MOIRAIN: MULTIMODAL ALIGNMENT AND PREFERENCE OPTIMIZATION FOR ZERO-SHOT CONDITIONAL RNA GENERATION

Anonymous authors
Paper under double-blind review

ABSTRACT

The design of RNA molecules that interact with specific proteins is a critical challenge in experimental and computational biology. Despite recent progress in natural language modeling and deep learning-based protein design, there remains significant room to improve the frequency of successful interactions and the authenticity of generated sequences for functional applications. In this work, we frame conditional RNA sequence generation as a multi-stage alignment problem, introducing Moirain: a suite of models optimized via multimodal supervised fine-tuning (SFT) and Direct Preference Optimization (DPO). Our approach begins with large-scale pretraining on diverse RNA corpora to capture the fundamental grammars of sequence plausibility. To achieve target-specific generation, we employ a multimodal SFT architecture that conditions RNA synthesis on protein structural and sequential features. Finally, we leverage DPO to refine the model using synthetic interaction data: taking advantage of DPO's unique ability to navigate non-aligned preference spaces, we improve functional fitness without collapsing the learned natural distribution. Extensive evaluation demonstrates that our framework consistently produces novel, diverse, and biologically plausible RNA sequences with superior binding affinities compared to existing baselines.

1 INTRODUCTION

Deep learning has transformed a broad spectrum of scientific and technical domains. In structural biology, it has redefined the prediction and design of proteins, as evidenced by AlphaFold2 (Jumper et al., 2021), Chroma (Ingraham et al., 2023), and ESM3 (Hayes et al., 2024). In parallel with biological developments, Transformer-based (Vaswani et al., 2023) autoregressive (Bengio et al., 2000) language models have established a new state of the art for sequence generation. Modern Large Language Models (LLMs) now serve as powerful generative engines that excel across a broad range of specialized tasks. Despite the broad progress across language modeling and structural biology, many areas involving biological sequences have yet to experience a comparable defining shift in performance and utility.

In particular, the de novo generation of ribonucleic acid (RNA) sequences binding to specific protein targets has not yet reached the same level of maturity as that of protein engineering. This task is central to understanding the interactions between RNAs and proteins (Li et al., 2024; Fasogbon et al., 2025), which govern essential biological processes such as gene regulation, splicing, and translation (Hentze et al., 2018). A primary objective in this field is the design of aptamers: short, single-stranded RNA sequences capable of binding specific proteins with high affinity. Because these molecules can function as inhibitors, probes, or delivery agents, they offer versatile applications in therapeutics and diagnostics (Guo et al., 2010; Thavarajah et al., 2021).

Several studies have explored RNA generation in this domain. More classical approaches exploited evolutionary signals and statistical models (Kim et al., 2007a;b; Aita & Husimi, 2010; Tseng et al., 2011; Zhang et al., 2025), molecular modeling (Torkamanian-Afshar et al., 2021), and Monte Carlo tree search (Lee et al., 2021; Wang et al., 2022; Shin et al., 2023; Obonyo et al., 2024). Recent works used long short-term memory models (Im et al., 2019; Park & Han, 2020), conditional variation autoencoders (Chen et al., 2022; Iwano et al., 2022; Andress et al., 2023), adversarial approach

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

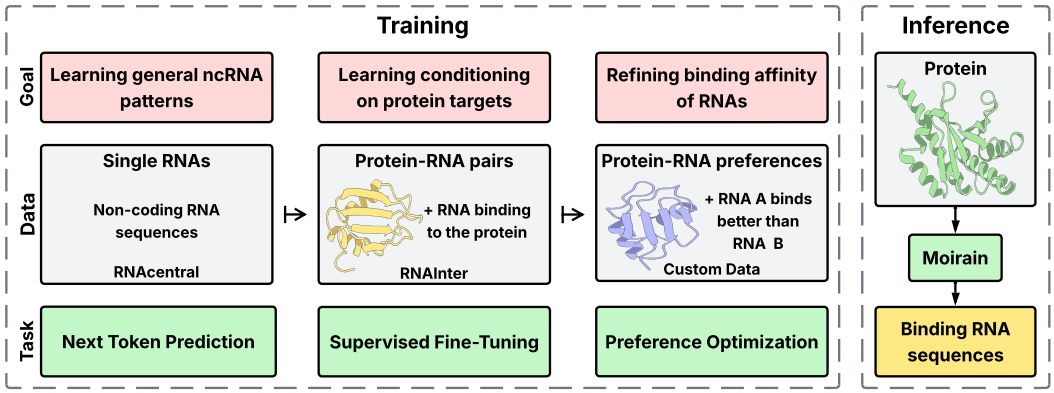


Figure 1: Overview of the Moirain framework. Moirain is developed in three successive stages: pretraining on non-coding RNA to learn general patterns; Supervised Fine-Tuning on protein–RNA pairs to learn conditioning on targets; and Preference Optimization to refine RNA binding affinity. At inference, the model generates RNA sequences conditioned on a protein target in a zero-shot manner.

(Ozden et al., 2023), diffusion processes (Wang et al., 2024; Zhang et al., 2024) and LLM fine-tuning (Zhao et al., 2024). Historically, RNA design was approached by training and running models independently for each target. A more compelling challenge lies in the immediate synthesis of binders for novel proteins, a zero-shot conditional task that has become the primary focus of recent research (Nori & Jin, 2024; Klypa et al., 2025; Tabrizi et al., 2025b;a; Favor et al., 2025). Despite the developments, there remains clear space to improve the frequency of successful interactions and the authenticity of the generated sequences required for functional biological applications.

To tackle the challenge, we adopt a methodological framework consistent with the development of modern LLMs, comprising large-scale pretraining to capture general patterns (Radford et al., 2019; Brown et al., 2020), followed by instruction tuning (Wei et al., 2022), also known as Supervised Fine-Tuning (SFT), and Preference Optimization (PO) alignment to refine the focus of the model on specific objectives (Bai et al., 2022; Ouyang et al., 2022). We begin with a pretraining stage on an extensive corpus of non-coding RNA sequences, establishing a broad biological context for the model. To enable protein-conditioned RNA generation, we perform multimodal supervised fine-tuning on a dataset of protein targets and their cognate RNA partners, aligning the model's output distribution with the subspace of interactive sequences. In a subsequent preference optimization phase, we train our model on preference pairs of RNA sequences ranked by experimentally measured binding affinities, tuning it toward specific performance objectives rather than mere sequence likelihood. Ultimately, our method (illustrated in Fig. 1) is designed to achieve protein-conditioned RNA sequence design with high binding affinity while preserving the essential biological authenticity acquired during earlier training stages. Our main contributions can be summarized as follows:

1. We frame target-conditioned RNA sequence design as a multimodal alignment problem, integrating LLM architectures with Supervised Fine-Tuning and Preference Optimization using a newly specifically curated preference dataset.

2. We present Moirain, a foundational generative model adapted to the multimodal nature of the task, capable of zero-shot RNA sequence generation conditioned on target protein binding.

3. We demonstrate through empirical evaluations that Moirain reaches state-of-the-art performance across key metrics, generating novel, diverse, highly interactive, and biologically plausible RNA sequences.

4. We conduct structured ablation analyses to isolate the key components driving performance, uncovering practical insights into Moirain training pipeline.

2 THEORETICAL PRELIMINARIES AND BACKGROUNDS

2.1 LARGE LANGUAGE MODELS

Large Language Models are trained as next-token predictors over a discrete vocabulary $\mathcal{V}$. Given a sequence of tokens $x = (x_1, ..., x_L)$, the model defines a conditional distribution $\pi_\theta(\cdot|x_{<l})$ over the next token x_l. Training consists of minimizing the Cross-Entropy (CE) between the predicted distribution and the data $\mathcal{D}$:

$$\mathcal{L}_{\mathrm{CE}}(\theta) = -\mathbb{E}_{x\sim\mathcal{D}}\left[\log \pi_\theta(x_l|x_{<l})\right]. \tag{1}$$

By applying the chain rule, the joint probability of an entire sequence $x_{1:L}$ under the model parameters θ is decomposed into the product of the conditional distributions:

$$\pi_\theta(x) = \prod_{l=1}^{L} \pi_\theta(x_l|x_{<l}), \tag{2}$$

thus making the training equivalent to maximizing the likelihood of the observed sequences.

2.2 SUPERVISED FINE-TUNING

To transition from general next-token prediction to reliable instruction following, the base model is trained via Supervised Fine-Tuning on a dataset $\mathcal{D}_{\mathrm{SFT}}$ of prompt-response pairs (x, y). The model conditions on the prompt as a fixed prefix but is optimized exclusively on the response:

$$\mathcal{L}_{\mathrm{CE}}(\theta) = \mathbb{E}_{(x,y)\sim\mathcal{D}_{\mathrm{SFT}}}\left[-\log \pi_\theta(y_l|y_{<l}, x)\right]. \tag{3}$$

This fine-tuning stage is often performed via the Low-Rank Adaptation (LoRA) (Hu et al., 2021) technique, which reduces the effective number of updated parameters, to improve efficiency and reduce overfitting as well as forgetting (Biderman et al., 2024).

The SFT has been observed to reduce generative diversity (O'Mahony et al., 2024; Klypa & Cherednichenko, 2026), thus undermining downstream exploration and subsequent alignment. To address this issue, recent work has explored modified loss functions, such as the Tempered Focal (TOFU) loss (Klypa & Cherednichenko, 2026). TOFU approach is based on weighting the Cross-Entropy gradients and applying temperature adjustment to the predicted distribution:

$$\mathcal{L}_{\mathrm{TOFU}}(\theta) = \mathbb{E}_{(x,y)\sim\mathcal{D}_{\mathrm{SFT}}}\left[-\mathrm{sg}\left[g(\pi_\theta(y_l|y_{<l}, x), \gamma)\right]\beta \log \pi_\theta^{\beta}(y_l|y_{<l}, x)\right], \tag{4}$$

where focal term $g(p, \gamma) = (1-p)^\gamma - \gamma p(1-p)^{\gamma-1}\log p$ is detached from the gradient computation.

2.3 MULTIMODALITY

Multimodal large language models integrate visual and textual representations through varying architectural strategies. CLIP (Radford et al., 2021) established this field by using contrastive pretraining to create a shared embedding space. To incorporate frozen vision encoders, Flamingo (Alayrac et al., 2022) utilized cross-attention layers, whereas BLIP-2 (Li et al., 2023) introduced a lightweight query transformer to bridge the modality gap. Further simplifying this paradigm, LLaVA (Liu et al., 2023) projects visual features directly into the LLM input space, achieving alignment through instruction tuning on image–text pairs.

2.4 PREFERENCE OPTIMIZATION

Preference Optimization alignment stage relies on a dataset $\mathcal{D}_{\mathrm{PO}}$ of human (or LLM)-annotated preference pairs (x, y^+, y^-), where x is the prompt, and y^+ and y^- are the preferred and dispreferred responses. The standard Reinforcement Learning from Human Feedback (RLHF) (Ziegler et al., 2020) paradigm addresses this setting in two steps. First, a scalar reward model $r_\phi(x, y)$ is fitted to $\mathcal{D}_{\mathrm{PO}}$. Second, the target policy π_θ is fine-tuned to maximize the expected reward while remaining close to the reference policy π_{ref} (the SFT model) via a KL-divergence constraint:

$$\max_\theta \mathbb{E}_{x\sim\mathcal{D}, y\sim\pi_\theta(x)}[r_\phi(x, y)] - \beta D_{\mathrm{KL}}\left(\pi_\theta(y \mid x) \,||\, \pi_{\text{ref}}(y \mid x)\right), \tag{5}$$

where $\beta > 0$ controls the strength of the KL penalty to prevent reward hacking and language degeneration. To solve this optimization problem, one approach is to employ Proximal Policy Optimization (PPO) (Schulman et al., 2017), an online, model-free policy gradient algorithm.

Although RLHF pipeline has achieved remarkable success in aligning with preferences, its complex multi-step nature makes it expensive in terms of computation time and memory usage. To alleviate these complexities, recent approaches derive closed-form expressions for the optimal policy, enabling direct optimization on preference data. Direct Preference Optimization (DPO) (Rafailov et al., 2023) served as a seminal work in this area. It leverages the analytical relationship between the reward and the policy, resulting in the objective:

$$\mathcal{L}_{\mathrm{DPO}}(\theta) := \mathbb{E}_{(x,y^+,y^-)\sim\mathcal{D}_{\mathrm{PO}}}\left[-\log\sigma\left(\beta\left(\log\frac{\pi_\theta(y^+ \mid x)}{\pi_{\mathrm{ref}}(y^+ \mid x)} - \log\frac{\pi_\theta(y^- \mid x)}{\pi_{\mathrm{ref}}(y^- \mid x)}\right)\right)\right], \quad (6)$$

where $\beta > 0$. DPO has recently demonstrated success in diverse biological applications, ranging from protein design to optimization of genomic sequences (Nguyen et al., 2024; Heinzinger & Rost, 2025; Listgarten & Jiang, 2026).

3 Proposed Method

3.1 Base Model Pretraining

The primary goal of our pretraining is to capture general RNA patterns, ensuring that generated sequences remain physiologically viable. We focus exclusively on ncRNAs, as they constitute the vast majority of RNA–protein interactions. Additionally, by avoiding translation, ncRNAs resist degradation better than coding RNAs and reach their targets more reliably.

Guided by these considerations, we utilized RNAcentral (The RNAcentral Consortium et al., 2019), the largest available database of non-coding RNA sequences, for the pretraining stage. To ensure data quality and reduce redundancy, we deduplicated the raw sequences via clustering, resulting in a final training set of 6.6 million unique RNAs. Naturally, RNA is composed of a four-letter alphabet. To enable a larger context window, we tokenized the sequences with Byte Pair Encoding (BPE) (Sennrich et al., 2016). This approach effectively compresses the data by reducing RNA length. We fix the vocabulary size to 256 to limit the long low-frequency tail and ensure sufficient training signal per embedding. The resulting train corpus size is 1.7 billion tokens.

Our base model, Moirain-Base, comprises 302 million parameters and follows the LLaMA (Touvron et al., 2023) architecture family, which we optimized for stable training. We trained it for one epoch (Muennighoff et al., 2023); additional data and training details are provided in Appendix A.1.

3.2 Multimodal Supervised Fine-Tuning

To adapt Moirain-Base for conditional RNA generation, we perform multimodal supervised fine-tuning using paired interaction data from RNAinter (Kang et al., 2022), comprising 203,811 protein-RNA pairs after processing. This approach mirrors instruction tuning, where the protein acts as a functional prompt guiding RNA generation. We encode protein features (sequences, three-dimensional structures from the AlphaFold Protein Structure Database (AFDB) (Fleming et al., 2025), and per-residue pLDDT scores) with the protein module from RNA-BAnG (Klypa et al., 2025), which integrates transformer layers with geometric Invariant Point Attention (IPA) (Jumper et al., 2021). Incorporating protein structures grounds RNA design in the target's physical reality while avoiding the massive computational burden of learning sequence-to-fold mappings from scratch.

As proteins represent a fundamentally distinct macromolecular class from RNA sequences, we considered projection-based conditioning unsuitable for our setup. We instead adopted a cross-attention framework inspired by Flamingo and BLIP-2, adapting it to be more parameter-efficient since the SFT dataset introduces no novel sequential patterns beyond those captured during pretraining. Specifically, we replaced the extra feed-forward layers with Low-Rank Adaptation (LoRA) on the pretrained weights. Furthermore, we reduced the number of cross-attention blocks to $M < N$, where N denotes the total number of transformer layers. These M modules are distributed uniformly throughout the stack, with the first and last ones positioned immediately before the 1st and

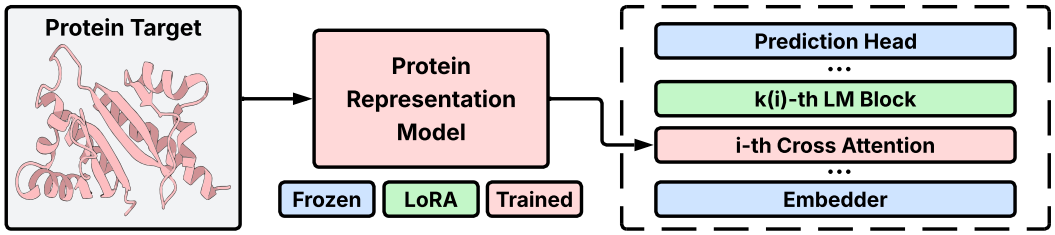


Figure 2: Schematic of the Moirain-SFT Cross-Attention Architecture training pipeline. The framework integrates protein features via M cross-attention blocks ($M < N$) uniformly distributed across the N-layer stack according to Equation 7. Parameter efficiency is maintained by utilizing LoRA of foundational RNA weights instead of adding fully trainable feed-forward layers.

N-th original layers, respectively. Under these constraints, the specific index $k(i)$ for each cross-attention block is defined by:

$$k(i) = \text{round}\left(1 + \frac{(i-1)(N-1)}{M-1}\right), \quad i = 1, \ldots, M. \tag{7}$$

This distribution of modules ensures a consistent injection of conditional features across all levels of the model's hierarchy, preventing signal degradation. The resulting architecture is schematically illustrated in Figure 2.

Integrating the multimodal architectural design described above into Moirain-Base results in the Moirain-SFT model, containing 331 million parameters, with 29 million being trainable and the remainder frozen. We fine-tuned it for three epochs (Ouyang et al., 2022; Zhou et al., 2023), using TOFU loss function to enhance output diversity and mitigate overconfidence. Additional data and training details are provided in Appendix A.2.

3.3 Preference Optimization

In many cases, RNA binding affinity to the target protein is determined by only a fraction of its total sequence, often confined to relatively small motifs (Ray et al., 2013). Consequently, training without additional conditioning signals while aiming for high-performance generation is equivalent to solving a needle-in-a-haystack problem. Unfortunately, deep neural networks in general (Geirhos et al., 2020) and autoregressive modeling in particular (Dziri et al., 2023) have been observed to struggle in this regime as models tend to prioritize global patterns over short localized subsequences. This limitation has been previously noted specifically in the context of RNA-protein interactions (Klypa et al., 2025).

We address this challenge by employing Preference Optimization. Unlike standard training, PO algorithms require each prompt to be paired with multiple candidate responses for evaluation or ranking. For our task, this format is provided by data from Ray et al. (2013); Sasse et al. (2025), which in total contains approximately 240,000 RNA sequences, experimentally scored against 420 protein targets. A critical feature of this dataset is that the sequences are synthetic and largely randomized. Consequently, using them as positive examples for supervised fine-tuning would be counterproductive, as we wish the model to maintain the natural RNA distribution learned during its previous training phases. Similarly, training a reward model on such data is risky, as it could bias the model to favor synthetic artifacts over genuine biological priors.

Given the nature of the data and the task requirements, Preference Optimization via a pairwise ranking loss, such as Direct Preference Optimization, emerges as the currently most suitable strategy. The DPO objective is specifically designed to shift the probability distribution by increasing the likelihood gap between preferred and dispreferred examples rather than simply maximizing the probability of the former. Crucially, the probability of a preferred response has been observed to actually decrease during optimization (Pal et al., 2024). Although potentially prone to unintentional unalignment (Razin et al., 2025), this effect serves our goal of prioritizing functional binding motifs while bypassing synthetic noise.

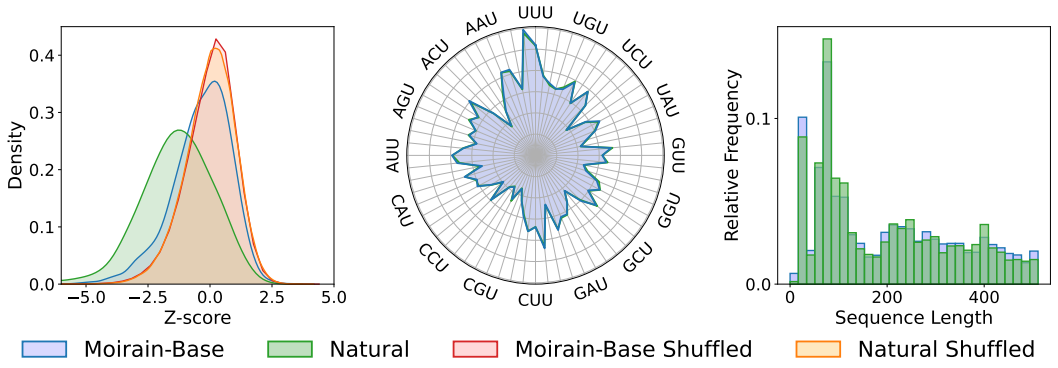


Figure 3: The comparison between generated (Moirain-Base) and natural RNA sequences. **(Left)** Z-score distributions for generated and natural sequences alongside their shuffled counterparts. **(Middle)** Radar plot of 3-mer frequency distributions (Fidelity $= 0.008$), where each angular axis represents a distinct 3-mer. **(Right)** Relative frequency distributions of RNA sequence lengths for sequences shorter than 512 nucleotides.

To implement this approach, we have constructed a preference dataset of 1000 pairs per 383 proteins and trained Moirain-SFT on it with LoRA (7.8 million active parameters) for 2 epochs (Ouyang et al., 2022; Zhou et al., 2023), resulting in our model's final version, Moirain. Additional data and training details are provided in Appendix A.3.

4 Evaluation Setup

To evaluate target binding affinity, we assessed generated RNA sequences using the RNA-BAnG benchmark (Klypa et al., 2025). This suite comprises 71 target proteins from the RNACompendium (Ray et al., 2013), with binding predictions evaluated by corresponding DeepCLIP models (Grønning et al., 2020). The test protein targets are originally divided into two distinct sets based on similarity to the RNA-BAnG training data: In-Domain (ID, 59 targets), which share varying degrees of similarity, and Out-of-Domain (OOD, 12 targets), which share none. We ensure that proteins in any Moirain training phase have minimal to no similarity with those in the OOD set.

Widely used for the task (Im et al., 2019; Zhao et al., 2024; Klypa et al., 2025; Tabrizi et al., 2025b), DeepCLIP is an interpretable, lightweight CNN architecture that scans for probabilistic sequential motifs and outputs a target binding probability score. To align with molecular design objectives, where the yield of high-performing candidates takes precedence over population averages, we introduce two threshold-based metrics: $Hits_x$ and Cov_x. Hits_x denotes the fraction of generated sequences exceeding a binding score of x. Analogous to the pass@k metric in LLM evaluation, it quantifies the model's efficiency in producing high-affinity binders. Complementing this, Cov_x measures the breadth of success across the target space, defined as the number of proteins for which at least x percent of generated sequences surpass a binding probability of 0.7 (Klypa et al., 2025).

As a baseline measure of biological plausibility, we compare the local patterns of generated sequences with those of natural RNAs. For this purpose, we employ the Total Variation distance between their respective 3-mer distributions, a metric we designate as *Fidelity*. A more computationally intensive evaluation relies on the observation that functional non-coding RNAs typically exhibit lower Minimum Free Energy (MFE) than random sequences of the same composition (Freyhult et al., 2005). To quantify this structural bias, we calculate an MFE Z-score (Freyhult et al., 2005), which measures the deviation of an RNA's MFE from the mean of its shuffled variants. Comparing these distributions between generated and natural sequences indicates whether the model generates thermodynamically realistic folds.

To demonstrate the ability of Moirain to produce novel sequences beyond its training data while avoiding mode collapse, we measure per target *Novelty* and *Diversity*. Novelty is defined as the proportion of generated sequences with no detectable hits against a reference RNA set of choice, serving as a primary indicator for memorization detection. To isolate the impact of different training

Table 1: Binding affinity and authenticity benchmarking results for tested models. Performance is quantified by: (i) Hits_x (0–1), representing target-specific binding success; (ii) Cov_x (0–1), denoting the coverage of the target manifold; and (iii) Fidelity (0–1), measuring the proximity to natural interacting RNA. † denotes models whose training data shares no or minimal similarity to OOD targets.

BENCH	METHOD	$\text{HITS}_{0.5}\uparrow$	$\text{HITS}_{0.7}\uparrow$	$\text{HITS}_{0.9}\uparrow$	$\text{COV}_{0.25}\uparrow$	$\text{COV}_{0.5}\uparrow$	FIDELITY $\downarrow$
ID	RNATRANSLATOR	0.46	0.37	0.26	0.59	0.32	**0.18**
	RNA-BANG	0.58	0.52	0.42	0.78	0.49	0.26
	MOIRAIN	**0.95**	**0.92**	**0.85**	**1.0**	**0.97**	0.25
OOD	RNAFLOW	0.13	0.10	0.06	0.17	0.08	0.75
	RNATRANSLATOR	0.39	0.31	0.21	0.42	0.25	**0.17**
	RNA-BANG†	0.53	0.46	0.34	**0.75**	0.42	0.22
	RFDPOLY	**0.67**	**0.61**	**0.54**	**0.75**	**0.67**	0.52
	MOIRAIN†	**0.67**	**0.61**	0.51	**0.75**	**0.67**	0.29

stages, we calculate this metric independently against each corresponding dataset used. Diversity is measured as the proportion of cluster representatives at fixed similarity threshold, quantifying the breadth of the generated sequence space. The details of tools and parameters used can be found in Appendix C.

5 RESULTS

We first evaluate whether Moirain-Base successfully captures the sequence and structural characteristics of natural non-coding RNAs. As shown in Figure 3, the model achieves near-perfect fidelity to natural 3-mer profiles, and the distribution of generated sequence lengths closely mirrors that of the natural dataset. Although generated sequences are less structured on average than natural RNAs, they nonetheless retain some folding signal.

Crucially, we benchmark Moirain's capacity to design high-affinity target-specific RNA sequences against state-of-the-art open-weight models capable of zero-shot generation (Nori & Jin, 2024; Klypa et al., 2025; Tabrizi et al., 2025b; Favor et al., 2025). Due to high inference latency, RNAFlow and RFDPoly were evaluated exclusively on the out-of-distribution set, using a tenfold lower sampling quota per target to remain within practical computational limits. The results of the comparison are summarized in Table 1.

In the ID interpolation setting, our model achieves near-maximal scores. On OOD targets, Moirain demonstrates strong generalization, outperforming RNA-BAnG under a fair data split. It also surpasses or matches existing methods such as RFDPoly, despite their use of undisclosed training data or the inclusion of OOD targets in their training sets. Notably, structural models such as RNAFlow and RFDPoly struggle with low Fidelity, whereas Moirain achieves competitive results alongside top-performing baselines.

Moirain demonstrates high sequence variance, achieving a Diversity score of 0.96 ± 0.09. Crucially, the model avoids mode collapse and training data memorization across all alignment phases. It maintains near-total Novelty relative to both the Base and SFT training sets (0.99 ± 0.01 and 0.99 ± 0.00, respectively), while achieving a maximum Novelty score of 1.0 against the Preference Optimization training set. The absence of matches within synthetic data validates our hypothesis regarding the suitability of DPO for this task.

To evaluate the impact of our architectural and training choices, we conducted ablation studies across each component of the pipeline, summarized in Table 2. The results validate the hypothesis that a purely autoregressive approach struggles with the needle-in-a-haystack nature of binding sites, confirming that an alignment stage is essential for reliably generating high-affinity binders. Regarding network architecture, layer positioning proves significant: uniform distribution of Cross-Attention

Table 2: Ablation studies on Moirain across out-of-distribution benchmark. Performance is quantified by: (i) Hits_x (0–1), representing target-specific binding success; (ii) Cov_x (0–1), denoting the coverage of the target manifold; and (iii) Fidelity (0–1), measuring the proximity to natural interacting RNA.

STAGE	CHANGE	$\text{HITS}_{0.5}\uparrow$	$\text{HITS}_{0.7}\uparrow$	$\text{HITS}_{0.9}\uparrow$	$\text{COV}_{0.25}\uparrow$	$\text{COV}_{0.5}\uparrow$	FIDELITY $\downarrow$
PO	NO PO	0.39	0.29	0.18	0.42	0.17	0.05
SFT	END CROSS-ATT	0.58	0.50	0.37	0.67	0.58	0.25
SFT	CE LOSS	0.62	0.55	0.43	**0.83**	0.58	0.22
PO	IPO LOSS	0.39	0.30	0.18	0.42	0.17	0.06
	MOIRAIN	**0.67**	**0.61**	**0.51**	0.75	**0.67**	0.29

blocks across the model outperforms clustering them at the end. Objective selection during the SFT stage is similarly critical, where integrating the TOFU loss is one of the drivers of model performance. Finally, comparing alternative PO objectives, Identity Preference Optimization (IPO) (Azar et al., 2023) proved less effective than DPO in our setup, appearing not to step away sufficiently from the SFT reference model.

6 CONCLUSION

In this work, we address the challenge of designing RNA molecules conditioned on specific target proteins, optimizing for both binding affinity and biological plausibility. We frame this task as a multimodal alignment problem and propose a multi-phase framework that integrates Large Language Model architectures with Supervised Fine-Tuning and Preference Optimization. Each phase serves a distinct purpose: pretraining captures fundamental RNA sequence patterns, SFT enables zero-shot target-conditioned generation, and PO refines the focus on binding sites. To enable the final alignment step, we repurposed existing protein-RNA interaction data into a novel preference dataset.

We introduce Moirain, a generative model capable of zero-shot RNA sequence design conditioned on binding to a target protein. Empirical evaluations demonstrate that Moirain achieves state-of-the-art performance, generating novel, diverse, and biologically plausible RNA sequences with high binding affinity. Crucially, the model generalizes well to target proteins unseen during training.

We conduct structured ablation analyses to isolate the key drivers of Moirain's performance, providing practical insights for both general generative modeling pipelines and biological sequence design. Our findings demonstrate that Preference Optimization stage is essential to overcome limitations of purely autoregressive approach on needle-in-a-haystack type of problem. Additionally, uniform Cross-Attention distribution alongside TOFU and DPO objectives serve as key factors in improving the binding affinity of generated sequences.

In summary, our work demonstrates that integrating previously untapped data sources into a multi-stage alignment pipeline provides a robust framework for protein-conditioned RNA design. By carefully pairing specific training data with phase-appropriate optimization objectives, Moirain establishes state-of-the-art performance across key computational metrics. While our model sets a new benchmark for zero-shot generation, limitations remain. Generated sequences are less structured than natural RNAs, and sequence plausibility can be further improved. The natural next step is experimental wet-lab validation to confirm the practical utility of the generated sequences. Ultimately, we hope that the methodologies introduced here provide a useful paradigm for biological sequence design.

AI USE STATEMENT

In this work, we used generative AI tools for assisting in writing and text polishing. The core method development in this research does not involve generative AI tools as any important, original, or non-

standard components. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

ETHICS STATEMENT

The main purpose of this work is to advance the field of generative models. However, the applications of this method may have social and industrial benefits. Potential applications include in-silico SELEX approaches, RNA vaccine design, the development of novel drugs, and some other therapeutic tasks.

REPRODUCIBILITY STATEMENT

The code and the models, along with the models weights, are available at `https://anonymous.4open.science/r/Moirain-D544`. The code allows for running the RNA sequences generation process for any correct protein structure as input.

REFERENCES

Takuyo Aita and Yuzuru Husimi. Biomolecular information gained through in vitro evolution. *Biophysical reviews*, 2:1–11, 2010.

Jean-Baptiste Alayrac, Jeff Donahue, Pauline Luc, Antoine Miech, Iain Barr, Yana Hasson, Karel Lenc, Arthur Mensch, Katie Millican, Malcolm Reynolds, Roman Ring, Eliza Rutherford, Serkan Cabi, Tengda Han, Zhitao Gong, Sina Samangooei, Marianne Monteiro, Jacob Menick, Sebastian Borgeaud, Andrew Brock, Aida Nematzadeh, Sahand Sharifzadeh, Mikolaj Binkowski, Ricardo Barreira, Oriol Vinyals, Andrew Zisserman, and Karen Simonyan. Flamingo: a Visual Language Model for Few-Shot Learning, November 2022. URL `http://arxiv.org/abs/2204.14198`. arXiv:2204.14198 [cs].

Stephen F. Altschul, Warren Gish, Webb Miller, Eugene W. Myers, and David J. Lipman. Basic local alignment search tool. *Journal of Molecular Biology*, 215(3):403–410, October 1990. ISSN 00222836. doi: 10.1016/S0022-2836(05)80360-2. URL `https://linkinghub.elsevier.com/retrieve/pii/S0022283605803602`.

Cameron Andress, Kalli Kappel, Marcus Elbert Villena, Miroslava Cuperlovic-Culf, Hongbin Yan, and Yifeng Li. DAPTEV: Deep aptamer evolutionary modelling for COVID-19 drug design. *PLOS Computational Biology*, 19(7):e1010774, July 2023. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1010774. URL `https://dx.plos.org/10.1371/journal.pcbi.1010774`.

Mohammad Gheshlaghi Azar, Mark Rowland, Bilal Piot, Daniel Guo, Daniele Calandriello, Michal Valko, and Rémi Munos. A General Theoretical Paradigm to Understand Learning from Human Preferences, 2023. URL `https://arxiv.org/abs/2310.12036`.

Yuntao Bai, Andy Jones, Kamal Ndousse, Amanda Askell, Anna Chen, Nova DasSarma, Dawn Drain, Stanislav Fort, Deep Ganguli, Tom Henighan, Nicholas Joseph, Saurav Kadavath, Jackson Kernion, Tom Conerly, Sheer El-Showk, Nelson Elhage, Zac Hatfield-Dodds, Danny Hernandez, Tristan Hume, Scott Johnston, Shauna Kravec, Liane Lovitt, Neel Nanda, Catherine Olsson, Dario Amodei, Tom Brown, Jack Clark, Sam McCandlish, Chris Olah, Ben Mann, and Jared Kaplan. Training a Helpful and Harmless Assistant with Reinforcement Learning from Human Feedback, April 2022. URL `http://arxiv.org/abs/2204.05862`. arXiv:2204.05862 [cs].

Yoshua Bengio, Réjean Ducharme, and Pascal Vincent. A Neural Probabilistic Language Model. In *Advances in Neural Information Processing Systems*, volume 13. MIT Press, 2000. URL `https://papers.nips.cc/paper_files/paper/2000/hash/728f206c2a01bf572b5940d7d9a8fa4c-Abstract.html`.

Dan Biderman, Jacob Portes, Jose Javier Gonzalez Ortiz, Mansheej Paul, Philip Greengard, Connor Jennings, Daniel King, Sam Havens, Vitaliy Chiley, Jonathan Frankle, Cody Blakeney, and John P. Cunningham. LoRA Learns Less and Forgets Less, May 2024. URL `https://arxiv.org/abs/2405.09673v2`.

Garth R. Brown, Vichet Hem, Kenneth S. Katz, Michael Ovetsky, Craig Wallin, Olga Ermolaeva, Igor Tolstoy, Tatiana Tatusova, Kim D. Pruitt, Donna R. Maglott, and Terence D. Murphy. Gene: a gene-centered information resource at NCBI. *Nucleic Acids Research*, 43(D1):D36–D42, January 2015. ISSN 0305-1048. doi: 10.1093/nar/gku1055. URL https://doi.org/10.1093/nar/gku1055.

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. Language Models are Few-Shot Learners, July 2020. URL http://arxiv.org/abs/2005.14165. arXiv:2005.14165 [cs].

Jonathan C. Chen, Jonathan P. Chen, Max W. Shen, Michael Wornow, Minwoo Bae, Wei-Hsi Yeh, Alvin Hsu, and David R. Liu. Generating experimentally unrelated target molecule-binding highly functionalized nucleic-acid polymers using machine learning. *Nature Communications*, 13(1): 4541, August 2022. ISSN 2041-1723. doi: 10.1038/s41467-022-31955-4. URL https://www.nature.com/articles/s41467-022-31955-4.

Mostafa Dehghani, Josip Djolonga, Basil Mustafa, Piotr Padlewski, Jonathan Heek, Justin Gilmer, Andreas Steiner, Mathilde Caron, Robert Geirhos, Ibrahim Alabdulmohsin, Rodolphe Jenatton, Lucas Beyer, Michael Tschannen, Anurag Arnab, Xiao Wang, Carlos Riquelme, Matthias Minderer, Joan Puigcerver, Utku Evci, Manoj Kumar, Sjoerd van Steenkiste, Gamaleldin F. Elsayed, Aravindh Mahendran, Fisher Yu, Avital Oliver, Fantine Huot, Jasmijn Bastings, Mark Patrick Collier, Alexey Gritsenko, Vighnesh Birodkar, Cristina Vasconcelos, Yi Tay, Thomas Mensink, Alexander Kolesnikov, Filip Pavetić, Dustin Tran, Thomas Kipf, Mario Lučić, Xiaohua Zhai, Daniel Keysers, Jeremiah Harmsen, and Neil Houlsby. Scaling Vision Transformers to 22 Billion Parameters, February 2023. URL http://arxiv.org/abs/2302.05442. arXiv:2302.05442 [cs].

Nouha Dziri, Ximing Lu, Melanie Sclar, Xiang (Lorraine) Li, Liwei Jiang, Bill Yuchen Lin, Sean Welleck, Peter West, Chandra Bhagavatula, Ronan Le Bras, Jena Hwang, Soumya Sanyal, Xiang Ren, Allyson Ettinger, Zaid Harchaoui, and Yejin Choi. Faith and Fate: Limits of Transformers on Compositionality. *Advances in Neural Information Processing Systems*, 36:70293–70332, December 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/deb3c28192f979302c157cb653c15e90-Abstract-Conference.html.

Ilemobayo Victor Fasogbon, Erick Nyakundi Ondari, Deusdedit Tusubira, Loganathan Rangasamy, Janarthanan Venkatesan, Angela Mumbua Musyoka, and Patrick Maduabuchi Aja. Recent focus in non-SELEX-computational approach for de novo aptamer design: A mini review. *Analytical Biochemistry*, 699:115756, April 2025. ISSN 1096-0309. doi: 10.1016/j.ab.2024.115756.

Andrew Favor, Riley Quijano, Elizaveta Chernova, Andrew Kubaney, Connor Weidle, Morgan A. Esler, Lilian McHugh, Ann Carr, Yang Hsia, David Juergens, Kenneth D. Carr, Paul T. Kim, Yuliya Politanska, Enisha Sehgal, Paul S. Kwon, Robert J. Pecoraro, Cameron Glasscock, Andrew J. Borst, Frank DiMaio, Barry L. Stoddard, and David Baker. De novo design of RNA and nucleoprotein complexes, October 2025. URL https://www.biorxiv.org/content/10.1101/2025.10.01.679929v1. ISSN: 2692-8205 Pages: 2025.10.01.679929 Section: New Results.

Jennifer Fleming, Paulyna Magana, Sreenath Nair, Maxim Tsenkov, Damian Bertoni, Ivanna Pidruchna, Marcelo Querino Lima Afonso, Adam Midlik, Urmila Paramval, Augustin Žídek, Agata Laydon, Oleg Kovalevskiy, Joshua Pan, Jun Cheng, Žiga Avsec, Clare Bycroft, Lai Hong Wong, Meera Last, Milot Mirdita, Martin Steinegger, Pushmeet Kohli, Mihály Váradi, and Sameer Velankar. AlphaFold Protein Structure Database and 3D-Beacons: New Data and Capabilities. *Journal of Molecular Biology*, 437(15):168967, August 2025. ISSN 0022-2836. doi: 10.1016/j.jmb.2025.168967. URL https://www.sciencedirect.com/science/article/pii/S0022283625000336.

Eva Freyhult, Paul P. Gardner, and Vincent Moulton. A comparison of RNA folding measures. *BMC Bioinformatics*, 6(1):241, October 2005. ISSN 1471-2105. doi: 10.1186/1471-2105-6-241. URL https://doi.org/10.1186/1471-2105-6-241.

Robert Geirhos, Jörn-Henrik Jacobsen, Claudio Michaelis, Richard Zemel, Wieland Brendel, Matthias Bethge, and Felix A. Wichmann. Shortcut Learning in Deep Neural Networks. *Nature Machine Intelligence*, 2(11):665–673, November 2020. ISSN 2522-5839. doi: 10.1038/s42256-020-00257-z. URL http://arxiv.org/abs/2004.07780. arXiv:2004.07780 [cs].

Alexander Gulliver Bjørnholt Grønning, Thomas Koed Doktor, Simon Jonas Larsen, Ulrika Simone Spangsberg Petersen, Lise Lolle Holm, Gitte Hoffmann Bruun, Michael Birkerod Hansen, Anne-Mette Hartung, Jan Baumbach, and Brage Storstein Andresen. DeepCLIP: predicting the effect of mutations on protein–RNA binding with deep learning. *Nucleic Acids Research*, pp. gkaa530, June 2020. ISSN 0305-1048, 1362-4962. doi: 10.1093/nar/gkaa530. URL https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkaa530/5859960.

Peixuan Guo, Oana Coban, Nicholas M. Snead, Joe Trebley, Steve Hoeprich, Songchuan Guo, and Yi Shu. Engineering RNA for Targeted siRNA Delivery and Medical Application. *Advanced Drug Delivery Reviews*, 62(6):650–666, April 2010. ISSN 0169409X. doi: 10.1016/j.addr.2010.03.008. URL https://linkinghub.elsevier.com/retrieve/pii/S0169409X10000773.

Thomas Hayes, Roshan Rao, Halil Akin, Nicholas J. Sofroniew, Deniz Oktay, Zeming Lin, Robert Verkuil, Vincent Q. Tran, Jonathan Deaton, Marius Wiggert, Rohil Badkundri, Irhum Shafkat, Jun Gong, Alexander Derry, Raul S. Molina, Neil Thomas, Yousuf Khan, Chetan Mishra, Carolyn Kim, Liam J. Bartie, Matthew Nemeth, Patrick D. Hsu, Tom Sercu, Salvatore Candido, and Alexander Rives. Simulating 500 million years of evolution with a language model, July 2024. URL http://biorxiv.org/lookup/doi/10.1101/2024.07.01.600583.

Michael Heinzinger and Burkhard Rost. Teaching AI to speak protein. *Current Opinion in Structural Biology*, 91:102986, April 2025. ISSN 0959-440X. doi: 10.1016/j.sbi.2025.102986. URL https://www.sciencedirect.com/science/article/pii/S0959440X25000041.

Dan Hendrycks and Kevin Gimpel. Gaussian Error Linear Units (GELUs), 2016. URL https://arxiv.org/abs/1606.08415. Version Number: 5.

Matthias W. Hentze, Alfredo Castello, Thomas Schwarzl, and Thomas Preiss. A brave new world of RNA-binding proteins. *Nature Reviews Molecular Cell Biology*, 19(5):327–341, May 2018. ISSN 1471-0080. doi: 10.1038/nrm.2017.130. URL https://www.nature.com/articles/nrm.2017.130.

Ari Holtzman, Jan Buys, Li Du, Maxwell Forbes, and Yejin Choi. The Curious Case of Neural Text Degeneration, February 2020. URL http://arxiv.org/abs/1904.09751. arXiv:1904.09751 [cs].

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. LoRA: Low-Rank Adaptation of Large Language Models, October 2021. URL http://arxiv.org/abs/2106.09685. arXiv:2106.09685 [cs].

Jinho Im, Byungkyu Park, and Kyungsook Han. A generative model for constructing nucleic acid sequences binding to a protein. *BMC genomics*, 20(Suppl 13):967, 2019.

John B. Ingraham, Max Baranov, Zak Costello, Karl W. Barber, Wujie Wang, Ahmed Ismail, Vincent Frappier, Dana M. Lord, Christopher Ng-Thow-Hing, Erik R. Van Vlack, Shan Tie, Vincent Xue, Sarah C. Cowles, Alan Leung, João V. Rodrigues, Claudio L. Morales-Perez, Alex M. Ayoub, Robin Green, Katherine Puentes, Frank Oplinger, Nishant V. Panwar, Fritz Obermeyer, Adam R. Root, Andrew L. Beam, Frank J. Poelwijk, and Gevorg Grigoryan. Illuminating protein space with a programmable generative model. *Nature*, 623(7989):1070–1078, November 2023. ISSN 0028-0836, 1476-4687. doi: 10.1038/s41586-023-06728-8. URL https://www.nature.com/articles/s41586-023-06728-8.

Natsuki Iwano, Tatsuo Adachi, Kazuteru Aoki, Yoshikazu Nakamura, and Michiaki Hamada. Generative aptamer discovery using RaptGen. *Nature Computational Science*, 2(6):378–386, June 2022. ISSN 2662-8457. doi: 10.1038/s43588-022-00249-6. URL https://www.nature.com/articles/s43588-022-00249-6.

Minghui Jiang, James Anderson, Joel Gillespie, and Martin Mayne. uShuffle: A useful tool for shuffling biological sequences while preserving the k-let counts. *BMC Bioinformatics*, 9(1):192, April 2008. ISSN 1471-2105. doi: 10.1186/1471-2105-9-192. URL https://doi.org/10.1186/1471-2105-9-192.

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, Martin Steinegger, Michalina Pacholska, Tamas Berghammer, Sebastian Bodenstein, David Silver, Oriol Vinyals, Andrew W. Senior, Koray Kavukcuoglu, Pushmeet Kohli, and Demis Hassabis. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873):583–589, August 2021. ISSN 0028-0836, 1476-4687. doi: 10.1038/s41586-021-03819-2. URL https://www.nature.com/articles/s41586-021-03819-2.

Juanjuan Kang, Qiang Tang, Jun He, Le Li, Nianling Yang, Shuiyan Yu, Mengyao Wang, Yuchen Zhang, Jiahao Lin, Tianyu Cui, Yongfei Hu, Puwen Tan, Jun Cheng, Hailong Zheng, Dong Wang, Xi Su, Wei Chen, and Yan Huang. RNAInter v4.0: RNA interactome repository with redefined confidence scoring system and improved accessibility. *Nucleic Acids Research*, 50(D1):D326–D332, January 2022. ISSN 0305-1048. doi: 10.1093/nar/gkab997. URL https://doi.org/10.1093/nar/gkab997.

Namhee Kim, Hin Hark Gan, and Tamar Schlick. A computational proposal for designing structured RNA pools for in vitro selection of RNAs. *RNA*, 13(4):478–492, 2007a.

Namhee Kim, Jin Sup Shin, Shereef Elmetwaly, Hin Hark Gan, and Tamar Schlick. RagPools: RNA-As-Graph-Pools: a web server for assisting the design of structured RNA pools for in vitro selection. *Bioinformatics*, 23(21):2959–2960, 2007b.

Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization, January 2017. URL http://arxiv.org/abs/1412.6980. arXiv:1412.6980 [cs].

Roman Klypa and Oleksandr Cherednichenko. Diversity in Large Language Models under Supervised Fine-Tuning, April 2026. URL http://arxiv.org/abs/2605.00195. arXiv:2605.00195 [cs] version: 1.

Roman Klypa, Alberto Bietti, and Sergei Grudinin. BAnG: Bidirectional Anchored Generation for Conditional RNA Design. In *Proceedings of the 42nd International Conference on Machine Learning*, pp. 31020–31043. PMLR, October 2025. URL https://proceedings.mlr.press/v267/klypa25a.html.

Ana Kozomara, Maria Birgaoanu, and Sam Griffiths-Jones. miRBase: from microRNA sequences to function. *Nucleic Acids Research*, 47(D1):D155–D162, January 2019. ISSN 0305-1048. doi: 10.1093/nar/gky1141. URL https://doi.org/10.1093/nar/gky1141.

Gwangho Lee, Gun Hyuk Jang, Ho Young Kang, and Giltae Song. Predicting aptamer sequences that interact with target proteins using an aptamer-protein interaction classifier and a Monte Carlo tree search approach. *PloS one*, 16(6):e0253760, 2021.

Danyu Li, Rubing Huang, Chenhui Cui, Dave Towey, Ling Zhou, Jinyu Tian, and Bin Zou. RNA-Protein Interaction Prediction Based on Deep Learning: A Comprehensive Survey. *arXiv preprint arXiv:2410.00077*, 2024.

Junnan Li, Dongxu Li, Silvio Savarese, and Steven Hoi. BLIP-2: Bootstrapping Language-Image Pre-training with Frozen Image Encoders and Large Language Models, June 2023. URL http://arxiv.org/abs/2301.12597. arXiv:2301.12597 [cs].

Jennifer Listgarten and Hanlun Jiang. How artificial intelligence is reengineering protein engineering. *Science*, 392(6794):159–166, April 2026. doi: 10.1126/science.aec8444. URL https://www.science.org/doi/10.1126/science.aec8444.

Haotian Liu, Chunyuan Li, Qingyang Wu, and Yong Jae Lee. Visual Instruction Tuning, December 2023. URL http://arxiv.org/abs/2304.08485. arXiv:2304.08485 [cs].

Ronny Lorenz, Stephan H. Bernhart, Christian Höner zu Siederdissen, Hakim Tafer, Christoph Flamm, Peter F. Stadler, and Ivo L. Hofacker. ViennaRNA Package 2.0. *Algorithms for Molecular Biology*, 6(1):26, November 2011. ISSN 1748-7188. doi: 10.1186/1748-7188-6-26. URL https://doi.org/10.1186/1748-7188-6-26.

Ilya Loshchilov and Frank Hutter. Decoupled Weight Decay Regularization, January 2019. URL http://arxiv.org/abs/1711.05101. arXiv:1711.05101 [cs.LG].

Niklas Muennighoff, Alexander M. Rush, Boaz Barak, Teven Le Scao, Aleksandra Piktus, Nouamane Tazi, Sampo Pyysalo, Thomas Wolf, and Colin Raffel. Scaling Data-Constrained Language Models, May 2023. URL https://arxiv.org/abs/2305.16264v5.

Eric Nguyen, Michael Poli, Matthew G. Durrant, Brian Kang, Dhruva Katrekar, David B. Li, Liam J. Bartie, Armin W. Thomas, Samuel H. King, Garyk Brixi, Jeremy Sullivan, Madelena Y. Ng, Ashley Lewis, Aaron Lou, Stefano Ermon, Stephen A. Baccus, Tina Hernandez-Boussard, Christopher Ré, Patrick D. Hsu, and Brian L. Hie. Sequence modeling and design from molecular to genome scale with Evo. *Science*, 386(6723):eado9336, November 2024. doi: 10.1126/science.ado9336. URL https://www.science.org/doi/10.1126/science.ado9336.

Divya Nori and Wengong Jin. RNAFlow: RNA Structure & Sequence Design via Inverse Folding-Based Flow Matching, June 2024. URL http://arxiv.org/abs/2405.18768. arXiv:2405.18768 [q-bio].

Stephen Obonyo, Nicolas Jouandeau, and Dickson Owuor. RNA Generative Modeling With Tree Search. In *2024 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, pp. 1–9. IEEE, 2024.

Nuala A. O'Leary, Mathew W. Wright, J. Rodney Brister, Stacy Ciufo, Diana Haddad, Rich McVeigh, Bhanu Rajput, Barbara Robbertse, Brian Smith-White, Danso Ako-Adjei, Alexander Astashyn, Azat Badretdin, Yiming Bao, Olga Blinkova, Vyacheslav Brover, Vyacheslav Chetvernin, Jinna Choi, Eric Cox, Olga Ermolaeva, Catherine M. Farrell, Tamara Goldfarb, Tripti Gupta, Daniel Haft, Eneida Hatcher, Wratko Hlavina, Vinita S. Joardar, Vamsi K. Kodali, Wenjun Li, Donna Maglott, Patrick Masterson, Kelly M. McGarvey, Michael R. Murphy, Kathleen O'Neill, Shashikant Pujar, Sanjida H. Rangwala, Daniel Rausch, Lillian D. Riddick, Conrad Schoch, Andrei Shkeda, Susan S. Storz, Hanzhen Sun, Francoise Thibaud-Nissen, Igor Tolstoy, Raymond E. Tully, Anjana R. Vatsan, Craig Wallin, David Webb, Wendy Wu, Melissa J. Landrum, Avi Kimchi, Tatiana Tatusova, Michael DiCuccio, Paul Kitts, Terence D. Murphy, and Kim D. Pruitt. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research*, 44(D1):D733–D745, January 2016. ISSN 0305-1048. doi: 10.1093/nar/gkv1189. URL https://doi.org/10.1093/nar/gkv1189.

Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul Christiano, Jan Leike, and Ryan Lowe. Training language models to follow instructions with human feedback, March 2022. URL http://arxiv.org/abs/2203.02155. arXiv:2203.02155 [cs].

Furkan Ozden, Sina Barazandeh, Dogus Akboga, Sobhan Shokoueian Tabrizi, Urartu Ozgur Safak Seker, and A. Ercument Cicek. RNAGEN: A generative adversarial network-based model to generate synthetic RNA sequences to target proteins, July 2023. URL http://biorxiv.org/lookup/doi/10.1101/2023.07.11.548246.

Laura O'Mahony, Leo Grinsztajn, Hailey Schoelkopf, and Stella Biderman. Attributing mode collapse in the fine-tuning of large language models. In *ICLR 2024 Workshop on Mathematical and Empirical Understanding of Foundation Models*, volume 2, 2024. URL https://openreview.net/forum?id=3pDMYjpOxk.

Arka Pal, Deep Karkhanis, Samuel Dooley, Manley Roberts, Siddartha Naidu, and Colin White. Smaug: Fixing Failure Modes of Preference Optimisation with DPO-Positive, July 2024. URL http://arxiv.org/abs/2402.13228. arXiv:2402.13228 [cs].

Byungkyu Park and Kyungsook Han. Discovering protein-binding RNA motifs with a generative model of RNA sequences. *Computational Biology and Chemistry*, 84:107171, February 2020. ISSN 14769271. doi: 10.1016/j.compbiolchem.2019.107171. URL https://linkinghub.elsevier.com/retrieve/pii/S1476927119305365.

Razvan Pascanu, Tomas Mikolov, and Yoshua Bengio. On the difficulty of training Recurrent Neural Networks, February 2013. URL http://arxiv.org/abs/1211.5063. arXiv:1211.5063 [cs.LG].

Alec Radford, Jeff Wu, R. Child, D. Luan, Dario Amodei, and I. Sutskever. Language Models are Unsupervised Multitask Learners, 2019. URL https://www.semanticscholar.org/paper/Language-Models-are-Unsupervised-Multitask-Learners-Radford-Wu/9405cc0d6169988371b2755e573cc28650d14dfe.

Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. Learning Transferable Visual Models From Natural Language Supervision, February 2021. URL http://arxiv.org/abs/2103.00020. arXiv:2103.00020 [cs].

Rafael Rafailov, Archit Sharma, Eric Mitchell, Stefano Ermon, Christopher D. Manning, and Chelsea Finn. Direct Preference Optimization: Your Language Model is Secretly a Reward Model, 2023. URL https://arxiv.org/abs/2305.18290.

Debashish Ray, Hilal Kazan, Kate B. Cook, Matthew T. Weirauch, Hamed S. Najafabadi, Xiao Li, Serge Gueroussov, Mihai Albu, Hong Zheng, Ally Yang, Hong Na, Manuel Irimia, Leah H. Matzat, Ryan K. Dale, Sarah A. Smith, Christopher A. Yarosh, Seth M. Kelly, Behnam Nabet, Desirea Mecenas, Weimin Li, Rakesh S. Laishram, Mei Qiao, Howard D. Lipshitz, Fabio Piano, Anita H. Corbett, Russ P. Carstens, Brendan J. Frey, Richard A. Anderson, Kristen W. Lynch, Luiz O. F. Penalva, Elissa P. Lei, Andrew G. Fraser, Benjamin J. Blencowe, Quaid D. Morris, and Timothy R. Hughes. A compendium of RNA-binding motifs for decoding gene regulation. *Nature*, 499(7457):172–177, July 2013. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature12311. URL https://www.nature.com/articles/nature12311.

Noam Razin, Sadhika Malladi, Adithya Bhaskar, Danqi Chen, Sanjeev Arora, and Boris Hanin. Unintentional Unalignment: Likelihood Displacement in Direct Preference Optimization, April 2025. URL http://arxiv.org/abs/2410.08847. arXiv:2410.08847 [cs].

Sashank J. Reddi, Satyen Kale, and Sanjiv Kumar. On the Convergence of Adam and Beyond, April 2019. URL http://arxiv.org/abs/1904.09237. arXiv:1904.09237 [cs].

Alexander Sasse, Debashish Ray, Kaitlin U. Laverty, Cyrus L. Tam, Mihai Albu, Hong Zheng, Yevgen Levdansky, Olga Lyudovyk, Taykhoom Dalal, Kate Nie, Cedrik Magis, Cedric Notredame, Eugene Valkov, Matthew T. Weirauch, Timothy R. Hughes, and Quaid Morris. A resource of RNA-binding protein motifs across eukaryotes reveals evolutionary dynamics and gene-regulatory function. *Nature Biotechnology*, 44(7):1226–1236, July 2025. ISSN 1546-1696. doi: 10.1038/s41587-025-02733-6. URL https://www.nature.com/articles/s41587-025-02733-6.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal Policy Optimization Algorithms, August 2017. URL http://arxiv.org/abs/1707.06347. arXiv:1707.06347 [cs].

Rico Sennrich, Barry Haddow, and Alexandra Birch. Neural Machine Translation of Rare Words with Subword Units, June 2016. URL http://arxiv.org/abs/1508.07909. arXiv:1508.07909 [cs].

Incheol Shin, Keumseok Kang, Juseong Kim, Sanghun Sel, Jeonghoon Choi, Jae-Wook Lee, Ho Young Kang, and Giltae Song. AptaTrans: a deep neural network for predicting aptamer-protein interaction using pretrained encoders. *BMC bioinformatics*, 24(1):447, 2023.

Martin Steinegger and Johannes Söding. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology*, 35(11):1026–1028, November 2017. ISSN 1546-1696. doi: 10.1038/nbt.3988. URL https://www.nature.com/articles/nbt.3988.

Martin Steinegger and Johannes Söding. Clustering huge protein sequence sets in linear time. *Nature Communications*, 9(1):2542, June 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-04964-5. URL https://www.nature.com/articles/s41467-018-04964-5.

Jianlin Su, Yu Lu, Shengfeng Pan, Ahmed Murtadha, Bo Wen, and Yunfeng Liu. RoFormer: Enhanced Transformer with Rotary Position Embedding, November 2023. URL http://arxiv.org/abs/2104.09864. arXiv:2104.09864 [cs].

Sobhan Shukueian Tabrizi, Helyasadat Hashemi Aghdam, and A. Ercument Cicek. RNA–X: Modeling RNA interactions to design binder RNA and simultaneously target multiple molecules of different types, November 2025a. URL https://www.biorxiv.org/content/10.1101/2025.11.24.690191v1. ISSN: 2692-8205 Pages: 2025.11.24.690191 Section: New Results.

Sobhan Shukueian Tabrizi, Sina Barazandeh, Helyasadat Hashemi Aghdam, and A. Ercument Cicek. RNAtranslator: Modeling protein-conditional RNA design as sequence-to-sequence natural language translation. *PLOS Computational Biology*, 21(10):e1013541, October 2025b. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1013541. URL https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1013541.

Walter Thavarajah, Laura M. Hertz, David Z. Bushhouse, Chloé M. Archuleta, and Julius B. Lucks. RNA Engineering for Public Health: Innovations in RNA-Based Diagnostics and Therapeutics. *Annual Review of Chemical and Biomolecular Engineering*, 12(1):263–286, June 2021. ISSN 1947-5438, 1947-5446. doi: 10.1146/annurev-chembioeng-101420-014055. URL https://www.annualreviews.org/doi/10.1146/annurev-chembioeng-101420-014055.

The UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2025. *Nucleic Acids Research*, 53(D1):D609–D617, January 2025. ISSN 1362-4962. doi: 10.1093/nar/gkae1010. URL https://doi.org/10.1093/nar/gkae1010.

The RNAcentral Consortium, Blake A Sweeney, Anton I Petrov, Boris Burkov, Robert D Finn, Alex Bateman, Maciej Szymanski, Wojciech M Karlowski, Jan Gorodkin, Stefan E Seemann, Jamie J Cannone, Robin R Gutell, Petra Fey, Siddhartha Basu, Simon Kay, Guy Cochrane, Kostantinos Billis, David Emmert, Steven J Marygold, Rachael P Huntley, Ruth C Lovering, Adam Frankish, Patricia P Chan, Todd M Lowe, Elspeth Bruford, Ruth Seal, Jo Vandesompele, Pieter-Jan Volders, Maria Paraskevopoulou, Lina Ma, Zhang Zhang, Sam Griffiths-Jones, Janusz M Bujnicki, Pietro Boccaletto, Judith A Blake, Carol J Bult, Runsheng Chen, Yi Zhao, Valerie Wood, Kim Rutherford, Elena Rivas, James Cole, Stanley J F Laulederkind, Mary Shimoyama, Marc E Gillespie, Marija Orlic-Milacic, Ioanna Kalvari, Eric Nawrocki, Stacia R Engel, J Michael Cherry, Silva Team, Tanya Z Berardini, Artemis Hatzigeorgiou, Dimitra Karagkouni, Kevin Howe, Paul Davis, Marcel Dinger, Shunmin He, Maki Yoshihama, Naoya Kenmochi, Peter F Stadler, and Kelly P Williams. RNAcentral: a hub of information for non-coding RNA sequences. *Nucleic Acids Research*, 47(D1):D221–D229, January 2019. ISSN 0305-1048, 1362-4962. doi: 10.1093/nar/gky1034. URL https://academic.oup.com/nar/article/47/D1/D221/5160993.

Mahsa Torkamanian-Afshar, Sajjad Nematzadeh, Maryam Tabarzad, Ali Najafi, Hossein Lanjanian, and Ali Masoudi-Nejad. In silico design of novel aptamers utilizing a hybrid method of machine learning and genetic algorithm. *Molecular diversity*, 25:1395–1407, 2021.

Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. LLaMA: Open and Efficient Foundation

Language Models, 2023. URL https://arxiv.org/abs/2302.13971. Version Number: 1.

Chih-Yuan Tseng, Md Ashrafuzzaman, Jonathan Y Mane, Janice Kapty, John R Mercer, and Jack A Tuszynski. Entropic Fragment-Based Approach to Aptamer Design. *Chemical Biology & Drug Design*, 78(1):1–13, 2011.

Michel van Kempen, Stephanie S. Kim, Charlotte Tumescheit, Milot Mirdita, Jeongjae Lee, Cameron L. M. Gilchrist, Johannes Söding, and Martin Steinegger. Fast and accurate protein structure search with Foldseek. *Nature Biotechnology*, 42(2):243–246, February 2024. ISSN 1546-1696. doi: 10.1038/s41587-023-01773-0. URL https://www.nature.com/articles/s41587-023-01773-0.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention Is All You Need, August 2023. URL http://arxiv.org/abs/1706.03762. arXiv:1706.03762 [cs].

Yue Wang, Bhaven A Mistry, and Tom Chou. Discrete stochastic models of SELEX: Aptamer capture probabilities and protocol optimization. *The Journal of Chemical Physics*, 156(24), 2022.

Zhen Wang, Ziqi Liu, Wei Zhang, Yanjun Li, Yizhen Feng, Shaokang Lv, Han Diao, Zhaofeng Luo, Pengju Yan, Min He, and others. AptaDiff: de novo design and optimization of aptamers based on diffusion models. *Briefings in Bioinformatics*, 25(6):bbae517, 2024.

Jason Wei, Maarten Bosma, Vincent Y. Zhao, Kelvin Guu, Adams Wei Yu, Brian Lester, Nan Du, Andrew M. Dai, and Quoc V. Le. Finetuned Language Models Are Zero-Shot Learners, February 2022. URL http://arxiv.org/abs/2109.01652. arXiv:2109.01652 [cs].

Biao Zhang and Rico Sennrich. Root Mean Square Layer Normalization, October 2019. URL http://arxiv.org/abs/1910.07467. arXiv:1910.07467 [cs].

Yaqing Zhang, Yuan Jiang, David Kuster, Qiwei Ye, Wenhao Huang, Simon Fürbacher, Jingye Zhang, Pia Doll, Wenjun Lin, Siwei Dong, Hui Wang, Zhipeng Tang, David Ibberson, Klemens Wild, Irmgard Sinning, Anthony A. Hyman, and Andres Jäschke. Single-step discovery of high-affinity RNA ligands by UltraSelex. *Nature Chemical Biology*, 21(7):1118–1126, July 2025. ISSN 1552-4469. doi: 10.1038/s41589-025-01868-6. URL https://www.nature.com/articles/s41589-025-01868-6.

Zaixi Zhang, Linlin Chao, RuoFan Jin, Yikun Zhang, Guowei Zhou, Yujie Yang, Yukang Yang, Kaixuan Huang, Qirong Yang, Ziyao Xu, and others. RNAGenesis: Foundation Model for Enhanced RNA Sequence Generation and Structural Insights. *bioRxiv*, pp. 2024–12, 2024.

Yichong Zhao, Kenta Oono, Hiroki Takizawa, and Masaaki Kotera. GenerRNA: A generative pre-trained language model for de novo RNA design. *PLOS ONE*, 19(10):e0310814, October 2024. ISSN 1932-6203. doi: 10.1371/journal.pone.0310814. URL https://dx.plos.org/10.1371/journal.pone.0310814.

Chunting Zhou, Pengfei Liu, Puxin Xu, Srini Iyer, Jiao Sun, Yuning Mao, Xuezhe Ma, Avia Efrat, Ping Yu, Lili Yu, Susan Zhang, Gargi Ghosh, Mike Lewis, Luke Zettlemoyer, and Omer Levy. LIMA: Less Is More for Alignment, May 2023. URL https://arxiv.org/abs/2305.11206v1.

Daniel M. Ziegler, Nisan Stiennon, Jeffrey Wu, Tom B. Brown, Alec Radford, Dario Amodei, Paul Christiano, and Geoffrey Irving. Fine-Tuning Language Models from Human Preferences, January 2020. URL http://arxiv.org/abs/1909.08593. arXiv:1909.08593 [cs].

APPENDIX

A TECHNICAL APPENDICES AND SUPPLEMENTARY MATERIAL

A.1 PRETRAINING DETAILS

Data Curation & Preprocessing The pre-training corpus was sourced from RNAcentral. To ensure sequence integrity and computational efficiency, we excluded any samples containing non-standard nucleotides or those with lengths exceeding 20,000 or falling below 16 nucleotides. The remaining dataset was clustered using MMseqs2 (Steinegger & Söding, 2018) with a minimum sequence identity of 0.8 and a coverage threshold of 0.8. Cluster representatives were designated for the training set, with 10,000 randomly selected reserved as a held-out validation set.

Tokenization We utilized a Byte Pair Encoding tokenizer. The vocabulary was learned from a subset of 1M sequences randomly sampled from the training set. The frequencies of the resulting tokens on the same subset are depicted in Figure A.1.

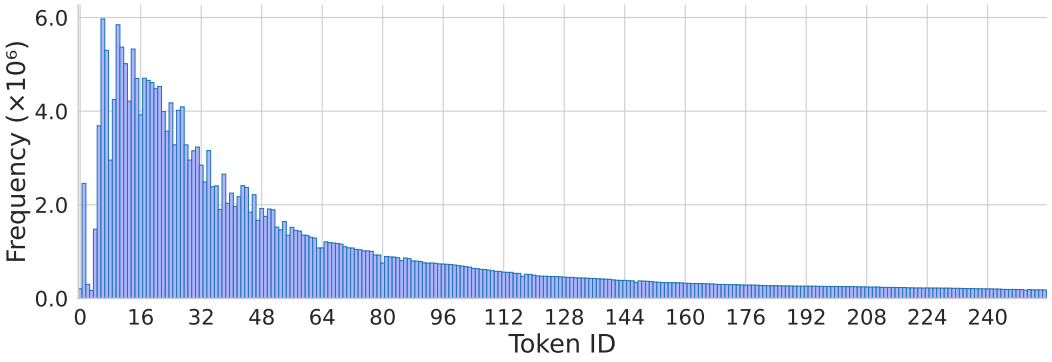


Figure A.1: The plot illustrates the frequency of tokens within the data subset used for tokenizer training. Tokens are ordered along the x-axis by their respective token IDs, which corresponds to their BPE merge order.

Architecture The model consists of $N = 24$ transformer blocks. Each block comprises self-attention layers and feed-forward networks utilizing GeLU activations (Hendrycks & Gimpel, 2016). To ensure stable training at scale, we employed RMSNorm (Zhang & Sennrich, 2019) for pre-attention normalization (Dehghani et al., 2023). Relative positional information was incorporated using Rotary Positional Encodings (RoPE) (Su et al., 2023), facilitating the handling of variable sequence lengths. Moirain-Base features a latent dimension of 1024, utilizing 16 attention heads with a hidden dimension of 64 each, and a feed-forward expansion factor (transition factor) of 4.

Training Configuration Moirain-Base was trained on eight NVIDIA A100 GPUs with a global batch size of 128. We utilized a maximum context length of 2,048 tokens, cropping sequences exceeding this limit.

Optimization We performed optimization using the AdamW (Loshchilov & Hutter, 2019) optimizer with hyperparameters $\beta_1 = 0.9$, $\beta_2 = 0.95$, $\epsilon = 10^{-6}$, $\lambda = 0$ and gradient norm clipping (Pascanu et al., 2013). The learning rate was set to 10^{-4}, governed by a schedule consisting of a linear warmup over the first 5,000 steps, followed by a sinusoidal decay for the remainder of the training duration.

A.2 MULTI-MODAL SFT DETAILS

Data Collection & Mapping Interaction data was sourced from RNAInter, which provides protein-RNA pairs via database identifiers. Protein IDs were mapped to their corresponding UniProt

(The UniProt Consortium, 2025) sequences, ensuring, where possible, that the UniProt gene nomenclature aligned with RNAInter records. For RNA, we restricted our selection to samples originating from NCBI (Brown et al., 2015) or miRBase (Kozomara et al., 2019) to ensure high-quality transcript mapping. These were mapped to RefSeq (O'Leary et al., 2016), filtered for non-coding RNA transcripts, and RNAcentral respectively. Any identifiers mapping to more than four distinct sequences were excluded to maintain data integrity. Successfully mapped proteins were grouped into FoldSeek (van Kempen et al., 2024) clusters, while RNA sequences were clustered using the same parameters as the pre-training data (MMseqs2, 0.8 identity, 0.8 coverage).

Data Curation & Preprocessing To prevent the model from collapsing into a "one-size-fits-all" generation strategy, we implemented a filtering pipeline to address highly promiscuous interactors. We observed that less than 1% of RNAs accounted for approximately 80% of protein cluster interactions. To mitigate this:

1. We removed RNA clusters that individually interacted with more than 40% of all protein clusters.
2. For each unique protein, we retained interactions only with RNA clusters that cover fewer than 64 protein clusters.
3. In cases where a protein had no such partners, we preserved only two of its paired RNA clusters, with the least number of interactions with distinct protein clusters .

To optimize computational throughput, we excluded any protein sequences longer than 512 amino acids and RNA sequences exceeding 512 tokens. The RNA-BAnG benchmark includes a "zero-similarity" subset, which contains proteins that share no detectable sequence similarity (via BLASTp (Altschul et al., 1990) with default parameters) with its own training data. It serves as a rigorous test for generalization across unseen target space. To maintain this independence in our own training, we withheld all protein clusters containing sequences with 50% or greater similarity to this subset, also using BLASTp with default parameters. The final processed training set comprised 203,811 interaction samples, spanning 10,648 protein clusters and 12,287 RNA clusters.

Architecture For the protein representation block, we utilize its default configuration from RNA-BAnG. The cross-attention architecture comprises $M = 4$ blocks with a hidden dimension of 64 and 16 heads. Each block applies pre-normalization to both keys and queries. Positional encodings are omitted, as the cross-attention operates across different modalities. The pLDDT scores representing protein structure confidence are encoded by discretizing the values into 32 equal bins.

Training Configuration TOFU loss parameters were set to recommended $\beta = 0.8, \gamma = 0.3$. Moirain-SFT was trained on four NVIDIA A100 GPUs with a global batch size of 32. During training, we sampled pairs sharing the same protein and RNA clusters at a rate of two pairs per epoch. We applied LoRA $(r = 32, \alpha = 32)$ to the keys, queries, projection, and feed-forward layers.

Optimization Optimization was performed using the AMSGrad (Reddi et al., 2019) variant of the Adam (Kingma & Ba, 2017) optimizer with default hyperparameters $(\beta_1 = 0.9, \beta_2 = 0.999)$. The learning rate was set to 10^{-4}, the linear warmup period was shortened to 1,000 steps.

A.3 Preference Optimization Details

Data Curation The preference dataset was constructed using experimental binding scores from (Ray et al., 2013) and (Sasse et al., 2025). To create high-contrast preference pairs, we selected the 1,000 RNA sequences with the highest binding scores and paired them with the 1,000 possessing the lowest scores. To maintain the integrity of our generalization benchmarks, we excluded any proteins belonging to MMseqs2 clusters that shared 40% or greater sequence similarity with the "zero-similarity" test subset. Furthermore, we restricted the dataset to proteins with lengths below 512 amino acids. The resulting training set comprised 383,000 preference samples.

Training & Optimization The DPO objective was optimized with a default $\beta = 0.1$ coefficient. We applied LoRA $(r = 16, \alpha = 16)$ to the keys, queries, projection, and feed-forward layers.

Moirain was trained on four NVIDIA A100 GPUs with a global batch size of 16. We utilized the same optimizer configuration as the SFT phase.

B Inference Details

During the generation of RNA sequences, we employed a temperature of $T = 1$ across all models. For Moirain-Base, we utilized full random sampling to explore the learned sequence space, performing unconstrained generation of 10,000 sequences with a maximum length of 512 tokens. For our tuned variants, Moirain-SFT and Moirain, we implemented Nucleus Sampling (Top-p) (Holtzman et al., 2020) with a threshold of $p = 0.9$ to balance diversity and coherence. In conditioned generation tasks, we generated 1000 sequences per target protein, with the decoding process constrained to a maximum of 50 tokens.

C Metrics

Z-score We used uShuffle (Jiang et al., 2008) to create 100 distorted versions of each RNA sequence via dinucleotide-preserving permutations. These were used to compare the original sequences against a shuffled background and to calculate MFE Z-scores. The MFE was computed using RNAfold (Lorenz et al., 2011), where we included only sequences shorter than 256 nucleotides to maintain adequate computational time. The resulting analysis was conducted on 3,690 generated RNA sequences and 3,272 natural RNA sequences.

Fidelity To evaluate Fidelity, we compared the 3-mer distributions of generated sequences against specific reference sets. For Moirain-Base, the comparison was performed against a held-out set of natural ncRNAs. For the remaining models, comparisons were made against the interacting RNA sequences from the multi-modal SFT training set. We restricted the analysis to sequences shorter than 512 nucleotides for the unconditional task and shorter than 51 nucleotides for conditional tasks to ensure that potentially cropped sequences were excluded from the evaluation. The resulting comparison sets comprise 5,568 samples for Moirain-Base, 5,864 natural ncRNAs, and 2,538 interacting RNAs, with approximately 12,000 (1,200 for RNAFlow and RFDPoly) for every other model variant.

Novelty & Diversity For the calculation of Novelty, we employed MMseqs2 search (Steinegger & Söding, 2017) with the search type parameter set to 3. Searches were performed individually against the respective training sets for Moirain-Base and Moirain-SFT, and against the preferred examples within the Preference Optimization phase training set. To determine Diversity clusters, we utilized MMseqs2 with a 0.8 similarity threshold and a 0.8 coverage.

Binding Affinity To compute binding affinity scores, we cropped all sequences to 50 nucleotides and excluded those with a length of less than 6 nucleotides. These constraints were applied to adhere to DeepCLIP size restrictions and because very short sequences lack biological relevance in this context.

D Additional Results

As illustrated in Figure D.1, Moirain generates a broad distribution of sequence lengths, centered primarily between 40–45 nucleotides. Importantly, an additional peak emerges at approximately 21-23 nucleotides, which matches the characteristic length of microRNAs. This suggests the model is successfully capturing specific biological archetypes within its broader length distribution.

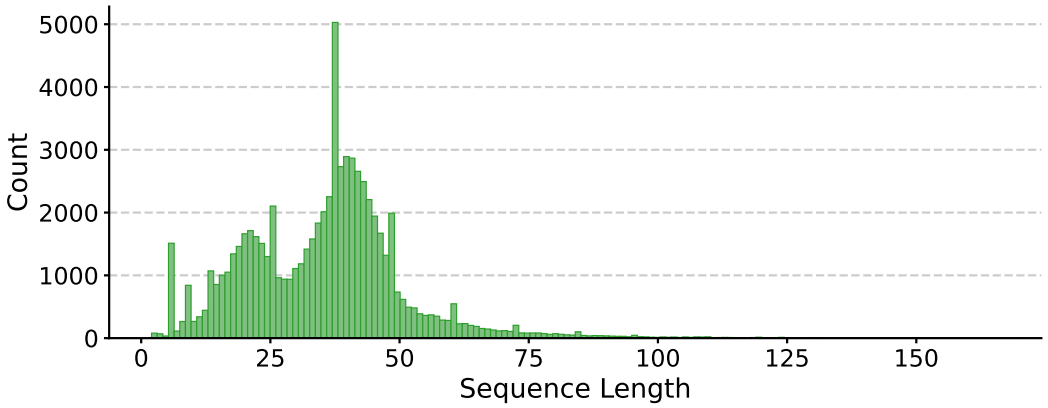


Figure D.1: Sequence length distribution (nucleotides) of Moirain (following TOFU SFT) generations across the complete test set.