LoRAFold : Leveraging Protein Foundation Models for RNA 3D Structure Prediction

Anselme Canivet, Defne Ozguven
Massinissa Hamidi, Jean-Christophe Gelly, Fariza Tahi

(Deep learning-based) Prediction of RNA 3D Structures is challenging

- Ribonucleic Acid (RNA) molecules play an important role in various biological processes. Predicting their 3D structures is essential to understand their function;
- Protein structure prediction has seen great advancements with the introduction of AlphaFold [1];
- However, RNA structure prediction still lags far behind protein prediction, because, compared to proteins:
  - there are 70 times less experimentally resolved RNA sequences;
  - RNA has greater conformational flexibility.

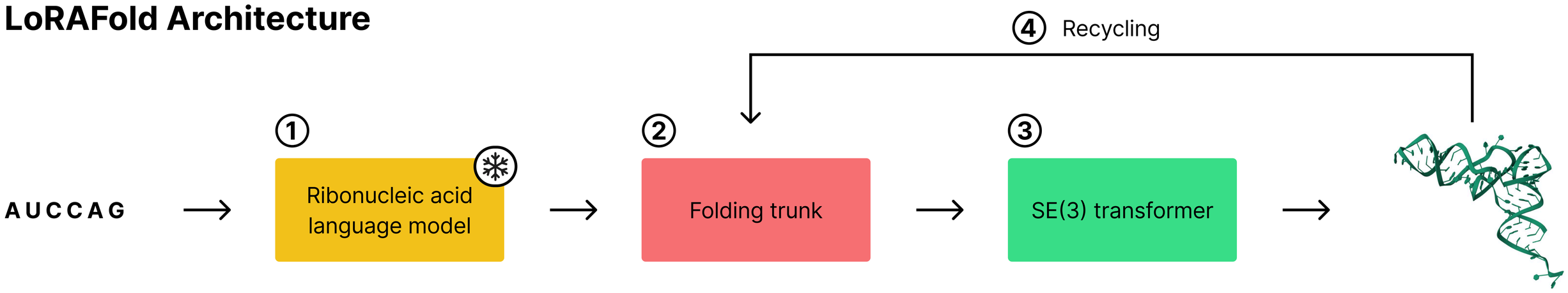
Can we leverage protein folding knowledge to predict RNA 3D Structure?

Recent Successes of Protein Structural Foundation Models are Promising

- AlphaFold [1] is the first deep learning model that can regularly predict protein structures with atomic accuracy even in cases in which no similar structure is known.
- ESMFold [2] is a large language model with 15B parameters that replaces the computationally expensive MSA module of AlphaFold.
- The third iteration of AlphaFold [3] and its other forks (Chai-1, Boltz-1) are able to predict the structure of different molecule types, with the same network components

This suggests the presence of high-level similarities in the folding of biomolecules, which could be used to overcome the lack of experimentally resolved RNA structures.

LoRAFold Architecture



① RNA Language Model

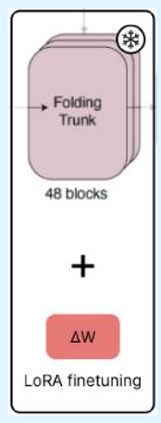
LoRAFold uses RiNALMo [4]:

- a large language model trained on RNA sequences;
- it is used as an encoder to embed the input RNA sequences;
- its weights are completely frozen during our training process.

② Folding Trunk

The folding trunk is based on ESMFold:

- Its blocks were pre-trained on protein molecules;
- Trainable parameters are quantized from 8-bit to 4-bit precision;
- Low-Rank Adaptation (LoRA) ensures that the pre-trained parameters remain frozen and that only low-rank matrices are fine-tuned.

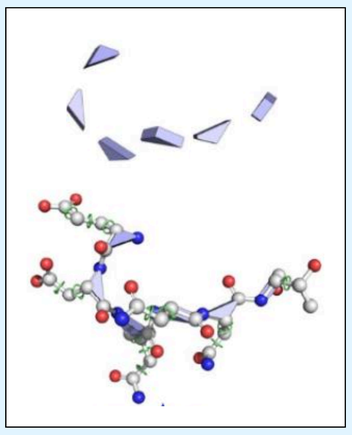


③ SE(3) Transformer (Structure Module)

SE(3) Transformer is a variant of self-attention module for 3D clouds.

RNA molecule is formed of nucleotides, which are formed of atoms. The structure module will therefore :

- Calculate rotation and translation matrices using Invariant Point Attention (IPA);
- Predict 3 atoms of a nucleotide, creating its “backbone” structure;
- Predict the rest of the atoms using 4 torsion angles.



④ Recycling

- The network is inferred multiple times iteratively, each time the outputs of the *i−1* iteration are embedded as inputs of the iteration *i;*
- It allows to make the network deeper without increasing the training time or the number of parameters.

Dataset

We use the RNA3DB [5] dataset:

- contains non-redundant RNA structures taken from Protein Data Bank;
- sequences are grouped according to sequence and structural similarity to minimize risks of overfitting;
- split into 9712 structures for training, 513 for validation and 1397 for testing.

Optimized Losses

The optimized loss is a weighted sum of multiple components :

$$\mathcal{L} = \lambda_{bb} \cdot \mathcal{L}_{bb} + \lambda_a \cdot \mathcal{L}_a + \lambda_{is} \cdot \mathcal{L}_{is} + \lambda_d \cdot \mathcal{L}_d + \lambda_{plddt} \cdot \mathcal{L}_{plddt}$$

- $\mathcal{L}_{bb}$ is frame aligned point error (FAPE) loss on backbone frames;
- $\mathcal{L}_a$ is loss for predicted torsion angles;
- $\mathcal{L}_{is}$ is the backbone and angles losses computed on intermediate structures;
- $\mathcal{L}_d$ and $\mathcal{L}_{plddt}$ are losses for predicted discretized pairwise distances and confidence.

- Bfloat16 mixed-precision training on A100 GPU with 80GB VRAM.
- 717 M total parameters with 14.8 M trainable.

References

1. Jumper & al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*
2. Lin & al. (2023). Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*.
3. Abramson & al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*.
4. Penić and al., (2024), RiNALMo: General-Purpose RNA Language Models Can Generalize Well on Structure Prediction Tasks, arXiv:2403.00043.
5. Szikszai & al. (2024). RNA3DB: A structurally-dissimilar dataset split for training and benchmarking deep learning models for RNA structure prediction. *Journal of molecular biology*.