason forward from [ENTITY_1] to [ENTITY_2]. Then reason backward from [ENTITY_2] to [ENTITY_1]. Check whether the two directions support a consistent biomedical relationship.`

Table 5: **Prompt templates for reasoning-view construction.** Placeholders are replaced with candidate entities, biomedical definitions, or KG paths. Gold labels are omitted from all inputs.

Hyperparameter	DDInter	DrugBank	PharmaDB
epochs	1	1	1
batch size	8	8	8
max length	768	768	768
lr_{s1}	3e-4	1e-4	5e-4
lr_{s2}	5e-4	1e-4	1e-4
lr_{misc}	1e-4	1e-4	1e-4
λ_2	0.01	0.10	0.00
λ_1	0.01	0.01	0.01
τ	1.0	1.0	1.0
seed	42	42	42

Table 6: **Llama3.2 training hyperparameters by dataset.** λ_1 controls routing regularization and λ_2 controls mutual distillation.

Hyperparameter	DDInter	DrugBank	PharmaDB
epochs	2	1	1
batch size	8	8	8
max length	768	768	768
lr_{s1}	3e-5	1e-4	1e-4
lr_{s2}	5e-5	1e-4	5e-4
lr_{misc}	1e-4	1e-4	1e-4
λ_2	0.10	0.01	0.10
λ_1	0.01	0.01	0.01
τ	1.0	1.0	1.0
seed	42	42	42

Table 7: **Qwen2.5 training hyperparameters by dataset.** λ_1 controls routing regularization and λ_2 controls mutual distillation.

B Implementation Details

We train DRUGREASON in two stages. In Stage 1, we optimize the routed supervised prediction objective together with routing regularization. In Stage 2, we keep the routed supervised objective and add cross-expert mutual distillation. The coefficient λ_1 controls routing load regularization, and λ_2 controls the mutual distillation loss. We use lr_{s1} and lr_{s2} for expert training in Stage 1 and Stage 2, respectively, and lr_{misc} for non-expert components such as the router.

C Additional Control Results

The main text reports compact control results to support RQ2. Here we provide the full control tables. These controls are designed to test whether the gains of DRUGREASON come merely from using more evidence or from using KG structure alone. They use the same candidate pairs and evaluation protocol as the main experiments.

C.1 Static Prompt-Fusion Controls

Table 8 reports the full prompt-control results. Single-prompt variants evaluate one reasoning view at a time. Concat-all performs input-level fusion by placing all reasoning views into one prompt. Majority vote and mean vote perform output-level fusion over view-level predictions. These controls complement the compact results in Table 2 of the main text.

C.2 Knowledge Graph Embedding Controls

Table 9 reports KGE baselines used to test whether KG structure alone is sufficient without language-level mechanistic reasoning. These results complement the compact KGE column in Table 2 of the main text.

D LLM Usage

We used `GPT-4o` as a teacher model to generate training-time mechanistic rationales, as described in the method. The teacher model is not queried during inference. Separately, large language models were used to assist with editing and polishing

Variant	DDInter		PharmaDB	
	Acc.	Macro-F1	Acc.	Macro-F1
Concat-all	20.40	8.61	43.25	20.13
Chain-of-Thought	3.80	6.03	41.27	22.33
Definition	15.80	11.73	43.25	30.64
Forward–Backward	20.50	14.88	42.86	29.24
K-Path	34.00	22.60	42.86	27.82
K-Paths direct prompt	34.00	23.36	42.86	27.82
Majority vote	18.40	11.55	42.46	23.24
Mean vote	8.10	7.40	46.43	35.44
Mechanism	4.60	5.15	46.03	34.52
Tree-of-Thought	31.10	23.47	49.60	37.63

Table 8: **Detailed static prompt-fusion and single-prompt controls.** Results are reported in percentage. These controls show that using more reasoning views without adaptive routing does not reliably improve performance.

Dataset	TransE		DistMult	
	Acc.	Macro-F1	Acc.	Macro-F1
DDInter	43.00	30.58	36.70	26.67
DrugBank	11.80	4.59	9.90	1.22
PharmaDB	37.70	36.43	32.94	32.05

Table 9: **Knowledge graph embedding baselines.** Results are reported in percentage. These controls evaluate graph structure alone, without language-level reasoning views or expert routing.

the manuscript, including improving clarity, grammar, and readability. They were not used to generate experimental results or alter reported numerical findings.