RNA Secondary Structure Prediction: Classical vs Quantum-Inspired Benchmarking

RNA Sequence Overview

- **Type**: Transfer RNA (tRNA Phe)
- **Organism**: *Saccharomyces cerevisiae* (Yeast)
- **Source**: PDB ID: 1EHZ, Chain A
- **Length**: 76 nucleotides
- **Known Base Pairs (Experimental)**: 30
- **Experimental Dot-Bracket**:
  (((((((.((((]{[.[<..))))(((((((..).)))))))...}((((((]>.)].)))))))))))..

Benchmarking Methods

We evaluated **four prediction methods**: three classical algorithms and one quantum-inspired approach based on QUBO formulation and simulated annealing. All predictions were compared to the experimentally determined structure using standard evaluation metrics.

1. ViennaRNA (RNAfold)

- Dynamic programming using Minimum Free Energy (MFE) model
- **No pseudoknot support**

2. IPknot (McCaskill + Refinement)

- Integer programming approach using Maximum Expected Accuracy (MEA)
- **Supports pseudoknots**

3. IPknot++ (NUPACK, No Refinement)

- Uses NUPACK model without parameter refinement
- **Supports pseudoknots**

4. QUBO + Simulated Annealing

- Encodes RNA folding constraints into a QUBO matrix:
    - Watson-Crick & wobble base pair validity
    - Thermodynamic pair strengths
    - No-overlap and stacking constraints
- Solved using classical simulated annealing (via neal)
- **Pseudoknot support: limited (by construction)**

Evaluation Metrics

- **TP**: Correct base pairs (in both prediction & experimental)
- **FP**: Incorrect base pairs (predicted but not present)
- **FN**: Missed base pairs (present in ground truth but not predicted)
- **Precision** = TP / (TP + FP)
- **Recall** = TP / (TP + FN)
- **F1 Score** = 2 × (Precision × Recall) / (Precision + Recall)

Results Summary

Method	TP	FP	FN	Precision	Recall	F1 Score
ViennaRNA (RNAfold)	21	0	9	**1.000**	0.700	0.824
IPknot (McCaskill)	11	9	9	0.550	0.550	0.550
IPknot++ (NUPACK)	0	8	30	0.000	0.000	0.000
QUBO + Simulated Annealing	**0**	**9**	**30**	**0.000**	**0.000**	**0.000**

Quantum-Inspired Method Insights

Despite enforcing biologically meaningful constraints like:

- Thermodynamic pair scoring
- Hard overlap penalties
- Stacking bonuses

...the QUBO model consistently failed to predict a **single true base pair** when benchmarked against either the **experimental** or **ViennaRNA-derived** structure. Multiple parameter adjustments, scoring tweaks, and biological filters were attempted, including:

- Varying top-k% pair selection (from 25% to 15%)
- Adjusting conflict and stacking penalties
- Bias scaling in the QUBO

Yet the final **QUBO-annealed structure** produced zero true positives: