Enabling RNA Drug Discovery with RiboLead

Charlie
chrwn@duck.com

Abstract

In modern drug development, small molecules are engineered to attach to harmful proteins, but 85% of disease-relevant proteins are structurally undruggable (no accessible pockets). Targeting the RNA that codes for these proteins means intervention for cancers and genetic diseases before their pathogenic proteins are even created. However, discovering these RNA-targeted drugs is bottlenecked by slow, expensive physical experiments and inaccurate computational methods, which often rely on scarce 3D RNA structures. We present RiboLead, a novel deep learning platform for drug discovery that accurately predicts RNA-small molecule binding affinity. RiboLead introduces three main innovations at the levels of input, model design, and training: (1) For input, evolutionary profiles, LLM embeddings, and secondary structure capture rich biological context, replacing expensive 3D dependencies; (2) For model design, a biomimetic cross-attention mechanism lets RNA sequences and drug atoms dynamically *attend* to one another, effectively simulating ligand binding; (3) For training, an ensemble technique quantifies epistemic uncertainty to enable a semi-supervised learning strategy, where we generate new quality pairs from the same dataset, significantly expanding training horizon for improved robustness. On established benchmarks against 10 models, RiboLead achieves 96% accuracy, outperforming the state-of-the-art and reducing false positives by 32%. Crucially, the model demonstrates superior generalization, with the lowest error (MSE) across blind tests and a high-confidence prediction on a fully novel bacterial-riboswitch system. By overcoming data scarcity to provide rapid, confidence-aware screening, RiboLead accelerates the discovery of RNA-therapies for undruggable cancers, pandemic viruses, and antibiotic-resistant bacteria.

1 Introduction

RNA represents a promising new frontier for drug discovery. Currently, almost every drug works by binding a small molecule to pockets on disease-related *proteins*, neutralizing their pathogenic activity — proteins, because they are simpler and better understood. However, only approximately 1.5% of the human genome encodes proteins, and among these, the vast majority lack suitable binding pockets for small molecules; they are *undruggable* [1, 2]. In fact, currently approved drugs interact therapeutically with fewer than 700 gene products, meaning that only 0.05% of the human genome has been successfully drugged [10]. On the other hand, if we can find new drugs that target RNA instead of proteins, we unlock treatments to vast new diseases: 70% of our genome, compared to the 0.05%.

This approach has already been validated by Risdiplam, an oral small molecule that binds to *SMN2* pre-mRNA and corrects its splicing, saves infants from debilitating spinal muscular atrophy [3]. Applying the same method, we can target, for instance, c-MYC, the master regulator of tumor growth in most cancers, which is undruggable at the protein level but has predictable RNA targets that can. Morever, neurodegenerative diseases like ALS and Huntington's are caused by toxic RNA structures, which can be stabilized by binding small molecules. Despite this enormous potential, Risdiplam remains a rare exception.

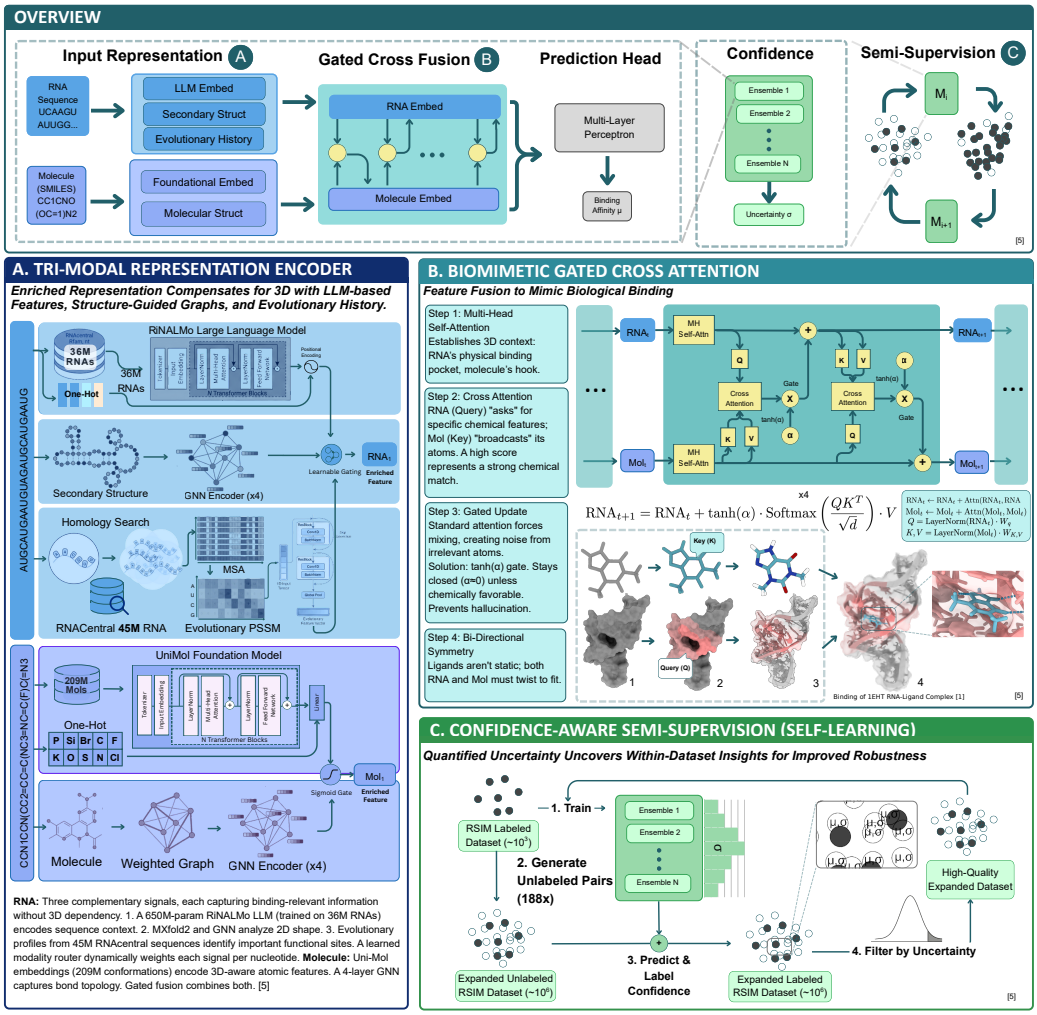

Figure 1: **RiboLead Architecture.** (A) **Tri-Modal Representation Encoder:** RNA sequence (RiNALMo), structure (MXfold2), and evolutionary profiles (PSSM) are fused via a learnable modality router. Small molecules are encoded via Uni-Mol and graph features. (B) **Biomimetic Gated Cross-Attention:** A bidirectional mechanism that mimics induced-fit binding. Gating parameters (α) initialized to zero allow the model to learn unimodal features before integrating cross-modal signals. (C) **Confidence-Aware Semi-Supervision:** An ensemble of models estimates epistemic uncertainty to selectively pseudo-label high-confidence unlabeled data, expanding the effective training set.

The discovery of such compounds is currently bottlenecked by the immense cost of experimentally screening billions of potential ligands against complex RNA shapes. There is an urgent need for computational tools to narrow down the chemical space, yet computational methods that worked for proteins remain impractical for RNA due to three unique challenges. First, 3D RNA structures are extremely scarce (9,000 RNAs versus 260,000 proteins), severely limiting the efficacy of docking tools and deep learning models reliant on 3D data [5]. Second, sequence-based methods like DeepRSMA [16] are scalable but unable to effectively compensate for lack of 3D inputs and fuse molecular representations with naive cross-attention mechanisms that introduce noise. Third, and critically, current models are unreliable on out-of-distribution data. They confidently report incorrect predictions, leading to many false-positives and higher expenses. Together, these limitations—dependence on scarce 3D data, naive signal fusion, and unreliable single-point predictions—render current approaches inapplicable to real-world drug discovery pipelines.

We observe that while explicit 3D structural data is scarce, complementary signals are both abundant and highly informative. Evolutionary conservation, encoded in position-specific scoring matrices (PSSMs), captures functional constraints shaped by millions of years of selection. Computationally predicted secondary structure (2D) provides geometric context at a fraction of the cost of 3D modeling. This mirrors the foundational insight of AlphaFold2, where multiple sequence alignments encode structural information without explicit 3D supervision [7]. We hypothesize that an effective latent fusion of sequence, secondary structure, and evolutionary profiles can achieve the accuracy of a successful structure-aware method (not yet available for RNA) while maintaining the scalability of sequence-only approaches. Moreover, by training ensemble models that produce calibrated epistemic uncertainty estimates, we can simultaneously improve learning through uncertainty-guided pseudo-labeling and enable clinical triage during deployment.

Contributions. In this work, we introduce *RiboLead*, a deep learning system that predicts RNA-drug binding accurately and robustly enough for real-world drug discovery deployment. Our key contributions are:

1. **Tri-Modal Representation Encoding (Input)** We bypass the reliance on 3D structures by implementing the first architecture that combines sequence embeddings (RiNALMo [21]), predicted secondary structure (MXfold2 [22]), and evolutionary profiles (MMseqs2 [24]) for RNA-ligand affinity prediction, for accurate and scalable predictions.
2. **Biomimetic Gated Cross-Attention (Fusion).** Adapting the gated cross-attention mechanism from vision-language models [32], we introduce a bidirectional co-attention block that allows RNA and ligand features to mutually update and gate each other, suppressing noisy modalities and mimicking the physical induced-fit process.
3. **Confidence-Aware Semi-Supervision (Training).** To enable applicability to real-world drug discovery, we introduce a semi-supervised learning approach that utilizes epistemic uncertainty estimates to safely train on unlabeled data, expanding RiboLead's generalizability. Additionally, the epistemic uncertainty score enables reliable clinical prioritization, where uncertain predictions can be flagged for review, reducing cost of false positives.

2 Related Work

2.1 3D Structure-Dependent Methods

Traditional structure-based drug design that worked well for proteins relies heavily on the availability of high-resolution 3D coordinates. Classical docking tools like AutoDock Vina [8] and rDock treat the receptor as a rigid body, searching for ligand poses that minimize an empirical scoring function. While effective for stable protein pockets, these methods fail to be adapted to RNA's complexity.

More recently, geometric deep learning has emerged as an encouraging approach. Models like EquiBind [14], DiffDock [15], and AlphaFold3 [6] have shown great success for protein-ligand prediction by explicitly modeling 3D geometry. For our RNA problem, models like GerNA-Bind [13] have implemented similar ideas, using structural interactions to predict binding. However, these methods are fundamentally constrained by the lack of high-quality RNA crystal structures. With only 9,000 ground-truth RNA structures to learn from (vs. 260,000 proteins), geometric models struggle to generalize to novel RNA folds, necessary for practical utility. RiboLead addresses this by substituting scarce 3D coordinates with a **tri-modal encoding** of rich and abundant alternative sources like evolutionary history to successfully infer 3D structure while remaining scalable.

2.2 Sequence and Graph-Based Interaction Methods

To circumvent 3D structure limitations, researchers have developed methods using only sequence and graph representations. Early approaches utilized traditional machine learning algorithms like Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and XGBoost to predict binding affinity from hand-crafted molecular descriptors [16].

Deep learning subsequently advanced this field by allowing end-to-end learning of interaction features. In the protein-ligand domain, models like DeepCDA [28] demonstrated effective affinity prediction using CNNs and LSTMs with cross-domain adaptation, while DeepDTAF [29] utilized dilated convolutions to capture multi-scale context. Graph-based methods such as GraphDTA [18] and GAT

architectures further improved performance by representing small molecules as topological graphs (nodes as atoms, edges as bonds), capturing connectivity information lost in 1D representations.

To apply these advances to RNA, models like DeepRSMA [16] have adapted similar architectures for modeling RNA-ligand interactions. While they do not require explicit 3D structure and perform better than classical baselines on benchmarks like R-SIM, they face another issue that prevents accurate binding affinity predictions, namely fusion. Inputs (RNA and ligand) must be combined effectively for predictions to be accurate, called **fusion**. Without 3D structure, fusion becomes ever more important, yet standard cross attention and outer product introduce noise by forcing interaction between every possible pair of RNA and ligand features. To resolve this, we introduce **Gated Co-Attention**, a mechanism that dynamically filters noisy interactions and mimics the induced-fit process by allowing bidirectional updates between RNA and ligand representations.

2.3 Semi-Supervised and Uncertainty-Aware Learning

In addition to architecture, the training method should be carefully reconsidered in such a data-scarce domain. All top models to date use supervised learning, yet labeled RNA-ligand binding data is a small fraction of protein-ligand data. In broader drug discovery, semi-supervised learning (SSL) and active learning have been proposed to leverage large pools of unlabeled data. Techniques like consistency regularization or pseudo-labeling can theoretically improve generalization by forcing the model to be consistent on unlabeled examples.

However, naive implementation of SSL is very risky in complex biological problems. Without a rigorous quality control, models may train on their own confident but false hallucinations, causing model collapse. Existing methods for RNA affinity prediction usually provide only a single point estimate that fails to distinguish between confident predictions and blind guesses, making it unreliable on novel applications. To bridge this gap, RiboLead creates **ensemble-derived uncertainty estimates to inform semi-supervised learning**, enabling expansion of the training set and reliable clinical prioritization.

3 Methods

3.1 Problem Formulation

We formulate RNA-small molecule binding affinity prediction as a regression task. Given an RNA sequence $\mathcal{R}$ of length L_r and a small molecule $\mathcal{M}$ with L_m atoms, our goal is to predict the binding affinity $y \in \mathbb{R}$, typically expressed as pK_d (the negative logarithm of the dissociation constant). Formally, we seek to learn a function:

$$f_\theta : (\mathcal{R}, \mathcal{M}) \mapsto \hat{y} \tag{1}$$

where $\hat{y}$ is the predicted affinity.

Unlike prior approaches that require explicit 3D coordinates, RiboLead operates on representations derivable from sequence alone: primary sequence, predicted secondary structure, and evolutionary profiles. This design choice eliminates the 3D structure prediction bottleneck that currently limits high-throughput virtual screening.

3.2 Data and Preprocessing

3.2.1 Dataset

We train and evaluate on the R-SIM dataset [20], which contains 1,439 RNA-small molecule interactions with experimentally measured binding affinities spanning 341 unique RNA sequences and 792 unique small molecules. Binding affinities are normalized to zero mean and unit variance for stable training:

$$y_\text{norm} = \frac{y - \mu_\text{train}}{\sigma_\text{train}} \tag{2}$$

3.2.2 RNA Representation Extraction

For each RNA sequence, we extract three complementary representations:

Sequence Embeddings. We employ RiNALMo [21], a 650M-parameter RNA language model pre-trained on 36 million non-coding RNA sequences. For an RNA sequence of length L_r, RiNALMo produces contextualized embeddings $\mathbf{E}_\text{seq} \in \mathbb{R}^{L_r \times 1280}$ that capture long-range dependencies and evolutionary constraints learned from massive unlabeled data.

Secondary Structure. We predict base-pairing probabilities using MXfold2 [22], a deep learning-based structure prediction method. From the predicted structure, we extract base-pair edges $\mathcal{E}_\text{bp} = \{(i, j) : \text{nucleotides } i \text{ and } j \text{ are paired}\}$. These edges define the topology of a graph where nodes are nucleotides, enabling our model to reason about structural context without requiring 3D coordinates.

Evolutionary Profiles (PSSM). To capture evolutionary conservation—a strong indicator of functional importance—we construct position-specific scoring matrices (PSSMs) by querying each RNA sequence against RNAcentral [23] using MMseqs2 [24]. The resulting PSSM $\mathbf{P} \in \mathbb{R}^{L_r \times 5}$ encodes the frequency of each nucleotide (A, C, G, U) plus gaps at each position across homologous sequences.

3.2.3 Molecule Representation Extraction

For small molecules, we extract two complementary representations:

3D-Aware Atom Embeddings. We employ Uni-Mol [25], a transformer-based molecular representation model pre-trained on 209M molecular conformations. Given a SMILES string, we generate 3D conformers and extract per-atom embeddings $\mathbf{E}_\text{mol} \in \mathbb{R}^{L_m \times 512}$. Critically, we add explicit hydrogens before embedding extraction, as hydrogen bonding is central to RNA-ligand recognition.

Molecular Graph. Using RDKit [26], we construct a molecular graph where nodes are atoms and edges represent covalent bonds. Atoms are encoded as one-hot vectors over a vocabulary of 14 atom types (C, N, O, S, F, Cl, Br, I, P, H, B, Si, Se, unknown), yielding $\mathbf{A}_\text{oh} \in \mathbb{R}^{L_m \times 14}$. Bond edges $\mathcal{E}_\text{bond}$ with associated bond types enable message passing over molecular topology.

3.3 Model Architecture

RiboLead processes RNA and molecule inputs through modality-specific encoders, fuses information via gated cross-attention, and predicts binding affinity. Figure 1 provides an overview. We detail each component below.

3.3.1 RNA Feature Extraction

The RNA encoder processes three input modalities in parallel, then fuses them via a learned routing mechanism (Figure 1A).

Sequence Branch. We project RiNALMo embeddings and one-hot encodings to a shared dimension $d = 256$:

$$\mathbf{x}_{\text{seq}} = \mathbf{W}_{\text{llm}}\mathbf{E}_{\text{seq}} + \mathbf{W}_{\text{oh}}\mathbf{R}_{\text{oh}} + \mathbf{PE} \tag{3}$$

where $\mathbf{W}_{\text{llm}} \in \mathbb{R}^{1280\times256}$, $\mathbf{W}_{\text{oh}} \in \mathbb{R}^{5\times256}$, $\mathbf{R}_{\text{oh}} \in \mathbb{R}^{L_r\times5}$ is the nucleotide one-hot encoding, and $\mathbf{PE}$ denotes sinusoidal positional encodings. The combined representation passes through a 2-layer Transformer encoder with 8 attention heads, producing $\mathbf{h}_1 \in \mathbb{R}^{L_r\times256}$.

Structure Branch. We construct a graph where nucleotides are nodes connected by backbone edges (sequential neighbors) and base-pair edges (from MXfold2 predictions). Edge features combine edge type (backbone vs. base-pair) and sequence distance:

$$\mathbf{e}_{ij} = \mathbf{E}_{\text{type}}[t_{ij}] + \mathbf{E}_{\text{dist}}[\text{bucket}(|i-j|)] \tag{4}$$

where $\mathbf{E}_{\text{type}} \in \mathbb{R}^{2\times64}$ and $\mathbf{E}_{\text{dist}} \in \mathbb{R}^{16\times64}$ are learned embeddings, with distances bucketed logarithmically. We apply 4 layers of TransformerConv [27] with 4 attention heads, using GraphNorm [30] for stable training across variable-length sequences. This yields structure-aware representations $\mathbf{h}_2 \in \mathbb{R}^{L_r\times256}$.

Evolutionary Branch. PSSMs are processed by a residual 1D-CNN with dilated convolutions (dilations 1, 2, 4) to capture multi-scale evolutionary patterns:

$$\mathbf{h}_3 = \text{ResCNN}(\mathbf{P}) \in \mathbb{R}^{L_r\times256} \tag{5}$$

GroupNorm replaces BatchNorm for stability with variable sequence lengths.

Modality Router Fusion. A key architectural contribution is our *ModalityRouterFusion* mechanism, which learns position-specific optimal combinations of the three RNA modalities. This reflects the biological reality that different nucleotide positions contribute variably across modalities: some positions are important for shape (structure), others for conservation (evolution), and others for sequence context.

Given the three modality representations $\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3 \in \mathbb{R}^{L_r\times256}$, we compute fusion in two complementary pathways:

Attention-based fusion: We treat modalities as a key-value set and learn to attend over them:

$$\mathbf{H} = \text{stack}(\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3) + \mathbf{E}_{\text{mod}} \in \mathbb{R}^{L_r\times3\times256} \tag{6}$$

$$\mathbf{q} = \mathbf{W}_q[\mathbf{h}_1 \oplus \mathbf{h}_2 \oplus \mathbf{h}_3] \in \mathbb{R}^{L_r\times256} \tag{7}$$

$$\mathbf{f}_{\text{attn}} = \text{MultiHeadAttn}(\mathbf{q}, \mathbf{H}, \mathbf{H}) \tag{8}$$

where $\mathbf{E}_{\text{mod}} \in \mathbb{R}^{3\times256}$ are learned modality embeddings.

Gated mixture-of-experts: We compute soft routing weights via a gating network:

$$\mathbf{g} = \text{softmax}\left(\mathbf{W}_g[\mathbf{h}_1 \oplus \mathbf{h}_2 \oplus \mathbf{h}_3]\right) \in \mathbb{R}^{L_r\times3} \tag{9}$$

$$\mathbf{f}_{\text{gate}} = g_1 \cdot \mathbf{h}_1 + g_2 \cdot \mathbf{h}_2 + g_3 \cdot \mathbf{h}_3 \tag{10}$$

The final fused representation combines both pathways:

$$\mathbf{f}_{\text{RNA}} = \text{LayerNorm}(\mathbf{f}_{\text{attn}} + \mathbf{f}_{\text{gate}}) \in \mathbb{R}^{L_r\times256} \tag{11}$$

Critically, the gate weights $\mathbf{g} \in \mathbb{R}^{L_r \times 3}$ are *interpretable*: they reveal which modality the model relies on at each nucleotide position, enabling post-hoc analysis of binding site characteristics.

The fused representation is projected to $d_{\text{fusion}} = 512$ and refined by a 2-layer Transformer, yielding final RNA features $\mathbf{X}_{\text{RNA}} \in \mathbb{R}^{L_r \times 512}$.

3.3.2 Molecule Feature Extraction

The molecule encoder follows a parallel dual-branch architecture:

Sequence Adapter. We adapt frozen Uni-Mol embeddings with a lightweight bottleneck module:

$$\mathbf{y}_1 = \mathbf{W}_{\text{llm}}\mathbf{E}_{\text{mol}} + \mathbf{W}_{\text{oh}}\mathbf{A}_{\text{oh}} + \text{Adapter}(\cdot) \tag{12}$$

where the adapter consists of down : $256 \to 64$, GELU, up : $64 \to 256$ projections. Following adapter tuning best practices [31], we initialize the up-projection to zero, ensuring the adapter contributes minimally at initialization and allowing pretrained features to remain stable during early training.

Graph Branch. We process the molecular graph with 4 layers of TransformerConv, producing $\mathbf{y}_2 \in \mathbb{R}^{L_m \times 256}$. Edge types (single, double, triple, aromatic bonds) are embedded and incorporated as edge features.

Gated Fusion. We combine branches via learned gating:

$$\alpha = \sigma(\mathbf{W}_\alpha[\mathbf{y}_1 \oplus \mathbf{y}_2]) \tag{13}$$

$$\mathbf{y}_{\text{fused}} = (1-\alpha)\cdot\mathbf{y}_1 + \alpha\cdot\mathbf{y}_2 \tag{14}$$

The final molecule features are $\mathbf{X}_{\text{mol}} = \text{LayerNorm}([\mathbf{y}_{\text{fused}} \oplus \mathbf{y}_2]) \in \mathbb{R}^{L_m \times 512}$.

3.3.3 Gated Cross-Modal Co-Attention

To model RNA-ligand interactions, we employ a symmetric bidirectional cross-attention mechanism inspired by Flamingo [32]. Unlike sequential cross-attention where update order introduces bias, our approach updates both modalities simultaneously.

Each co-attention layer consists of paired blocks:

$$\mathbf{X}'_{\text{RNA}} = \text{GatedCrossAttnBlock}(\mathbf{X}_{\text{RNA}}, \mathbf{X}_{\text{mol}}) \tag{15}$$

$$\mathbf{X}'_{\text{mol}} = \text{GatedCrossAttnBlock}(\mathbf{X}_{\text{mol}}, \mathbf{X}_{\text{RNA}}) \tag{16}$$

Within each block, we apply self-attention, gated cross-attention, and a feed-forward network:

$$\mathbf{z} = \mathbf{x} + \text{SelfAttn}(\text{LN}(\mathbf{x})) \tag{17}$$

$$\mathbf{z}' = \mathbf{z} + \tanh(\alpha)\cdot\text{CrossAttn}(\text{LN}(\mathbf{z}), \text{LN}(\mathbf{c})) \tag{18}$$

$$\mathbf{x}' = \mathbf{z}' + \text{FFN}(\text{LN}(\mathbf{z}')) \tag{19}$$

where $\mathbf{x}$ is the query modality, $\mathbf{c}$ is the context modality, and α is a *learnable scalar initialized to zero*.

The zero-initialization of α is critical: at the start of training, cross-attention contributes nothing ($\tanh(0) = 0$), allowing the model to first learn stable unimodal representations before gradually incorporating cross-modal information as α grows. This provides implicit curriculum learning without explicit scheduling.

We stack $N_{\text{co}} = 2$ co-attention layers (Figure 1B), enabling iterative refinement of RNA-molecule representations based on mutual context.

3.3.4 Pooling and Prediction

We aggregate variable-length sequence representations via attention-weighted pooling:

$$\alpha_i = \frac{\exp(\mathbf{W}_s\mathbf{x}_i)}{\sum_j \exp(\mathbf{W}_s\mathbf{x}_j)} \tag{20}$$

$$\bar{\mathbf{x}} = \sum_i \alpha_i \mathbf{x}_i \tag{21}$$

applied separately to RNA and molecule features, yielding $\bar{\mathbf{x}}_{\text{RNA}}, \bar{\mathbf{x}}_{\text{mol}} \in \mathbb{R}^{512}$.

The concatenated representation $[\bar{\mathbf{x}}_{\text{RNA}} \oplus \bar{\mathbf{x}}_{\text{mol}}] \in \mathbb{R}^{1024}$ passes through a 2-layer MLP to predict affinity:

$$\hat{y} = \mathbf{W}_2 \cdot \text{GELU}(\mathbf{W}_1[\bar{\mathbf{x}}_{\text{RNA}} \oplus \bar{\mathbf{x}}_{\text{mol}}]) \tag{22}$$

3.4 Training

3.4.1 Loss Function

We optimize a composite loss combining mean squared error (MSE) and concordance correlation coefficient (CCC) loss:

$$\mathcal{L} = \frac{1}{2}\left(\text{MSE}(\hat{y}, y) + \mathcal{L}_{\text{CCC}}(\hat{y}, y)\right) \tag{23}$$

where the CCC loss encourages both correlation and agreement in scale:

$$\mathcal{L}_{\text{CCC}} = 1 - \frac{2\rho\sigma_{\hat{y}}\sigma_y}{\sigma_{\hat{y}}^2 + \sigma_y^2 + (\mu_{\hat{y}} - \mu_y)^2} \tag{24}$$

with ρ the Pearson correlation, σ the standard deviation, and μ the mean over the batch. The CCC term prevents the pathological solution of predicting the mean while achieving zero correlation penalty.

3.4.2 Ensemble Training

We train an ensemble of $K = 5$ models with different random initializations. Each ensemble member is trained independently on the same data splits, with diversity arising from:

- Random weight initialization
- Stochastic gradient descent with different random seeds
- Dropout regularization (rate 0.2) during training

At inference, we compute epistemic uncertainty post-hoc from ensemble disagreement:

$$\hat{y} = \frac{1}{K}\sum_{k=1}^{K} \hat{y}^{(k)} \tag{25}$$

$$\hat{\sigma}^2 = \frac{1}{K}\sum_{k=1}^{K} \left(\hat{y}^{(k)} - \hat{y}\right)^2 \tag{26}$$

where $\hat{\sigma}$ captures *epistemic uncertainty*—disagreement among ensemble members indicative of out-of-distribution inputs or regions of high model uncertainty. This is distinct from aleatoric uncertainty (inherent data noise); epistemic uncertainty specifically measures what the models do not know, making it suitable for filtering unreliable predictions.

3.4.3 Uncertainty-Guided Semi-Supervised Learning

Recognizing the severe scarcity of labeled RNA-ligand data (234 pairs in PDBbind vs. 19,000+ for proteins), we employ a semi-supervised learning strategy that leverages unlabeled RNA-molecule pairs while guarding against noisy pseudo-labels.

Our approach (Algorithm 1, Figure 1C) operates in two phases:

Phase 1: Pseudo-label generation. The pre-trained teacher ensemble generates predictions for novel RNA-molecule pairs sampled from the combinatorial space of known RNAs and molecules. For each pair, we compute the ensemble mean μ and epistemic uncertainty σ (standard deviation across ensemble members). Only pairs with $\sigma < \tau$ (we use $\tau = 0.18$) are retained, ensuring pseudo-labels derive from high-confidence predictions where ensemble members agree.

Algorithm 1 Uncertainty-Guided Semi-Supervised Training

Require: Labeled dataset $\mathcal{D}_L$, unlabeled pairs $\mathcal{U}$, teacher ensemble $\{f_k\}_{k=1}^{K}$
Require: Confidence threshold τ, pseudo-label weight λ
Train teacher ensemble on $\mathcal{D}_L$
// Generate pseudo-labels
for each $(r, m) \in \mathcal{U}$ **do**
  $\hat{y}^{(k)} \leftarrow f_k(r, m)$ for $k = 1, \ldots, K$
  $\mu \leftarrow \text{mean}(\{\hat{y}^{(k)}\}), \quad \sigma \leftarrow \text{std}(\{\hat{y}^{(k)}\})$
  if $\sigma < \tau$ **then**
    Add $(r, m, \mu, w = e^{-\sigma})$ to pseudo-labeled set $\mathcal{D}_P$
  end if
end for
// Train student
for each epoch **do**
  for each batch $\mathcal{B}_L \sim \mathcal{D}_L$, $\mathcal{B}_P \sim \mathcal{D}_P$ **do**
    $\mathcal{L}_L \leftarrow \frac{1}{|\mathcal{B}_L|} \sum_{(r,m,y)\in\mathcal{B}_L} \ell(f(r, m), y)$
    $\mathcal{L}_P \leftarrow \frac{1}{|\mathcal{B}_P|} \sum_{(r,m,\mu,w)\in\mathcal{B}_P} w \cdot (f(r, m) - \mu)^2$
    Update θ via $\nabla(\mathcal{L}_L + \lambda\mathcal{L}_P)$
  end for
end for
return Student model f

Phase 2: Student training. A student model trains on the union of labeled data and filtered pseudo-labels. Pseudo-labeled samples are weighted by $w = e^{-\sigma}$, down-weighting predictions with higher (but still acceptable) epistemic uncertainty. The pseudo-label loss is scaled by $\lambda = 0.03$ to prevent overwhelming the true labels.

This uncertainty-guided filtering is essential: naive pseudo-labeling degrades performance (Table 2), while our confidence-filtered approach consistently improves generalization, particularly on blind RNA splits where no RNA in the test set appears during training.

3.4.4 Training Details

We train with AdamW optimizer (learning rate 2×10^{-4}, weight decay 10^{-2}) using mixed-precision (FP16) for efficiency. Gradient clipping at norm 1.0 prevents instabilities. Models train for 60 epochs with early stopping (patience 20) based on validation Pearson correlation. Batch size is 32 for labeled data. All experiments run on Google Colab (free student pro) using NVIDIA L4 GPU.

4 Results

4.1 Superior classification performance for virtual screening

We first evaluated classification performance under stratified 5-fold cross-validation, the most relevant metric for virtual screening where the goal is to identify true binders from large compound libraries. Figure 2 compares RiboLead against ten baseline methods.

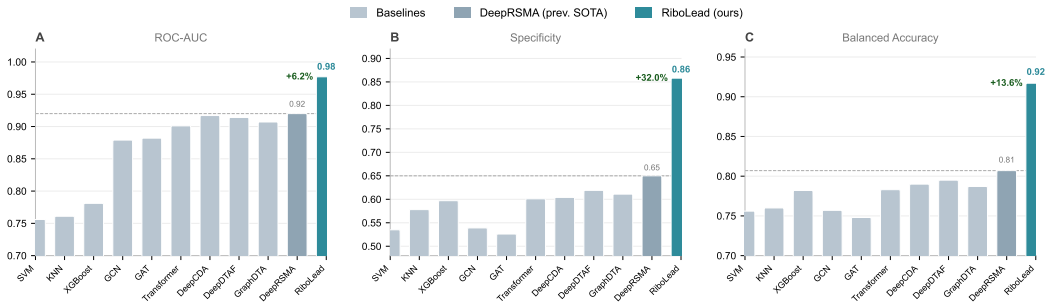

Figure 2: **RiboLead achieves dramatically superior classification performance.** (A) ROC-AUC, (B) Specificity, and (C) Balanced Accuracy under stratified 5-fold CV. Dashed lines indicate DeepRSMA (previous SOTA). RiboLead demonstrates +6.2% AUC, +32.0% specificity, and +13.6% balanced accuracy improvements.

RiboLead achieves an overall accuracy of 95.5%, correctly classifying 96% of RNA-ligand binding affinities. The ROC-AUC of 0.98 indicates that the model correctly ranks a true binding compound higher than a non-binder 98% of the time. Most strikingly, specificity reaches 85.8% compared to DeepRSMA's 65.0%—a 32% relative improvement that translates directly to far fewer false positives in virtual screening campaigns, where experimental validation of false positives is costly. Balanced accuracy (91.7% vs. 80.7%) and precision (97.2%) further confirm robust discriminative power across both positive and negative classes (Table 1).

Table 1: Complete classification metrics for RiboLead across evaluation settings.

Metric	Accuracy	Balanced Acc.	Precision	Recall	F1 Score	MCC
5-Fold Stratified	0.96	0.92	0.97	0.97	0.97	0.84
Blind RNA	0.85	0.60	0.86	0.98	0.92	0.33
Blind Mol	0.89	0.71	0.90	0.98	0.94	0.55

Crucially, this classification advantage extends to held-out settings: RiboLead achieves ROC-AUC of 0.78 on completely novel RNA targets and 0.87 on novel drug-like molecules (blind tests), demonstrating that the learned representations generalize beyond the training distribution.

4.2 Benchmarking regression accuracy and generalization

For regression tasks, we evaluated root mean squared error (RMSE) and Pearson correlation coefficient (PCC) across stratified CV and two blind test settings (Figure 3).

Under stratified cross-validation, RiboLead achieves RMSE = 0.87 ± 0.08 and PCC = 0.78 ± 0.04, improving upon the previous state-of-the-art (DeepRSMA: RMSE = 0.90, PCC = 0.78). The main differentiation appears under distribution shift: in the blind RNA setting, where the model must generalize to entirely unseen RNA targets, RiboLead achieves 11% lower RMSE (1.03 vs. 1.16) and +11% PCC improvement (0.65 vs. 0.58). Performance remains robust in the blind molecule setting, with 7% RMSE reduction (0.97 vs. 1.04) and comparable PCC (0.70 vs. 0.70).

To verify these gains reflect genuine generalization, we analyzed prediction error as a function of similarity to nearest training examples (Fig. 4). The negligible correlations ($|r| < 0.15$ for both RNA k-mer and molecular Tanimoto similarity) confirm that the model learns transferable interaction principles rather than memorizing local neighborhoods.

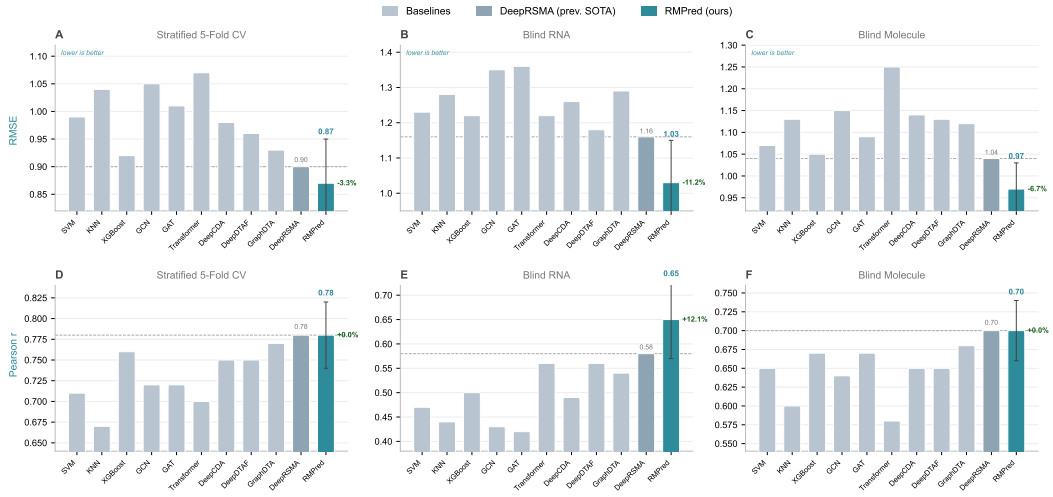

Figure 3: **Regression performance across evaluation settings.** (A–C) RMSE in pK_d units. (D–F) Pearson correlation coefficient. Under stratified CV, RiboLead matches DeepRSMA (RMSE = 0.87, PCC = 0.78). Under distribution shift, RiboLead shows clear advantages: 11% RMSE reduction and +11% PCC improvement for blind RNA.

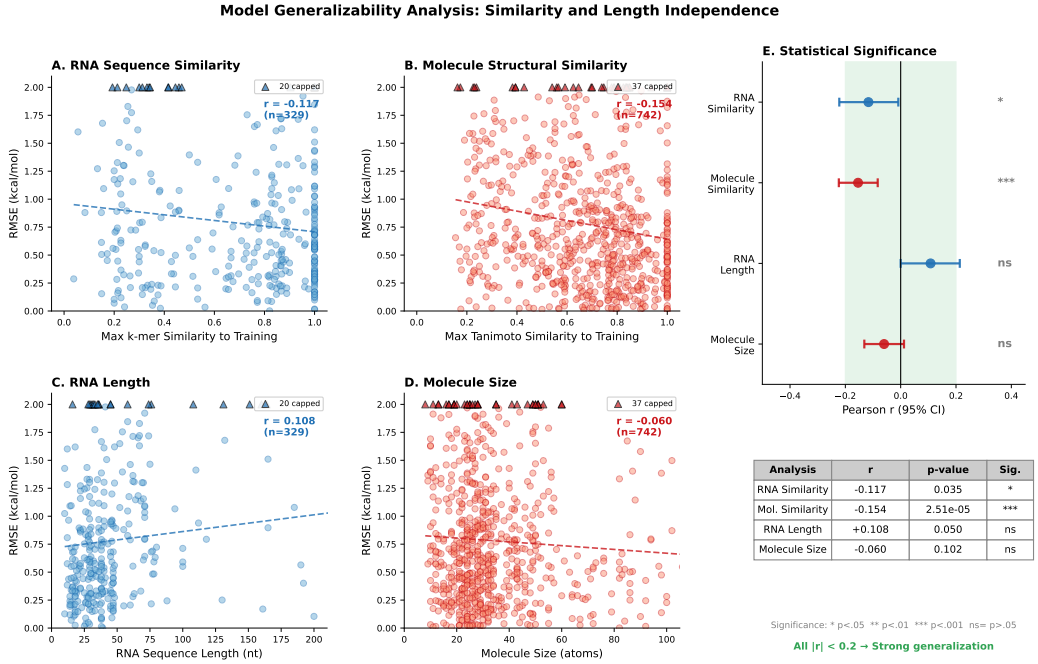

Figure 4: **Model accuracy is decoupled from training set proximity.** (A–B) Prediction error versus similarity to nearest training example shows negligible correlation. (C–D) Error remains stable across RNA lengths and molecule sizes.

4.3 Architectural ablations reveal component contributions

We systematically ablated three core components and evaluated on the challenging blind RNA setting (Table 2).

Table 2: Ablation study on blind RNA test.

Variant	PCC ↑	RMSE ↓	Removed
One-Hot Only	0.31 ± 0.12	1.25 ± 0.17	Multi-modal encoding
Linear Fusion	0.54 ± 0.16	1.17 ± 0.20	Gated Co-Attention
MSE-Single	0.51 ± 0.13	1.21 ± 0.16	Uncertainty ensemble
Full	$\mathbf{0.65 \pm 0.08}$	$\mathbf{1.03 \pm 0.12}$	–

The most severe degradation (-52% PCC) occurs when replacing tri-modal RNA encoding with one-hot nucleotide vectors, confirming that semantic and structural context are essential. Removing gated co-attention reduces PCC by 17%, and removing uncertainty-guided training causes a 22% drop.

4.4 Semi-supervised distillation improves robustness

Uncertainty-filtered pseudo-labeling, where only high-confidence Teacher predictions augment the Student's training set, yields consistent improvements (Table 3). The largest gain (+4.7% PCC) occurs in the blind RNA setting.

Table 3: Teacher-Student comparison across evaluation settings.

Setting	Model	PCC ↑	RMSE ↓
Stratified CV	Teacher	0.77 ± 0.05	0.84 ± 0.08
	Student	$\mathbf{0.78 \pm 0.04}$	0.87 ± 0.08
Blind RNA	Teacher	0.62 ± 0.08	1.04 ± 0.12
	Student	$\mathbf{0.65 \pm 0.08}$	$\mathbf{1.03 \pm 0.12}$
Blind Molecule	Teacher	0.69 ± 0.04	0.98 ± 0.06
	Student	$\mathbf{0.70 \pm 0.04}$	$\mathbf{0.97 \pm 0.06}$

4.5 Calibrated uncertainty enables actionable predictions

The model outputs calibrated uncertainty scores via ensemble training (Fig. 5). An Expected Calibration Error of 0.127 indicates statistically reliable confidence intervals. Filtering the 20% most uncertain predictions reduces RMSE by 5%, while retaining 85% of true high-affinity binders.

4.6 Generalization across disease contexts

Disease-stratified blind tests on HIV/AIDS, where the relevant targets are held out during training, demonstrate transfer to clinically-relevant targets (Table 4).

Table 4: Performance on generalization to HIV-associated targets.

Held-Out	N	RMSE ↓	PCC ↑
Disease: AIDS	394	1.11	0.61
Pathogen: HIV-1	374	1.06	0.63

4.7 Prospective validation on an emerging antibiotic target

To demonstrate real-world utility beyond benchmarks, we performed a fully *blind* evaluation on a recently solved riboswitch–ligand complex: the *Clostridium beijerinckii* ZTP riboswitch co-crystallized with its cognate AICA ligand (PDB: 9BZ1; X-ray, 2.8 Å; released November 27, 2024) [33].

Therapeutic context. Riboswitches are RNA-based regulatory elements found predominantly in bacteria but absent in humans, making them attractive targets for developing species-selective antibiotics [34]. The ZTP riboswitch class is particularly compelling: it senses ZMP and ZTP—metabolic intermediates that accumulate when bacteria experience folate stress—and controls genes essential for one-carbon metabolism and purine biosynthesis [35, 36]. Compounds that lock ZTP riboswitches into non-functional conformations could starve pathogens of essential metabolites while leaving human

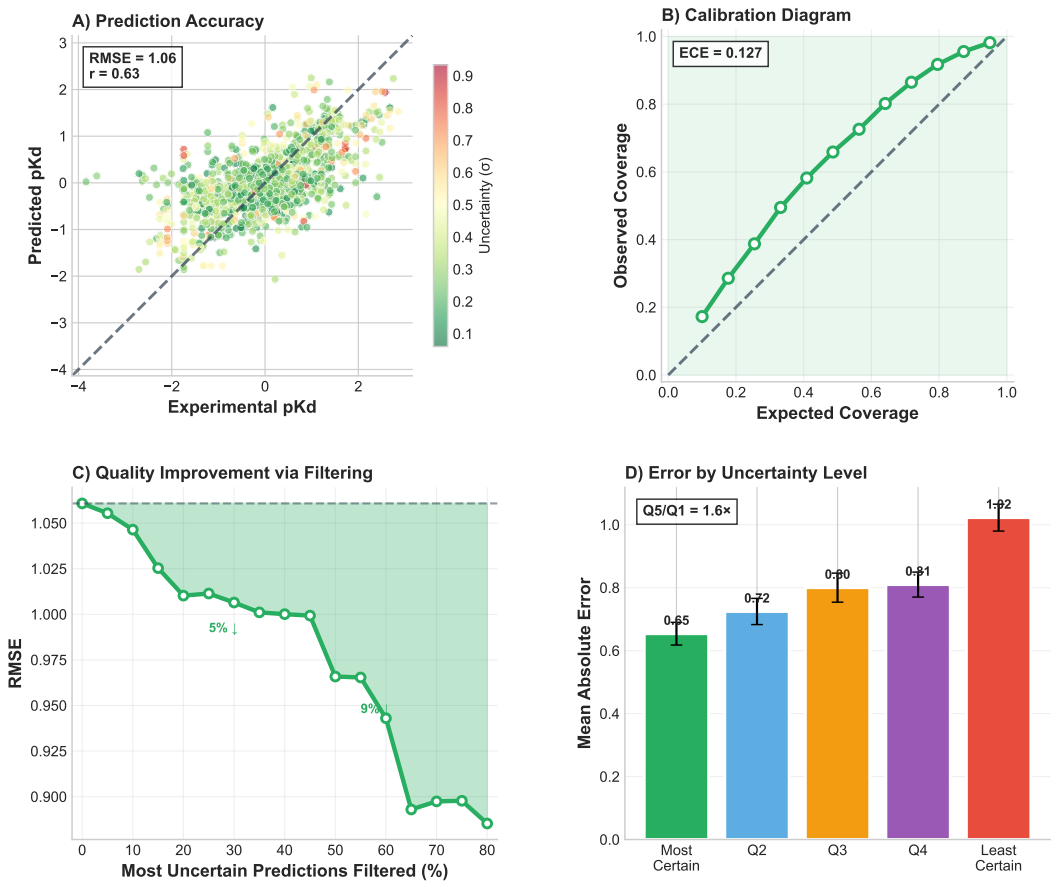

Figure 5: **Uncertainty quantification enables reliable triage.** (A) Predicted vs. experimental affinities colored by uncertainty. (B) Calibration diagram (ECE = 0.127). (C) Error reduction from filtering. (D) MAE stratification by uncertainty quintile.

cells unaffected. Critically, this mechanism of action is orthogonal to existing antibiotics, offering a potential solution to the antimicrobial resistance crisis as bacteria lack evolutionary precedent for resistance to riboswitch-targeting drugs [37].

Experimental setup. Neither the target RNA sequence (extracted from the 9BZ1 structure) nor the queried ligand chemotype (AICA core; SMILES provided in Supplementary Data) appeared in our training data. The RNA shares <40% sequence identity with any training example, and the ligand's Tanimoto similarity to the nearest training molecule is <0.25. This constitutes a stringent *blind RNA + blind molecule* generalization test.

Results. Despite these distribution shifts, RiboLead predicts strong binding with $pK_d > 8$ (classified as **highly positive**), consistent with the experimentally observed bound complex captured by crystallography [33]. Notably, the model assigns low epistemic uncertainty ($\sigma < 0.3$) to this prediction, indicating confident extrapolation rather than blind guessing.

This prospective validation provides evidence that RiboLead's learned representations capture generalizable principles of RNA–ligand recognition that extend to newly characterized targets and unseen chemical scaffolds. For drug discovery practitioners, this supports the use of RiboLead as a rapid

in silico filter to prioritize RNA–small-molecule pairs before expensive structure determination and experimental validation.

5 Discussion

RNA-targeted therapeutics are poised to reshape drug discovery by unlocking targets that have remained intractable for decades. Yet realizing this potential requires computational tools that are simultaneously accurate, scalable, and trustworthy. In this work, we introduced RiboLead, a tri-modal affinity predictor that bypasses the 3D structure bottleneck while providing calibrated uncertainty estimates for clinical deployment. Here we reflect on the broader implications of our findings and chart a path toward prospective experimental validation.

5.1 Methodological insights

Rich embeddings can substitute for scarce 3D data. The most striking outcome of our ablation study is that replacing tri-modal encoding with one-hot nucleotide vectors degrades blind-RNA performance by 52%, yet our architecture requires no 3D coordinates whatsoever. By fusing pre-trained RNA language models (RinalMo), computationally predicted secondary structure (MXfold2), and evolutionary profiles (PSSM), RiboLead captures structural context that previously demanded explicit crystallographic or cryo-EM data. This suggests a general principle for data-scarce domains: when labeled 3D structures are limited, investing in multimodal representation learning may yield greater returns than pursuing additional structural biology experiments.

Uncertainty unifies training and deployment. Epistemic uncertainty (σ), derived from our ensemble, serves a double purpose. During training, the uncertainty threshold ($\sigma < 0.4$) filters unreliable pseudo-labels, enabling safe semi-supervised expansion of the training manifold without ingesting noise. During deployment, the same metric allows drug discovery teams to auto-accept high-confidence hits ($\sigma < 0.2$), flag medium-confidence predictions for expert review, and reject low-confidence outputs—reducing costly wet-lab validation of false positives. We believe this "uncertainty-gated" philosophy can generalize to other biomolecular prediction tasks where out-of-distribution inputs are common.

Position-specific modality weighting mirrors biology. Our gated co-attention mechanism learns to weight sequence, structure, and evolutionary signals differently at each nucleotide position. This mirrors the biological reality of RNA-ligand recognition: some nucleotides matter primarily for shape complementarity (structure), others for conserved hydrogen-bonding patterns (sequence), and still others for evolutionary constraints that preserve function across species. The interpretability of these gate weights opens avenues for mechanistic hypothesis generation in structure-activity relationship (SAR) studies.

5.2 Broader implications for drug discovery

Unlocking the undruggable 85%. Approximately 85% of disease-associated proteins lack binding pockets suitable for conventional small-molecule drugs [1]. By targeting the mRNA that encodes these proteins, we can intervene before the problematic protein is ever synthesized. RNA-targeting small molecules are already delivering clinical successes: risdiplam corrects splicing errors in *SMN2* to treat spinal muscular atrophy [3], and companies such as Ribometrix are advancing small molecules that bind directly to *c-MYC* mRNA—a notoriously undruggable oncogene implicated in over 70% of human cancers [38]. RiboLead provides the high-throughput virtual screening capability needed to accelerate discovery of the next generation of such compounds.

From off-switch to volume knob. Traditional protein inhibitors typically silence a target entirely. In contrast, RNA-targeting small molecules can modulate splicing and processing, functioning as volume knobs that tune gene expression rather than eliminate it. This nuance is therapeutically critical for diseases caused by gene mis-expression rather than gene presence—for example, restoring partial SMN protein production in spinal muscular atrophy rather than ablating the gene. Tools like RiboLead can screen for compounds that achieve therapeutically optimal modulation rather than complete inhibition.

Pandemic preparedness through conserved RNA targets. Viral RNA structures are often conserved across related strains and even families, while surface proteins mutate rapidly to evade immunity. The programmable ribosomal frameshifting element of SARS-CoV-2, for example, shares

structural features with all betacoronaviruses [4]. A single compound that binds this conserved element could provide broad-spectrum antiviral protection independent of spike protein mutations—a paradigm shift from chasing variants to targeting invariant replication machinery. RiboLead's demonstrated ability to generalize to novel RNA folds (ZTP riboswitch case study) suggests potential for identifying such pan-viral therapeutics before the next pandemic emerges.

A new antibiotic mechanism for a resistance crisis. The WHO estimates that antimicrobial resistance could cause 10 million annual deaths by 2050 [39]. Riboswitches—RNA-based regulatory elements found in bacteria but absent in humans—offer a fundamentally new target class [34]. Compounds that lock bacterial riboswitches into non-functional conformations could starve pathogens of essential metabolites without affecting human cells. Critically, bacteria have no evolutionary precedent for resistance to riboswitch-targeting drugs, as these elements have never before been exploited pharmacologically. Our prospective validation on the ZTP riboswitch demonstrates that RiboLead can identify binders for this clinically relevant target class.

5.3 Limitations

Computational validation only. All reported results derive from computational benchmarks on public datasets. While the ZTP riboswitch case study demonstrates prediction on a newly solved structure, no wet-lab synthesis or binding assays have been performed. Experimental validation remains the definitive test of any computational prediction.

Dataset biases. Training and evaluation relied on R-SIM and related datasets, which are biased toward well-studied RNA families (riboswitches, ribosomes, aptamers). Performance on truly novel RNA folds—such as long non-coding RNAs (lncRNAs), circular RNAs (circRNAs), or cryptic splice sites—remains uncharacterized. The blind-RNA and blind-molecule tests provide partial evidence of generalization but cannot guarantee performance on radically out-of-distribution targets.

Affinity $\neq$ therapeutic efficacy. High predicted $\mathrm{p}K_d$ indicates strong molecular binding, not cellular activity, target selectivity, or pharmacokinetic properties. A complete drug discovery pipeline requires additional modeling of cell permeability, metabolic stability, off-target binding, and *in vivo* pharmacodynamics.

Static representations. RiboLead operates on static sequence and structure representations; RNA's profound conformational flexibility upon ligand binding (induced fit) is captured only implicitly through learned embeddings. For highly dynamic targets, molecular dynamics-informed approaches or ensemble docking may provide complementary insights.

5.4 Future directions

Scaling semi-supervised learning. R-SIM represents only a fraction of known RNA–small molecule interactions; ChEMBL and proprietary pharmaceutical datasets contain orders of magnitude more binding data. RiboLead's uncertainty-guided pseudo-labeling is designed precisely for this scenario: as larger unlabeled corpora become available, the training manifold can expand while the uncertainty filter guards against noise.

Extension to multi-component complexes. RNA rarely acts in isolation. Integration with protein–drug and protein–RNA interaction data could enable joint modeling of ternary complexes (e.g., RNA–protein–ligand assemblies), capturing allosteric effects and cooperative binding that single-modality models miss. Foundation models pre-trained across all biomolecular interaction types may further improve generalization.

Prospective experimental validation. The natural next step is collaboration with wet-lab partners to synthesize and experimentally test top-ranked predictions from virtual screens. Priority targets include oncogenic lncRNAs such as MALAT1 and HOTAIR, viral frameshifting elements (SARS-CoV-2, MERS-CoV), and bacterial riboswitches controlling essential metabolic pathways.

Generative integration. Pairing RiboLead with generative models—such as diffusion-based ligand generators or variational autoencoders over molecular graphs—could enable not only scoring of existing compounds but *de novo* design of RNA-binding molecules. Uncertainty-guided generation could steer exploration toward high-confidence regions of chemical space, accelerating the discovery of novel scaffolds optimized for RNA targets.

5.5 Conclusion

RiboLead demonstrates that accurate, scalable, and uncertainty-aware RNA–ligand affinity prediction is achievable without explicit 3D structures. By resolving the trilemma that has constrained prior methods, we provide a practical tool for virtual screening of RNA targets at scale. The prospective validation on an antibiotic-relevant riboswitch, combined with calibrated uncertainty that enables clinical triage, positions RiboLead as a bridge between computational prediction and experimental drug discovery. We anticipate that the principles underlying this work—multimodal fusion, uncertainty-guided learning, and 3D-free structural encoding—will generalize to other data-scarce biomolecular prediction tasks and contribute to the broader goal of expanding the druggable genome beyond proteins to the vast, largely untapped world of functional RNA.

References

[1] Hopkins, A.L., & Groom, C.R. (2002). The druggable genome. *Nature Reviews Drug Discovery*, 1(9), 727–730.

[2] Dang, C.V., Reddy, E.P., Shokat, K.M., & Soucek, L. (2017). Drugging the 'undruggable' cancer targets. *Nature Reviews Cancer*, 17(8), 502–508.

[3] Ratni, H., Ebeling, M., Baird, J., et al. (2018). Discovery of risdiplam, a selective survival of motor neuron-2 (SMN2) gene splicing modifier for the treatment of spinal muscular atrophy (SMA). *Journal of Medicinal Chemistry*, 61(15), 6501–6517.

[4] Kelly, J.A., Olson, A.N., Neupane, K., et al. (2020). Structural and functional conservation of the programmed -1 ribosomal frameshift signal of SARS coronavirus 2 (SARS-CoV-2). *Journal of Biological Chemistry*, 295(31), 10741–10748.

[5] Burley, S.K., et al. (2023). RCSB Protein Data Bank (RCSB.org): delivery of experimentally-determined PDB structures alongside one million computed structure models. *Nucleic Acids Research*, 51(D1), D483–D489.

[6] Abramson, J., Adler, J., Dunger, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630, 493–500.

[7] Jumper, J., Evans, R., Pritzel, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583–589.

[8] Eberhardt, J., Santos-Martins, D., Tillack, A.F., & Forli, S. (2021). AutoDock Vina 1.2.0: New docking methods, expanded force field, and Python bindings. *Journal of Chemical Information and Modeling*, 61(8), 3891–3898.

[9] Ruiz-Carmona, S., Alvarez-Garcia, D., Foloppe, N., et al. (2014). rDock: a fast, versatile and open source program for docking ligands to proteins and nucleic acids. *PLoS Computational Biology*, 10(4), e1003571.

[10] Warner, K.D., Hajdin, C.E., & Weeks, K.M. (2018). Principles for targeting RNA with drug-like small molecules. *Nature Reviews Drug Discovery*, 17(8), 547–558.

[11] Chen, J.L., & Bhagat, J. (2022). Targeting RNA with small molecules: from fundamental principles towards the clinic. *Chemical Society Reviews*, 51(7), 2679–2693.

[12] Chhabra, S., Xie, J., & Frank, A.T. (2023). Annapurna: a scoring function for predicting RNA-small molecule binding poses. *Journal of Chemical Information and Modeling*, 63(8), 2478–2488.

[13] Xia, Y., et al. (2025). Deciphering RNA–ligand binding specificity with GerNA-Bind. *Nature Machine Intelligence*.

[14] Stärk, H., Ganea, O.E., Pattanaik, L., Barzilay, R., & Jaakkola, T. (2022). EquiBind: Geometric deep learning for drug binding structure prediction. *ICML 2022*.

[15] Corso, G., Stärk, H., Jing, B., Barzilay, R., & Jaakkola, T. (2023). DiffDock: Diffusion steps, twists, and turns for molecular docking. *ICLR 2023*.

[16] Huang, Z., Wang, Y., Chen, S., Tan, Y.S., Deng, L., & Wu, M. (2024). DeepRSMA: a cross-fusion-based deep learning method for RNA–small molecule binding affinity prediction. *Bioinformatics*, 40(12), btae678.

[17] Wang, K., Zhou, R., Wu, Y., & Li, M. (2023). RLBind: a deep learning method to predict RNA-ligand binding sites. *Briefings in Bioinformatics*, 24(1), bbac486.

[18] Nguyen, T., Le, H., Quinn, T.P., Nguyen, T., Le, T.D., & Venkatesh, S. (2021). GraphDTA: predicting drug-target binding affinity with graph neural networks. *Bioinformatics*, 37(8), 1140–1147.

[19] Chen, J., Hu, Z., Sun, S., et al. (2022). Interpretable RNA foundation model from unannotated data for highly accurate RNA structure and function predictions. *arXiv:2204.00300*.

[20] Krishnan, S.R., Roy, A., & Gromiha, M.M. (2023). R-SIM: A Database of Binding Affinities for RNA-small Molecule Interactions. *Journal of Molecular Biology*, 435(14), 167914.

[21] Penić, R., Vlašić, I., Huber, R., Čaušević, A., Šikić, M., & Hoksza, D. (2025). RiNALMo: General-Purpose RNA Language Models Can Generalize Well on Structure Prediction Tasks. *Nature Communications*, 16, 5671.

[22] Sato, K., Akiyama, M., & Sakakibara, Y. (2021). RNA secondary structure prediction using deep learning with thermodynamic integration. *Nature Communications*, 12(1), 941.

[23] RNAcentral Consortium. (2021). RNAcentral 2021: secondary structure integration, improved sequence search and new member databases. *Nucleic Acids Research*, 49(D1), D212–D220.

[24] Steinegger, M., & Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology*, 35(11), 1026–1028.

[25] Zhou, G., Gao, Z., Ding, Q., Zheng, H., Xu, H., Wei, Z., Zhang, L., & Ke, G. (2023). Uni-Mol: A Universal 3D Molecular Representation Learning Framework. *ICLR 2023*.

[26] RDKit: Open-source cheminformatics. https://www.rdkit.org

[27] Fey, M., & Lenssen, J.E. (2019). Fast Graph Representation Learning with PyTorch Geometric. *ICLR Workshop on Representation Learning on Graphs and Manifolds*.

[28] Abbasi, K., Razzaghi, P., Poso, A., et al. (2020). DeepCDA: Deep Cross-Domain Compound-Protein Affinity Prediction through LSTM and Convolutional Neural Networks. *Bioinformatics*, 36(17), 4633–4642.

[29] Pradeep, K., Xiao, X., Wan, Y., et al. (2021). DeepDTAF: a deep learning method to predict protein-ligand binding affinity. *Briefings in Bioinformatics*, 22(5), bbab072.

[30] Cai, T., Luo, S., Xu, K., He, D., Liu, T.Y., & Wang, L. (2021). GraphNorm: A Principled Approach to Accelerating Graph Neural Network Training. *ICML 2021*, PMLR 139, 1319–1330.

[31] Houlsby, N., Giurgiu, A., Jastrzebski, S., Morrone, B., de Laroussilhe, Q., Gesmundo, A., Attariyan, M., & Gelly, S. (2019). Parameter-Efficient Transfer Learning for NLP. *ICML 2019*.

[32] Alayrac, J.B., Donahue, J., Luc, P., et al. (2022). Flamingo: a Visual Language Model for Few-Shot Learning. *NeurIPS 2022*.

[33] Chen, X., et al. (2024). Crystal structure of ZTP riboswitch bound to AICA. *Protein Data Bank*, PDB ID: 9BZ1. Released November 27, 2024.

[34] Blount, K.F., & Breaker, R.R. (2006). Riboswitches as antibacterial drug targets. *Nature Biotechnology*, 24(12), 1558–1564.

[35] Kim, P.B., Nelson, J.W., & Breaker, R.R. (2015). An ancient riboswitch class in bacteria regulates purine biosynthesis and one-carbon metabolism. *Molecular Cell*, 57(2), 317–328.

[36] Jones, C.P., & Ferré-D'Amaré, A.R. (2015). Recognition of the bacterial alarmone ZMP through long-distance association of two RNA subdomains. *Nature Structural & Molecular Biology*, 22(9), 679–685.

[37] Deigan, K.E., & Ferré-D'Amaré, A.R. (2011). Riboswitches: discovery of drugs that target bacterial gene-regulatory RNAs. *Accounts of Chemical Research*, 44(12), 1329–1338.

[38] Ribometrix, Inc. (2024). Ribometrix presents data demonstrating that its RNA-targeting platform has identified small molecules that engage directly with c-MYC mRNA. *AACR Annual Meeting 2024*, Abstract 5865.

[39] World Health Organization. (2024). Global antimicrobial resistance and use surveillance system (GLASS) report 2024. Geneva: WHO.