

RNARL: reinforcement learning-driven unified generative framework for multi-objective RNA codon design

Received: 21 Oct 2025

Accepted: 08 Jul 2026

Published online: 21 July 2026

Shenggeng Lin, Hong Tan, Keyao Wang, Ruixuan Wang, Hongxia Wang, Tong Zhu & Yi Xiong

Cite this article as: Lin, S., Tan, H., Wang, K. *et al.* RNARL: reinforcement learning-driven unified generative framework for multi-objective RNA codon design. *Genome Biol* (2026). <https://doi.org/10.1186/s13059-026-04203-x>

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

RNARL: reinforcement learning-driven unified generative framework for multi-objective RNA codon design

Shenggeng Lin ^{1, 2, 3}, Hong Tan ^{1, 3}, Keyao Wang ^{1, 3}, Ruixuan Wang ^{1, 3}, Hongxia Wang ⁴, Tong Zhu ^{2, 5, *}, Yi Xiong ^{1, 2, 3, *}

¹ State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China

² Shanghai Innovation Institute, Shanghai, China

³ Artificial Intelligence Biomedical Center, Zhangjiang Institute for Advanced Study, Shanghai Jiao Tong University, Shanghai, China

⁴ Department of Medical Oncology, Fudan University Shanghai Cancer Center; Department of Oncology, Shanghai Medical College, Fudan University, Shanghai, China

⁵ Shanghai Engineering Research Center of Molecular Therapeutics and New Drug Development, School of Chemistry and Molecular Engineering, East China Normal University, Shanghai, China

* Corresponding author(s): Yi Xiong (xiongyi@sjtu.edu.cn) or Tong Zhu (tongzhu.work@gmail.com)

Abstract

Current RNA codon design methods are limited by inefficient long-sequence processing and poor generalizability, often relying on a decoupled “generate-or-optimize” paradigm. We introduce RNARL, a reinforcement learning-driven framework that unifies sequence generation with multi-objective optimization. RNARL directly learns to generate high-performance sequences, effectively optimizing sequences over 3,900 nucleotides and demonstrating superior performance and universality across six species and five RNA types. RNARL thus establishes an effective and generalizable framework for RNA codon design. Finally, a user-friendly web platform is freely available to facilitate its application for RNA therapeutic design.

Key words : RNA Codon Optimization, Reinforcement Learning, Generative Model, RNA vaccine

Background

RNA vaccines deliver RNA that encodes antigens, enabling host cells to autonomously synthesize these antigens and activate immune responses [1-4]. However, due to the degeneracy of the genetic code, a single protein sequence can be encoded by multiple distinct RNA sequences [5, 6]. This codon degeneracy poses a critical challenge: although encoding identical proteins, synonymous RNA variants can exhibit significant differences in translational efficiency, structural stability and expression levels [7]. These differences primarily stem from species-specific codon usage bias and RNA structural features [8, 9]. Therefore, RNA codon design constitutes a critical strategy for achieving optimal vaccine efficacy [10-14]. This process enhances sequence compatibility with host translational machinery while stabilizing RNA secondary structures, thereby amplifying antigen expression to potentiate both prophylactic and therapeutic applications [15-17].

Despite the demonstrated importance of codon design [18], current implementation strategies encounter multifaceted challenges. Traditional strategies, such as “host-preferred codon substitution” and “natural codon distribution matching”, are conceptually straightforward but primarily optimize for translational efficiency, often at the expense of RNA structural stability [19, 20]. These methods operate by replacing codons based on host frequency statistics. For example, the “host-preferred substitution” strategy uses only the single most frequent codon for each amino acid. This emphasis on codon frequency inherently neglects the sequence-specific nucleotide (nt) context that is essential for optimal folding, often resulting in suboptimal trade-offs between translational efficiency and RNA structural stability. Moreover, such frequency-based codon designs can lead to the depletion of specific tRNA pools, further limiting their practical utility [5]. More recent approaches, such as LinearDesign [15], aim to achieve a balance between translational efficiency and structural stability, demonstrating promising solutions. However, their reliance on heuristic search strategies limits their efficiency in achieving high-throughput optimization of long RNA sequences. Deep learning-based methods [21-26],

including those employing LSTM architectures, struggle to capture long-range dependencies in RNA sequences and are unsuitable for handling exceptionally long sequences [27]. Furthermore, pretrained models, often trained on natural sequence distributions, tend to generate sequences whose properties closely resemble their natural counterparts, thereby hindering their capacity to produce RNA sequences with substantially superior performance [5, 28-33].

Based on the aforementioned facts, current RNA codon design algorithms encounter three unresolved challenges: First, efficiently processing and optimizing long RNA sequences ($>3,000$ nts) remains a critical bottleneck [15]. Second, existing methods demonstrate limited generalizability across species and RNA types [27]. Third, while generative models enable high throughput design, they rarely produce RNAs with properties superior to natural RNA sequences. In contrast, optimization algorithms can generate sequences that surpass natural ones, but are often too computationally intensive for large-scale design. The conventional “generate-or-optimize” paradigm imposes inherent efficiency and performance limitations, necessitating deep integration of sequence generation with optimization objectives to directly generate high-performance sequences.

To address these three challenges, we propose RNA sequence design via reinforcement learning (named RNARL), a unified framework that integrates a mixture-of-experts (MoE) Transformer-based generative model [34] with group relative policy optimization (GRPO) [35] for RNA codon design. The core idea is to leverage the high-throughput sequence generation capabilities of generative models, while employing reinforcement learning to iteratively refine the generative policy in accordance with predefined multi-objective design criteria. In RNARL, the generative model is directly optimized based on a reward function that evaluates desired properties (e.g., high translational efficiency, enhanced structural stability), thereby refining its policy to generate RNA sequences that satisfy these objectives. This approach tightly couples sequence generation with optimization, distinguishing it from the conventional “generate-or-optimize” paradigm and enabling the direct design of high-performance RNA sequences.

Extensive computational experiments validate the framework’s efficacy and

universality. Notably, by training on datasets containing RNA sequences up to ~3,900 nts in length and incorporating length-aware distribution sampling (LADS) techniques, our framework overcomes existing limitations in processing long sequences, enabling efficient generation and optimization. Furthermore, evaluations across six major biological categories demonstrate its exceptional adaptability to diverse host environments and heterologous expression scenarios, while assessments on five RNA types confirm its broad applicability to diverse RNA molecular designs. Successful optimization of two publicly available RNA vaccine patent sequences by RNARL underscores the framework's practical potential.

Results

Overview of the RNARL framework

RNARL integrates the GRPO [35] reinforcement learning algorithm with generative models to optimize RNA sequences that encode target proteins. The framework comprises an actor model and a reward model (**Fig. 1(a)**). The actor model serves as a generator, taking protein sequences as input and producing candidate RNA sequences. The reward model acts as an evaluator, taking an RNA sequence as input and assigning it a quantitative score. The two models are trained jointly via reinforcement learning: the reward model defines a multi-objective reward signal, and the parameters of the actor model are iteratively updated to refine its generative policy, thereby favoring the generation of RNA sequences that maximize this reward.

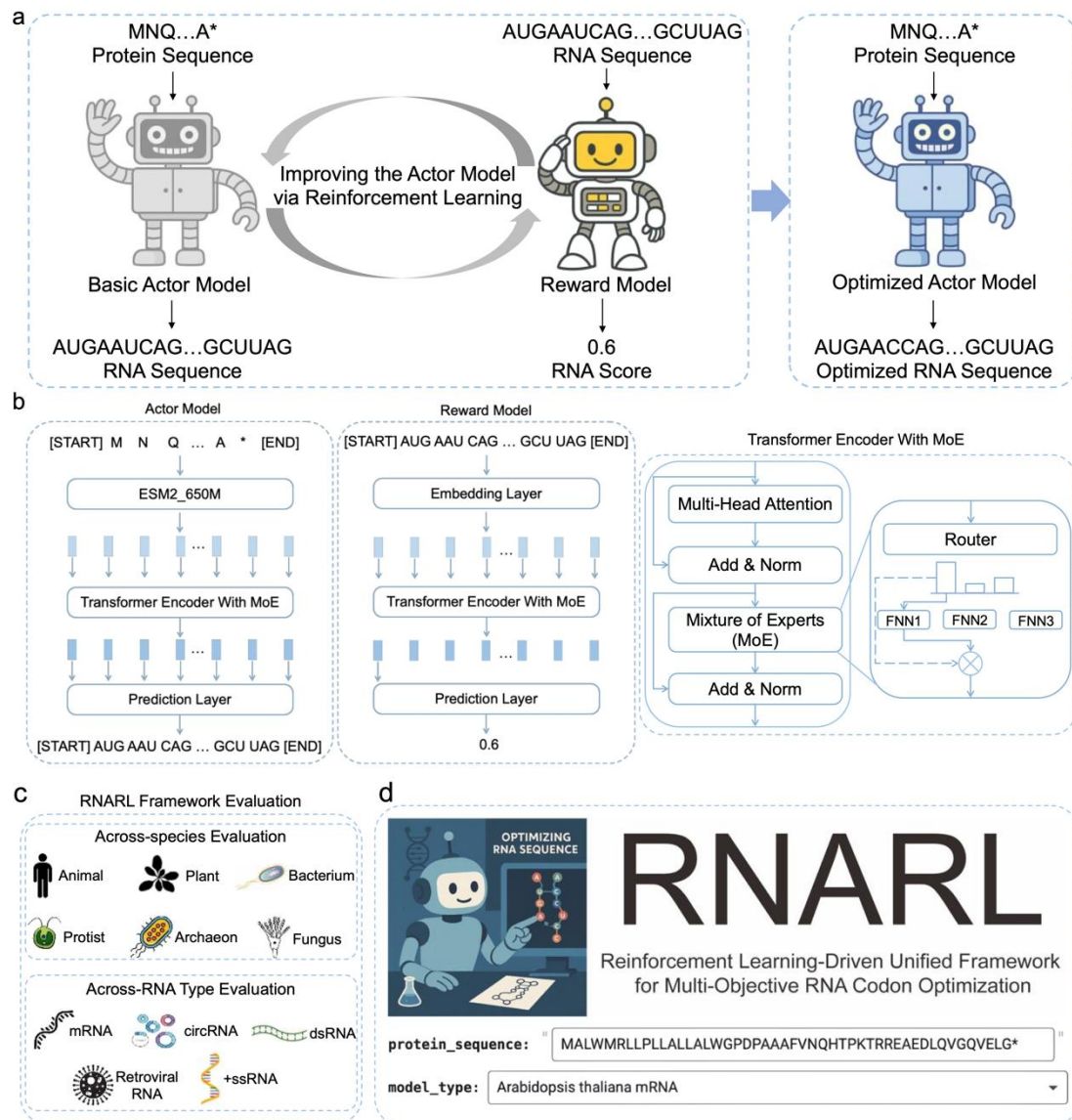


Fig. 1. Overview of the RNARL framework. **a**, The reinforcement learning optimization process, where an actor model (generator) is iteratively refined based on reward signals from a reward model (evaluator). **b**, The core model architecture incorporating Transformer and MoE layers. **c**, Validation of the RNARL framework's generalizability across six species and five RNA types. **d**, The user-friendly RNARL online design platform.

Both the actor and reward models are based on a Transformer encoder architecture enhanced with MoE layers (**Fig. 1(b)**). The framework's performance was rigorously evaluated across six diverse biological categories: *Homo sapiens* (animals), *Arabidopsis thaliana* (plants), *Penicillium chrysogenum* (fungi), *Chlamydomonas reinhardtii* (protists), *Escherichia coli* (bacteria), and *Thermococcus kodakarensis*

KOD1 (archaea). Its generalizability was further tested on five distinct RNA types, including mRNA, circular RNA (circRNA), positive-sense single-stranded RNA (+ssRNA), double-stranded RNA (dsRNA), and retroviral RNA (**Fig. 1(c)**). Furthermore, a user-friendly web platform was developed to facilitate RNA therapeutic design (**Fig. 1(d)**).

Dataset benchmarking and ablation analysis

To comprehensively evaluate the capabilities of the RNARL framework, we constructed six datasets spanning distinct biological categories. These datasets cover protein lengths ranging from 10 to 1,305 amino acids (corresponding to RNA lengths of 30–3,915 nts), encompassing most potential antigen length ranges. Each dataset was rigorously partitioned into training, validation, and test sets (**Fig. 2(a)**, see Materials and Methods). We performed detailed analyses across the datasets and present the *Homo sapiens* dataset as an illustrative example of our findings.

A pronounced “long-tail distribution” was observed in RNA sequence lengths: when binned into 100-nt intervals, sequence counts varied by up to an order of magnitude across intervals (**Fig. 2(b)**). This imbalance poses challenges for existing models in effectively optimizing long sequences. To mitigate this issue, we employed LADS techniques during the model training. RNA structural features play a crucial role in RNA structural stability. We quantified the frequency of major RNA secondary structures in the dataset (**Fig. 2(c)**). Stem-loop and hairpin loop structures were dominated, accounting for 31.4% and 28.4% of the total ones, respectively, consistent with their roles in maintaining RNA stability and regulating translation processes.

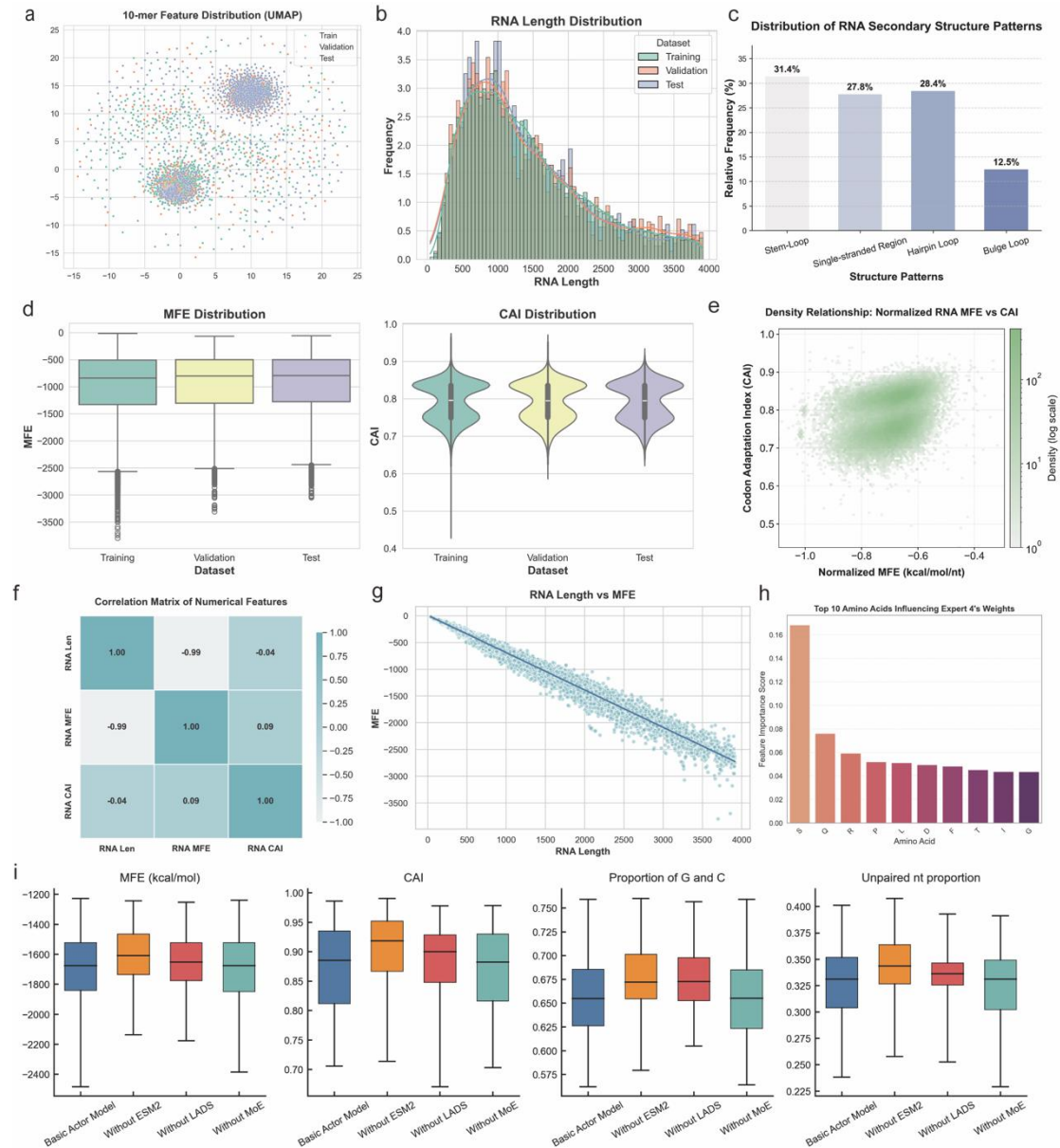


Fig. 2. Characterization of the *Homo sapiens* dataset and ablation study of the RNARL framework. **a**, Distribution of protein sequences based on 10-mer frequency vectors across training, validation, and test sets. **b**, The distribution of RNA lengths (in nucleotides) for sequences on the training, validation, and test datasets. **c**, The relative frequency of different RNA secondary structure patterns (Stem-Loop, Single-stranded region, Hairpin Loop, Bulge Loop) within the dataset. **d**, Distributions of MFE and CAI for RNA sequences across the training, validation, and test sets. **e**, The joint distribution of normalized MFE and CAI values; color intensity indicates density. **f**, Pearson correlation coefficients between RNA length (RNA Len), RNA MFE, and RNA CAI values. **g**, The correlation between RNA length and MFE, with a fitted polynomial regression line. **h**, The top 10 amino acid types influencing

the activation status (weights) of Expert 4 in the MoE gating network, based on feature importance scores. **i**, Ablation experiments on the RNARL framework, comparing the performance of the basic actor model with variants lacking key components (without ESM2, without LADS, without MoE).

Four core metrics were used to evaluate the generated RNA sequences: minimum free energy (MFE; lower values correlate with higher structural stability), codon adaptation index [36] (CAI; higher values correlate with enhanced expression potential), proportion of G (guanine) and C (cytosine) [37] (a measure of structural stability; higher proportion of G and C may favor stable secondary structures), and unpaired nucleotide proportion (a measure of flexibility/degradation risk; lower values reduce susceptibility to ribonuclease degradation). We further analyzed two key optimization metrics, that are MFE and CAI, for RNA sequences on the dataset. Overall, MFE and CAI distributions were similar across subsets (**Fig. 2(d)**). Joint analysis of CAI and normalized MFE revealed a weak positive correlation ($r = 0.333$, $p < 0.01$; **Fig. 2(e)**), consistent with a trade-off between RNA structural stability (lower MFE) and translational efficiency (higher CAI), which poses a key challenge for RNA design algorithms. Additionally, a negative correlation between RNA length and MFE supports the use of normalized MFE (MFE divided by sequence length) as a length-agnostic stability metric (**Figs. 2(f) and 2(g)**).

Analysis of the MoE [38] gating network revealed its ability to activate specific experts based on input sequence features, enhancing the model's adaptability to diverse protein characteristics (**Additional file 1: Note 1, Additional file 1: Fig. S1**). For instance, Expert 4 specializes in processing serine-rich sequences or serine-associated patterns (**Fig. 2(h)**). Hyperparameter tuning demonstrated the robustness of the RNARL model to parameter variations (**Additional file 1: Table S1**). Ablation experiments showed that both the ESM-2 [39] encoder and the LADS strategy are essential for multi-objective optimization in RNARL (**Fig. 2(i)**). Removal of either component results in over-optimization for translation efficiency at the expense of structural stability, leading to an improved CAI but also an increased MFE and unpaired nucleotide proportion.

Cross-species RNA codon design capacity of the RNARL framework

To validate the universality of the RNARL framework across species, we trained six species-specific models based on the RNARL framework using the previously described datasets. The performance comparison between a unified model and six species-specific models is presented in **Additional file 1: Note 2** and **Additional file 1: Fig. S2**. RNARL-generated RNA sequences were benchmarked against five methods (GEMORNA [32], LinearDesign [15], CodonTransformer [5], CAI-optimized, and Random) and natural sequences from all six species (**Fig. 3**, **Additional file 1: Fig. S3**). Specifically, GEMORNA and CodonTransformer represent advanced approaches based on multi-species pre-trained large language models; LinearDesign represents a state-of-the-art heuristic search method; CAI-optimized represents a traditional codon optimization method; and Random serves as the baseline.

ARTICLE IN PRESS

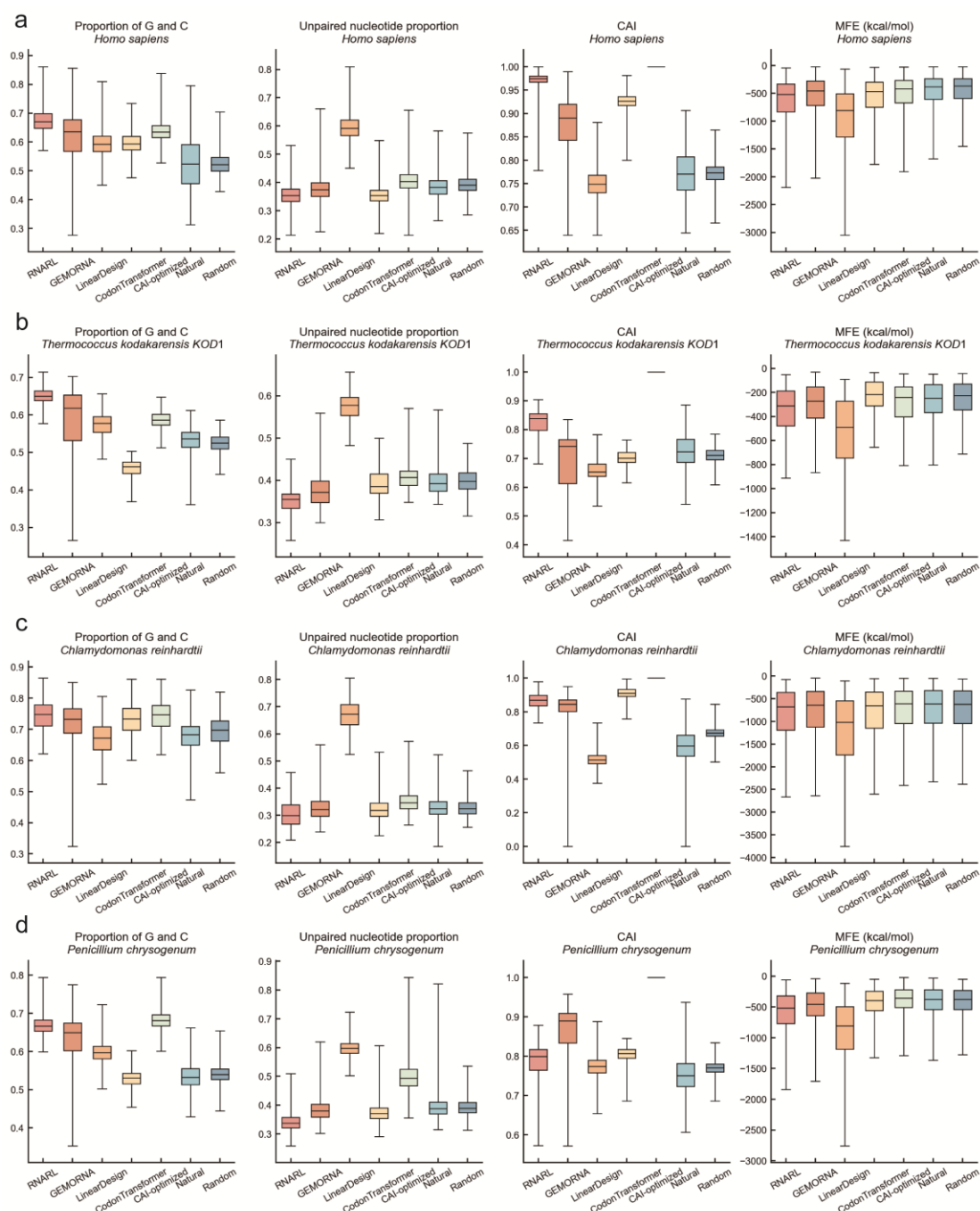


Fig. 3. Cross-species evaluation of codon design methods. The codon design methods compared include GEMORNA, RNARL, LinearDesign, Codon Transformer, CAI-optimized, Random sequences, and Natural (representing the wild-type RNA sequence). **a**, Results for *Homo sapiens*. **b**, Results for *Thermococcus kodakarensis KOD1*. **c**, Results for *Chlamydomonas reinhardtii*. **d**, Results for *Penicillium chrysogenum*.

RNARL outperformed other methods and natural sequences across multiple metrics,

including proportion of G and C, unpaired nucleotide proportion, and CAI metrics. For example, on *Homo sapiens*, RNARL-generated RNA sequences exhibited the proportion of G and C of 0.676 ± 0.040 , compared to 0.617 ± 0.090 for GEMORNA, 0.596 ± 0.045 for LinearDesign and 0.598 ± 0.036 for CodonTransformer. RNARL also achieved a CAI of 0.967 ± 0.027 , outperforming GEMORNA's CAI of 0.877 ± 0.057 , LinearDesign's CAI of 0.749 ± 0.030 and CodonTransformer's CAI of 0.926 ± 0.015 . Although the CAI-optimized method attained the maximum CAI scores (1.0) through optimal codon selection, it demonstrated inferior performance across other metrics. In contrast, RNARL maintained high CAI values while achieving a more balanced performance across all evaluated metrics. LinearDesign showed lowest (optimal) MFE values, while RNARL exhibited marginally higher (less optimal) MFE values than LinearDesign but still outperformed all other methods. The superior MFE of LinearDesign-generated sequences is expected, as LinearDesign is a heuristic search algorithm that can, in principle, find the lowest MFE sequence when its lambda hyperparameter is set to 0 and the search width is sufficiently large. In addition, a comparison of sequence similarity revealed that RNA sequences generated by both RNARL and CodonTransformer exhibit >70% similarity to their corresponding natural sequences (**Additional file 1: Note 3, Additional file 1: Fig. S4**). Overall, RNARL's reinforcement learning-driven multi-objective design achieved well-balanced performance across all metrics, demonstrating its effectiveness in generating high-quality RNA sequences.

To quantify the influence of protein properties on RNARL performance, we analyzed the relationship between the optimization gains, defined as normalized MFE Gain and CAI Gain, and the intrinsic physicochemical properties of the corresponding proteins. A strong positive correlation between normalized MFE Gain and CAI Gain (Spearman's $r = 0.64$, $p < 0.01$, **Additional file 1: Figs. S5(a) and S5(b)**) indicates that RNARL can simultaneously enhance RNA stability and translational efficiency, underscoring its capability for multi-objective optimization. Furthermore, "easy-to-optimize" proteins are more hydrophobic and exhibit lower instability indices, whereas "hard-to-optimize" proteins are less hydrophobic and more unstable

(**Additional file 1: Fig. S5(f)**). A detailed description of the methodology and extended analyses are provided in **Additional file 1: Note 4**.

T-SNE visualization of the protein sequence embeddings derived from the actor model revealed species-specific clustering, with analogous patterns observed in RNARL-generated RNA sequences (**Fig. 4(a) and 4(b)**). These results confirm that RNARL effectively learns host-specific codon preferences and preserves phylogenetically conserved sequence features. Codon pairs have been demonstrated to significantly influence protein expression levels by modulating translational kinetics [40, 41]. Quantitative analysis of undesirable codon pair frequencies demonstrated that RNARL and LinearDesign generated sequences with fewer inhibitory codon pairs compared to other methods, supporting their enhanced translation efficiency (**Fig. 4(c)**) [42]. While thermophilic organisms typically exhibit high proportion of G and C in genomic DNA to enhance thermostability, excessive GC pairs in RNA may form overly stable secondary structures that hinder ribosomal translocation [43-45]. RNARL selectively suppressed GC-rich codon pairs of *Thermococcus kodakarensis* *KOD1*. The proportion of GC-rich codon pairs in RNARL-generated RNA sequences was comparable to those in LinearDesign-generated sequences and 36.7% lower than that in natural sequences, potentially mitigating the formation of RNA secondary structures. This adaptive codon optimization strategy mirrors natural evolutionary pressures in hyperthermophiles, where reduced GC content in coding sequences (CDSs) minimizes thermodynamically stable RNA structures that could impede ribosome progression during translation at elevated temperatures [46, 47]. To validate heterologous expression compatibility, the RNARL model trained on *Homo sapiens* data was utilized to optimize RNA sequences encoding *E. coli* proteins. Remarkably, the CAI (calculated based on human codon usage preferences) of *E. coli* RNA sequences optimized by the *Homo sapiens*-trained RNARL model was approximately 0.182 higher than that of sequences optimized by the *E. coli*-trained RNARL model, while maintaining comparable structural stability metrics. This demonstrates RNARL's robust alignment with target host codon preferences during cross-species optimization (**Fig. 4(d)**). These results highlight the versatility of the RNARL framework in RNA codon optimization across diverse species.



Fig. 4. The capability of RNARL for cross-species RNA sequence representation and

design. **a.** t-SNE visualization of protein sequence embeddings extracted using the actor model. Each point represents a protein sequence, colored by its species origin, demonstrating clear species-specific clustering. **b.** t-SNE visualization of RNARL-generated RNA sequences using 4-mer frequency vectors. Each point represents an RNA sequence, colored by its species origin, demonstrating clear species-specific clustering. **c.** Comparison of undesirable codon pair frequencies in RNA sequences optimized by RNARL versus other methods across six species. **d.** Evaluation of RNA sequences optimized by the RNARL-*Homo sapiens* model for *E. coli* RNAs (cross-species optimization) and the RNARL-*E. coli* model for *E. coli* RNAs (intra-species optimization).

Cross-type RNA codon design capacity of the RNARL framework

While mRNA remains the predominant focus for vaccine development, other RNA types exhibit substantial therapeutic potential in gene therapy, immune modulation, and innovative vaccine design, owing to their distinct biological characteristics (e.g., circRNA) [48-50]. To systematically validate the generalizability of the RNARL framework across RNA types, we further evaluated its performance on four additional RNA types: circRNA, +ssRNA, dsRNA, and retroviral RNA. These RNA types differ significantly from linear mRNA in both structure and function, thereby posing unique challenges and requirements for optimization strategies. For instance, circRNA exhibits inherent resistance to nuclease degradation due to its closed-loop structure and absence of free termini, making it an ideal vector for sustained expression. +ssRNA and retroviral RNA, as viral genomes or intermediates, rely critically on their structural and sequence features for replication, gene expression, and packaging.

As shown in **Fig. 5(a)**, the results indicate that RNARL-generated sequences consistently outperformed other methods in circRNA optimization with respect to the proportion of G and C, unpaired nucleotide proportion, and CAI compared to other methods. For instance, RNARL-generated sequences exhibited the proportion of G and C of 0.661 ± 0.047 , compared to 0.588 ± 0.111 (GEMORNA), 0.579 ± 0.054 (LinearDesign) and 0.584 ± 0.040 (CodonTransformer), respectively. The optimized proportion of G and C and a favorable unpaired nucleotide proportion are vital for the overall stability and functional integrity of circRNA, as well as for enhanced resistance to enzymatic degradation [51-53]. Furthermore, across the three viral RNA

types (+ssRNA, dsRNA, and retroviral RNA), RNARL-generated sequences consistently demonstrated superior performance in both proportion of G and C and CAI compared to other methods (**Figs. 5(b) and 5(c), Additional file 1: Fig. S6**). For +ssRNA, the CAI of RNARL-generated sequences was approximately 0.104 higher than that of GEMORNA-generated sequences, 0.239 higher than that of LinearDesign-generated sequences and 0.043 higher than that of CodonTransformer-generated sequences. A higher CAI is essential for efficient viral replication and robust protein expression in viral vectors [54-56]. For +ssRNA, elevated GC content and reduced UA/UU dinucleotide content significantly mitigate the likelihood of degradation by RNase L. For instance, a comparative analysis of the wild-type RNA sequence encoding the AAT98583.1 protein of *Human coronavirus HKU1* and its RNARL-generated counterpart revealed that the wild-type sequence contained a substantially higher abundance of UA and UU dinucleotides, which are prone to RNase L cleavage, rendering it more susceptible to RNase L-mediated degradation (**Fig. 5(d)**).

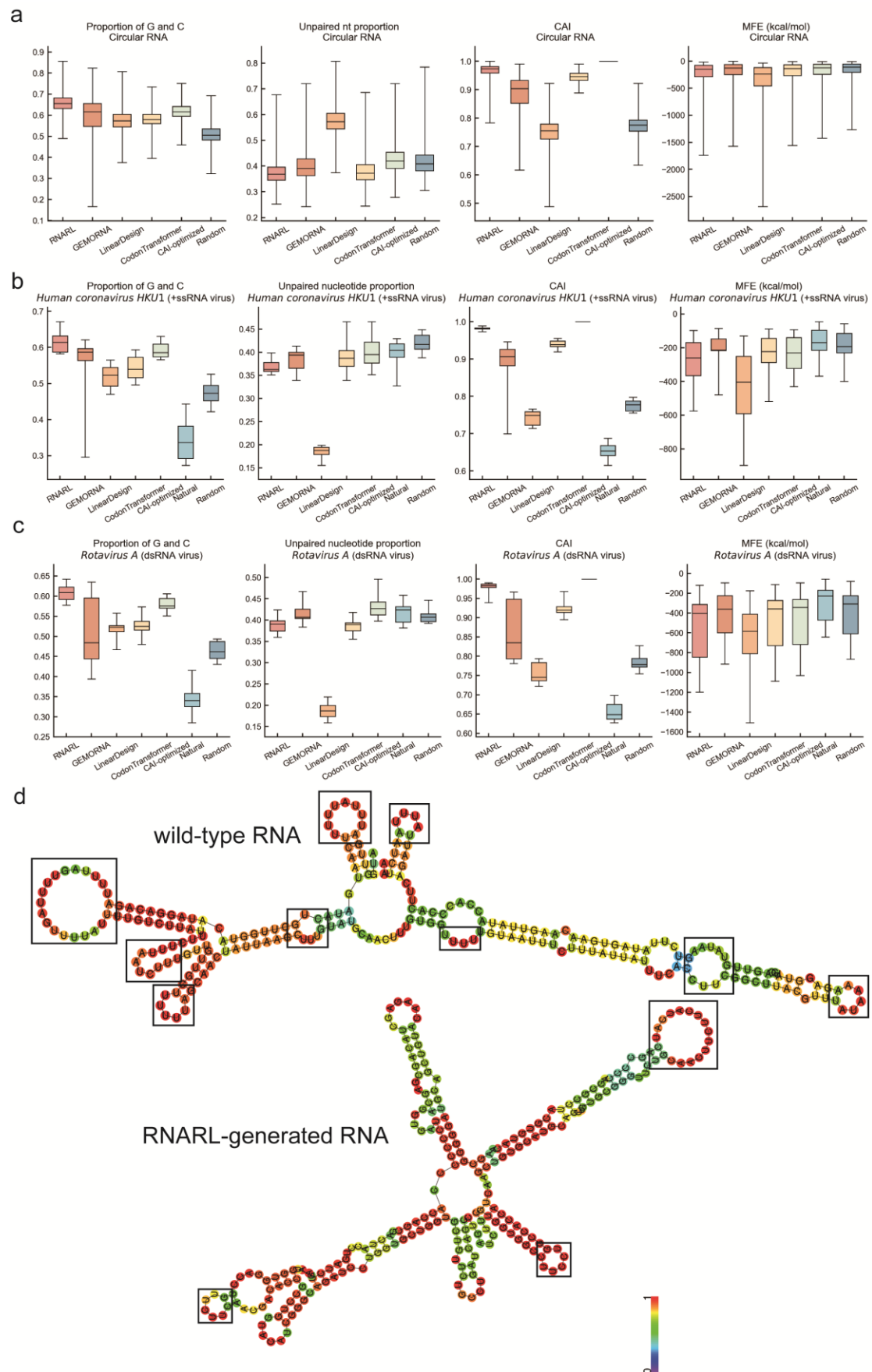


Fig. 5. Performance evaluation of codon design methods across diverse RNA types. **a**,

Results for circRNA. **b**, Results for +ssRNA. **c**, Results for dsRNA. **d**, Secondary structure comparison and RNase L susceptibility. Predicted secondary structure of the wild-type RNA sequence encoding *human coronavirus HKU1* protein AAT98583.1, compared to the corresponding sequence generated by RNARL. Regions indicated by black boxes are susceptible to RNase L-mediated degradation. The color of each nucleotide reflects its base-pairing probability, where red indicates a high probability of being paired and blue indicates a low probability.

In summary, these results underscore the RNARL framework's robust and generalizable optimization capacity across diverse RNA types. By consistently improving key metrics such as proportion of G and C, CAI, and structural stability, RNARL effectively addresses the distinct biological requirements of each RNA type, resulting in enhanced functional properties such as translational efficiency, structural stability, and nuclease resistance.

Application of the RNARL framework in RNA vaccine design

To demonstrate the utility of our RNARL framework in RNA vaccine design, we redesigned antigen-encoding RNA sequences derived from two clinically relevant RNA vaccine patents: (1) an mRNA vaccine targeting pathogenic gene regions of human papillomavirus (HPV) types 16 and 18, connected by a flexible peptide linker (Patent No.: CN116019904A); and (2) an mRNA vaccine encoding four critical glycoproteins (gp220, gp42, gL, gH) of the Epstein-Barr virus (EBV) (Patent No.: CN117529335A) (**Fig. 6(a)**).

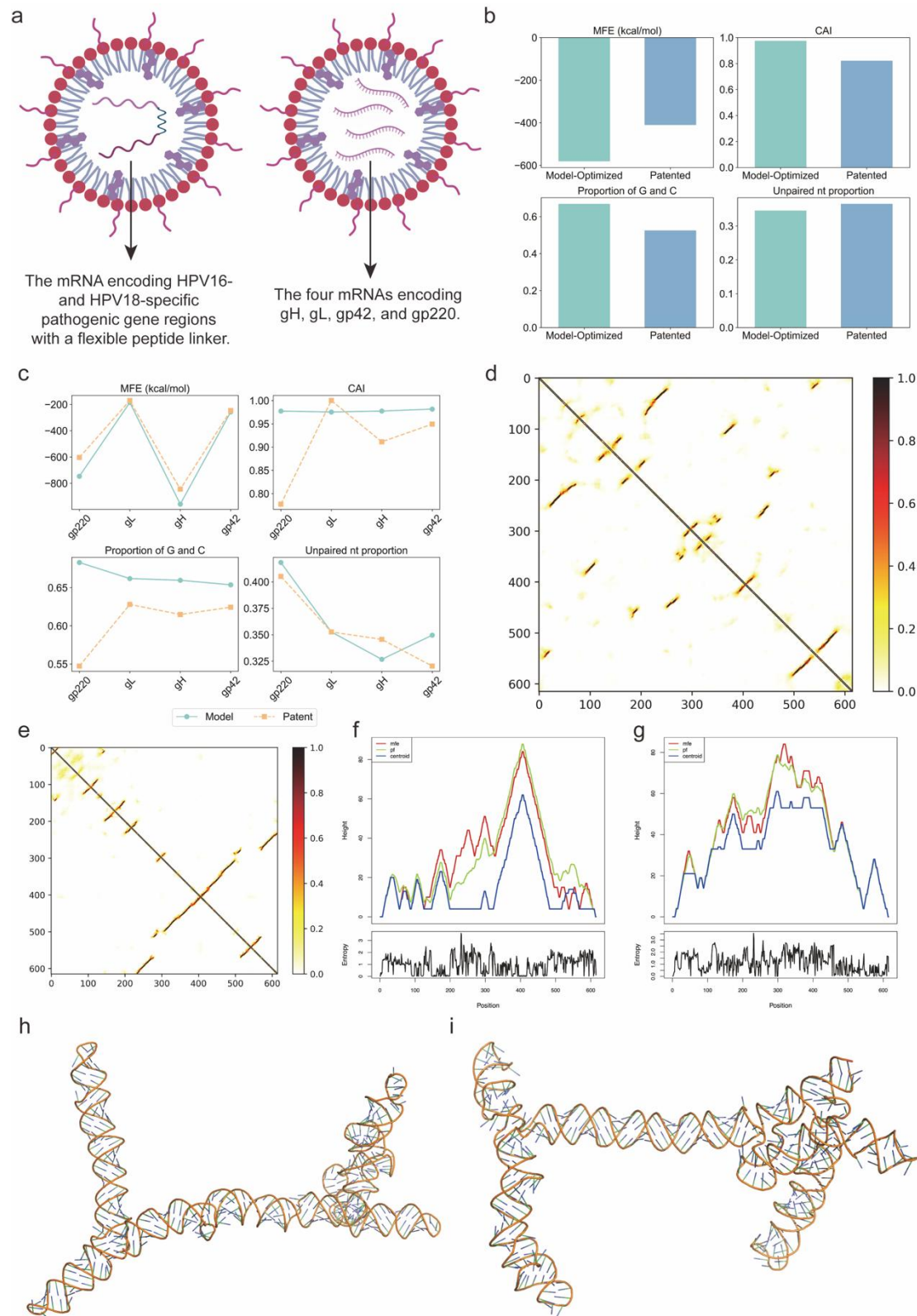


Fig. 6. Structural and biophysical evaluation of patented and RNARL-optimized mRNA vaccine designs. **a**, Schematic diagrams illustrating the two types of mRNA vaccines. The vaccine described in patent CN116019904A contains a single mRNA

encoding concatenated pathogenic gene regions of HPV types 16 and 18. The vaccine from patent CN117529335A contains four distinct mRNA sequences, each encoding one of four critical glycoproteins (gp220, gp42, gL, gH). **b, c**, Comparison of key RNA properties between RNARL-optimized and patented RNA sequences for the HPV vaccine (b) and the EBV vaccine (c). **d, e**, Structural contact maps of the RNA sequence encoding gp42 for the RNARL-optimized design (d) and the patented design (e). **f, g**, Comprehensive secondary structure analysis of the gp42-encoding RNA sequences for the patented (f) and RNARL-optimized design (g). Each panel contains two plots: (Top) Mountain plot representations of the secondary structure, where “mfe” (red) is the MFE structure, “centroid” (blue) is the most representative structure of the folding ensemble, and “pf” (green) represents the average structure derived from the partition function calculation. (Bottom) Positional entropy, where higher values indicate greater structural uncertainty. **h, i**, Tertiary structures of the RNA sequence encoding gL for the RNARL-optimized design (h) and the patented design (i).

Using the RNARL framework, we generated optimized mRNA sequences for the antigen proteins derived from these patents. We then computationally compared these optimized sequences against their patent-provided counterparts. For the HPV vaccine, RNARL-optimized sequences consistently outperformed the patent sequences across all four key computational metrics (**Fig. 6(b)**). The MFE of the RNARL-generated sequence was -580.8 kcal/mol, compared to -410.4 kcal/mol for the patent RNA sequence. The CAI of the RNARL-generated sequence was 0.669, versus 0.526 for the patent RNA sequence. Similarly, for the EBV vaccine, the RNARL-generated sequences for gL and gH demonstrated superior or comparable performance in MFE, CAI, and proportion of G and C when compared to the patent sequences (**Fig. 6(c)**).

Beyond computational metrics, we conducted further structural analyses of the RNARL-optimized sequences for gp42 and gL antigens from the EBV vaccine patent (limited by the length constraints of current structural analysis tools). For gp42-encoding mRNAs, structural contact maps showed a higher number of contact points in RNARL-optimized sequences, suggesting the formation of more compact and stable tertiary structures (**Figs. 6(d) and 6(e), Additional file 1: Note 5, Additional file 1: Fig. S7**). The secondary structure and detailed visualization of the

energy landscape further confirm these findings. While specific regions of the patent gp42-encoding mRNA sequence (e.g., 300–450 nt positions) exhibited significant structural energy fluctuations (**Fig. 6(f)**), RNARL-optimized sequences maintained significantly higher structural stability in these critical regions (**Fig. 6(g)**). These intrinsically unstable regions in the patent sequences could serve as vulnerable targets for nuclease attacks in vivo, leading to reduced mRNA expression levels and duration. A comparative analysis of tertiary structures of gL-encoding mRNAs revealed that the RNARL-optimized sequence (**Fig. 6(h)**) folds into a more compact and tightly packed conformation compared to the patent-provided sequence (**Fig. 6(i)**), particularly at the central junction of the optimized structure. In complex cellular environments, such enhanced structural stability enables RNA to better maintain its functional conformation, reduce nonspecific interactions, and potentially improve overall stability.

In summary, the comprehensive evaluations demonstrate that the RNARL framework effectively generates RNA codon sequences with superior biophysical properties and enhanced structural stability across multiple dimensions. These strongly validate the practical value and broad application potential of the RNARL framework in accelerating the rational design of high-quality RNA vaccine sequences.

Discussion

The RNARL framework presented in this study addresses a fundamental limitation in RNA codon design: the decoupled “generate-or-optimize” paradigm. Existing methods often excel at either high-throughput sequence generation or computationally intensive optimization, but struggle to achieve both goals simultaneously. RNARL bridges this gap by deeply unifying these two stages through reinforcement learning, enabling high-throughput generation of high-performance RNA sequences.

Furthermore, it is worth emphasizing that RNARL is designed as a universal computational framework rather than a pretrained large model. The framework demonstrates high flexibility and modularity, and researchers can readily replace or adjust components such as the actor model, reward model, or reinforcement learning algorithms according to specific experimental or design requirements.

RNARL addresses at least three critical challenges in RNA codon design. First, by training on long RNA sequence datasets and incorporating LADS, RNARL effectively handles RNA sequences up to ~3,900 nts in length. Second, computational validation across six distinct species and five RNA types demonstrates RNARL's broad applicability, which is crucial for developing RNA therapies tailored to diverse host environments. Third, unlike single-objective methods, RNARL leverages reinforcement learning to more effectively jointly optimize multiple objectives, such as translation efficiency and structural stability, directly generating RNA sequences with enhanced overall performance.

Despite its strengths, the current RNARL framework has certain limitations. Its primary focus remains on CDS optimization, while non-coding regions of RNA molecules (e.g., 5'/3' UTRs of mRNA, IRES elements) play equally critical roles in RNA stability, translation efficiency, and subcellular localization. Given the high diversity of sequence and structural features in non-coding regions across RNA types, their unified incorporation into the design model presents complex challenges. Future work could explore the development of modular optimization components that design specific optimization strategies tailored to specific non-coding regions, which could then be integrated into the RNARL framework to enable global optimization of full-length RNA sequences.

Conclusions

This study introduces and validates RNARL, a reinforcement-learning-driven framework for unified RNA sequence generation and multi-objective optimization. The results demonstrate that RNARL offers a practical and generalizable solution for designing RNA codon sequences with improved biophysical properties, such as enhanced translational efficiency and structural stability. The framework effectively handles long sequences across diverse species and RNA types.

By efficiently exploring the vast RNA sequence space, the RNARL framework holds the potential to accelerate the rational design of high-quality RNA vaccines and nucleic acid drugs. Its demonstrated ability to improve upon patented vaccine sequences highlights its translational relevance. Finally, the deployment of RNARL via a user-friendly web platform provides the community with an accessible tool to

support future advances in the development of RNA-based medicine.

Methods

Dataset construction and preprocessing

To validate the generality and effectiveness of the RNARL framework across species and RNA types, we constructed datasets encompassing multi-species and multi-type RNA sequences.

Across-species mRNA dataset

We downloaded and extracted the protein sequences and corresponding CDSs of six representative organisms from the National Center for Biotechnology Information (NCBI) database [57]: *Chlamydomonas reinhardtii* (protist), *Thermococcus kodakarensis* KOD1 (archaeon), *Homo sapiens* (animal), *Arabidopsis thaliana* (plant), *Escherichia coli* (bacterium), and *Penicillium chrysogenum* (fungus). To build a high-quality, non-redundant dataset, we clustered protein sequences using the Linclust algorithm [58]. Parameters were set to a minimum sequence identity threshold of 90% and a coverage threshold of 80%, ensuring that 80% of the length of one sequence aligned with another at 90% identity. One representative protein sequence and its corresponding RNA sequence were selected from each cluster to minimize bias from sequence homology. Sequences were filtered to retain proteins with lengths between 10 and 1,305 amino acids (corresponding to RNA lengths of 30–3,915 nts), covering most potential antigen length ranges. Each species dataset was randomly partitioned into training set (90%), validation set (5%), and test set (5%). Detailed sequence counts for each species are provided in **Additional file 1: Table S2**. Codon usage bias information for each species was referenced from the Kazusa Codon Usage Database (<https://www.kazusa.or.jp/codon/>). The impact of sequence clustering identity thresholds on model performance is presented in **Additional file 1: Note 6**, **Additional file 1: Fig. S8** and **Additional file 1: Table S3**.

Circular RNA dataset

Human circRNAs and their corresponding protein sequences were downloaded from the circRNADb database [59]. Data processing followed a similar pipeline to the mRNA dataset, including clustering-based redundancy reduction, length filtering and

dataset partitioning.

Viral RNA dataset

Protein sequences and their corresponding CDS sequences for three representative human viruses were downloaded from the NCBI database: *Human coronavirus HKU1* (+ssRNA virus), *Rotavirus A* (dsRNA virus), and *Human immunodeficiency virus 2* (retroviral RNA). Due to the limited number of viral proteins and CDSs, all collected data were used as a test set to evaluate RNARL's optimization capabilities for viral RNA types. As these viruses infect human hosts, the RNARL model trained on human mRNA data was employed for the optimization evaluations.

RNA vaccine patent dataset

We retrieved two publicly available RNA vaccine patents (CN116019904A and CN117529335A) via Google Patents. Antigen protein sequences and their corresponding optimized mRNA sequences described in the patents were extracted for RNARL practical validation.

The RNARL framework

The RNARL framework is an RNA codon design model based on generative models and reinforcement learning, designed to generate RNA sequences that translate into target proteins while exhibiting superior performance. The core of the RNARL framework is a reinforcement learning loop composed of an actor model (generator) and a reward model (evaluator) (**Fig. 1**).

The actor model generates candidate RNA sequences for a given protein sequence. Its architecture combines a Transformer Encoder with a MoE network. The protein sequences are first encoded into the embeddings using the pretrained protein language model ESM2_650M. These embeddings are then fed into the Transformer Encoder-MoE module, where they undergo multi-layer processing to generate nucleotide probability distributions at each position. Finally, the model generates RNA sequences via sampling.

The reward model evaluates the quality of generated RNA sequences. It takes RNA sequences as input, first encoding each nucleotide into a vector via a learned

embedding layer. These vector sequences are then processed through a Transformer Encoder-MoE architecture similar to the actor model. Subsequently, a fully connected layer in the reward model generates quality scores, which serve as the reward signal during reinforcement learning training. Each RNA sequence is scored as follows:

$$NegStandMFE = -\frac{MFE}{len(CDS)} \quad (1)$$

$$ScaledMFE = \frac{(NegStandMFE - \min(NegStandMFE))}{(\max(NegStandMFE) - \min(NegStandMFE))} \quad (2)$$

$$ScaledCAI = \begin{cases} 1.0 & \text{if } CAI \geq TargetCAI \\ \frac{CAI}{TargetCAI} & \text{if } CAI < TargetCAI \end{cases} \quad (3)$$

$$CombinedReward = WMFE \times ScaledMFE + WCAI \times ScaledCAI \quad (4)$$

Here, $len(CDS)$ denotes the length of CDS, MFE and CAI are the MFE and CAI values of CDS. $WMFE$, $WCAI$ and $TargetCAI$ are Hyperparameters. For the computational experiments presented in this study, we configured the following parameters: $WMFE=2.0$, $WCAI=1.0$, $TargetCAI=0.8$. The detailed selection process for these hyperparameters is presented in **Additional file 1: Note 7** and **Additional file 1: Fig. S9**.

GRPO reinforcement learning algorithm

We employed the GRPO reinforcement learning algorithm to train the actor model to generate RNA sequences with high reward values based on input protein sequences. The GRPO algorithm achieves optimization of the actor network through policy gradient methods. GRPO incorporates the core principles of Proximal Policy Optimization (PPO), aiming to maximize rewards while maintaining policy stability by constraining the step size of policy updates. GRPO also introduces a KL divergence regularization term against a fixed reference policy ($\pi_{\theta_{ref}}$), as well as a mechanism for estimating the advantage function based on grouped sampling.

Additional file 1: Note 8 and **Additional file 1: Fig. S10** present the impact of

GRPO reinforcement learning on the actor model.

GRPO algorithm objective function

The core of the GRPO algorithm lies in maximizing an objective function that combines reward signals, policy stability constraints, and reference policy regularization. For a given state s and action a , the optimization objective is defined as:

$$\mathcal{L}^{GRPO}(\theta) = \max_{\theta} \hat{\mathbb{E}}_{(s,a) \sim \pi_{\theta_{old}}} [\min(r(s,a)\hat{A}(s,a), \text{clip}(r(s,a), 1 - \epsilon, 1 + \epsilon)\hat{A}(s,a)) - \beta \log r_{ref}(s,a)] \quad (5)$$

, where $\hat{\mathbb{E}}_{(s,a) \sim \pi_{\theta_{old}}}$ denotes the empirical expectation computed over data sampled under the old policy $\pi_{\theta_{old}}$. $r(s,a) = \frac{\pi_{\theta}(a|s)}{\pi_{\theta_{old}}(a|s)}$ is the policy ratio between the current policy π_{θ} and the old policy $\pi_{\theta_{old}}$. $\hat{A}(s,a)$ is the estimated advantage function for action a in state s . $\text{clip}(r(s,a), 1 - \epsilon, 1 + \epsilon)$ truncates the policy ratio within the range $[1 - \epsilon, 1 + \epsilon]$, where ϵ is the clipping coefficient (set to 0.2 in experiments). This term, borrowed from PPO's design, restricts the magnitude of policy updates to prevent drastic policy changes. $\log r_{ref}(s,a) = \log \pi_{\theta}(a|s) - \log \pi_{ref}(a|s)$ represents the log policy ratio between the current policy π_{θ} and a fixed reference policy π_{ref} . This term is equivalent to a sample-based approximation of the KL divergence $D_{KL}(\pi_{\theta} || \pi_{ref})$. β (set to 0.01 in experiments) is the coefficient for the KL divergence regularization term, encouraging the current policy to remain close to the reference policy. The detailed selection process for these hyperparameters (ϵ, β) is presented in **Additional file 1: Note 9** and **Additional file 1: Fig. S11**.

GRPO training procedure

In each training iteration, the GRPO algorithm executes the following steps to maximize the aforementioned objective function:

Data sampling. For a batch of N input protein sequences $\{s_i\}_{i=1}^N$, sample G candidate RNA sequences $\{a_{i,j}\}_{j=1}^G$ for each protein sequence s_i using the current

old policy $\pi_{\theta_{old}}$, where G is the group size. This results in a dataset of $N \times G$ (protein, RNA) pairs $\{(s_i, a_{i,j})\}_{i=1\dots N, j=1\dots G}$.

Reward evaluation. Compute the reward value $R(s_i, a_{i,j})$ for each sampled sequence $(s_i, a_{i,j})$.

Advantage function estimation. For each protein sequence s_i , compute the advantage function for its G sampled sequences by normalizing rewards within the group: Calculate the group mean $\mu_i = \frac{1}{G} \sum_{j=1}^G R(s_i, a_{i,j})$ and standard deviation $\sigma_i = \sqrt{\frac{1}{G} \sum_{j=1}^G (R(s_i, a_{i,j}) - \mu_i)^2}$. Compute the normalized advantage function: $\hat{A}(s_i, a_{i,j}) = \frac{R(s_i, a_{i,j}) - \mu_i}{\sigma_i}$.

Objective function calculation and policy update. For each sampled pair $(s_i, a_{i,j})$, compute the probability ratios: $r_{i,j} = \frac{\pi_{\theta}(a_{i,j}|s_i)}{\pi_{\theta_{old}}(a_{i,j}|s_i)}$. Compute the KL divergence approximation between the current policy and the fixed reference policy: $\log r_{ref,i,j} = \log \pi_{\theta}(a_{i,j}|s_i) - \log \pi_{\theta_{ref}}(a_{i,j}|s_i)$. Use the sampled data $\{(s_i, a_{i,j})\}$, ratios $r_{i,j}$, KL terms $\log r_{ref,i,j}$, and advantages $\hat{A}(s_i, a_{i,j})$ to compute the empirical objective function (a Monte Carlo approximation of the expectation in $\mathcal{L}^{GRPO}(\theta)$). Compute gradients of the objective function with respect to the actor model parameters θ . In practice, maximizing the objective is equivalent to minimizing its negative as a loss function. Update parameters θ via one step of gradient descent using an optimizer (AdamW).

Old policy update. At fixed intervals (every 50 gradient steps in experiments), copy the current actor model parameters θ to the old policy parameters θ_{old} ($\theta_{old} \leftarrow \theta$) for subsequent sampling. The fixed reference policy parameters θ_{ref} remain unchanged throughout training.

Length-aware distribution sampling technique

The training dataset exhibited a pronounced imbalance in protein sequence and RNA

sequence length distribution. This imbalance inherently biased the model toward shorter sequences and compromised its performance on longer RNA sequences. To mitigate these issues, we implemented the LADS techniques. Protein sequences were binned into non-overlapping intervals of 100 amino acids (e.g., 1–100, 101–200, etc.), and adaptive sampling weights were assigned to each interval to equalize their contributions during training. Specifically, for each length bin, the sampling weight was calibrated inversely proportional to its occupancy, ensuring that the product of the weight and the number of proteins in each bin remained approximately constant across all intervals. This approach dynamically prioritizes underrepresented length categories during data sampling, thereby mitigating bias while preserving the diversity of sequences.

Training details

The actor and reward models were configured with the following hyperparameters: the Transformer Encoder utilized a hidden dimension of 768 and 8 layers, while the MoE component incorporated 6 experts with 2 experts activated per input token. Optimization was performed using the AdamW optimizer with a learning rate of 3.0×10^{-5} , and the models were trained for 6 epochs to ensure convergence. Comprehensive hyperparameter tuning procedures are documented in the **Additional file 1: Note 10** and **Additional file 1: Table S1**.

Evaluation metrics and tools

We evaluated generated RNA sequences using the following metrics and bioinformatics tools:

MFE: Computed using LinearFold Package [60]. Lower MFE indicates higher structural stability. Normalized MFE (MFE / Sequence Length) is used in training to eliminate length bias.

CAI: Measures codon usage bias for species-specific translation efficiency. Codon usage bias information for each species was referenced from the Kazusa Codon Usage Database (<https://www.kazusa.or.jp/codon/>).

Proportion of G and C: Proportion of guanine (G) and cytosine (C) nucleotides.

Unpaired nucleotide proportion: Fraction of nucleotides in single-stranded regions.

Secondary structure prediction and analysis were performed using RNAfold [61].

This included the generation of the MFE structure, the positional entropy for each nucleotide position and the mountain plot representation of the MFE structure, the thermodynamic ensemble of RNA structures, and the centroid structure.

Tertiary structure prediction and analysis were performed using trRosettaRNA [62] to assess RNA compactness and stability.

Implementation details of competing methods

The performance of the proposed RNARL was evaluated against five competing methods: GEMORNA [32], LinearDesign [15], CodonTransformer [5], a CAI-optimized method and a random method. For a fair and reproducible comparison, these five models were benchmarked by using their publicly available implementations and default configurations. The specific setup for each method is as follows:

GEMORNA: The evaluation was performed using the official implementation and pre-trained model weights from its GitHub repository (<https://github.com/RainaBio/GEMORNA.git>), with the provided inference scripts used without modification.

LinearDesign: The program was executed with its default parameter settings, using the software downloaded from its official repository (<https://github.com/LinearDesignSoftware/LinearDesign.git>).

CodonTransformer: Performance testing utilized the official code and pre-trained weights available on GitHub (<https://github.com/Adibvafa/CodonTransformer.git>).

CAI-optimized Method: This method involves the deterministic selection of the most frequent synonymous codon for each amino acid.

Random Method: This approach selects a synonymous codon for each amino acid through weighted random sampling, where weights are proportional to the natural codon usage frequencies of the target species.

Web interface

To enhance accessibility and facilitate the use of the RNARL framework, a user-friendly web interface was developed and hosted on Google Colab (https://colab.research.google.com/drive/1Z_t_Lt9CjqA0aygoNcdRzrThF1MOGBRb). This interface is designed to require no environment setup or programming knowledge, enabling users to readily input protein sequences and obtain optimized RNA sequences. The platform supports two modes of operation: the optimization of RNA sequences for a single protein by directly inputting its protein sequence, and the high-throughput optimization for multiple proteins simultaneously via the upload of CSV files. Additionally, users can select different species or RNA type model to perform targeted optimization based on their specific needs.

Availability of data and materials

The multi-species mRNA and multi-type RNA datasets are available on Zenodo (<https://zenodo.org/records/15486992>) [63] under the CC-BY 4.0 License. The source code is publicly accessible on GitHub (<https://github.com/xlab-BioAI/RNARL>) [64] and archived on Zenodo (<https://zenodo.org/records/17212245>) [65] under the MIT License. An interactive web application hosted on Google Colab is available at https://colab.research.google.com/drive/1Z_t_Lt9CjqA0aygoNcdRzrThF1MOGBRb [66] for efficient codon optimization of target RNA sequences.

Acknowledgments

We thank the editor and the two anonymous reviewers for their inspiring and constructive comments and suggestions to greatly improve the quality and presentation of the work.

Peer review information

Albert Cheng and Tim Sands were the primary editors of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team. The peer-review history is available in the online version of this article.

Funding

This work was supported by National Key Research and Development Program of China (Grant Nos. 2023YFC2506400, 2023YFC2506402 and 2024YFA1306902);

National Natural Science Foundation of China (Grant No. 32570779) and Shanghai Municipal Education Commission (Grant Nos. 2024AIZD008).

Ethics declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Authors' contributions

S.L. conceptualized the project, developed the model, performed the validation, and wrote the original draft. H.T., K.W., and R.W. participated in the literature review and model discussions. H.W. handled manuscript revision and acquired funding. T.Z. handled project administration and manuscript revision. Y.X. conceived and supervised the project, acquired funding, handled project administration, and revised the manuscript. All authors read and approved the final manuscript.

Supplementary information

Additional file 1: Supplementary materials. Supplementary Notes 1–10, Figs. S1–S11, and Tables S1–S3. Supplementary Notes include additional results.

References

1. Pardi N, Krammer F: **mRNA vaccines for infectious diseases — advances, challenges and opportunities.** *Nature Reviews Drug Discovery* 2024, **23**:838-861.
2. Zhang G, Tang T, Chen Y, Huang X, Liang T: **mRNA vaccines in disease prevention and treatment.** *Signal Transduction and Targeted Therapy* 2023, **8**:365.

3. Chaudhary N, Weissman D, Whitehead KA: **mRNA vaccines for infectious diseases: principles, delivery and clinical translation.** *Nature Reviews Drug Discovery* 2021, **20**:817-838.
4. Pardi N, Hogan MJ, Porter FW, Weissman D: **mRNA vaccines—a new era in vaccinology.** *Nature Reviews Drug Discovery* 2018, **17**:261-279.
5. Fallahpour A, Gureghian V, Filion GJ, Lindner AB, Pandi A: **CodonTransformer: a multispecies codon optimizer using context-aware neural networks.** *Nature Communications* 2025, **16**:3205.
6. Leppek K, Byeon GW, Kladwang W, Wayment-Steele HK, Kerr CH, Xu AF, Kim DS, Topkar VV, Choe C, Rothschild D, et al: **Combinatorial optimization of mRNA structure, stability, and translation for RNA-based therapeutics.** *Nature Communications* 2022, **13**:1536.
7. Kim Y-A, Mousavi K, Yazdi A, Zwierzyna M, Cardinali M, Fox D, Peel T, Collier J, Aggarwal K, Maruggi G: **Computational design of mRNA vaccines.** *Vaccine* 2024, **42**:1831-1840.
8. Mo O, Zhang Z, Cheng X, Zhu L, Zhang K, Zhang N, Li J, Li H, Fan S, Li X, Hao P: **mRNA designer: an integrated web server for optimizing mRNA design and protein translation in eukaryotes.** *Nucleic Acids Research* 2025:gkaf410.
9. Mauger DM, Cabral BJ, Presnyak V, Su SV, Reid DW, Goodman B, Link K, Khatwani N, Reynders J, Moore MJ: **mRNA structure regulates protein expression through changes in functional half-life.** *Proceedings of the National Academy of Sciences* 2019, **116**:24075-24083.

10. Ren Z, Jiang L, Di Y, Zhang D, Gong J, Gong J, Jiang Q, Fu Z, Sun P, Zhou B, Ni M:
CodonBERT: a BERT-based architecture tailored for codon optimization using the cross-attention mechanism. *Bioinformatics* 2024, **40**:btae330.
11. Qin S, Tang X, Chen Y, Chen K, Fan N, Xiao W, Zheng Q, Li G, Teng Y, Wu M:
mRNA-based therapeutics: powerful and versatile tools to combat diseases. *Signal Transduction and Targeted Therapy* 2022, **7**:166.
12. Rohner E, Yang R, Foo KS, Goedel A, Chien KR: **Unlocking the promise of mRNA therapeutics.** *Nature Biotechnology* 2022, **40**:1586-1600.
13. Metkar M, Pepin CS, Moore MJ: **Tailor made: the art of therapeutic mRNA design.** *Nature Reviews Drug Discovery* 2024, **23**:67-83.
14. To KK, Cho WC: **An overview of rational design of mRNA-based therapeutics and vaccines.** *Expert opinion on drug discovery* 2021, **16**:1307-1317.
15. Zhang H, Zhang L, Lin A, Xu C, Li Z, Liu K, Liu B, Ma X, Zhao F, Jiang H, et al:
Algorithm for optimized mRNA design improves stability and immunogenicity. *Nature* 2023, **621**:396-403.
16. Zeng C, Hou X, Yan J, Zhang C, Li W, Zhao W, Du S, Dong Y: **Leveraging mRNA sequences and nanoparticles to deliver SARS-CoV-2 antigens in vivo.** *Advanced Materials* 2020, **32**:2004452.
17. Sparmann A, Vogel J: **RNA-based medicine: from molecular mechanisms to therapy.** *The EMBO Journal* 2023, **42**:e114760.
18. Gamble CE, Brule CE, Dean KM, Fields S, Grayhack EJ: **Adjacent codons act in concert to modulate translation efficiency in yeast.** *Cell* 2016, **166**:679-690.

19. Fu H, Liang Y, Zhong X, Pan Z, Huang L, Zhang H, Xu Y, Zhou W, Liu Z: **Codon optimization with deep learning to enhance protein expression.** *Scientific Reports* 2020, **10**:17617.
20. Bae H, Collier J: **Codon optimality-mediated mRNA degradation: Linking translational elongation to mRNA stability.** *Molecular Cell* 2022, **82**:1467-1476.
21. Castillo-Hair SM, Seelig G: **Machine Learning for Designing Next-Generation mRNA Therapeutics.** *Accounts of Chemical Research* 2022, **55**:24-34.
22. Pathak S, Lin G: **ppLM-CO: Pre-trained protein language model for codon optimization.** *bioRxiv* 2024:2024. 2012. 2012. 628267.
23. Nikolados E-M, Oyarzún DA: **Deep learning for optimization of protein expression.** *Current Opinion in Biotechnology* 2023, **81**:102941.
24. Jain R, Jain A, Mauro E, LeShane K, Densmore D: **ICOR: improving codon optimization with recurrent neural networks.** *BMC Bioinformatics* 2023, **24**:132.
25. Akiyama M, Sakakibara Y: **Informative RNA base embedding for RNA structural alignment and clustering by deep representation learning.** *NAR Genomics and Bioinformatics* 2022, **4**:lqac012.
26. Li Y, Wang F, Yang J, Han Z, Chen L, Jiang W, Zhou H, Li T, Tang Z, Deng J, et al: **Deep generative optimization of mRNA codon sequences for enhanced mRNA translation and therapeutic efficacy.** *Nature Communications* 2025, **16**:9957.
27. Gong H, Wen J, Luo R, Feng Y, Guo J, Fu H, Zhou X: **Integrated mRNA sequence optimization using deep learning.** *Briefings in Bioinformatics* 2023, **24**(1):bbad001.
28. Chen J, Hu Z, Sun S, Tan Q, Wang Y, Yu Q, Zong L, Hong L, Xiao J, Shen T:

- Interpretable RNA foundation model from unannotated data for highly accurate RNA structure and function predictions.** *arXiv preprint arXiv:220400300* 2022.
29. Zhang Y, Lang M, Jiang J, Gao Z, Xu F, Litfin T, Chen K, Singh J, Huang X, Song G: **Multiple sequence alignment-based RNA language model and its application to structural inference.** *Nucleic Acids Research* 2024, **52**:e3-e3.
 30. Wang N, Bian J, Li Y, Li X, Mumtaz S, Kong L, Xiong H: **Multi-purpose RNA language modelling with motif-aware pretraining and type-guided fine-tuning.** *Nature Machine Intelligence* 2024, **6**:548-557.
 31. Chen K, Zhou Y, Ding M, Wang Y, Ren Z, Yang Y: **Self-supervised learning on millions of pre-mRNA sequences improves sequence-based RNA splicing prediction.** *BioRxiv* 2023:2023.2001. 2031.526427.
 32. Zhang H, Liu H, Xu Y, Huang H, Liu Y, Wang J, Qin Y, Wang H, Ma L, Xun Z, et al: **Deep generative models design mRNA sequences with enhanced translational capacity and stability.** *Science* 2025, **390**:eadr8470.
 33. Yu H, Yang H, Sun W, Yan Z, Yang X, Zhang H, Ding Y, Li K: **An interpretable RNA foundation model for exploring functional RNA motifs in plants.** *Nature Machine Intelligence* 2024, **6**:1616-1625.
 34. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I: **Attention is all you need.** *Advances in Neural Information Processing Systems* 2017, **30**.
 35. Guo D, Yang D, Zhang H, Song J, Zhang R, Xu R, Zhu Q, Ma S, Wang P, Bi X: **Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning.**

arXiv preprint arXiv:250112948 2025.

36. Sharp PM, Li W-H: **The codon adaptation index-a measure of directional synonymous codon usage bias, and its potential applications.** *Nucleic Acids Research* 1987, **15**:1281-1295.
37. Courel M, Clément Y, Bossevain C, Foretek D, Vidal Cruchez O, Yi Z, Bénard M, Benassy M-N, Kress M, Vindry C: **GC content shapes mRNA storage and decay in human cells.** *Elife* 2019, **8**:e49708.
38. Cai W, Jiang J, Wang F, Tang J, Kim S, Huang J: **A survey on mixture of experts.** *arXiv preprint arXiv:240706204* 2024.
39. Lin Z, Akin H, Rao R, Hie B, Zhu Z, Lu W, Smetanin N, Verkuil R, Kabeli O, Shmueli Y, et al: **Evolutionary-scale prediction of atomic-level protein structure with a language model.** *Science* 2023, **379**:1123-1130.
40. Alexaki A, Kames J, Holcomb DD, Athey J, Santana-Quintero LV, Lam PVN, Hamasaki-Katagiri N, Osipova E, Simonyan V, Bar H: **Codon and codon-pair usage tables (CoCoPUTs): facilitating genetic variation analyses and recombinant gene design.** *Journal of Molecular Biology* 2019, **431**:2434-2441.
41. Kames J, Alexaki A, Holcomb DD, Santana-Quintero LV, Athey JC, Hamasaki-Katagiri N, Katneni U, Golikov A, Ibla JC, Bar H: **TissueCoCoPUTs: novel human tissue-specific codon and codon-pair usage tables based on differential tissue gene expression.** *Journal of Molecular Biology* 2020, **432**:3369-3378.
42. Meyer D, Kames J, Bar H, Komar AA, Alexaki A, Ibla J, Hunt RC, Santana-Quintero LV, Golikov A, DiCuccio M: **Distinct signatures of codon and codon pair usage in 32**

- primary tumor types in the novel database CancerCoCoPUTs for cancer-specific codon usage. *Genome Medicine* 2021, **13**:122.
43. Nakashima H, Fukuchi S, Nishikawa K: **Compositional changes in RNA, DNA and proteins for bacterial adaptation to higher and lower temperatures.** *Journal of Biochemistry* 2003, **133**:507-513.
44. Trivedi S, Rao SR, Gehlot HS: **Nucleic acid stability in thermophilic prokaryotes: a review.** *J Cell Mol Biol* 2005, **4**:61-69.
45. Grosjean H, Oshima T: **How nucleic acids cope with high temperature.** *Physiology and Biochemistry of Extremophiles* 2007:39-56.
46. De Tarafder A: **Adaptive Evolution of the Bacterial Translation Machinery.** Acta Universitatis Upsaliensis, 2024.
47. Forterre P: **Strategies of hyperthermophiles in nucleic acids adaptation to high temperature.** *Extremophiles: Life Under Extreme Conditions* 2009, **1**:189-208.
48. Niu D, Wu Y, Lian J: **Circular RNA vaccine in disease prevention and treatment.** *Signal Transduction and Targeted Therapy* 2023, **8**:341.
49. Chen R, Wang SK, Belk JA, Amaya L, Li Z, Cardenas A, Abe BT, Chen C-K, Wender PA, Chang HY: **Engineering circular RNA for enhanced protein production.** *Nature Biotechnology* 2023, **41**:262-272.
50. Wesselhoeft RA, Kowalski PS, Anderson DG: **Engineering circular RNA for potent and stable translation in eukaryotic cells.** *Nature Communications* 2018, **9**:2629.
51. Geng X, Jia Y, Zhang Y, Shi L, Li Q, Zang A, Wang H: **Circular RNA: Biogenesis, degradation, functions and potential roles in mediating resistance to anticarcinogens.**

- Epigenomics* 2020, **12**:267-283.
52. Wang C, Liu H: **Factors influencing degradation kinetics of mRNAs and half-lives of microRNAs, circRNAs, lncRNAs in blood in vitro using quantitative PCR.** *Scientific Reports* 2022, **12**:7259.
 53. Ren L, Jiang Q, Mo L, Tan L, Dong Q, Meng L, Yang N, Li G: **Mechanisms of circular RNA degradation.** *Communications Biology* 2022, **5**:1355.
 54. Herrera-Carrillo E, Liu YP, Berkhout B: **Improving miRNA delivery by optimizing miRNA expression cassettes in diverse virus vectors.** *Human Gene Therapy Methods* 2017, **28**:177-190.
 55. Lico C, Chen Q, Santi L: **Viral vectors for production of recombinant proteins in plants.** *Journal of Cellular Physiology* 2008, **216**:366-377.
 56. Spunde K, Korotkaja K, Zajakina A: **Recombinant viral vectors for therapeutic programming of tumour microenvironment: Advantages and limitations.** *Biomedicines* 2022, **10**:2142.
 57. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D: **Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation.** *Nucleic Acids Research* 2016, **44**:D733-D745.
 58. Steinegger M, Söding J: **Clustering huge protein sequence sets in linear time.** *Nature Communications* 2018, **9**:2542.
 59. Chen X, Han P, Zhou T, Guo X, Song X, Li Y: **circRNADb: A comprehensive database for human circular RNAs with protein-coding annotations.** *Scientific Reports* 2016,

6:34985.

60. Huang L, Zhang H, Deng D, Zhao K, Liu K, Hendrix DA, Mathews DH: **LinearFold: linear-time approximate RNA folding by 5'-to-3'dynamic programming and beam search.** *Bioinformatics* 2019, **35**:i295-i304.
61. Lorenz R, Bernhart SH, Höner zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL: **ViennaRNA Package 2.0.** *Algorithms for Molecular Biology* 2011, **6**:1-14.
62. Wang W, Feng C, Han R, Wang Z, Ye L, Du Z, Wei H, Zhang F, Peng Z, Yang J: **trRosettaRNA: automated prediction of RNA 3D structure with transformer network.** *Nature Communications* 2023, **14**:7266.
63. Lin S, Tan H, Wang K, Wang R, Wang H, Zhu T, Xiong Y: **The dataset for the training, validation, and testing of RNARL.** Zenodo; 2026.
<https://zenodo.org/records/15486992>.
64. Lin S, Tan H, Wang K, Wang R, Wang H, Zhu T, Xiong Y: **RNARL.** GitHub; 2026.
<https://github.com/xlab-BioAI/RNARL>.
65. Lin S, Tan H, Wang K, Wang R, Wang H, Zhu T, Xiong Y: **RNARL: reinforcement learning-driven unified generative framework for multi-objective RNA codon design.** Zenodo; 2026. <https://zenodo.org/records/17212245>.
66. Lin S, Tan H, Wang K, Wang R, Wang H, Zhu T, Xiong Y: **Interactive web application for efficient codon optimization of target RNA sequences.** Google Colab; 2026.
https://colab.research.google.com/drive/1Z_t_Lt9CjqA0aygoNcdRzrThF1MOGBRb.