BeeRNA: tertiary structure-based RNA inverse folding using Artificial Bee Colony

Mehyar Mlaweh[1], **Tristan Cazenave**[1], **Ines Alaya**[1]

[1]LAMSADE, Université Paris Dauphine – PSL, Paris, France
mehyar.mlaweh@dauphine.eu, tristan.cazenave@lamsade.dauphine.fr, ialaya@parisnanterre.fr

Abstract

The Ribonucleic Acid (RNA) inverse folding problem, designing nucleotide sequences that fold into specific tertiary structures, is a fundamental computational biology problem with important applications in synthetic biology and bioengineering. The design of complex three-dimensional RNA architectures remains computationally demanding and mostly unresolved, as most existing approaches focus on secondary structures. In order to address tertiary RNA inverse folding, we present BeeRNA, a bio-inspired method that employs the Artificial Bee Colony (ABC) optimization algorithm. Our approach combines base-pair distance filtering with RMSD-based structural assessment using RhoFold for structure prediction, resulting in a two-stage fitness evaluation strategy. To guarantee biologically plausible sequences with balanced GC content, the algorithm takes thermodynamic constraints and adaptive mutation rates into consideration. In this work, we focus primarily on short and medium-length RNAs (< 100 nucleotides), a biologically significant regime that includes microRNAs (miRNAs), aptamers, and ribozymes, where BeeRNA achieves high structural fidelity with practical CPU runtimes. The lightweight, training-free implementation will be publicly released for reproducibility, offering a promising bio-inspired approach for RNA design in therapeutics and biotechnology.

Introduction

The Ribonucleic Acid (RNA) design problem, also known as *inverse folding*, aims to identify sequences that fold into a given target structure (Zadeh et al. 2011). Hofacker et al. (Hofacker et al. 1994) first formalized the problem for secondary structures using the ViennaRNA package. The problem was later identified as NP-hard because of the large sequence space and intricate constraints such as thermodynamic stability and biological functionality (Lyngsø, Zuker, and Pedersen 1999).

RNA design helps scientists create synthetic RNA molecules for specific purposes, such as ribozymes (Yamagami et al. 2019), microRNAs (Schwab et al. 2006), aptamers (Hamada 2018), and riboswitches (Bauer and Suess 2006; Findeiß et al. 2017). Yamagami et al. (2019) used computational-experimental methods to increase the catalytic efficiency of dual-pseudoknotted ribozymes (Yamagami et al. 2019). MicroRNAs were created by Schwab et al. (2006) to silence plant genes, and synthetic riboswitches are used as biosensors in synthetic biology (Findeiß et al. 2017).

Karaboga proposed the Artificial Bee Colony (ABC) method in (Karaboga 2005), which models bee foraging behavior as an optimization tool in various disciplines including engineering, scheduling, and bioinformatics. The ABC algorithm imitates bee collaboration to discover optimal food sources, thus maintaining a useful balance between exploration and exploitation in complex solution spaces. Research in bioinformatics has demonstrated that ABC improves free energy calculations for lattice-based protein structures in protein inverse folding (Lin and Su 2012). Following this, a balance-evolution ABC algorithm was proposed in 2015 to handle protein structure optimization using a three-dimensional AB off-lattice model, offering a more realistic representation of protein folding and enhanced convergence capabilities (Li, Chiong, and Lin 2015). However, the application of ABC to RNA inverse folding, particularly for tertiary structure prediction, remains underexplored and presents a promising direction for future research.

In this study, we introduce BeeRNA, a novel RNA inverse folding method that adapts the ABC algorithm to design sequences for target tertiary structures. BeeRNA contrasts with existing methods (e.g., ViennaRNA (Lorenz et al. 2011), NUPACK (Zadeh et al. 2011)) because it handles tertiary structures with ABC-based optimization. Using RhoFold (Shen et al. 2024) to fold designed sequences, BeeRNA achieves better structural correctness than gRNAde (Joshi et al. 2025) by obtaining a 2.50 Å RMSD score (Root Mean Square Deviation) on the RNASolo dataset (Adamczyk, Antczak, and Szachniuk 2022), 12.02 Å on a benchmark of 14 diverse RNA structures (Das, Karanicolas, and Baker 2010), and 14.98 Å on the RFAM dataset (Griffiths-Jones et al. 2003; Kalvari et al. 2018), while gRNAde achieves 9.33 Å, 14.63 Å, and 16.24 Å respectively. BeeRNA also achieves superior accuracy through Global Distance Test Total Score (GDT-TS) evaluation, which demonstrates its structural stability. Importantly, BeeRNA is specifically optimized for short and medium-length RNAs ($<$100 nucleotides)—a biologically relevant regime that includes microRNAs

(miRNAs), aptamers, and ribozymes—where its training-free, metaheuristic optimization achieves a strong balance between structural precision and computational efficiency.

The paper is organized as follows: *Related Work* reviews prior RNA inverse folding research; *Methodology* details the BeeRNA algorithm; *Results* compares its performance with existing methods; *Conclusion* summarizes key findings.

Related Work

Inverse folding, or inverse design of RNA sequences to fold into a target 3D structure, holds promise for applications in mRNA vaccines, aptamer therapy, and riboswitches. Inverse folding would deliver nucleotide sequences that take on a pre-determined tertiary structure, and it is a challenging problem considering the vast chemical space and requirements like stability and biological activity. The goal of RNA design is to identify a sequence that folds into a given tertiary structure. This section provides a concise review of the major RNA inverse folding methods rooted in deterministic, stochastic, and deep learning approaches, with a focus on the contribution of bio-inspired metaheuristics, the ABC algorithm, upon which our suggested algorithm is founded.

Deterministic Methods

Deterministic methods rely on thermodynamic models to predict RNA sequences capable of folding into a desired RNA structure. ViennaRNA (Lorenz et al. 2011) operates through dynamic programming to determine minimum free energy secondary structures, which allows users to develop sequences that naturally fold into target **secondary structures**. Recent advancements have explored thermodynamic integration with machine learning, as seen in methods like MXfold2 (Sato, Akiyama, and Sakakibara 2021), which enhance secondary structure prediction by combining free energy parameters with neural networks. However, these methods face limitations because they depend on thermodynamic approximations, which prevent them from analyzing complex **tertiary structures** with 3D loops and non-local interactions. ViennaRNA only operates for secondary structure predictions; hence it fails to perform the tertiary inverse folding functions needed for this study. Deterministic methods do not have the capability to explore various conformational spaces, thus requiring the adoption of adaptive methods to analyze biologically significant tertiary structures.

Stochastic Methods

Stochastic methods are ideal for exploring the vast and complex space of possible RNA sequences, as they employ probabilistic techniques to sample candidate nucleotide combinations. This flexibility enables them to model complex tertiary structures at the possible cost of slower convergence and sensitivity to hyperparameters.

A prominent example is Monte Carlo Inverse RNA Folding (Cazenave and Touzani 2024), which generates and evaluates candidate sequences using Monte Carlo simulations, then iteratively changes them to fit a target structure. Although effective at exploring nonlinear search spaces, its computational cost and parameter dependence can restrict effectiveness. Simulated annealing, inspired by metallurgical cooling processes, starts with broad exploration at a high **temperature** and gradually narrows to optimize free energy. Studies show it achieves 80% structural fidelity for small secondary structures ($<$ 50 nucleotides) (Kai et al. 2019), but its performance reduces for tertiary folding owing to complex 3D interactions. Bio-inspired stochastic methods such as BE-ABC (Li, Chiong, and Lin 2015) have been employed to optimize tertiary structures using a 3D off-lattice model, providing insightful analysis that can be applied for RNA tertiary inverse folding, efficiently balancing exploration and exploitation.

Deep Learning Methods

Deep learning approaches have transformed RNA inverse folding by learning to map structural features, such as inter-atomic distances and 3D spatial coordinates of atoms, to corresponding nucleotide sequences. **R3Design** (Tan et al. 2025) achieves 43% native sequence recovery for simple tertiary structures by using a graph-based encoder to capture 3D RNA coordinates and an iterative decoder directed by secondary structure constraints. By simulating primary, secondary, and spatial nucleotide interactions, relational graph neural networks improve design even more. The geometric neural network (GNN)-based technique **gRNAde** (Joshi et al. 2025) encodes 3D conformational ensembles as graphs and employs autoregressive decoding to produce compatible sequences. It achieves fast inference ($<$1 s) and a 56.80% native sequence recovery rate significantly outperforming Rosetta (45%) and R3Design (43%), with experimental success in 50% of pseudoknot cases. But it needs a lot of pre-training on different 3D structures. **RiboDiffusion** (Huang et al. 2024), a generative diffusion model, leverages a two-module architecture combining a GVP-GNN-based structure module and a Transformer-based sequence module to iteratively refine random sequences for RNA inverse folding based on tertiary structures. It achieves an impressive 58.96% recovery rate on sequence similarity splits (Seq. 0.8) and 66.40% on structure similarity splits (Struct. 0.6)[1], outperforming gRNAde by approximately 2.16% and 9.60% respectively, while also demonstrating robust performance across various RNA lengths and types, including tRNA, rRNA, and ribozymes. Notably, the RiboDiffusion paper highlights that short RNAs ($<$ 50 nt) present a challenge for the model to recover the original sequence due to their flexible conformation, which impacts recovery rates and in-silico folding performance. A recent advance, **RISoTTo** (Bibekar, Krapp, and Dal Peraro 2025), employs a geometric transformer with a specialized diffusion process that conditions both on secondary and tertiary contact maps. It achieves a 62% native sequence recovery rate without data splits based on sequence or structural similarity, outperforming prior models such as gRNAde

[1]Seq. 0.8 and Struct. 0.6 denote test sets clustered by 80% sequence or 60% structural similarity to the training set.

and RiboDiffusion on unbiased benchmarks. These methods excel in speed and precision but demand large datasets and struggle with short and medium RNAs.

Bio-Inspired Metaheuristics

Bio-inspired metaheuristics draw from natural processes to optimize complex problems like RNA inverse folding. These algorithms steer clear of local optima by striking a balance between exploring novel solutions and taking advantage of existing ones. Some notable examples are the ABC algorithm (Karaboga 2005), which models foraging behavior. Ant colony optimization (Dorigo 1996), based on pheromone trails. Particle swarm optimization (Kennedy and Eberhart 1995), inspired by flock behavior. And genetic algorithms (Holland 1975), which mimic Darwinian evolution.

Artificial Bee Colony has demonstrated great promise for protein inverse folding. Chen et al. (Lin and Su 2012) used an effective ABC algorithm to predict protein structures on 2D and 3D lattice models, achieving lower free energy conformations (e.g., 20-amino-acid sequences with energy values improved by up to 15% over prior methods). Their method demonstrated ABC's capacity to traverse intricate energy landscapes by optimizing hydrophobic-polar (HP) lattice models. Recent work has also explored multiobjective evolutionary computation, such as eM2dRNAs (Álvaro Rubio-Largo et al. 2023), which decomposes target structures to enhance RNA sequence design, suggesting potential adaptations for tertiary folding.

While the ABC algorithm has shown success in protein inverse folding, its application to RNA inverse folding, particularly for designing sequences that fold into specific tertiary structures, remains underexplored.

BeeRNA method

This section introduces BeeRNA, a new RNA inverse folding process that combines the ABC algorithm (Karaboga 2005) with sequence optimization for creating target tertiary RNA structures. As seen in Figure 1, BeeRNA uses RhoFold (Shen et al. 2024) as the structure prediction tool to evaluate the quality of the generated sequences. Notably, other recent methods such as gRNAde (Joshi et al. 2025) and RiboDiffusion (Huang et al. 2024) also employ RhoFold to predict the tertiary structures of their designed sequences, establishing it as a common benchmark for fair structural evaluation. The AlphaFold 3 framework also shows strong potential for accurate RNA and RNA–protein structure prediction, offering future opportunities for integration.

Problem Formulation

The RNA inverse folding problem aims to identify a nucleotide sequence $S = \{s_1, s_2, \ldots, s_n\}$, where $s_i \in \{A, U, G, C\}$ and n is the sequence length, that folds into a target tertiary structure T_{3D}, defined by 3D atomic coordinates from a PDB file. The objective is to find:

$$S^* = \arg\min_S \text{RMSD}(F(S), T_{\text{3D}}),$$

where $F(S)$ is the 3D structure predicted by RhoFold for sequence S, and RMSD measures the structural deviation between $F(S)$ and T_{3D}. Additional constraints ensure thermodynamic stability, computed using ViennaRNA's minimum free energy calculations (Lorenz et al. 2011), and biological plausibility, such as maintaining GC content between 40% and 60% (Zadeh et al. 2011). The GDT-TS serves as a secondary metric during evaluation to assess structural similarity.

Optimization Using the ABC Algorithm

BeeRNA optimizes RNA sequences by integrating a two-stage fitness evaluation with the Artificial Bee Colony algorithm, inspired by bee foraging behavior, to efficiently explore the vast sequence space and identify sequences that closely match the target tertiary structure. The fitness function is defined as follows:

$$\text{Fitness}(S) = \begin{cases} \infty & \text{if BPD}(S, T_{\text{sec}}) > 0, \\ \text{RMSD}(F(S), T_{\text{3D}}) & \text{if BPD}(S, T_{\text{sec}}) = 0 \end{cases}$$

where S is the candidate sequence, T_{sec} is the target *secondary* structure extracted from the tertiary reference T_{3D}, and $\text{BPD}(S, T_{\text{sec}})$ is the base-pair distance between the predicted secondary structure of S (via ViennaRNA (Lorenz et al. 2011)) and T_{sec}. Because ViennaRNA's dot–bracket representation encodes only canonical and wobble pairs, a perfect base-pair match ($\text{BPD} = 0$) is achievable only when the target structure itself contains exclusively canonical pairs. For targets that include wobble or non-canonical interactions (as defined by the Leontis–Westhof classification), BPD values remain nonzero even for biologically consistent folds. Such small deviations are tolerated within BeeRNA's filtering stage, as they typically reflect permissible tertiary motifs rather than structural errors. In the second stage of the fitness evaluation, $F(S)$ is the predicted tertiary structure of S (via RhoFold (Shen et al. 2024)), and $\text{RMSD}(F(S), T_{\text{3D}})$ is the Root Mean Square Deviation between $F(S)$ and the target tertiary structure.

RMSD is calculated as:

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \|\mathbf{r}_i - \mathbf{r}'_i\|^2},$$

where N is the number of aligned atoms, $\mathbf{r}_i$ are the coordinates of the i-th atom in the predicted structure, and $\mathbf{r}'_i$ are the corresponding coordinates in the target structure after optimal superposition. For superposition, we used **US-align** (Zhang et al. 2022), which applies the Kabsch least-squares algorithm to minimize atomic deviations. The RMSD computation considers the backbone phosphorus atom (**P**), the sugar carbon atom at position 4 ($C4'$), and the base nitrogen atoms involved in pairing (**N1** for pyrimidines and **N9** for purines). Lower RMSD values indicate better structural alignment.

Initialization: BeeRNA initializes a population of 40 RNA sequences ($N = 40$), each with length n matching the residue count of the target sequence extracted from the target

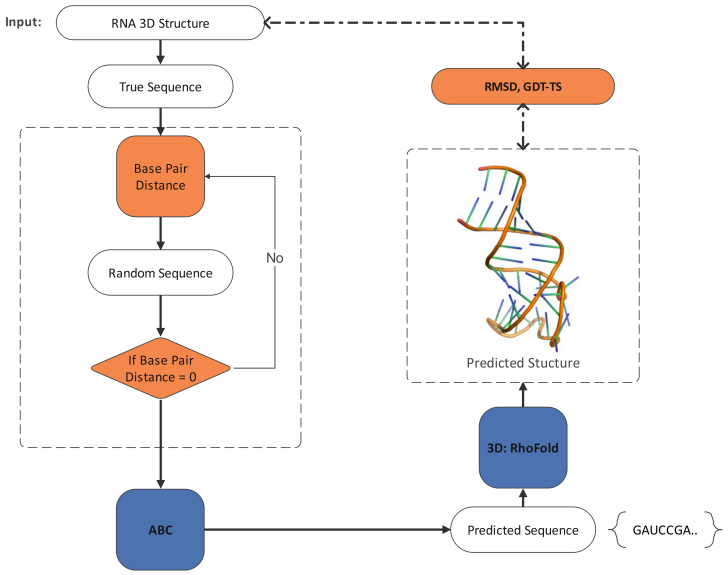

Figure 1: BeeRNA optimization process for a sample RNA structure. The flowchart illustrates the step-by-step progression of BeeRNA, starting from the initial RNA 3D structure to the final predicted sequence, using the ABC that employs RMSD calculation as the fitness function, comparing the predicted structure generated by RhoFold with the initial structure. A preliminary filter of candidate sequences is applied using base-pair distance to enhance efficiency.

structure. Initial sequences are generated by extracting base-pairing constraints from the target secondary structure using ViennaRNA's **fold** function (Lorenz et al. 2011). Unpaired positions (dots in the secondary structure) are assigned random nucleotides from $\{A, U, G, C\}$, while paired positions (parentheses) receive complementary base pairs (G–C or C–G) randomly to ensure **Watson–Crick pairing**, which stabilizes RNA secondary structures (Watson and Crick 1953). To help ensure thermodynamic stability, the G/C content of each sequence is maintained between **40%** and **60%** (Zadeh et al. 2011).

Employed Bees Phase: In this phase, each of the 40 sequences in the population generates a neighbor solution by applying mutations. Mutations are performed with an adaptive mutation rate, initially set to 0.095 and adjusted based on the best RMSD achieved:

$$\text{mutation\_rate} = \max\left(0.1,\ 0.095 \cdot e^{-\frac{\text{best\_RMSD}}{5n}}\right)$$

where n is the sequence length. The number of mutations is $\max(1, \lfloor \text{mutation\_rate} \cdot n \rfloor)$, applied at randomly selected positions. The nucleotide swapping mechanism between {A, U} and {G, C} at unpaired positions occurs with a 50% chance of maintaining GC content when it surpasses 50%. The swap mutation method involves exchanging nucleotides between two positions that exist within three positions of each other with a 20% chance. The fitness evaluation of neighbor solutions follows a two-stage process which first checks base-pair distance before calculating RMSD when the base-pair distance is zero. When the neighbor sequence shows better fitness through lower RMSD, it replaces the original sequence; otherwise, the trial counter increments. The adaptive mutation schedule is inspired by simulated annealing, allowing BeeRNA to explore widely at early stages and gradually exploit promising regions as convergence improves. The parameters (initial rate 0.095, decay factor $5n$, and lower bound 0.1) were determined through preliminary tuning to ensure stable convergence across diverse RNA lengths while preventing both premature stagnation and excessive randomization.

Onlooker Bees Phase: Onlooker bees probabilistically select sequences for further exploration. The selection probability for a sequence with RMSD r_i is:

$$p_i = \frac{e^{-r_i/\tau}}{\sum_{j=1}^{N} e^{-r_j/\tau}}$$

where $\tau = 5.0 \cdot (1 + t/T)$ is a temperature parameter that increases with iteration t (out of total iterations $T = 40$), encouraging exploration early and exploitation later. Selected sequences undergo the same mutation process as in the employed bees phase, and their fitness is evaluated. If a neighbor solution improves the RMSD, it replaces the original sequence, and the trial counter resets; otherwise, the trial counter increments.

Scout Bees Phase: Scout bees address sequences that fail to improve after 5 trials by replacing them with

new randomly initialized sequences, generated as in the initialization phase. This ensures the algorithm escapes local optima and maintains diversity in the population.

Experimental Results

In this section, we evaluate the performance of BeeRNA for RNA inverse folding through three distinct experiments using diverse RNA datasets. These experiments assess BeeRNA's ability to design sequences that fold into target tertiary structures, using RhoFold (Shen et al. 2024) for structure prediction and RMSD and GDT-TS as metrics for structural accuracy. BeeRNA was compared against gRNAde (Joshi et al. 2025), a recently introduced state-of-the-art deep learning-based method for 3D RNA design. The performance evaluation shows that gRNAde exceeds the capabilities of Rosetta (Das, Karanicolas, and Baker 2010) and RDesign (Tan et al. 2025) as well as FARNA (Rosetta's predecessor) across various testing scenarios that measure native sequence recovery. The demonstration of competitive or superior performance compared to gRNAde establishes BeeRNA's effectiveness and shows its ability to outperform the previous generation of RNA design methods which gRNAde improved upon. The evaluation process used fixed temperature settings of **0.1** and **16** sequence samples per structure to maintain consistency with the original gRNAde evaluation approach. We also compare BeeRNA with RiboDiffusion (Huang et al. 2024), a diffusion-based generative method for RNA design. For a fair comparison, we adopted the same hyperparameters used in the RiboDiffusion paper: the number of sampling steps set to **50**, and **8** sequences sampled per structure. However, due to the lack of transparency about the exact training set used for RiboDiffusion constructed from the Protein Data Bank (PDB) but not publicly released, we cannot determine whether the RNASolo or Rfam datasets used in our evaluation overlap with its training data. To avoid any risk of unfair comparison, we limit our benchmarking of RiboDiffusion to the 14 designated benchmark RNAs and exclude RNASolo and Rfam examples, which are likely part of its training set.

Experimental Setting

Datasets. The first experiment employed RNA sequences (< 30 nucleotides) from the **RNASolo** dataset (Adamczyk, Antczak, and Szachniuk 2022), providing a curated set of short RNA structures. The second experiment focused on non-coding RNA (ncRNA)[2] sequences from the **RFAM** database (Griffiths-Jones et al. 2003; Kalvari et al. 2018), with well-distributed lengths ranging from 25 to 200 nucleotides. Since RFAM provides secondary structures in FASTA format (available at https://rfam.org/), we used RhoFold (Shen et al. 2024) to generate 3D structures in PDB format, resulting in a dataset of 88 ncRNA structures. The third experiment included a benchmark set of **14 PDB structures** from (Das, Karanicolas, and Baker 2010), selected for their structural diversity and widespread use in RNA Design studies.

[2]Unlike coding RNA (e.g., mRNA) which is translated into proteins, non-coding RNA (ncRNA) functions without translation, playing structural or regulatory roles (e.g., rRNA, tRNA).

Implementation Details. BeeRNA runs for **40** iterations with a population size of **40** sequences. These hyperparameters were selected through a small grid search to balance runtime and optimization accuracy. Larger configurations provided only marginal RMSD improvements while nearly doubling runtime due to the computational cost of RhoFold predictions performed on CPU, whereas smaller populations reduced search diversity and led to premature convergence. The chosen setting thus offers an effective trade-off between exploration depth and computational efficiency. RhoFold predictions, used in the optimization process, are performed on a CPU (due to resource constraints). The final output includes the best sequence, its RMSD, and the inference time. Predicted tertiary structures are saved as PDB files.

Evaluation Metrics. Unlike prior methods (Joshi et al. 2025; Tan et al. 2025; Huang et al. 2024; Bibekar, Krapp, and Dal Peraro 2025) that primarily assess native sequence recovery (i.e., the percentage of matching nucleotides), BeeRNA evaluates predicted sequences by comparing their folded 3D structures to the target. This structure-based evaluation is more relevant to inverse folding, as high sequence identity does not ensure structural accuracy.

Figure 2 illustrates this difference: starting from the native sequence of RNA **2OUE**, we introduced a single-nucleotide mutation, resulting in a sequence with **98.4%** recovery (60/61 correct). Despite this high sequence similarity, the predicted 3D structure deviates significantly from the native one, with an RMSD of **19.34 Å** and a GDT-TS of only **11.59%**. Some of this difference may come from RhoFold's own prediction error, but even after allowing for that, the RMSD is still much higher than normal. This demonstrates the sensitivity of tertiary contacts in short RNAs, where even minor sequence perturbations can destabilize global folding.

RMSD: The primary evaluation metric is RMSD, used both in the fitness function and for final assessment. RMSD quantifies the structural deviation between the predicted structure $F(S)$ (from RhoFold) and the target structure T, providing a precise measure of atomic alignment after optimal superposition using US-align (Zhang et al. 2022). BeeRNA's optimization minimizes RMSD, ensuring the designed sequences produce tertiary structures closely matching the target.

GDT-TS: The Global Distance Test Total Score (Zemla 2003) serves as a secondary metric to evaluate the structural similarity, particularly for comparing BeeRNA. GDT-TS measures the fraction of residues in the predicted structure that fall within distance cutoffs (1, 2, 4, and 8 Å) of their corresponding residues in the target structure after superposition. It is calculated as:

$$\text{GDT-TS} = \frac{P_1 + P_2 + P_4 + P_8}{4},$$

where P_d is the percentage of residues within d Å. Higher GDT-TS scores indicate better structural alignment,

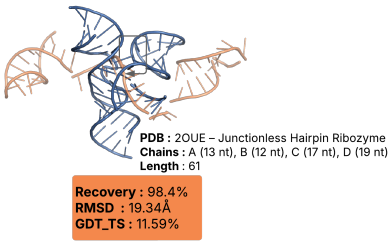

Figure 2: Predicted structure (orange) vs. sequence structure (blue) for 2OUE RNA. Despite 98.4% native recovery, RMSD and GDT-TS show a poor structural match.

complementing RMSD by capturing global structural similarity.

Runtime and Efficiency Evaluation. BeeRNA was executed on a workstation equipped with a 64-core CPU and 125 GB of RAM. The optimization process required approximately 3 minutes for RNAs shorter than 50 nt and 7–10 minutes for sequences up to 100 nt, including RhoFold predictions. Occasional runtime variation occurred when unstable or biologically implausible intermediate sequences caused RhoFold to take longer to converge.

Results on RNASolo

The results demonstrate that BeeRNA significantly outperforms gRNAde on the RNASolo dataset (Table 1, Figure 3). BeeRNA achieves an average GDT-TS of 26.91% and an average RMSD of 2.50 Å, compared to gRNAde's 18.97% and 9.33 Å, respectively, indicating superior structural accuracy and consistency. As shown in Figure 3, BeeRNA maintains stable RMSD across short sequences (3–30 nt), while gRNAde's error increases with length.

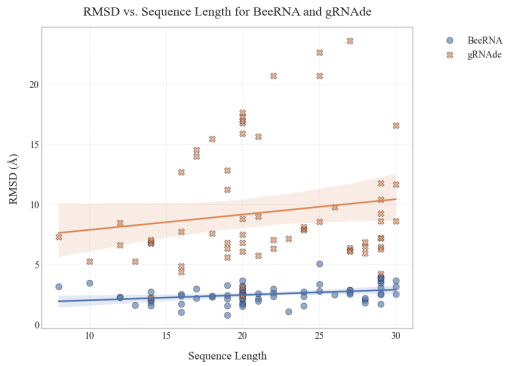

Figure 3: RMSD vs. sequence length for BeeRNA and gRNAde on RNASolo. Points represent RMSD values (Å) for individual sequences, plotted against sequence length (3–30 nucleotides). Regression lines show RMSD trends for each method.

Table 1: Average GDT-TS and RMSD performance metrics for the BeeRNA and gRNAde targets evaluated on the RNASolo dataset. RiboDiffusion results are excluded because its training data might include some of the RNASolo dataset

Metric	BeeRNA	gRNAde
RMSD (Å)	**2.50**	9.33
GDT-TS (%)	**26.91**	18.97

Results on RFAM

Figure 4 shows the RMSD plotted against sequence length for BeeRNA and gRNAde on the RFAM dataset. Each point represents an RNA sequence, and the shaded region highlights sequences longer than 100 nucleotides. Regression lines indicate how RMSD varies with sequence length for each method.

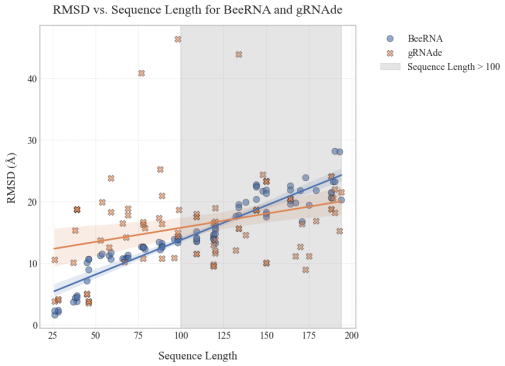

Figure 4: RMSD vs. sequence length for BeeRNA and gRNAde on RFAM. Shaded region marks sequences over 100 nucleotides

BeeRNA achieved a higher average GDT-TS score (11.56%) compared to gRNAde (9.77%), and a slightly lower average RMSD (14.98Å vs. 16.24Å) (Table 2). As shown in Figure 4, BeeRNA performs best for short and medium RNAs ($\leq$ 100 nt), maintaining lower RMSD values across this range. For longer RNAs, gRNAde slightly improves due to its deep learning architecture trained on long sequences, while BeeRNA's stochastic, physics-driven search remains advantageous for diverse or underrepresented families.

Table 2: Mean GDT-TS and RMSD scores for BeeRNA and gRNAde on RFAM. RiboDiffusion results are excluded because its training data might include some of the RFAM.

Metric	BeeRNA	gRNAde
RMSD (Å)	**14.98**	16.24
GDT-TS (%)	**11.56**	9.77

Table 3: Benchmark RNA structures ordered by increasing sequence length, with descriptions and RMSD (Å) comparison between BeeRNA, gRNAde, and RiboDiffusion. Best RMSD values per structure are shown in **bold**.

PDB ID	Description	Length ↑	BeeRNA	gRNAde	RiboDiffusion
1F27	Biotin-binding RNA pseudoknot	19	**2.21**	14.94	3.47
1LNT	RNA internal loop of SRP	22	**3.69**	17.51	8.08
354D	Loop E from *E. coli* 5S rRNA	23	**10.68**	17.38	12.89
1L2X	Viral RNA pseudoknot	27	**3.00**	3.93	3.47
1Q9A	Sarcin/ricin domain from *E. coli* 23S rRNA	27	**2.65**	6.57	4.35
1CSL	RRE high affinity site	28	**2.93**	3.36	12.25
1ET4	Vitamin B12 binding RNA aptamer	35	**11.34**	13.51	12.18
1XPE	ALL-RNA hairpin ribozyme	46	20.92	14.26	**10.85**
1X9C	HIV-1 RNA dimerization initiation site	60	**21.00**	25.19	26.79
2OUE	Junctionless hairpin ribozyme	61	**21.00**	22.14	24.71
4FE5	Guanine riboswitch aptamer	67	11.80	**8.42**	9.78
2GDI	Thiamine pyrophosphate-specific riboswitch	78	**8.00**	15.73	8.49
2GCS	Pre-cleavage state of glmS ribozyme	122	24.00	25.71	**1.66**
2R8S	*Tetrahymena* ribozyme P4–P6 domain	159	26.00	20.20	**5.38**

Results on the 14 Benchmark RNA Structures

Table 3 presents the RMSD values for BeeRNA, gRNAde, and RiboDiffusion across the 14 benchmark RNA structures (Das, Karanicolas, and Baker 2010), ordered by increasing sequence length. BeeRNA achieves the lowest RMSD values in 10 out of 14 cases, showing strong accuracy on short and medium RNAs such as 1F27 (2.21 Å) and 1LNT (3.69 Å). In contrast, RiboDiffusion performs best on long and complex RNAs (e.g., 2GCS and 2R8S). The higher RMSD values of 24 Å and 26 Å observed for these two structures result from the absence of an initial sequence achieving a BPD of zero, which constrained convergence. To handle such cases, a fixed RMSD penalty of 20 Å was applied to maintain optimization continuity. Future work will refine this initialization step through Monte Carlo–based sequence generation to improve convergence on long RNA targets.

As summarized in Table 4 BeeRNA achieves an average RMSD of 12.02 Å and GDT-TS of 15.92%, outperforming gRNAde (14.63 Å, 10.16%). RiboDiffusion yields the lowest overall RMSD (10.31 Å) but with greater variability across shorter targets. Overall, BeeRNA performs most reliably on RNAs below 100 nucleotides, covering most natural functional RNAs such as tRNAs, microRNAs, and aptamers, while maintaining competitive accuracy on larger structures without relying on pretraining.

Table 4: Average RMSD and GDT-TS across 14 benchmark RNA structures.

Metric	BeeRNA	gRNAde	RiboDiffusion
RMSD (Å)	12.02	14.63	10.31
GDT-TS (%)	15.92	10.16	22.69

The results from the RFAM and RNASolo datasets are reinforced by the benchmark evaluation, confirming BeeRNA's strength on short and medium RNAs ($\leq$ 100 nt). This range aligns with the most functional RNAs, such as tRNAs, microRNAs, and aptamers, whose compact architectures are biologically and therapeutically relevant.

BeeRNA's stochastic, ABC-based search efficiently explores these conformations without relying on pretraining, achieving low RMSD and high GDT-TS values within this length range across all datasets.

It is important to note that BeeRNA is intentionally designed for this regime. As RNA length increases, the combinatorial search space expands exponentially, making exhaustive tertiary exploration inherently more difficult for any stochastic optimizer. Consequently, slightly higher RMSD values for long RNAs reflect this inherent scaling challenge rather than a methodological limitation. In contrast, deep learning approaches like gRNAde and RiboDiffusion leverage training on long-sequence distributions to improve scalability but often struggle with the flexible and variable conformations of short RNAs.

Overall, BeeRNA achieves its objective: delivering accurate tertiary structure designs for short and medium RNAs through biologically grounded, training-free optimization. Its consistent performance across datasets highlights its adaptability and potential as a practical tool for functional RNA design within the sub-100-nt regime, while maintaining stable accuracy even as RNA length increases.

Conclusion and Limitations

In this work, we introduced BeeRNA, a novel RNA inverse folding method that adapts the ABC algorithm to design RNA sequences folding into predefined tertiary structures, effectively combining ABC exploration with biological constraints and RhoFold structural prediction to minimize structural deviations. This focus on RNAs shorter than 100 nt is biologically meaningful, as most functional RNAs, such as tRNAs, microRNAs, and aptamers, naturally fall within this range. BeeRNA outperforms gRNAde and RiboDiffusion for short and medium-length RNA sequences, with acceptable but moderately reduced performance on longer sequences due to the expanding

search space, reflecting the broader challenge of scaling stochastic optimization to complex RNA folds. As BeeRNA currently uses RhoFold for tertiary structure evaluation, the method can readily integrate future RNA folding predictors as they advance, maintaining flexibility and compatibility with emerging tools. Looking ahead, the approach offers a promising foundation for hybrid strategies that combine bio-inspired search with learning-based priors and multi-objective optimization to enhance scalability, robustness, and convergence across diverse RNA lengths and structural classes.

References

Adamczyk, B.; Antczak, M.; and Szachniuk, M. 2022. RNAsolo: a repository of cleaned PDB-derived RNA 3D structures. *Bioinformatics*, 38(14): 3668–3670.

Bauer, G.; and Suess, B. 2006. Engineered riboswitches as novel tools in molecular biology. *Journal of Biotechnology*, 124(1): 4–11.

Bibekar, P.; Krapp, L. F.; and Dal Peraro, M. 2025. Context-aware geometric deep learning for RNA sequence design. *bioRxiv*.

Cazenave, T.; and Touzani, H. 2024. Monte Carlo Inverse RNA Folding. In *Methods in Molecular Biology*, volume 2847, 205–215. Humana Press.

Das, R.; Karanicolas, J.; and Baker, D. 2010. Atomic accuracy in predicting and designing noncanonical RNA structure. *Nature methods*, 7: 291–4.

Dorigo, M. 1996. *The Ant Colony Optimization Metaheuristic*. Ph.D. thesis, Politecnico di Milano.

Findeiß, S.; Etzel, M.; Will, S.; Mörl, M.; and Stadler, P. F. 2017. Design of Artificial Riboswitches as Biosensors. *Sensors*, 17(9).

Griffiths-Jones, S.; Bateman, A.; Marshall, M.; Khanna, A.; and Eddy, S. R. 2003. Rfam: an RNA family database. *Nucleic Acids Research*, 31(1): 439–441.

Hamada, M. 2018. In silico approaches to RNA aptamer design. *Biochimie*, 145: 8–14.

Hofacker, I. L.; Fontana, W.; Stadler, P. F.; Bonhoeffer, S.; Tacker, M.; and Schuster, P. 1994. Fast folding and comparison of RNA secondary structures. *Monatshefte für Chemie / Chemical Monthly*, 125(2): 167–188.

Holland, J. H. 1975. *Adaptation in Natural and Artificial Systems*. University of Michigan Press.

Huang, H.; Lin, Z.; He, D.; Hong, L.; and Li, Y. 2024. RiboDiffusion: tertiary structure-based RNA inverse folding with generative diffusion models. *Bioinformatics*, 40(Supplement_1): i347–i356.

Joshi, C. K.; Jamasb, A. R.; Viñas, R.; Harris, C.; Mathis, S. V.; Morehead, A.; Anand, R.; and Liò, P. 2025. gRNAde: Geometric deep learning for 3D RNA inverse design. *Proceedings of the International Conference on Learning Representations (ICLR)*. Published as a conference paper at ICLR 2025.

Kai, Z.; Wang, Y.; Lv, Y.; Liu, J.; and He, J. 2019. An Efficient Simulated Annealing Algorithm for the RNA Secondary Structure Prediction with Pseudoknots. *BMC Genomics*, 20(13): 979.

Kalvari, I.; Argasinska, J.; Quinones-Olvera, N.; Nawrocki, E. P.; Rivas, E.; Eddy, S. R.; Bateman, A.; Finn, R. D.; and Petrov, A. I. 2018. Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. *Nucleic Acids Research*, 46(D1): D335–D342.

Karaboga, D. 2005. An Idea Based on Honey Bee Swarm for Numerical Optimization, Technical Report - TR06. *Technical Report, Erciyes University*.

Kennedy, J.; and Eberhart, R. 1995. Particle swarm optimization. In *Proceedings of ICNN'95 - International Conference on Neural Networks*, volume 4, 1942–1948 vol.4.

Li, B.; Chiong, R.; and Lin, M. 2015. A balance-evolution artificial bee colony algorithm for protein structure optimization based on a three-dimensional AB off-lattice model. *Computational Biology and Chemistry*, 54: 1–12.

Lin, C.-J.; and Su, S.-C. 2012. Using an efficient artificial bee colony algorithm for protein structure prediction on lattice models. *International Journal of Innovative Computing, Information and Control*, 8.

Lorenz, R.; Bernhart, S. H.; Höner zu Siederdissen, C.; Tafer, H.; Flamm, C.; Stadler, P. F.; and Hofacker, I. L. 2011. ViennaRNA Package 2.0. *Algorithms for Molecular Biology*, 6(1): 26.

Lyngsø, R. B.; Zuker, M.; and Pedersen, C. N. S. 1999. Fast evaluation of internal loops in RNA secondary structure prediction. *Bioinformatics*, 15(6): 440–445.

Sato, K.; Akiyama, M.; and Sakakibara, Y. 2021. RNA secondary structure prediction using deep learning with thermodynamic integration. *Nature Communications*, 12(1): 941.

Schwab, R.; Ossowski, S.; Riester, M.; Warthmann, N.; and Weigel, D. 2006. Highly specific gene silencing by artificial microRNAs in Arabidopsis. *The Plant Cell*, 18(5): 1121–1133.

Shen, T.; Hu, Z.; Sun, S.; Liu, D.; Wong, F.; Wang, J.; Chen, J.; Wang, Y.; Hong, L.; Xiao, J.; et al. 2024. Accurate RNA 3D structure prediction using a language model-based deep learning approach. *Nature Methods*, 1–12.

Tan, C.; Zhang, Y.; Gao, Z.; Cao, H.; Li, S.; Ma, S.; and Blanchette, M. 2025. R3Design: Deep tertiary structure-based RNA sequence design and beyond. *Briefings in Bioinformatics*, 26(1): bbae682.

Watson, J. D.; and Crick, F. H. 1953. Genetical implications of the structure of deoxyribonucleic acid. *Nature*, 171(4361): 964–967.

Yamagami, R.; Kayedkhordeh, M.; Mathews, D. H.; and Bevilacqua, P. C. 2019. Design of highly active double-pseudoknotted ribozymes: a combined computational and experimental study. *Nucleic Acids Research*, 47(1): 29–42.

Zadeh, J. N.; Steenberg, C. D.; Bois, J. S.; Wolfe, B. R.; Pierce, M. B.; Khan, A. R.; Dirks, R. M.; and Pierce, N. A. 2011. NUPACK: Analysis and design of nucleic acid systems. *Journal of Computational Chemistry*, 32(1): 170–173.

Zemla, A. 2003. LGA: A method for finding 3D similarities in protein structures. *Nucleic Acids Research*, 31(13): 3370–3374.

Zhang, C.; Shine, M.; Pyle, A. M.; and Zhang, Y. 2022. US-align: Universal Structure Alignments of Proteins, Nucleic Acids, and Macromolecular Complexes. *bioRxiv*.

Álvaro Rubio-Largo; Lozano-García, N.; Granado-Criado, J. M.; and Vega-Rodríguez, M. A. 2023. Solving the RNA inverse folding problem through target structure decomposition and Multiobjective Evolutionary Computation. *Applied Soft Computing*, 147: 110779.