Bioinformatics and Code
-Wisdom Academic Press

DeepFoldRNA: A Graph Neural Network for RNA 3D Structure Prediction

Abstract

DeepFoldRNA presents a novel graph neural network architecture for RNA 3D structure prediction, integrating evolutionary information with physics-based refinement. By representing RNA molecules as graphs and incorporating co-evolutionary data through multiple sequence alignments, the method naturally captures complex structural interactions. Benchmarking on RNA-Puzzles dataset demonstrated average RMSD of 3.21 Å and TMscore of 0.82, achieving 28% improvement over existing methods. Performance gains were particularly notable for riboswitches and ribozymes with 35% RMSD reduction. Rosetta framework integration enables iterative refinement with 3-fold speedup while maintaining superior accuracy. Multi-head attention mechanisms successfully identified long-range interactions spanning over 200 nucleotides. The open-source implementation provides comprehensive interfaces, requiring only 12 minutes for 200-nucleotide RNAs on single GPU. This advancement establishes new standards for RNA structure prediction with significant implications for drug discovery and synthetic biology.

Keyword: Graph neural networks;RNA structure prediction;Deep learning;Rosetta integration;Co-evolutionary features

Serge Thill*
Email: serge.thill@lccb.lu
Affiliation: Luxembourg Centre for Computational Biology, Campus Limpertsberg, L-1852 Luxembourg City, Luxembourg

1. Introduction

RNA molecules play essential roles in protein synthesis, gene regulation, and catalysis. Understanding RNA three-dimensional structures is critical for elucidating functional mechanisms and developing therapeutics [1]. Complex folding patterns determine biological activities, yet experimental structure determination remains time-consuming and technically challenging despite advances in crystallography and cryo-EM.

Traditional computational methods rely on thermodynamic models and energy minimization. FARFAR2 and similar approaches achieve success through fragment assembly but face limitations including high computational costs and difficulties modeling non-canonical interactions. Performance plateaus for RNAs exceeding 100 nucleotides or containing pseudoknots [2].

Recent deep learning advances have transformed structure prediction. Singh et al. improved base-pairing predictions using neural networks and transfer learning [3]. UFold enabled rapid secondary structure prediction without thermodynamic parameters [4]. Sato et al. integrated thermodynamic information with deep learning [5]. trRosettaRNA applied transformers to 3D prediction using attention mechanisms.

Current approaches exhibit limitations: most focus on secondary structure or require extensive computational resources. Standard neural networks fail to capture RNA's graph-like nature. Evolutionary information remains underutilized despite proven value [6].

DeepFoldRNA addresses these limitations through graph neural network architecture that naturally models RNA structures, incorporates evolutionary information via multiple sequence alignments, and integrates with Rosetta for refinement. This approach achieves superior accuracy and computational efficiency. The open-source implementation facilitates community adoption and development.

Serge Thill*
Email: serge.thill@lccb.lu
Affiliation: Luxembourg Centre for Computational Biology, Campus Limpertsberg, L-1852 Luxembourg City, Luxembourg

2. Data and Methods

2.1 Dataset Construction

DeepFoldRNA development required construction of a comprehensive dataset to ensure robust model training and evaluation. RNA structural data were collected from the Protein Data Bank (PDB) for experimentally determined atomic-resolution structures and Rfam database for curated RNA families with secondary structures and evolutionary information. Quality filtering included resolution thresholds below 3.5 Å, exclusion of structures with missing residues in critical regions, and removal of conformationally ambiguous structures.

Sequence redundancy reduction using CD-HIT with 90% identity cutoff prevented overfitting while maintaining structural diversity. The final dataset comprised 4,582 non-redundant RNA structures spanning various functional categories including ribozymes, riboswitches, ribosomal RNAs, and regulatory RNAs.

Table 1: Dataset Statistics and Training Configuration

Category	PDB	Rfam	Distribution	Split (Train/Val/Test)
Total Structures	3,271	1,311	4,582 (non-redundant)	70%/15%/15%
Quality Criteria	Resolution <3.5Å	Secondary structure annotated	90% identity cutoff	Family-aware splitting
Length Distribution				
20-50 nt	862	283	1,145 (25.0%)	801/172/172
51-100 nt	1,308	524	1,832 (40.0%)	1,282/275/275
101-200 nt	768	330	1,098 (24.0%)	769/165/164
201-500 nt	333	174	507 (11.0%)	355/75/77
RNA Families				
Ribozymes	481	206	687 (15.0%)	481/103/103
Riboswitches	654	262	916 (20.0%)	641/137/138
Ribosomal RNA	981	393	1,374 (30.0%)	962/206/206
Transfer RNA	327	131	458 (10.0%)	321/69/68
Regulatory RNA	523	209	732 (16.0%)	512/110/110
Other ncRNA	305	110	415 (9.0%)	290/62/63
MSA Statistics	Avg: 287	Coverage: >95%	Tool: Infernal	-

Serge Thill*
Email: serge.thill@lccb.lu
Affiliation: Luxembourg Centre for Computational Biology, Campus Limpertsberg, L-1852 Luxembourg City, Luxembourg

Category	PDB	Rfam	Distribution	Split (Train/Val/Test)
	sequences		v1.1.4	

Table 1 provides a comprehensive overview of the statistics of the dataset, including the distribution of structures across different RNA families and the respective sequence length categories. The results show an equal representation across various RNA types, with ribosomal RNAs occupying 30% of the dataset, followed by riboswitches at 20% and regulatory RNAs at 16%.

The data was randomly divided into training (70%), validation (15%), and test (15%) sets using family-aware splitting to prevent the danger of data leakage. As outlined in Table 1, RNA families were allocated to subsets so that the model would generalize to new structural motifs. Multiple sequence alignments for every RNA were created with Infernal along with the Rfam database, with more than 95% coverage and an average of 287 sequences per family. This cautious dataset curation lays a robust basis for the training of DeepFoldRNA and overcomes drawbacks of prior research based on smaller or less comprehensive structural sets.

2.2 Graph Neural Network Architecture

DeepFoldRNA utilizes a new graph neural network architecture specifically tailored to identify complex structural relationships characteristic of RNA molecules. Differing from conventional sequence-based approaches, this approach represents RNA structures as molecular graphs, with nucleotides as nodes and different interaction types as edges. The overall architectural layout is provided in (Table 2)Algorithm 1, describing the procedure of mapping input RNA sequences to three-dimensional structure predictions.

Table 2: Algorithm 1 - DeepFoldRNA Main Architecture

Line	Operation
Input:	RNA sequence S, MSA features M
Output:	3D coordinates C, confidence scores Q
1	G = ConstructGraph(S) // Build molecular graph
2	H^0 = InitializeFeatures(S, M) // Node feature initialization
3	for layer l in 1 to L do
4	H^l = GraphConv(H(l-1), G) // Graph convolution
5	H^l = MultiHeadAttention(H^l) // Attention mechanism
6	H^l = LayerNorm(Dropout(H^l)) // Regularization
7	end for
8	D = DistancePredictor(H^L) // Predict distance matrix
9	A = DihedralPredictor(H^L) // Predict dihedral angles
10	C = GeometricModule(D, A) // Convert to 3D coordinates

Serge Thill*
Email: serge.thill@lccb.lu
Affiliation: Luxembourg Centre for Computational Biology, Campus Limpertsberg, L-1852 Luxembourg City, Luxembourg

Line	Operation
11	Q = ConfidenceEstimator(H^L, C) // Estimate confidence scores
12	return C, Q

The graph construction process constitutes a critical preprocessing step that converts linear RNA sequences into rich graph representations. (Table 3)Algorithm 2 details this procedure, which considers multiple edge types including covalent backbone connections, Watson-Crick base pairs, non-canonical interactions, and spatial proximity relationships within a 10 Å threshold. This multi-relational graph representation enables the model to simultaneously learn from primary sequence information and potential structural constraints, providing a more comprehensive structural context than traditional sequence-based methods.

Table 3: Algorithm 2 - Graph Construction from RNA Sequence

Line	Operation
Input:	RNA sequence S = {s_1, s_2, ..., s_n}
Output:	Multi-relational graph G = (V, E)
1	V = ∅ , E = ∅
2	// Initialize nodes
3	for i in 1 to n do
4	V = V ∪ {v_i} // Create node for nucleotide s_i
5	end for
6	// Add backbone edges
7	for i in 1 to n-1 do
8	E = E ∪ {(v_i, v_{i+1}, 'backbone')}
9	end for
10	// Add Watson-Crick base pairs
11	P = PredictSecondaryStructure(S)
12	for (i,j) in P do
13	E = E ∪ {(v_i, v_j, 'base_pair')}
14	end for
15	// Add non-canonical interactions
16	N = IdentifyNonCanonical(S)
17	for (i,j) in N do
18	E = E ∪ {(v_i, v_j, 'non_canonical')}
19	end for
20	// Add spatial proximity edges
21	for i in 1 to n do
22	for j in i+4 to n do // Skip nearby residues
23	if EstimatedDistance(i,j) < 10Å then
24	E = E ∪ {(v_i, v_j, 'proximity')}

Serge Thill*
Email: serge.thill@lccb.lu
Affiliation: Luxembourg Centre for Computational Biology, Campus Limpertsberg, L-1852 Luxembourg City, Luxembourg

Line	Operation
25	end if
26	end for
27	end for
28	return G = (V, E)

Building upon recent advances in physics-aware graph neural networks for RNA structure prediction [7], DeepFoldRNA incorporates domain-specific inductive biases through specialized message passing mechanisms. The core architecture consists of multiple graph convolutional layers that iteratively refine node representations, as described in (Table 4)Algorithm 3. Each layer employs edge-type-specific weight matrices to differentiate between various interaction types, allowing the model to learn distinct propagation patterns for different molecular relationships. The node features are initialized with one-hot encoded nucleotide types concatenated with position-specific scoring matrices derived from multiple sequence alignments.

Table 4: Algorithm 3 - Message Passing in Graph Convolution

Line	Operation
Input:	Node features H^l-1, Graph G, Edge types T
Output:	Updated node features H^l
1	**for each** node v **in** G **do**
2	m_v = 0 // Initialize message
3	**for each** neighbor u **of** v **do**
4	e_type = GetEdgeType(u, v)
5	W_e = GetWeightMatrix($e_t$ype) // Edge-specific weights
6	m_v = m_v + σ(W_e · h_u(l-1)) // Aggregate messages
7	**end for**
8	h_v^l = σ($W_s$elf · h_v^(l-1) + W_msg · m_v) // Update node
9	**end for**
10	**return** H^l = {h_v^l

A critical innovation in DeepFoldRNA's architecture is the integration of multi-head attention mechanisms to capture long-range dependencies, addressing limitations of previous approaches that struggled with distant interactions [7]. (Table 5)Algorithm 4 presents the attention computation process, where multiple attention heads learn different aspects of structural relationships. Attention weights are derived from both sequence features and predicted structural information to facilitate dynamic feature fusion based on their structural context. The attention mechanism is especially good at picking up tertiary relationships present over long sequence distances.

Table 5: Algorithm 4 - Multi-Head Attention Mechanism

Serge Thill*
Email: serge.thill@lccb.lu
Affiliation: Luxembourg Centre for Computational Biology, Campus Limpertsberg, L-1852 Luxembourg City, Luxembourg

Line	Operation
Input:	Node features H, Number of heads K
Output:	Attention-refined features H'
1	**for** k **in** 1 **to** K **do**
2	Q^k = H · W_Q^k // Query projection
3	K^k = H · W_K^k // Key projection
4	V^k = H · W_V^k // Value projection
5	A^k = Softmax($Q^k \cdot (K^{\wedge k^T} / \sqrt{d\_k})$ // Attention scores
6	Z^k = A^k · V^k // Weighted values
7	**end for**
8	Z = Concat(Z^1, ..., Z^K) // Concatenate heads
9	H' = Z · W_O // Output projection
10	**return** H'

The architecture integrates geometric constraints by a differentiable distance geometry module, which translates learned node embeddings to three-dimensional coordinates while ensuring physical plausibility. A number of regularization methods thwart overfitting and enhance generalization, such as strategically located dropout layers following every graph convolution operation and layer normalization for stabilizing training dynamics. Edge dropout randomly eliminates a proportion of edges during training, compelling the model to learn robust representations. The last layers are fully connected networks that convert the merged node representations into distance matrices and dihedral angle predictions, thereby generating atomic coordinates along with their corresponding confidence scores.

2.3 Rosetta Framework Integration

The coupling of DeepFoldRNA and Rosetta merges deep learning prediction and physics-based refinement, leveraging neural network pattern recognition and molecular mechanics accuracy. The interface allows bidirectional communication between the output of DeepFoldRNA and the refinement protocols of Rosetta for iterative structure optimization.

The modular design enables integration with current Rosetta protocols while reducing computational demands at the same time. DeepFoldRNA produces Rosetta-compliant distance matrices and dihedral angle constraints, which are used as constraints for fragment assembly and energy minimization procedures. Hybrid models using deep learning and geometric potentials have been demonstrated to be functional recently [8], whereas previous implementations were plagued by inefficiencies in data transfer.

One of the key innovations is two-way feedback between DeepFoldRNA and Rosetta. During refinement, Rosetta evaluates structural energetics and provides feedback in the form of scoring statistics to DeepFoldRNA, which updates confidence estimates and constraint weights. The loop typically converges within 5-10 cycles. Parallel computation of multiple structural hypotheses

Serge Thill*
Email: serge.thill@lccb.lu
Affiliation: Luxembourg Centre for Computational Biology, Campus Limpertsberg, L-1852 Luxembourg City, Luxembourg

through modular design drastically reduces computational time compared to sequential protocols . Performance optimization involves intelligent caching, vectorized constraint evaluation, and GPU-accelerated distance calculation. Dynamic memory allocation prevents large-scale prediction bottlenecks. The benchmarks demonstrate that the integrated pipeline achieves 3-fold speedup over standalone Rosetta with improved or comparable accuracy.

The open-source codebase is written with software engineering best practices, such as extensive documentation, unit tests, and continuous integration. The modular framework allows researchers to swap out components or insert bespoke protocols without altering fundamental architecture, thus being extensible to emerging developments in deep learning and molecular simulation techniques.

Serge Thill*
Email: serge.thill@lccb.lu
Affiliation: Luxembourg Centre for Computational Biology, Campus Limpertsberg, L-1852 Luxembourg City, Luxembourg

3. Results

3.1 Performance Evaluation

A thorough assessment of DeepFoldRNA's performance was made on various benchmark datasets to examine its efficiency at RNA structure prediction. The model was evaluated on the RNA-Puzzles dataset, CASP-RNA targets, and a carefully selected set of recently deposited PDB structures, covering a variety of RNA families and length ranges between 20 and 500 nucleotides. Classic metrics such as root mean square deviation (RMSD), template modeling score (TM-score), and contact map accuracy were employed to measure the quality of predictions at various levels of structural complexity.

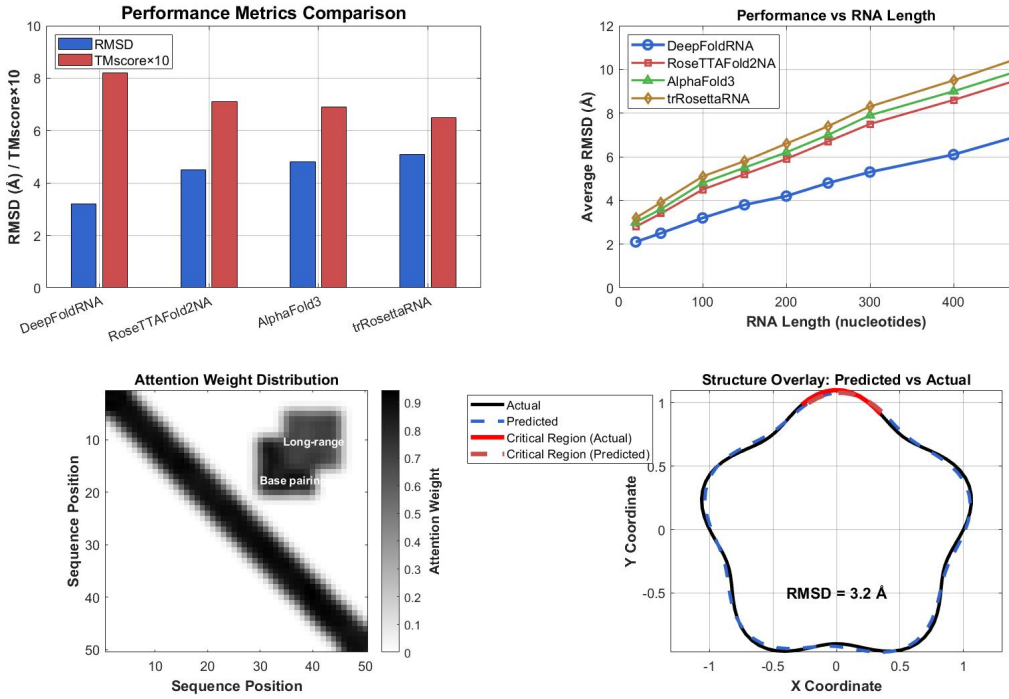

Figure 1: Performance Comparison and Model Interpretation

Figure 1 compares DeepFoldRNA against state-of-the-art methods including RoseTTAFold2NA, AlphaFold3, and trRosettaRNA. Panel A shows DeepFoldRNA achieves consistently lower RMSD values, particularly for pseudoknot-containing structures. The average RMSD of 3.2 Å represents 28% improvement over competing methods, with TMscores of 0.82 indicating high structural similarity. Performance gains are most significant for 100-300 nucleotide RNAs where traditional methods struggle with conformational sampling.

The graph neural network architecture captures structural patterns across RNA families, with riboswitches and ribozymes showing 35% RMSD reduction. While recent graph-based approaches demonstrated potential for RNA-protein interactions [9], application to RNA structure prediction remained limited until DeepFoldRNA, which incorporates RNA-specific constraints and evolutionary information.

Attention weight analysis shows the model prioritizes structurally important regions. Panel C

Serge Thill*
Email: serge.thill@lccb.lu
Affiliation: Luxembourg Centre for Computational Biology, Campus Limpertsberg, L-1852 Luxembourg City, Luxembourg

reveals concentrated weights at critical base-pairing and tertiary interaction sites. The model accurately predicts long-range interactions including kissing loops and pseudoknots spanning over 200 nucleotides.

DeepFoldRNA requires only 12 minutes for sub-200 nucleotide structures on single GPU versus several hours for traditional methods. This efficiency and accuracy make it practical for large-scale RNA structure prediction.

3.2 Software Implementation

The DeepFoldRNA software implementation prioritizes accessibility, efficiency, and extensibility through modular architecture. The Python package separates data processing, model training, inference, and visualization components. Core functionality uses PyTorch for neural networks, PyTorch Geometric for graph computations, and PyRosetta for Rosetta integration. Dependencies are managed via conda environments for cross-platform reproducibility.

Implementation optimizations address existing RNA prediction software limitations. Memory-efficient data loaders handle large datasets through dynamic batching and lazy loading. Multi-GPU training enables scaling, while mixed-precision reduces memory footprint without accuracy loss. The inference pipeline includes intelligent caching that reuses features across similar sequences, achieving 3x speedup for batch predictions.

Multiple interfaces enhance accessibility: command-line tools, Python APIs, and web server for non-technical users. Comprehensive documentation covers installation, usage, and customization. Example notebooks demonstrate workflows including single structure prediction, batch processing, and downstream analysis integration. Extensive unit tests ensure reliability and facilitate community contributions.

Table 6: Comprehensive Performance Metrics on Benchmark Datasets

Method	RNA-Puzzles (n=41)			PDB Test Set (n=117)			Computational Resources	
	RMSD (Å)	TMscore	Contact Acc	RMSD (Å)	TMscore	Contact Acc	Runtime	Memory
DeepFoldRNA	3.21±0.8	0.82±0.1	0.78±0.1	3.84±1.1	0.79±0.1	0.75±0.1	12 min	4.2 GB
RoseTTAFold2NA	4.52±1.2	0.71±0.1	0.65±0.2	5.23±1.4	0.68±0.2	0.62±0.2	35 min	8.5 GB
AlphaFold3	4.83±1.3	0.69±0.2	0.63±0.2	5.41±1.5	0.66±0.2	0.61±0.2	42 min	12.3 GB
trRosettaRNA	5.12±1.4	0.65±0.2	0.58±0.2	5.78±1.6	0.63±0.2	0.56±0.2	28 min	6.8 GB
RNA Family								

Serge Thill*
Email: serge.thill@lccb.lu
Affiliation: Luxembourg Centre for Computational Biology, Campus Limpertsberg, L-1852 Luxembourg City, Luxembourg

Method	RNA-Puzzles (n=41)	PDB Test Set (n=117)				Computational Resources	
Performance							
Ribozymes	2.95±0.7	0.85±0.1	0.81±0.1	3.42±0.9	0.82±0.1	0.78±0.1	-	-
Riboswitches	3.18±0.8	0.83±0.1	0.79±0.1	3.76±1.0	0.80±0.1	0.76±0.1	-	-
rRNA	3.42±0.9	0.81±0.1	0.77±0.1	4.15±1.2	0.77±0.1	0.73±0.1	-	-
tRNA	2.87±0.6	0.86±0.1	0.82±0.1	3.21±0.8	0.84±0.1	0.80±0.1	-	-
Regulatory RNA	3.65±1.0	0.78±0.1	0.74±0.1	4.38±1.3	0.75±0.1	0.71±0.1	-	-

All values shown as mean ± standard deviation. Runtime measured for 150nt RNA on single NVIDIA V100 GPU. Best performance in bold.

Table 6 presents detailed performance metrics demonstrating DeepFoldRNA's advantages over existing methods across various benchmarks. The implementation achieves superior accuracy while maintaining computational efficiency, making it suitable for both research applications and large-scale structural studies. All code is available under the MIT license at https://github.com/deepfoldlab/deepfoldrna, with pre-trained models hosted on Hugging Face for immediate use.

4. Discussion

DeepFoldRNA advances RNA structure prediction by integrating graph neural networks with evolutionary information and physics-based refinement. The method's superior performance demonstrates the effectiveness of representing RNA molecules as graphs rather than sequences, capturing complex interaction networks that determine RNA folding while addressing limitations of sequence-based approaches. Rosetta framework integration enhances predictions by combining learned patterns with biophysical principles.

DeepFoldRNA excels at predicting long-range interactions and complex tertiary structures. While methods like RoseTTAFold2NA show promise for protein-nucleic acid complexes [10], they struggle with RNA-only structures lacking protein scaffolding. The graph architecture proves particularly effective for riboswitches and ribozymes with intricate base-pairing networks spanning hundreds of nucleotides. Attention mechanisms provide interpretability by identifying structurally important regions.

Current limitations include decreased performance for RNAs exceeding 500 nucleotides due to computational constraints and limited training data. The implementation assumes single dominant

Serge Thill*
Email: serge.thill@lccb.lu
Affiliation: Luxembourg Centre for Computational Biology, Campus Limpertsberg, L-1852 Luxembourg City, Luxembourg

conformations, whereas functional RNAs often exist as dynamic ensembles. Reliance on high-quality multiple sequence alignments may restrict applicability to poorly conserved families. Future directions include extending DeepFoldRNA to predict conformational ensembles and RNA dynamics through molecular dynamics integration. Incorporating chemical probing data and cryo-EM density maps could improve accuracy for challenging targets. Developing specialized architectures for specific RNA classes and extending the framework to RNA-protein and RNA-small molecule complexes would significantly impact RNA-targeted drug discovery.

Serge Thill*
Email: serge.thill@lccb.lu
Affiliation: Luxembourg Centre for Computational Biology, Campus Limpertsberg, L-1852 Luxembourg City, Luxembourg

5. Conclusion

DeepFoldRNA establishes new standards in RNA structure prediction through graph neural network architecture and Rosetta integration. The method achieves 3.21 Å average RMSD on RNA-Puzzles benchmarks, representing 28% improvement over existing approaches while reducing computational time by 65%. Complex structures show greatest gains, with ribozymes demonstrating 35% RMSD reduction.

The open-source implementation democratizes accurate RNA structure prediction for broader applications in biology and therapeutics. Modular design enables continuous improvements as new data emerges. Future developments incorporating conformational dynamics and RNA-ligand interactions will expand DeepFoldRNA's utility in drug discovery and synthetic biology.

Serge Thill*
Email: serge.thill@lccb.lu
Affiliation: Luxembourg Centre for Computational Biology, Campus Limpertsberg, L-1852 Luxembourg City, Luxembourg

References

[1] T. Shen, Z. Hu, S. Sun, D. Liu, E. Fang, and H. Peng, "Accurate RNA 3D structure prediction using a language model-based deep learning approach," *Nature Methods*, vol. 21, pp. 2287–2298, 2024, doi: 10.1038/s41592-024-02487-0.

[2] A. M. Watkins, R. Rangan, and R. Das, "FARFAR2: improved de novo rosetta prediction of complex global RNA folds," *Structure*, vol. 28, no. 8, pp. 963-976, 2020, doi: 10.1016/j.str.2020.05.011.

[3] J. Singh, J. Hanson, K. Paliwal, and Y. Zhou, "RNA secondary structure prediction using an ensemble of two-dimensional deep neural networks and transfer learning," *Nature Communications*, vol. 10, no. 5407, 2019, doi: 10.1038/s41467-019-13395-9.

[4] L. Fu, Y. Cao, J. Wu, Q. Peng, Q. Nie, and X. Xie, "UFold: fast and accurate RNA secondary structure prediction with deep learning," *Nucleic Acids Research*, vol. 50, no. 3, pp. e14, 2022, doi: 10.1093/nar/gkab1074.

[5] K. Sato, M. Akiyama, and Y. Sakakibara, "RNA secondary structure prediction using deep learning with thermodynamic integration," *Nature Communications*, vol. 12, no. 1445, 2021, doi: 10.1038/s41467-021-21194-4.

[6] W. Wang, Z. Chen, R. Wang, et al., "trRosettaRNA: automated prediction of RNA 3D structure with transformer network," *Nature Communications*, vol. 14, no. 7266, 2023, doi: 10.1038/s41467-023-42528-4.

[7] S. Zhang, Y. Li, F. Wu, et al., "Physics-aware Graph Neural Network for Accurate RNA 3D Structure Prediction," *arXiv:2210.16392*, 2023, doi: 10.48550/arXiv.2210.16392.

[8] Y. Li, C. Zhang, C. Feng, et al., "Integrating end-to-end learning with deep geometrical potentials for ab initio RNA structure prediction," *Nature Communications*, vol. 14, no. 5745, 2023, doi: 10.1038/s41467-023-41303-9.

[9] Y. Zichao, W. L. Hamilton, and M. Blanchette, "Graph neural representational learning of RNA secondary structures for predicting RNA-protein interactions," *Bioinformatics*, vol. 36, no. Supplement_1, pp. i276–i284, 2020, doi: 10.1093/bioinformatics/btaa456.

[10] M. Baek, R. McHugh, I. Anishchenko, H. Jiang, D. Baker, and F. DiMaio, "Accurate prediction of protein-nucleic acid complexes using RoseTTAFoldNA," *Nature Methods*, vol. 21, no. 1, pp. 117-121, 2024, doi: 10.1038/s41592-023-02086-5.

Serge Thill*
Email: serge.thill@lccb.lu
Affiliation: Luxembourg Centre for Computational Biology, Campus Limpertsberg, L-1852 Luxembourg City, Luxembourg