HYPE-DTI: VALIDATION-GUARDED KNOWLEDGE EXCHANGE FOR LOW-LABEL DRUG–TARGET INTERACTION PREDICTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Predicting drug–target interactions (DTIs) from few labeled examples requires models to combine molecular, sequence, and relational evidence without amplifying errors during cross-view exchange. Existing dual-view methods integrate these sources, but often couple information transfer with model updating and final fusion, making it difficult to reject unreliable transfers. We introduce Hype-DTI, a dual-expert framework that validates each stage of collaboration. A relation expert provides selected pseudo-labels to a hybrid self expert, the reverse-distillation update is retained only after validation, and the learned gate is deployed only when it outperforms the selected expert. Across three five-fold low-label benchmarks, with each fold containing 10 positive and 10 negative training pairs, Hype-DTI improves mean AUPR over MoseDTI by 6.10 points on AGO, 0.49 on Blocker, and 4.60 on E-. It also outperforms MoseDTI in every matched semi-inductive fold involving unseen drugs, targets, or both. On DrugCentral, Hype-DTI reaches 86.78 ± 0.38% AUPR, the highest score among evaluated methods that do not use additional task-related DTI labels. ConPLex reaches 88.41 ± 0.58% with labeled DUD-E pretraining, illustrating the benefit of stronger external supervision. These results support validation-guarded exchange in the evaluated low-label and semi-inductive DTI settings, while highlighting the effect of external supervision on cross-dataset performance.

1 INTRODUCTION

Accurately predicting drug–target interactions (DTIs) can narrow the experimental search for promising candidate pairs during drug discovery. Experimental validation is costly, and only a small fraction of possible pairs are labeled. Models therefore draw on molecular structure, protein sequence, and biomedical context. Sequence-based, molecular-graph, and protein language models encode information intrinsic to each drug–target pair (Öztürk et al., 2018; Nguyen et al., 2021; Huang et al., 2021; Bai et al., 2023; Singh et al., 2023). Yet the most informative source may vary across pairs.

When labeled examples are scarce, models can combine pair-intrinsic evidence from molecular structure and protein sequence with relational context from a biomedical knowledge graph. Such graphs encode typed relations among drugs, proteins, and other biomedical entities. Pair-intrinsic evidence may generalize beyond observed links, whereas graph context can provide information unavailable from an isolated pair. However, neither source is reliable for every pair: molecular similarity need not imply interaction, and graph proximity may reflect unrelated associations. MoseDTI treats the two sources as complementary experts (Zhai et al., 2025), but deciding when to exchange information between them remains challenging.

This challenge arises because exchanging information between the two experts can introduce several reliability risks. Pseudo-labels, or predicted labels for unlabeled pairs, may be inaccurate or concentrated around a few entities. Distillation can update one expert using signals from the other, but gains on selected pairs may coincide with lower overall validation performance. A learned gate assigns pair-specific weights, but it may perform worse than the stronger of the two individual experts.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

When transfer, model updating, and fusion are coupled, identifying the source of a performance loss can be difficult.

To address these problems, we introduce Hype-DTI, which separates dual-expert collaboration into stages with distinct reliability checks. The relation expert encodes graph context, whereas the hybrid self expert combines evidence from molecular structure and protein sequence. Relation-to-self transfer filters pseudo-labels by confidence and entity diversity. Self-to-relation (reverse) distillation retains an update only when validation criteria are met; otherwise, the previous checkpoint is restored. Finally, Hype-DTI deploys the learned gate only when validation criteria favor it over the selected reference expert (Figure 1).

We evaluate this design on three low-label DTI benchmarks, three semi-inductive AGO tests, the larger DrugCentral benchmark, and a fixed-expert ablation that varies only whether the learned gate is deployed. Hype-DTI improves mean AUPR over MoseDTI on all three low-label benchmarks and in every matched semi-inductive fold, and attains the highest mean AUPR on DrugCentral among methods without additional pretraining on task-related DTI labels. These results support the effectiveness of validation-guarded exchange under low-label and semi-inductive conditions.

2 Related Work

Drug–target interaction prediction rests on two complementary sources of evidence: pair-intrinsic information from molecular structure and protein sequence, and relational context from biomedical knowledge graphs. A third line of work combines heterogeneous predictors using knowledge distillation or mixture-of-experts models. These three directions motivate the hybrid self expert, the relation expert, and the validation-guarded exchange introduced by Hype-DTI, respectively.

Pair-intrinsic drug and protein representations. Pair-intrinsic models infer interactions from drug and protein features alone. Early neural approaches encoded strings directly (Öztürk et al., 2018); later systems introduced molecular graphs (Nguyen et al., 2021), cross-molecule interactions (Huang et al., 2021), and bilinear attention under domain shift (Bai et al., 2023). Pretrained chemistry and protein encoders further transfer patterns from large unlabeled corpora (Chithrananda et al., 2020; Rives et al., 2021; Lin et al., 2023), while ConPLex adds supervised contrastive learning (Singh et al., 2023). These methods strengthen evidence intrinsic to a pair, but they do not directly encode the pair's biomedical neighborhood.

Relational context under scarce supervision. Relational models instead use associations surrounding each entity. DRKG integrates multiple biomedical sources into a multi-relational graph (Zeng et al., 2020), and TransE embeds its entities in a shared latent space (Bordes et al., 2013). These embeddings provide useful context when the target relation has few labels, but proximity in a heterogeneous graph can conflate multiple relation types rather than isolate evidence for a particular DTI pair. MoseDTI combines this relational view with a pair-intrinsic self view (Zhai et al., 2025). Hype-DTI retains both evidence sources, but makes the conditions for information exchange explicit.

Cross-expert learning and deployment. Knowledge distillation transfers predictions between models (Hinton et al., 2015), and mixtures of experts learn input-dependent combinations of specialized predictors (Jacobs et al., 1991). In a 10-shot setting, however, neither a transferred label nor a learned combination should be trusted by default. The better teacher can change across candidate regions, and the best deployable predictor can change across folds. Hype-DTI therefore separates pseudo-label eligibility, update acceptance, and mixture deployment into separate decisions that can each be rejected. We apply the same distinction to external supervision: because full ConPLex uses labeled active and decoy pairs from DUD-E (Mysinger et al., 2012), Section 4.4 reports it separately from methods without additional task-related DTI labels.

3 Method

3.1 Problem Formulation and Design Rationale

We consider binary DTI prediction from a small labeled set $D_{\mathrm{tr}} = \{(d_i, t_i, y_i)\}_{i=1}^{N}$. Here $d_i \in V_D$ is a drug, $t_i \in V_T$ is a protein target, and $y_i \in \{0, 1\}$ is the interaction label. Separate sets D_{val}

and D_{te} are used for model selection and final evaluation. A relation expert f_r and a hybrid self expert f_s produce logits $z_r(d,t)$ and $z_s(d,t)$, with positive-class probabilities p_r and p_s. The final predictor is either a learned convex mixture of these logits or one validation-selected expert.

The central design principle is that reliability is local to a prediction and a learning stage. Biomedical context may be informative for one pair, while molecular and sequence evidence may be more reliable for another. Moreover, a useful pseudo-label does not imply that an entire model update is beneficial, and a well-trained gate does not imply that the gate should be deployed. Hype-DTI therefore decomposes learning into three nested decisions: which predictions may cross between experts, whether a reverse update should be retained, and whether the resulting mixture should replace a single expert. These three decisions are embedded in the four stages of Figure 1: the first governs pseudo-label selection within S1, the second controls distillation within S3, and the third authorizes fusion within S4.

Figure 1 instantiates this idea as a four-stage pipeline. In S1, the relation expert is fitted on observed labels and proposes a small, balanced, entity-diverse pseudo-labeled set. In S2, the hybrid self expert is adapted using those pairs. In S3, the self expert returns soft targets only where it is confident and the relation expert is uncertain; validation then keeps or rolls back the update. In S4, both experts are frozen while a small gate learns pair-specific weights. The gate remains only a candidate predictor until a final validation rule authorizes its deployment. The implemented selection and authorization rules use only training or validation metrics; test outcomes are not optimization inputs.

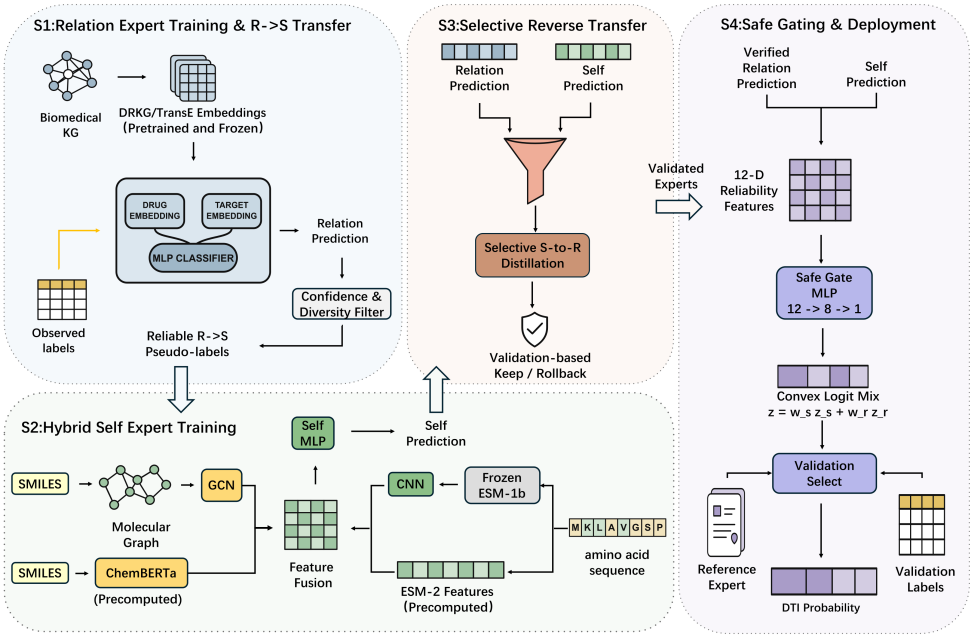

Figure 1: Overview of Hype-DTI. The relation expert encodes biomedical graph context, and the hybrid self expert encodes molecular structure and protein sequence. Pseudo-labels move in two directed stages. Validation first accepts or rejects relation tuning, then decides whether the frozen-expert gate or a reference expert is deployed. Runtime protected-pair checks precede every pseudo-labeling stage that feeds an optimizer.

3.2 Relation Expert: Biomedical Context

The relation expert serves as the initial teacher because its biomedical context is available before the larger self expert has adapted to the low-label task. It retrieves frozen 1,000-dimensional drug and target embeddings from a TransE model pretrained on DRKG (Bordes et al., 2013; Zeng et al., 2020). Rather than reuse a TransE relation vector or triple score, the downstream expert learns a task-specific classifier from the two entity embeddings:

$$z_r = W_{r2}\ \mathrm{ReLU}(W_{r1}[e_d; e_t] + b_{r1}) + b_{r2}. \tag{1}$$

W_{r1} maps the 2,000-dimensional concatenation to 1,000 hidden units, and W_{r2} produces two logits. We first train this classifier only on observed positive and negative pairs:

$$L_{r,0} = \frac{1}{N}\sum_{i=1}^{N} \mathrm{CE}(z_r(d_i, t_i), y_i)\,. \tag{2}$$

The entity table remains frozen, so only the two-layer classifier is updated using the 20 observed training pairs rather than rewriting the upstream graph geometry. Consequently, f_r contributes a task-specific decision boundary in the pretrained biomedical embedding space, not a reused TransE link score. We preserve this initial checkpoint for two roles: it teaches the self expert during forward transfer, and it provides the recoverable state if reverse transfer later fails validation.

3.3 Hybrid Self Expert: Pair-Intrinsic Evidence

The self expert asks a complementary question: does the drug–protein pair appear compatible based on its own molecular and sequence evidence? It does not require a known DTI edge for the candidate pair. To balance task adaptation against overfitting, the expert combines a trainable local route with a frozen global route. The local route learns drug neighborhoods and residue patterns for the current task, whereas the global route projects representations learned from large unlabeled corpora of chemical structures and protein sequences.

For the local drug route, a SMILES string becomes a molecular graph with bidirectional bond edges and self-loops. Each atom carries a 78-dimensional feature vector. Two graph-convolution layers map $78 \to 1,000 \to 1,000$ with ReLU activations, followed by an MLP node readout and max aggregation (Kipf & Welling, 2017). This produces $g_d \in \mathbb{R}^{1000}$.

For the protein in each pair, frozen ESM-1b representations provide one vector per residue (Rives et al., 2021); three one-dimensional convolutions transform the channel dimension from 1,280 to 96 to 128 to 1,000. Each convolution is followed by a ReLU activation, and adaptive max pooling yields $s_t \in \mathbb{R}^{1000}$. The graph layers, protein convolutions, and readout remain trainable, while ESM-1b is fixed.

The global route supplies broader priors without adding a large fine-tuning burden. We compute representations once with ChemBERTa-zinc-base-v1 and ESM-2-t33-650M (Chithrananda et al., 2020; Lin et al., 2023). Attention-mask-aware mean pooling yields a 768-dimensional drug vector and a 1,280-dimensional protein vector. Separate linear, layer-normalization, ReLU, and dropout blocks project both vectors into a shared width:

$$u_d = P_d(\bar{c}_d), \qquad u_t = P_t(\bar{a}_t), \qquad u_d, u_t \in \mathbb{R}^{1000}. \tag{3}$$

Only the projection blocks are optimized. Freezing both pretrained encoders prevents a 10-shot training set from estimating hundreds of millions of language-model parameters.

The two routes are combined into the pair representation

$$h_{d,t} = \big[g_d; s_t; u_d; u_t; u_d \odot u_t; \\ |u_d - u_t|\big] \in \mathbb{R}^{6000}. \tag{4}$$

The product and absolute-difference terms make coordinate-wise agreement and discrepancy directly available to the classifier. An MLP maps $h_{d,t}$ to z_s. Dropout is 0.2 throughout. This hybrid design lets the self expert adapt local patterns while retaining broad pretrained priors.

3.4 Decision I (S1): Selective Relation-to-Self Transfer

The first decision is which relation predictions deserve to become supervision. Transferring the entire candidate space would amplify teacher error, while retaining only the largest scores could

repeatedly select the same well-connected entities. We therefore filter for confidence, class balance, and endpoint diversity. After initial training, f_r scores a random 2% sample of the Cartesian product of drug–target pairs for which features are available, denoted C_r. A candidate is provisionally positive when $p_r \geq 0.9$ and negative when $p_r \leq 0.1$.

Within each class, candidates are ranked by confidence and initially limited to three pairs per drug and per target. The cap spreads supervision across endpoints rather than allowing a few drugs or proteins to dominate. If a pass cannot fill its quota, selection relaxes the cap deterministically while preserving class balance and diversity.

The positive and negative quotas are each set to 3×10^{-5} of the scored list. The 2% sample limits teacher evaluations, whereas these quotas limit the size and class composition of the augmented set. Let $\widetilde{D}_{r\to s}$ be the retained pairs and $A_s = D_{\mathrm{tr}} \uplus \widetilde{D}_{r\to s}$. Real examples receive weight one, while pseudo-labeled examples receive weight 0.5:

$$L_s = \frac{\sum_{i=1}^{|A_s|} w_i \ \mathrm{CE}(z_s(d_i, t_i), \widehat{y}_i)}{\sum_{i=1}^{|A_s|} w_i}. \tag{5}$$

Here $\widehat{y}_i = y_i$ and $w_i = 1$ for observed pairs, while $\widehat{y}_i = \widetilde{y}_i \in \{0, 1\}$ and $w_i = 0.5$ for transferred pairs. Downweighting prevents a large pool of pseudo-labeled examples from overwhelming the observed labels. Before optimization, a protected-pair guard removes candidates whose pair identity matches any record in the split inventory. A second assertion inspects the constructed DataLoader and requires zero residual overlap before the first optimizer update.

3.5 Decision II (S3): Reversible Self-to-Relation Distillation

The second decision asks whether a self prediction is informative and whether the resulting model update is globally acceptable. Candidate selection is deliberately asymmetric: the self expert must be confident where the relation expert is uncertain. The self expert scores a random 0.02% sample of the feature-complete Cartesian product, and the relation expert scores the same pairs. We define two uncertainty pools:

$$\begin{aligned} C^{+}_{s\to r} &= \{(d,t) : p_s \geq 0.85,\ |p_r - 0.5| \leq 0.25\}, \\ C^{-}_{s\to r} &= \{(d,t) : p_s \leq 0.15,\ |p_r - 0.5| \leq 0.25\}. \end{aligned} \tag{6}$$

Positive candidates are ranked by $p_s(1 - 2|p_r - 0.5|)$ and negatives by $(1 - p_s)(1 - 2|p_r - 0.5|)$. Each score increases with self confidence and relation uncertainty. Reverse transfer therefore targets information missing from the relation expert rather than merely copying agreement. A smaller agreement pool stabilizes the update by retaining pairs with $p_s, p_r \geq 0.85$ or $p_s, p_r \leq 0.15$, ranked by $\max\{p_s p_r, (1 - p_s)(1 - p_r)\}$. Each uncertainty pool receives 1.5×10^{-4} of the scored list, and agreement receives 5×10^{-5}. All pools use the same endpoint cap, deterministic fill rule, and protected-pair guard as in forward transfer.

The selected self probabilities remain soft targets, preserving confidence information that hard labels would discard. Let $m_r = z_{r,1} - z_{r,0}$ be the relation logit margin. Each mini-batch combines the observed-label loss, probability matching, and a ranking term over $K = \min(|P|, |N|, 128)$ matched selected positive–negative pairs:

$$\begin{aligned} L_r = L_{\mathrm{CE,hard}} &+ \lambda \ \mathrm{MSE}(p_r, p_s) \\ &+ \frac{0.5\lambda}{K} \sum_{j=1}^{K} \left[0.2 - m^{+}_{r,j} + m^{-}_{r,j}\right]_{+}, \qquad K \leq 128. \end{aligned} \tag{7}$$

The ranking term requires a selected positive margin to exceed a selected negative margin by 0.2. The distillation weight begins at $\lambda = 0.2$. If validation AUPR does not decrease, we multiply λ by 1.2. If validation AUPR regresses, we restore the best epoch and halve λ, clipping it to [0.02,0.6]. This inner loop adapts how strongly the relation expert follows the self expert.

Adaptive optimization still cannot guarantee that the resulting relation expert is better than its starting point. We therefore compare the tuned model with the checkpoint saved before distillation. If p is the pre-tuning validation AUPR and $n = |D_{\text{val}}|$, the required improvement is

$$\delta_r = \max\Big(0.005,\ 0.35\sqrt{p(1-p)/n}\Big). \tag{8}$$

The tuned state is accepted only when validation AUPR rises by at least δ_r, AUROC does not decrease, and accuracy falls by no more than 0.05. Otherwise Hype-DTI restores the pre-tuning checkpoint. Thus candidate selection identifies where reverse transfer is informative, the inner loop controls its strength, and the outer comparison decides whether the update becomes part of the model.

3.6 Decision III (S4): Validation-Authorized Fusion

The final decision is not how to mix the experts, but whether mixing is warranted at all. We distinguish a *training anchor* from a *deployment reference*. Before gate training, a validation hierarchy selects the anchor by AUPR, then AUROC, then accuracy, with additional accuracy and saturation checks. The anchor determines the gate's initial bias and reference losses. After training, a separate rule selects the expert that will be deployed if the gate is rejected. This separation prevents the optimization starting point from determining the final fallback in advance.

The gate is intentionally small and uses disagreement signals without relearning either expert. Let p_k and $m_k = z_{k,1} - z_{k,0}$ denote the positive probability and logit margin of expert $k \in \{s, r\}$. We define predictive entropy as $H(p) = -p\log p - (1-p)\log(1-p)$. The 12 inputs summarize confidence, disagreement, margin geometry, predictive entropy, and KG embedding geometry:

$$\begin{aligned}\phi(d,t) = \big[&p_s, p_r, 2|p_s - 0.5|, 2|p_r - 0.5|, |p_s - p_r|,\\ &\tanh(m_s), \tanh(m_r), \tanh(m_s - m_r), H(p_s), H(p_r),\\ &\cos(e_d, e_t), |\tanh(\|e_d\|/10) - \tanh(\|e_t\|/10)|\big].\end{aligned} \tag{9}$$

A $12 \to 8 \to 1$ tanh MLP turns these signals into a relation weight. With anchor-dependent prior logit $\rho_a \in \{-1, +1\}$,

$$w_r = \sigma(q_\theta(\phi) + \rho_a), \qquad z_g = (1 - w_r)z_s + w_r z_r. \tag{10}$$

Here $\rho_a = +1$ for a relation anchor and -1 for a self anchor. The output layer of q_θ is initialized to zero, so optimization begins near the selected anchor rather than an arbitrary equal mixture. Both experts remain frozen, and gradients update only the gate. The gate therefore learns bounded corrections to a validated starting point instead of jointly moving all three models.

The active gate objective is

$$\begin{aligned}L_g = L_{\text{cls}} + L_{\text{safe}} + 0.5L_{\text{rank}} + 0.2L_{\text{anchor}} + 0.01L_{\text{sparse}}\\ + L_{\text{budget}} + 0.05L_{\text{contrast}} + 0.1L_{\text{protect}}.\end{aligned} \tag{11}$$

The objective gives the non-anchor expert a narrow, evidence-dependent role. L_{cls} fits real and pseudo-labeled pairs, with weight 0.1 for pseudo-labels. L_{safe} and L_{rank} preserve real examples and the anchor's relative ordering. L_{anchor} identifies cases where the other expert helps, while L_{sparse} and L_{budget} limit its overall allocation. L_{contrast} separates helpful from harmful non-anchor cases, and L_{protect} suppresses non-anchor weight where the anchor is better. Gate pseudo-labels pass through the protected-pair guard before optimization. Appendix B gives the complete definitions and inactive alternatives.

Gate training produces a candidate, not a deployment decision. We select the reference (fallback) expert primarily by validation AUPR, with accuracy and saturation checks. The best gate checkpoint must then clear an inverse-standard-error-weighted AUPR–AUROC gain of at least $\max\left(0.005, \frac{0.1}{\sqrt{n}}\right)$. Its AUROC may fall by at most 0.01, and its accuracy must remain within $\frac{0.2}{\sqrt{n}}$

of the reference. Non-inferiority clauses apply when the reference expert dominates or the ranking metrics are near saturation, but accuracy and saturation checks may still reject the gate. If authorization fails, the reference expert becomes the deployed predictor. The entire pipeline therefore ends with an explicit option not to fuse.

4 EXPERIMENTS

4.1 EXPERIMENTAL SETUP

Datasets and evaluation protocols. We evaluate Hype-DTI across three regimes that differ in label availability and entity coverage. The principal low-label evaluation uses the AGO, Blocker, and E- relation subsets introduced in MoseDTI (Zhai et al., 2025), retaining the supplied splits and negative pairs. Each of the five folds contains 10 positive and 10 negative training pairs (20 labeled training pairs in total; we refer to this setting as 10-shot), followed by 20 positive and 20 negative validation pairs. The balanced test sets contain 1,308, 383, and 1,371 pairs per class for AGO, Blocker, and E-, respectively. We report accuracy, AUROC, and area under the precision–recall curve (AUPR), with AUPR as the primary metric because ranking positive interactions is central to candidate prioritization.

Three AGO variants withhold labeled training interactions involving the test drug, the test target, or both. Held-out entities may still occur in auxiliary non-DTI graph relations, so we call this semi-inductive rather than fully inductive evaluation. DrugCentral (Ursu et al., 2017) adds a separate five-fold benchmark with more labels.

Baselines and supervision regimes. We compare Hype-DTI with MoseDTI (Zhai et al., 2025), DrugBAN (Bai et al., 2023), and ConPLex (Singh et al., 2023). We use ConPLex without DUD-E as the matched-supervision setting (no external DUD-E pairs) and report full ConPLex with labeled active and decoy pairs from DUD-E as a higher-supervision comparator on DrugCentral (Appendix Figure 2).

Leakage safeguards and audit scope. To prevent evaluation pairs from re-entering pseudo-label-driven training, protected-pair checks precede every optimizer stage that uses pseudo-labels. Separate audited reruns generated 45 runtime-guard records, and a post hoc identity audit found zero failures across 150 within-fold checks (details in Appendix E).

Optimization and result selection. All runs use seed 0 and Adam; the full hyperparameters are in Appendix C.

Tables 1–3 report means and sample standard deviations over five folds; Table 4 reports the five-fold means for the matched deployment ablation. For the 15 principal runs, the reported metrics come from the gate checkpoint that passed validation or from the validation-selected fallback; no checkpoint is selected by maximizing test AUPR.

4.2 LOW-LABEL DTI PREDICTION

With the evaluation protocol fixed, we first test Hype-DTI when each fold contains 10 positive and 10 negative labeled training pairs (20 pairs in total). Table 1 compares mean AUPR across AGO, Blocker, and E-.

Table 1: 10-shot test AUPR (%, 10 positive + 10 negative training pairs per fold; mean ± SD over five folds).

Method	AGO	Blocker	E-
DrugBAN	60.46 ± 4.02	75.61 ± 4.07	67.84 ± 4.67
ConPLex without DUD-E	67.11 ± 2.37	81.23 ± 3.25	68.04 ± 5.55
ConPLex with DUD-E	49.07 ± 1.78	49.57 ± 2.38	48.61 ± 1.27
MoseDTI	75.22 ± 5.44	97.19 ± 0.88	87.59 ± 2.66
Hype-DTI	**81.32 ± 5.11**	**97.68 ± 0.81**	**92.19 ± 2.21**

Hype-DTI achieves the highest mean AUPR among the evaluated methods on all three benchmark relations, exceeding MoseDTI by 6.10 points on AGO, 0.49 on Blocker, and 4.60 on E-. The difference is largest on AGO and smallest on Blocker, where MoseDTI already reaches 97.19% AUPR.

Accuracy and AUROC show the same overall pattern. Hype-DTI reaches accuracies of 70.82 ± 8.93%, 87.94 ± 8.04%, and 75.00 ± 9.00%, with AUROC values of 79.39 ± 8.65%, 97.91 ± 0.78%, and 91.76 ± 2.01%. Full ConPLex with DUD-E yields lower AUPR on these relation splits; we treat this as a protocol-specific observation rather than evidence against contrastive learning in general. Appendix Table 5 identifies the predictor deployed in each fold.

4.3 Semi-Inductive Cold-Start Evaluation

We next evaluate a DTI-label cold-start setting by withholding labeled training interactions involving the test drug, the test target, or both. This setting examines whether pair-intrinsic features and auxiliary graph context remain informative when DTI supervision does not cover one or both endpoints.

Table 2: AGO semi-inductive AUPR (%, mean ± SD over five folds).

Method	Unseen drug	Unseen target	Unseen both
MoseDTI	70.09 ± 7.05	73.37 ± 4.51	72.39 ± 7.27
Hype-DTI	**78.86 ± 2.79**	**79.39 ± 2.86**	**81.76 ± 4.67**

Hype-DTI improves mean AUPR over MoseDTI by 8.76 points for unseen drugs, 6.02 for unseen targets, and 9.38 when both endpoints are unseen (Table 2). It scores higher in all five matched folds of each setting (Appendix Table 6). The largest mean gain occurs when both endpoints are unseen. This pattern is consistent with pair-intrinsic features providing information when labeled DTI links are unavailable, although held-out entities may still receive context through non-DTI graph relations. The experiment therefore supports semi-inductive rather than fully inductive generalization.

4.4 Large-Scale Evaluation on DrugCentral

We next evaluate Hype-DTI on DrugCentral, which provides a larger labeled training set than the low-label benchmarks. To separate comparisons by supervision, we first consider methods without additional pretraining on task-related DTI labels and report full ConPLex with DUD-E as a higher-supervision comparator.

Table 3: DrugCentral test (%, mean ± SD over five folds; bold = best matched-supervision result).

Method	Acc.	AUROC	AUPR
Without additional pretraining on task-related DTI labels			
DrugBAN	77.68 ± 0.77	85.43 ± 0.84	83.22 ± 1.16
MoseDTI	79.67 ± 0.51	87.28 ± 0.30	84.79 ± 0.63
ConPLex without DUD-E	70.82 ± 0.41	80.86 ± 0.30	79.81 ± 0.44
Hype-DTI	**81.41 ± 0.39**	**89.00 ± 0.15**	**86.78 ± 0.38**
With external DUD-E pretraining			
ConPLex with DUD-E	79.38 ± 0.96	89.46 ± 0.55	88.41 ± 0.58

Among methods without additional pretraining on task-related DTI labels, Hype-DTI achieves the highest mean accuracy, AUROC, and AUPR in Table 3. Its AUPR exceeds MoseDTI by 1.99 points, DrugBAN by 3.56, and ConPLex without DUD-E by 6.97.

Additional supervision changes the ordering. With labeled DUD-E pretraining, full ConPLex reaches 89.46 ± 0.55% AUROC and 88.41 ± 0.58% AUPR, exceeding Hype-DTI by 0.46 and 1.63 points, respectively. Hype-DTI retains higher accuracy (81.41 ± 0.39% versus 79.38 ± 0.96%). These results indicate that Hype-DTI maintains strong performance on the larger DrugCentral benchmark under matched supervision, while full ConPLex attains higher ranking metrics with access to additional labeled DUD-E pairs.

4.5 Ablation of Validation-Authorized Fusion

Finally, we isolate the deployment decision while holding the expert checkpoints fixed. We compare the full system with a variant that always deploys the learned gate; both variants use the same archived expert checkpoints within each fold.

Table 4: Archived-run test AUPR (%, five-fold mean).

Variant	AGO	Blocker	E-
Always deploy gate	79.55	**97.68**	87.94
Full Hype-DTI	**81.32**	**97.68**	**92.19**

The archived-run bundle does not include reruns for the historical representation and transfer variants, so this ablation isolates deployment authorization only. We do not combine results from incompatible training states into broader component-level claims.

Relative to always deploying the gate, validation authorization changes mean test AUPR by $+1.77$, 0.00, and $+4.25$ points on AGO, Blocker, and E-, respectively. It deploys the gate in 9 of 15 folds and falls back to the reference expert in six folds. Among those six folds, the fallback performs better on five test folds and is 0.02 points lower on one. Because selection uses validation data rather than test outcomes, authorization acts as a safeguard against unsupported fusion but cannot guarantee higher test performance in every fold. Figure 3 shows the fold-level decisions.

5 Discussion and Limitations

Reliable drug–target interaction prediction from only a handful of labeled pairs remains difficult because neither molecular structure nor graph context is trustworthy for every candidate pair. Hype-DTI addresses this challenge by treating knowledge exchange as a set of decisions that can each be rejected, rather than assuming that every transfer is beneficial. Across the three 10-shot benchmarks, it improves mean AUPR over MoseDTI by 0.49–6.10 points, and it scores higher in every matched semi-inductive cold-start fold. Notably, validation authorizes fusion in only 9 of 15 folds, yet the model still matches or exceeds its teacher. This result suggests that allowing the system to reject unsupported fusion may improve robustness, although the ablation does not isolate every transfer decision.

The strength of supervision affects how these results should be interpreted. The semi-inductive gains are consistent with pair-intrinsic features compensating for absent labeled DTI links, even though held-out entities may still appear in auxiliary non-DTI relations. On DrugCentral, Hype-DTI achieves the best accuracy, AUROC, and AUPR among methods without additional task-related DTI labels, while full ConPLex, which benefits from labeled DUD-E contrastive pretraining, remains ahead by 0.46 AUROC and 1.63 AUPR points. This highlights the performance benefit of stronger external supervision.

Limitations and Broader Impact. Additional modalities, such as three-dimensional structure and textual descriptions of biological entities, could further improve low-label DTI prediction. We leave their evaluation to future work. The staged design may be useful when developing robust models for data-scarce biological tasks.

Taken together, this work points to a broader principle: when labeled supervision is scarce, what makes a model robust is not how much information it extracts from multiple sources, but how reliably it can decide when to reject a particular transfer. By recasting knowledge exchange as a sequence of individually testable, individually vetoable decisions, Hype-DTI outperforms the baselines across all three 10-shot benchmarks and every matched semi-inductive fold, and it activates fusion only when validation clearly supports it. This "reject-first" design offers a promising template for multi-expert collaboration under data scarcity.

AI Use Statement

Generative AI tools assisted with code implementation, manuscript language polishing, and LaTeX document preparation. AI-generated code was reviewed and tested by the authors, and AI-assisted text was checked against the underlying methods and results. These tools did not generate or alter experimental measurements. The authors reviewed all AI-assisted work and take responsibility for the final content of this manuscript.

Ethics Statement

This study uses publicly available or internally assembled molecular and interaction records from non-human sources and involves no human participants, personal data, or animals. Its predictions are computational hypotheses, not clinical recommendations.

Reproducibility Statement

The anonymous package contains source, frozen configuration, all 15 split directories, archived full-run logs, and extraction utilities. It also contains a separate audited-rerun bundle covering all 45 stages that feed an optimizer, the selected deployment state, and post-filter protected-pair overlap. Appendix C gives configuration details, and Appendix E gives audit details.

References

Peizhen Bai, Filip Miljković, Bino John, and Haiping Lu. Interpretable bilinear attention network with domain adaptation improves drug–target prediction. *Nature Machine Intelligence*, 5 (2): 126–136, 2023. doi: 10.1038/s42256-022-00605-1.

Antoine Bordes, Nicolas Usunier, Alberto Garcia-Durán, Jason Weston, and Oksana Yakhnenko. Translating embeddings for modeling multi-relational data. In *Advances in Neural Information Processing Systems*, volume 26, pp. 2787–2795, 2013.

Seyone Chithrananda, Gabriel Grand, and Bharath Ramsundar. ChemBERTa: Large-scale self-supervised pretraining for molecular property prediction. *arXiv preprint arXiv:2010.09885*, 2020. URL https://arxiv.org/abs/2010.09885.

Geoffrey Hinton, Oriol Vinyals, and Jeff Dean. Distilling the knowledge in a neural network. *arXiv preprint arXiv:1503.02531*, 2015. URL https://arxiv.org/abs/1503.02531.

Kexin Huang, Cao Xiao, Lucas M. Glass, and Jimeng Sun. MolTrans: Molecular interaction transformer for drug–target interaction prediction. *Bioinformatics*, 37 (6): 830–836, 2021. doi: 10.1093/bioinformatics/btaa880.

Robert A. Jacobs, Michael I. Jordan, Steven J. Nowlan, and Geoffrey E. Hinton. Adaptive mixtures of local experts. *Neural Computation*, 3 (1): 79–87, 1991. doi: 10.1162/neco.1991.3.1.79.

Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations*, 2017. URL https://openreview.net/forum?id=SJU4ayYgl.

Zeming Lin, Halil Akin, Roshan Rao, Brian Hie, Zhongkai Zhu, Wenting Lu, Nikita Smetanin, Robert Verkuil, Ori Kabeli, Yaniv Shmueli, Allan dos Santos Costa, Maryam Fazel-Zarandi, Tom Sercu, Salvatore Candido, and Alexander Rives. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, 379 (6637): 1123–1130, 2023. doi: 10.1126/science.ade2574.

Michael M. Mysinger, Michael Carchia, John J. Irwin, and Brian K. Shoichet. Directory of useful decoys, enhanced (DUD-E): Better ligands and decoys for better benchmarking. *Journal of Medicinal Chemistry*, 55 (14): 6582–6594, 2012. doi: 10.1021/jm300687e.

Thin Nguyen, Hang Le, Thomas P. Quinn, Tri Nguyen, Thuc Duy Le, and Svetha Venkatesh. GraphDTA: Predicting drug–target binding affinity with graph neural networks. *Bioinformatics*, 37 (8): 1140–1147, 2021. doi: 10.1093/bioinformatics/btaa921.

Hakime Öztürk, Arzucan Özgür, and Elif Ozkirimli. DeepDTA: Deep drug–target binding affinity prediction. *Bioinformatics*, 34 (17): i821–i829, 2018. doi: 10.1093/bioinformatics/bty593.

Alexander Rives, Joshua Meier, Tom Sercu, Siddharth Goyal, Zeming Lin, Jason Liu, Demi Guo, Myle Ott, C. Lawrence Zitnick, Jerry Ma, and Rob Fergus. Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proceedings of the National Academy of Sciences*, 118 (15): e2016239118, 2021. doi: 10.1073/pnas.2016239118.

Rohit Singh, Samuel Sledzieski, Bryan Bryson, Lenore Cowen, and Bonnie Berger. Contrastive learning in protein language space predicts interactions between drugs and protein targets. *Proceedings of the National Academy of Sciences*, 120 (24): e2220778120, 2023. doi: 10.1073/pnas.2220778120.

Oleg Ursu, Jayme Holmes, Jeffrey Knockel, Cristian G. Bologa, Jeremy J. Yang, Stephen L. Mathias, Stuart J. Nelson, and Tudor I. Oprea. DrugCentral: Online drug compendium. *Nucleic Acids Research*, 45 (D1): D932–D939, 2017. doi: 10.1093/nar/gkw993.

Xiangxiang Zeng, Xiang Song, Tengfei Ma, Xiaoqin Pan, Yadi Zhou, Yuan Hou, Zheng Zhang, Kenli Li, George Karypis, and Feixiong Cheng. Repurpose open data to discover therapeutics for COVID-19 using deep learning. *Journal of Proteome Research*, 19 (11): 4624–4636, 2020. doi: 10.1021/acs.jproteome.0c00316.

Xinlong Zhai, Chunchen Wang, Ruijia Wang, Jiazheng Kang, Shujie Li, Boyu Chen, Tengfei Ma, Zikai Zhou, Cheng Yang, and Chuan Shi. Blend the separated: Mixture of synergistic experts for data-scarcity drug–target interaction prediction. Proceedings of the AAAI Conference on Artificial Intelligence, 39 (21): 22336–22344, 2025. doi: 10.1609/aaai.v39i21.34389.

A Fold-Level Results

Table 5: Runtime-guarded full-model test AUPR (%) by fold. G denotes the gate and F the validation-selected fallback expert. Authorization uses validation outcomes only.

Dataset	F0	F1	F2	F3	F4
AGO	79.73 (G)	81.37 (F)	88.61 (G)	82.48 (G)	74.42 (F)
Blocker	96.57 (G)	97.36 (G)	98.78 (G)	98.01 (G)	97.68 (F)
E-	92.04 (F)	88.52 (F)	92.93 (F)	94.36 (G)	93.09 (G)

Table 6: AGO semi-inductive AUPR (%) by fold.

Setting	Method	0	1	2	3	4
Unseen drug	Hype-DTI	79.30	77.24	75.50	79.30	82.95
	MoseDTI	63.16	76.23	71.88	62.21	76.99
Unseen target	Hype-DTI	81.16	78.91	74.89	79.57	82.41
	MoseDTI	77.25	69.43	67.57	76.64	75.94
Unseen both	Hype-DTI	84.21	76.86	86.83	76.70	84.22
	MoseDTI	75.74	59.38	75.34	75.51	75.96

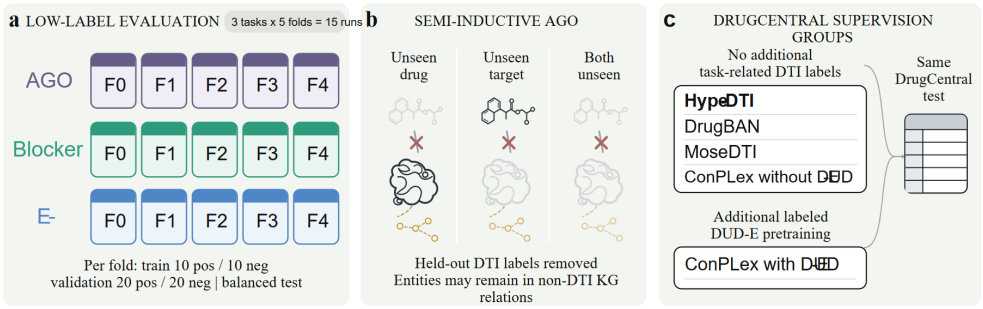

Figure 2: Evaluation and supervision protocol. The low-label benchmark comprises five folds each for AGO, Blocker, and E- (15 runs total). Semi-inductive AGO evaluation holds out drugs, targets, or both from training DTI labels while allowing their occurrence in auxiliary non-DTI graph relations. DrugCentral methods are grouped by whether they use additional labeled DUD-E interactions; all groups share the same DrugCentral test.

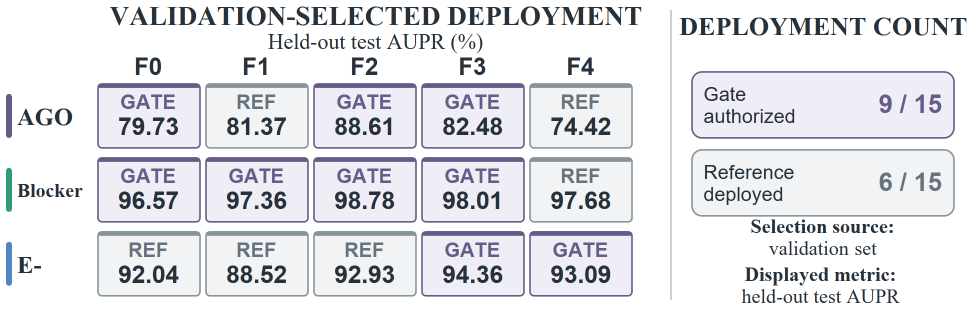

Figure 3: Validation-authorized deployment across the 15 low-label folds. Each cell reports the validation-selected deployed predictor and its held-out test AUPR; the learned gate is deployed in 9 folds and the reference expert in 6.

B Active Gate Objective

Gate pseudo-label source. The gate receives the real training pairs and a hard pseudo-labeled pool derived from both earlier scoring stages. The longer of the relation-scored and self-scored candidate lists is randomly downsampled to the length of the shorter list; the two lists are concatenated and sorted by their teacher probabilities. For a combined list S, $k = \lfloor 0.003|S| \rfloor$ top-scoring pairs are labeled positive. Negative pseudo-labels are sampled from the lower-scoring half. With n_+ real positive training pairs, the active balancing rule uses $b = 1$ when $k < n_+$ and $b = \lfloor k/n_+ \rfloor$ otherwise, and samples $k + (b-1)n_+$ pseudo-negatives while repeating the real positives b times. Every pseudo-example enters the classification loss with weight 0.1; the configured cap is 0.2.

Notation. For a gate-training minibatch B, let R and P denote its real and pseudo-labeled subsets. Let a be the validation-selected training anchor and b the other expert. Define per-example cross-entropies $\ell_i^x = \mathrm{CE}(z_{x,i}, y_i)$ and margins $m_i^x = z_{x,i,1} - z_{x,i,0}$ for $x \in \{g, a, b\}$. The expert logits, losses, and all targets derived from them are detached. Let u_i be the non-anchor mixture weight: $u_i = w_{r,i}$ for a self anchor and $u_i = 1 - w_{r,i}$ for a relation anchor. Empty-set loss terms below are set to zero in the implementation.

Classification, safety, and ranking. Writing $\omega_i = 0.1$ for a pseudo-example, the classification and real-example safety losses are

$$\begin{aligned} L_{\mathrm{cls}} &= \frac{1}{|R|}\sum_{i\in R}\ell_i^g + \frac{1}{|P|}\sum_{i\in P}\min(\omega_i, 0.2)\ell_i^g, \\ L_{\mathrm{safe}} &= \frac{1}{|R|}\sum_{i\in R}[\ell_i^g - \ell_i^a]_+ . \end{aligned} \tag{12}$$

For real positive–negative pairs, define $\Delta_{ij}^x = m_i^x - m_j^x$ and $Q = \{(i,j) : i \in R, y_i = 1;\ j \in R, y_j = 0;\ \Delta_{ij}^a > 0\}$. The active zero-margin ranking loss is

$$L_{\mathrm{rank}} = \frac{1}{|Q|}\sum_{(i,j)\in Q}\left[\Delta_{ij}^a - \Delta_{ij}^g\right]_+ . \tag{13}$$

Anchor supervision, sparsity, and budget. The detached target and its real-example upweighting are

$$\begin{aligned} t_i &= \sigma\big((\ell_i^a - \ell_i^b)/0.2\big), \qquad h_i = \mathbf{1}\big[\ell_i^b < \ell_i^a\big], \\ q_i &= 1 + 2h_i. \end{aligned} \tag{14}$$

The remaining terms are

$$
\begin{aligned}
L_{\text{anchor}} &= \frac{\sum_{i\in R} q_i \ \mathrm{BCE}(u_i, t_i)}{\sum_{i\in R} q_i}, \\
L_{\text{sparse}} &= \frac{1}{|B|}\sum_{i\in B} u_i, \qquad L_{\text{budget}} = \left[L_{\text{sparse}} - 0.2\right]_+ .
\end{aligned}
\tag{15}
$$

Contrast and anchor protection. Let $A = \{i \in R : \ell_i^a < \ell_i^b\}$ and $W = \{i \in R : \ell_i^b < \ell_i^a\}$. Then

$$
\begin{aligned}
L_{\text{protect}} &= \frac{1}{|A|}\sum_{i\in A} u_i, \\
L_{\text{contrast}} &= \left[0.05 - \frac{1}{|W|}\sum_{i\in W} u_i + \frac{1}{|A|}\sum_{i\in A} u_i\right]_+ .
\end{aligned}
\tag{16}
$$

The contrast term is zero unless both sets are nonempty. Substituting these terms into Eq. 11 yields the active coefficients 1,1,0.5,0.2,0.01,1,0.05, and 0.1 for classification, safety, ranking, anchor, sparsity, budget, contrast, and protection, respectively. Adaptive regularization is disabled; the adaptive local-cap term has zero weight; and selective-complement, supervised-complement, and local-competence branches are inactive in the archived configuration.

Anchor, reference, and deployment authorization. The training anchor is selected by the validation hierarchy, which orders AUPR, AUROC, and accuracy; exact ties default to the relation expert. An accuracy guard uses tolerance $\frac{0.2}{\sqrt{n}}$, and a selected expert whose AUPR and AUROC both reach 0.99 is replaced when its accuracy is lower. The deployment reference is selected separately by validation AUPR, with an exact-tie accuracy check and the same saturation guard. Let A_x, U_x, and C_x denote validation AUPR, AUROC, and accuracy for $x \in \{g, r\}$, where r is the reference. For $M \in \{A, U\}$, define $\widetilde{M} = \mathrm{clip}((M_g + M_r)/2, 10^{-6}, 1 - 10^{-6})$, $s_M = \sqrt{\widetilde{M}(1-\widetilde{M})/n}$, $v_M = 1/\max(s_M, 10^{-8})$, and $\alpha_M = v_M/(v_A + v_U)$. The combined gain is $G = \alpha_A(A_g - A_r) + \alpha_U(U_g - U_r)$.

The primary rule authorizes the gate only when $G \geq \max\left(0.005, \frac{0.1}{\sqrt{n}}\right)$, $U_g \geq U_r - 0.01$, and $C_g \geq C_r - \frac{0.2}{\sqrt{n}}$. A Pareto clause can waive only the combined-gain test when the two experts differ in AUPR by at least $\frac{0.2}{\sqrt{n}}$ and the gate is noninferior in reference AUPR and AUROC. A saturation clause can do the same when at least one gate or reference ranking metric reaches 0.95. The accuracy guard always remains active; if both gate ranking metrics reach 0.99, the gate must also have strictly higher accuracy than the reference.

C Frozen Configuration

Table 7: Principal active settings reconstructed from the 15 runtime-guarded AGO, Blocker, and E- runs.

Component	Setting
Relation expert	Frozen 1,000-D entity vectors; MLP $2000 \to 1000 \to 2$
Self expert	Graph/ESM-1b + ChemBERTa/ESM-2; dropout 0.2
Optimization	Adam; module and gate learning rate 10^{-3}
Weight decay	Experts 0; gate 10^{-2}
Batch size	Downstream train/validation/test 16
R-to-S transfer	Thresholds 0.9/0.1; pseudo weight 0.5; entity cap 3
S-to-R transfer	Thresholds 0.85/0.15; relation uncertainty width 0.25
Distillation	Initial 0.2; range 0.02–0.6; multipliers 1.2/0.5
Ranking term	Weight 0.5λ; margin 0.2; up to 128 sampled pairs
Gate	12 inputs; hidden width 8; at most 15 epochs
Patience	Relation stages 2; self and gate stages 5
Selection	Validation early stopping, relation rollback, gate/fallback authorization
Randomness	Seed 0; deterministic cuDNN requested

The archived JSON contains 236 non-runtime fields shared by the 15 main runs. Dataset names, devices, and output locations are runtime fields. Several inherited KGE options, including a nominal relation batch size of 1,024, are unused in the downstream Hype-DTI path and are not reported as active hyperparameters.

The relation acceptance threshold is defined in Eq. 8. It requires no decrease in validation AUROC and at most a 0.05 decrease in validation accuracy.

D Fold Dependence

Under the strict offline identity definition, which uses full InChIKey with canonical isomeric-SMILES fallback for drugs and sequence identity for targets, all 150 within-fold integrity checks report zero failures. The labeled training, validation, and test split files are pairwise disjoint within each fold, including across positive and negative labels. Across folds of the same dataset, test sets overlap strongly:

Table 8: Mean pairwise cross-fold test-set overlap.

Dataset	Overlap coefficient	Jaccard index
AGO	97.76%	95.62%
Blocker	92.89%	86.72%
E-	97.87%	95.82%

The folds mainly vary the 10-shot supervision and validation samples while evaluating largely shared pair populations. We therefore do not apply a paired t-test or treat folds as independent experimental units. With five paired outcomes, even five wins yield a two-sided sign-test value of $p = 0.0625$.

E Runtime and Strict-Identity Leakage Audit

Runtime protected-pair guard. A separate audited rerun of each of the 15 folds emitted a diagnostic record containing all required fields for self-training, relation retuning, and gate training. These audited reruns verify guard behavior; the primary metrics in Tables 1, 4, and 5 are transcribed from the archived full-run logs supplied in the anonymous package. The resulting 45 stage records show that the final assertion occurs after DataLoader construction and before the first optimizer update. Self training retained 6,780 pseudo-labeled records after removing 1,270 protected candidates; relation retuning retained 345 after removing 30; and gate training retained 20,160 with no removal. Across all stages, 27,285 pseudo-labeled records entered 35,535 total optimizer samples, with zero post-filter protected-pair overlap. A dry run also injected a violating optimizer input, produced a failure, and stopped before any optimizer step. Runtime drug matching used the conservative InChIKey connectivity block, while target matching used sequence hashes and exact identifiers. The guard protects labeled pair identities rather than excluding every held-out entity occurrence.

Post hoc strict-identity audit. A separate offline audit used full InChIKey identity with canonical isomeric-SMILES fallback and sequence-based target identity. It did not retrain the model. All 150 within-fold checks found no duplicates, split crossings, or positive–negative conflicts. The scan covered 11,326,934 triples, 81,272 unique entities, and 146 relations present in the supplied KG. It found no malformed row, direct Compound–Gene edge, validation/test direct-pair hit, or pseudo-label overlap. Across the fold-specific inventories, 21,538 evaluation-only entity records were represented through auxiliary KG relations, but none was incident to a direct Compound–Gene edge. This exposure is why the cold-start setting remains semi-inductive rather than fully inductive.

Decision boundary. The final audit decision is *CAUTION*. No DTI-label leakage was demonstrated in the supplied KG or recorded runtime stages, and runtime coverage is complete for the 15 revised main runs. However, no verified KGE training manifest was supplied. The scanned KG and checkpoint checksums identify the artifacts that were inspected, but they do not establish the checkpoint's training corpus or end-to-end command lineage. The evidence therefore does not support the stronger labels of "audit passed" or "leakage-free". The strict offline audit was post hoc, and the cold-start and DrugCentral experiments were not rerun as part of this audit bundle.

F Efficiency Evidence Boundary

Historical end-to-end wall-clock times are 135.90 ± 2.79, 130.87 ± 5.42, and 138.74 ± 4.02 minutes per fold for AGO, Blocker, and E-, respectively. These measurements combine all five stages and were not normalized to a common hardware and software stack across methods. Parameter counts, peak GPU memory, standardized throughput, and isolated test latency are unavailable, so the measurements provide evidence about practical runtime without supporting an efficiency comparison.