

# FR3D/BGSU under400 Baseline Comparison

## Scope

This report compares the current pipeline against three single-sequence baselines on the FR3D/BGSU representative non-rRNA set (**under 400 nt**). Metrics focus on **top-1** (F1@1) and **best-of-100** (F1@100).

## Methods (high level)

- **RNAneal-ss (pipeline)**: existing CaCoFold + refinement + MCMC pipeline outputting 100 candidates.
- **RNAstructure**: Fold MFE structure (padded to 100 for consistent @100 evaluation).
- **LinearFold**: LinearFold-V with `--zucker` suboptimal enumeration (first 100 structures used).
- **EternaFold**: MEA prediction plus additional structures from `contrafold sample` to reach 100 total candidates.

## Results

Table 1: Overall performance on FR3D/BGSU under400 non-rRNA benchmark (N=804; metrics @1 and best-of-100).

Method	Mean F1@1	Med F1@1	Mean F1@100	Med F1@100	$\Delta$ F1	Fail@100
RNAneal-ss (pipeline)	0.636	0.701	0.784	0.872	0.148	1.2%
RNAstructure (Fold MFE)	0.627	0.688	0.627	0.688	0.000	5.7%
LinearFold-V (Zuker)	0.622	0.682	0.794	0.846	0.172	0.6%
EternaFold (MEA + samples)	0.656	0.733	0.820	0.872	0.163	0.1%

## Empirical CDFs

Table 2: Mean F1@1 / F1@100 by truth-length bucket (N=804).

Bucket	Method	N	Mean F1@1	Mean F1@100
30-80	RNAnneal-ss (pipeline)	587	0.660	0.820
30-80	RNAstructure (Fold MFE)	587	0.647	0.647
30-80	LinearFold-V (Zuker)	587	0.637	0.815
30-80	EternaFold (MEA + samples)	587	0.670	0.845
81-150	RNAnneal-ss (pipeline)	149	0.625	0.756
81-150	RNAstructure (Fold MFE)	149	0.633	0.633
81-150	LinearFold-V (Zuker)	149	0.650	0.803
81-150	EternaFold (MEA + samples)	149	0.675	0.815
151-300	RNAnneal-ss (pipeline)	53	0.441	0.533
151-300	RNAstructure (Fold MFE)	53	0.439	0.439
151-300	LinearFold-V (Zuker)	53	0.429	0.603
151-300	EternaFold (MEA + samples)	53	0.496	0.628
301-400	RNAnneal-ss (pipeline)	15	0.468	0.538
301-400	RNAstructure (Fold MFE)	15	0.468	0.468
301-400	LinearFold-V (Zuker)	15	0.475	0.563
301-400	EternaFold (MEA + samples)	15	0.502	0.546

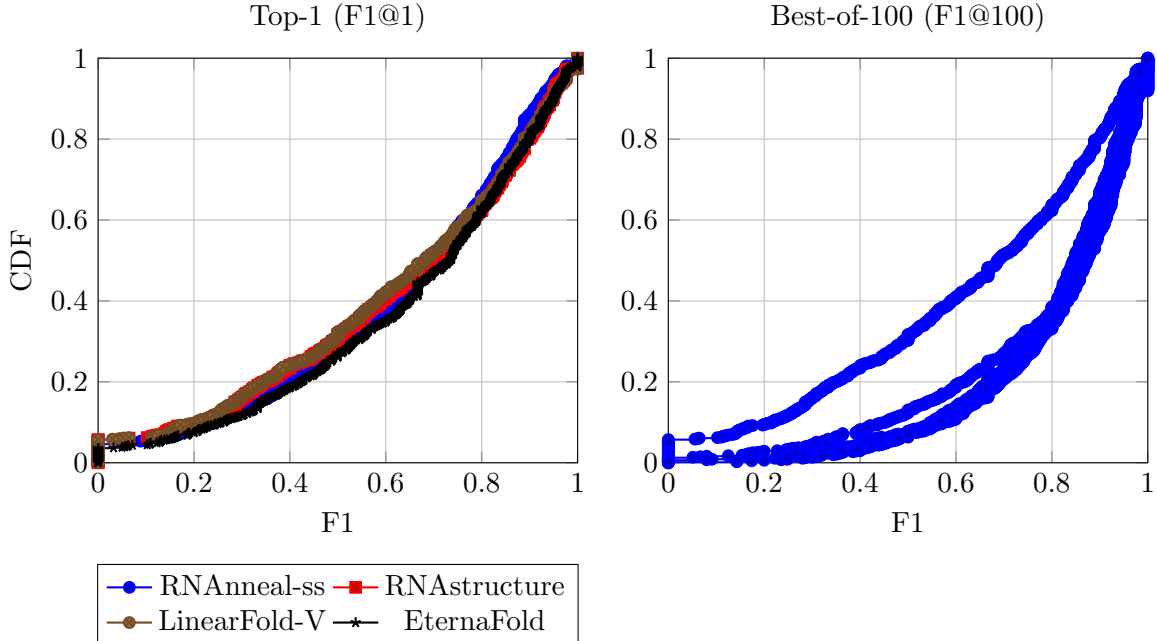


Figure 1: Empirical CDFs of per-target F1. Curves that shift right indicate higher accuracy.