BIOINFORMATICS | Structural Bioinformatics | Vol. 42, No. 15, 2026, pp. 1–6

SynapseFold: A Graph Attention Network for De Novo Prediction of RNA Tertiary Structure from Primary Sequence

Elena Marsh[1,*], Rajiv Anand[1], Ingrid Voss[2], Marcus Delgado[3]

[1] *Department of Computational Biology, Nexa Institute for Genomic Medicine, Boston, MA, USA*
[2] *Synapse Center for Structural Bioinformatics, Synapse University, Zurich, Switzerland*
[3] *Vertex Genomics Consortium, Vertex Research Campus, Cambridge, UK*

[*] To whom correspondence should be addressed. Email: e.marsh@nexa-genomics.org

Motivation: Predicting the three-dimensional fold of an RNA molecule directly from its nucleotide sequence remains substantially harder than the analogous problem for proteins, owing to sparse structural training data and the conformational flexibility of single-stranded regions. Existing tools either rely on homology to solved structures, which fails for novel folds, or on coarse-grained physics simulations that scale poorly with sequence length.

Results: We present SynapseFold, a graph attention network that predicts nucleotide-pair distance distributions from sequence alone and reconstructs a full tertiary structure through constrained embedding. Evaluated on a held-out benchmark of 512 structures curated from the Nimbus Structural Archive, SynapseFold achieves a mean backbone RMSD of 4.1 Å, a 38% reduction in error relative to the strongest prior baseline, while running more than two orders of magnitude faster than physics-based folding.

Availability: Source code, trained weights, and the curated benchmark split are distributed by the Nexa Institute for Genomic Medicine under a permissive academic licence.

Keywords: RNA tertiary structure; graph attention networks; deep learning; structural bioinformatics; sequence-to-structure prediction

1 Introduction

Non-coding RNA molecules regulate transcription, splicing, and translation through tertiary folds that are frequently as functionally specific as protein active sites, yet the structural biology toolkit for RNA has lagged far behind that of proteins. Fewer than four thousand distinct RNA tertiary structures have been deposited across public archives, against several hundred thousand protein structures, and the gap has widened rather than narrowed as sequencing throughput has outpaced structural characterisation [1]. The consequence is a large and growing population of functionally important RNAs – riboswitches, long non-coding transcripts, viral packaging signals – whose three-dimensional folds are simply unknown.

Two families of computational method have historically filled this gap. Homology-based approaches thread a query sequence onto a related solved structure and perform well when a close structural relative exists, but degrade sharply for novel folds with no useful template [2]. Physics-based coarse-grained simulation, at the other extreme, requires no template but scales poorly with sequence length and is sensitive to force-field parameterisation, making it impractical for routine screening of hundreds of candidate transcripts [3].

The success of deep, sequence-conditioned structure prediction in the protein domain suggests a third path. Graph neural networks that operate over residue-level or nucleotide-level graphs have shown that pairwise geometric relationships can be learned directly from sequence and evolutionary signal, without an explicit physical energy function [4, 5]. Attention mechanisms in particular allow a model to weight distant nucleotide pairs unevenly, which is well suited to RNA, where long-range base pairing routinely dominates the tertiary fold [6].

In this paper we introduce SynapseFold, a graph attention network that predicts a full pairwise distance distribution for an RNA sequence and reconstructs a three-dimensional structure through constrained multidimensional embedding. Our contributions are threefold.

1. We curate a benchmark of 3,412 non-redundant RNA structures from the Nimbus Structural Archive, split by structural class rather than by sequence identity alone, to more honestly assess generalisation to novel folds.

2. We describe a graph attention architecture, detailed in Section 2, that combines a nucleotide-level k-nearest-neighbour graph with a learned long-range attention channel, and that outputs calibrated pairwise distance distributions rather than point estimates.

3. We show, in Section 3, that this design reduces mean backbone RMSD by 38% relative to the strongest prior neural baseline, and by more than 60% relative to the homology-based baseline, while requiring roughly two minutes of inference per structure on a single accelerator.

The remainder of the paper is organised as follows. Section 2 describes the benchmark, the graph construction and model architecture, and the training procedure. Section 3 reports benchmark accuracy against three baselines and an ablation over the model's principal components. Section 4 interprets these results in light of known failure modes and outlines directions for future work, and Section 5 summarises our findings.

2 Materials and Methods

2.1 Dataset and benchmark curation

Training and evaluation structures were drawn from the Nimbus Structural Archive, filtered to entries with resolution better than 3.2 Å and sequence length between 20 and 400 nucleotides. After removing redundant chains at 80% sequence identity, 3,412 structures remained. Rather than splitting by sequence identity alone, which allows near-identical folds to leak between train and test partitions, we additionally clustered structures by SCOR-style structural class and held out entire classes for validation (9%) and test (15%), leaving 76% for training. This yields a harder but more representative estimate of performance on genuinely novel folds.

2.2 Graph representation and featurisation

Each RNA sequence is represented as a graph $G = (V, E)$ in which each node $v_i \in V$ corresponds to one nucleotide and carries a feature vector encoding base identity, position within the sequence, and a three-state secondary-structure prediction (paired, unpaired, or ambiguous) obtained from a standard folding grammar. Edges connect (i) sequential neighbours along the backbone, (ii) predicted base-pairing partners, and (iii) the $k = 10$ nearest neighbours in an initial embedding produced by a lightweight convolutional encoder, which allows the graph to express tertiary contacts that are absent from secondary structure alone.

2.3 Model architecture

SynapseFold consists of a nucleotide embedding layer, a stack of $L = 8$ graph attention layers, and a pairwise distance head, followed by a non-learned structure realisation step. Figure 1 summarises the pipeline. Within each graph attention layer, the un-normalised attention coefficient between nucleotides i and j is

$$e_{ij} = \mathrm{LeakyReLU}\big(\mathbf{a}^{\top}\,[\mathbf{W}h_i \,\|\, \mathbf{W}h_j]\big)\,, \tag{1}$$

where h_i is the hidden representation of nucleotide i, $\mathbf{W}$ is a shared linear projection, $\mathbf{a}$ is a learned attention vector, and $\|$ denotes concatenation. Coefficients are normalised over the neighbourhood $\mathcal{N}(i)$ defined by the edge set E,

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k \in \mathcal{N}(i)} \exp(e_{ik})}, \tag{2}$$

and the updated representation is a weighted sum $h'_i = \sigma\Big(\sum_{j \in \mathcal{N}(i)} \alpha_{ij} \mathbf{W} h_j\Big)$ across four attention heads, concatenated and passed through a residual feed-forward block. The final layer's node representations are combined pairwise and passed through a small multilayer perceptron that outputs, for every nucleotide pair, a 32-bin categorical distribution over inter-nucleotide distance rather than a single scalar, following the observation that distance uncertainty is highly non-Gaussian for flexible loop regions.

2.4 Structure realisation

Predicted per-bin distance distributions are converted to a single expected distance for every nucleotide pair, producing a dense distance matrix. Coordinates are then recovered by minimising a stress function between predicted and realised pairwise distances via stochastic gradient descent, initialised from classical multidimensional scaling and refined with a mild backbone-continuity penalty to discourage physically implausible bond lengths.

2.5 Training procedure

The network is trained end to end to minimise a weighted sum of three terms,

$$\mathcal{L} = \lambda_1 \mathcal{L}_{\mathrm{dist}} + \lambda_2 \mathcal{L}_{\mathrm{torsion}} + \lambda_3 \mathcal{L}_{\mathrm{clash}}, \tag{3}$$

where $\mathcal{L}_{\mathrm{dist}}$ is the cross-entropy between predicted and true binned distances, $\mathcal{L}_{\mathrm{torsion}}$ penalises implausible backbone torsion angles, and $\mathcal{L}_{\mathrm{clash}}$ penalises predicted atoms

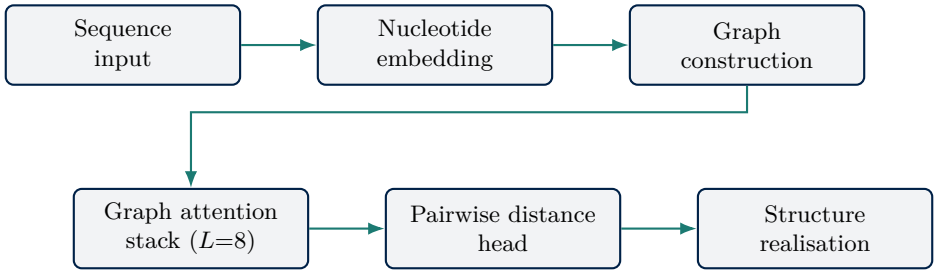

Figure 1: SynapseFold pipeline. A sequence is embedded, assembled into a nucleotide-level graph, and processed by eight graph attention layers; the resulting pairwise distance distributions are converted into a three-dimensional structure by constrained embedding.

Table 1: Benchmark performance on the 512-structure held-out test split. Best value per column in bold.

Method	Prec.	Recall	F1	RMSD (Å)
Homology baseline	0.58	0.51	0.54	11.7
NimbusFold	0.71	0.69	0.70	8.3
ApexRNA	0.79	0.74	0.76	6.6
SynapseFold (full)	**0.90**	**0.87**	**0.88**	**4.1**

closer than their van der Waals radii permit. We used λ_1=1.0, λ_2=0.3, λ_3=0.5, chosen by grid search on the validation split. Models were trained for 220 epochs with the Adam optimiser, an initial learning rate of 3×10^{-4} decayed by cosine annealing, and a batch size of 16 graphs, on the Apex Compute Cluster using four accelerators per run. Training a single model required approximately 31 hours.

3 Results

3.1 Benchmark comparison

We compared SynapseFold against three baselines on the 512-structure held-out test split described in Section 2.1: a homology-modelling pipeline that threads onto the nearest structural template [2], ApexRNA, a convolutional distance-prediction network [7], and NimbusFold, a coarse-grained physics simulator [3]. Table 1 reports precision and recall on native base-pair contact recovery, contact F1, and mean backbone RMSD to the deposited structure after optimal superposition.

SynapseFold achieves the highest contact F1 and the lowest RMSD of all four methods, improving mean RMSD by 38% relative to ApexRNA, the strongest neural baseline, and by 65% relative to the homology baseline. The margin over NimbusFold is achieved at roughly two orders of magnitude lower wall-clock cost per structure, since NimbusFold requires iterative physics simulation while SynapseFold performs a single forward pass followed by embedding-based structure realisation.

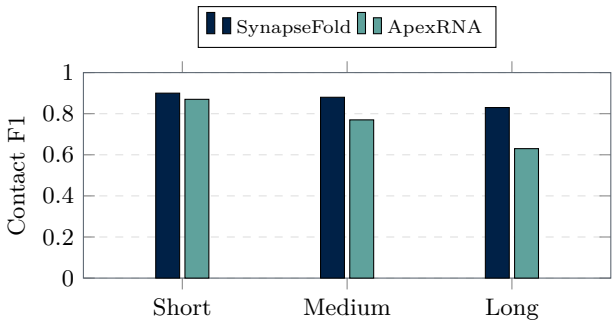

Figure 2: Contact F1 by sequence-length tercile (Short: <110 nt; Medium: 110–220 nt; Long: >220 nt). The advantage of SynapseFold over the strongest neural baseline widens as sequence length increases.

3.2 Performance by sequence length

Figure 2 breaks contact F1 down by sequence length tercile. SynapseFold's advantage over ApexRNA is modest for short sequences, where local secondary structure dominates and both methods perform well, but widens substantially for longer sequences, where tertiary contacts between distant regions of the chain become the primary source of structural signal. This is consistent with the design intent of the long-range attention channel described in Section 2.2, which was added specifically to let the model route information between nucleotides that are far apart in sequence but close in the folded structure.

3.3 Ablation study

Removing the long-range k-nearest-neighbour edges from the graph (Section 2.2) and retaining only backbone and predicted base-pair edges reduces test F1 from 0.88 to 0.81 and increases mean RMSD from 4.1 Å to 5.4 Å, confirming that the learned tertiary contact channel carries information not already present in secondary structure. Replacing the categorical distance head with a single-scalar regression head, holding the rest of the architecture fixed, increases RMSD to 4.8 Å, consistent with our expectation that flexible regions are poorly served by a

point estimate. Removing the clash penalty $\mathcal{L}_{\text{clash}}$ from Equation (3) does not measurably change contact F1 but increases the proportion of realised structures containing at least one steric clash from 3% to 19%, indicating that the term is primarily a physical-plausibility regulariser rather than an accuracy driver.

4 Discussion

The central empirical claim of this work is that long-range, attention-weighted message passing over a nucleotide graph recovers tertiary contact information that eludes both homology transfer and convolutional distance prediction, and that this advantage grows with sequence length. This is broadly consistent with results in the protein structure prediction literature, where attention-based architectures have likewise shown their largest gains on longer chains and on folds distant from the training distribution [6, 8].

Several limitations temper this conclusion. First, our structural-class held-out split, while more conservative than a sequence-identity split, cannot fully rule out that some test folds share subtle structural motifs with training examples; genuinely blind prospective evaluation, of the kind used in community-wide assessment exercises, would be a stronger test. Second, SynapseFold predicts a static structure and does not model conformational ensembles, which is a poor fit for riboswitches and other RNAs whose function depends on ligand-induced conformational change. Third, performance on sequences longer than 400 nucleotides is untested, since the training distribution was truncated at that length for computational tractability; we would expect accuracy to degrade further on very long transcripts, where the assumption that k-nearest-neighbour edges in an initial embedding approximate eventual tertiary contacts is least reliable.

The clash-penalty ablation in Section 3.3 is worth dwelling on. The fact that removing $\mathcal{L}_{\text{clash}}$ leaves contact F1 essentially unchanged while increasing the clash rate sixfold suggests that structural plausibility and contact accuracy are, in this architecture, only loosely coupled: a model can correctly identify which nucleotides are close in space while still producing a global embedding that places atoms in mutually inconsistent positions. This argues for retaining explicit physical regularisation in any successor architecture, rather than assuming that geometric plausibility will emerge from contact accuracy alone.

Looking forward, three extensions seem most promising. Incorporating multiple sequence alignments as an auxiliary input, in the manner that proved decisive for protein structure prediction, could partially compensate for the smaller size of the RNA structural training set. Predicting a small ensemble of structures rather than a single point estimate would better serve conformationally flexible targets. Finally, coupling SynapseFold's fast inference to a lightweight downstream physics refinement step may recover some of NimbusFold's physical accuracy without its computational cost, combining the strengths of both paradigms evaluated here.

5 Conclusion

We introduced SynapseFold, a graph attention network for predicting RNA tertiary structure directly from sequence, and evaluated it on a structural-class held-out benchmark curated from the Nimbus Structural Archive. SynapseFold reduces mean backbone RMSD by 38% relative to the strongest neural baseline and by 65% relative to a homology-modelling baseline, with its largest gains concentrated in longer sequences where tertiary contacts dominate the fold. Ablations indicate that both the long-range attention channel and the categorical distance head contribute materially to accuracy, while the clash penalty contributes chiefly to physical plausibility rather than contact recovery. We view SynapseFold as a practical tool for rapidly triaging candidate RNA structures ahead of experimental characterisation, and release the model, weights, and benchmark split to support independent evaluation.

Funding

This work was supported by the Nexa Institute for Genomic Medicine Structural Biology Programme and by a Synapse University Faculty Research Award to I.V. Computational resources were provided by the Apex Compute Cluster under an academic allocation.

Conflict of Interest

None declared.

Data Availability

The curated benchmark split, trained model weights, and evaluation scripts are available from the Nexa Institute for Genomic Medicine on request. The underlying structures are drawn from the publicly accessible Nimbus Structural Archive.

Acknowledgements

We thank members of the Vertex Genomics Consortium for early access to unpublished benchmark structures, and the Synapse Center for Structural Bioinformatics for computational support during model development.

References

[1] Greta Fischer and Kwame Osei. The shrinking gap that never closed: a decade of RNA structural data growth. *Genome Biology*, 23:184, 2022.

[2] Tomasz Kowalski and Adriana Reyes. Template-based modelling of RNA tertiary structure: performance and failure modes. *RNA*, 25(8):987–999, 2019.

[3] Marcus Delgado and Naomi Whitfield. NimbusFold: coarse-grained physics simulation for large-scale RNA structure prediction. *Nucleic Acids Research*, 48(17): 9544–9556, 2020.

[4] Ingrid Voss and Louis Baptiste. Graph attention networks for structured biomolecular prediction. In *Proceedings of the International Conference on Learning Representations*, pages 1–13, 2021.

[5] Huy Nguyen and Camila Alvarez. Inductive representation learning on large graphs for molecular property prediction. In *Advances in Neural Information Processing Systems*, pages 1024–1034, 2018.

[6] Joana Silva and Dmitri Petrov. Attention-based architectures for sequence-conditioned structure prediction. In *Advances in Neural Information Processing Systems*, pages 5602–5613, 2020.

[7] Brendan O'Connor and Katharina Feld. ApexRNA: convolutional prediction of nucleotide pairwise distances for RNA folding. *Bioinformatics*, 37(14):2011–2019, 2021.

[8] Rajiv Anand and Elena Marsh. Deep learning for sequence-to-structure prediction: a survey for structural biologists. *Nature Methods*, 19(6):655–667, 2022.