Nucleic Acids Research, 2025, **53**, W415–W426
https://doi.org/10.1093/nar/gkaf410
Advance access publication date: 19 May 2025
Web Server issue

mRNAdesigner: an integrated web server for optimizing mRNA design and protein translation in eukaryotes

Ouyang Mo[1,2,†], Zhuo Zhang[1,2,†], Xiang Cheng[2,3,†], Liqi Zhu[1,2], Kaixiang Zhang[3,4], Niubing Zhang[3], Justin Li[5], Honglin Li[6], Shixin Fan[7], Xuan Li[2,3,*], Pei Hao[1,2,*]

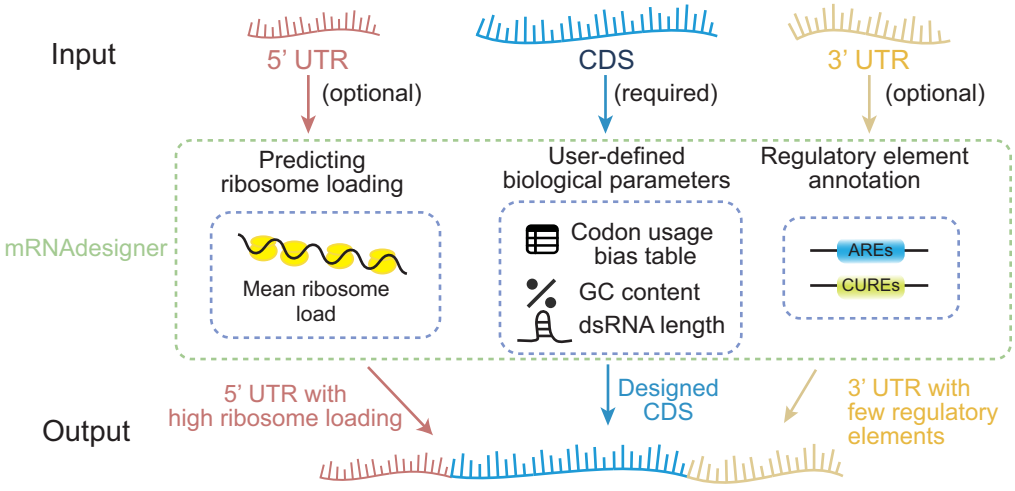
[1]Key Laboratory of Molecular Virology and Immunology, Shanghai Institute of Immunity and Infection, Chinese Academy of Sciences, Shanghai 200031, China
[2]College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
[3]Key Laboratory of Synthetic Biology, State Key Laboratory of Plant Trait Design, CAS Center for Excellence in Molecular Plant Sciences, Chinese Academy of Sciences, Shanghai 200032, China
[4]State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, Shanghai 200237, China
[5]Department of Mathematics, University of Connecticut, 352 Mansfield Road, Storrs, CT 06269, United States
[6]Innovation Center for AI and Drug Discovery, School of Pharmacy, East China Normal University, Shanghai 200237, China
[7]ChemPartner PharmaTech Co., Ltd, Shanghai 200237, China

[*]To whom correspondence should be addressed. Email: phao@siii.cas.cn
Correspondence may also be addressed to Xuan Li. Email: lixuan@sippe.ac.cn
[†]The first three authors should be regarded as Joint First Authors.

Abstract

Messenger RNA (mRNA) therapy has revolutionized modern medicine through its rapid development capabilities and ability to induce effective immune responses, becoming a powerful weapon against infectious diseases. The expression level of target proteins from mRNA sequences is primarily influenced by translational efficiency and stability, which can be significantly enhanced by modifying the 5′ and 3′ untranslated regions (UTRs), codon adaptation index, GC content, and secondary structure. To address the challenges of optimizing mRNA design, we have developed mRNAdesigner (https://www.biosino.org/mRNAdesigner/), a web server specifically designed to improve mRNA stability and translational efficiency in eukaryotes. Users can input a coding sequence (CDS) along with optional 5′ UTR and 3′ UTR, and the tool optimizes the CDS by reducing unpaired regions, minimizing complex stem-loop structures, and mitigating the use of rare codons while adhering to user-defined GC content preferences. Additionally, mRNAdesigner identifies optimal UTR sequences to enhance translation efficiency and stability. As an open-access computational resource, mRNAdesigner supports full-length mRNA design, enabling researchers to generate high-expression mRNA sequences for efficient protein production in eukaryotic expression systems, providing extra support for vaccine development and protein therapeutics. This is the first such tool that was made open accessible to the public.

Graphical abstract



Received: March 19, 2025. **Revised:** April 19, 2025. **Editorial Decision:** April 26, 2025. **Accepted:** May 2, 2025

Introduction

Messenger RNA (mRNA)-based therapeutics have emerged as a revolutionary platform for the prevention and treatment of various diseases, demonstrating unprecedented success in recent global health challenges and offering new possibilities for cancer therapy [1, 2]. Their modular nature, potential for rapid and scalable production, high efficacy, minimal off-target effects, and cost-effectiveness have catalyzed extensive research and clinical applications in infectious diseases and oncology [3]. By delivering synthetic mRNA that encodes therapeutic proteins to host cells, these treatments enable endogenous protein production directly within the body. Upon cellular uptake, the host translational machinery decodes the mRNA to synthesize the desired therapeutic proteins, which may temporarily function as immunomodulatory agents to activate the immune system, tumor-associated antigens to induce immune memory, or replacement proteins to restore physiological functions in cases of deficient or dysfunctional endogenous proteins [4, 5].

The therapeutic efficacy and duration of mRNA-based therapies are highly dependent on the ability of the delivered mRNA to achieve sufficient and accurate translation into functional target proteins [6]. Multiple interrelated factors influence this process, including mRNA stability, translational efficiency, and immunogenicity [4, 7–9]. Researchers face several key challenges: enhancing mRNA stability, improving translation efficiency, reducing immunogenicity, and maximizing protein expression levels—all of which must be addressed simultaneously to advance the effectiveness of mRNA-based therapeutics.

Several sequence optimization strategies have been developed to improve *in vivo* mRNA expression levels, each with specific advantages and limitations. Codon optimization, a widely adopted approach, involves the synonymous substitution of codons to match the target species' codon usage bias, thereby increasing translational efficiency [8, 9]. However, simply optimizing codon approach insufficiently addresses the complexities of mRNA translation. Structural optimization tools such as CDSfold [10] aim to minimize free energy to improve sequence stability, while RiboTree [11] innovatively proposed optimizing sequences by reducing average unpaired probability, thereby lowering RNA multi-loop structures to indirectly improve mRNA half-life. More recent integrative approaches such as LinearDesign [12] have combined codon optimization with structural considerations, balancing codon adaptation index (CAI) and free energy minimization to achieve higher protein expression levels.

Despite these advances, existing tools face challenges in addressing the multifaceted nature of mRNA expression *in vivo*. As evidenced by recent research, mRNA expression is not governed by isolated factors but rather coordinated through a dynamic regulatory network encompassing transcription efficiency, mRNA stability, translation elongation, and co-translational folding [13], rendering unilateral optimization strategies fundamentally inadequate. For instance, optimization strategies focusing on mRNA stability can lead to increased GC content, resulting in regions that may affect polymerase chain reaction (PCR) amplification and sequencing. Additionally, the presence of secondary structures in mRNA sequences can influence recognition by intracellular RNA sensors, which may impact cellular responses and mRNA persistence [14]. Furthermore, codon usage patterns can influence translation dynamics, as certain optimization approaches may inadvertently introduce rare codons that affect protein synthesis rates [15, 16]. Importantly, while many current tools concentrate on the coding sequence (CDS), the untranslated regions (UTRs) also play essential roles. UTRs and their structural features contribute significantly to mRNA stability and translation efficiency, which in turn affect protein expression levels [17–19]. Also, recent studies have shown that synergistic effects between 5′ UTR and 3′ UTR sequences can dramatically enhance translation efficiency. For instance, recent studies have demonstrated that when the 5′ UTR from rabbit β-globin was paired with the 3′ UTR containing mitochondrial RNA regulator and encephalomyocarditis virus elements, or when the 5′ UTR from adenoviral terminal protein leader sequences was combined with BioNTech's optimized 3′ UTR, these configurations had significantly increased luciferase expression [20]. These findings underscore the importance of optimizing UTR pairs for maximal protein production in mRNA-based therapeutics. There remains an opportunity to develop integrated solutions that consider both CDS and UTR optimization in a user-friendly format for comprehensive mRNA design.

To address these challenges, we have developed an optimization framework for mRNA sequence design. Building on established algorithms and incorporating recent advances in RNA biology, our method simultaneously optimizes multiple parameters known to influence mRNA stability and translation efficiency. These parameters include CAI, rare codon usage, GC content, unpaired nucleotide ratios, and structural features such as maximum stem length. Through iterative multi-parameter optimization, we generate CDSs with balanced and desirable properties. Importantly, our approach extends beyond CDS optimization to include upstream and downstream UTRs using a pre-constructed UTR library, enhancing translation initiation and full-length mRNA stability. To make this approach widely accessible to users, we have developed mRNAdesigner, a web server that incorporates the optimization strategies for mRNA sequence design in eukaryotes into a user-friendly platform. The mRNAdesigner web server is freely available at https://www.biosino.org/mRNAdesigner/.

Materials and methods

Implementation of CDS optimization module

The implementation of the CDS optimization module involves three main steps.

(i) Optimal codon substitution: The initial step involves synonymous substitution of all codons in the input CDS with the highest-frequency codons specific to the target species. This process generates a "root sequence" with a CAI of 1, serving as the foundational sequence for subsequent optimization.

(ii) Sequence exploration: Starting from the root sequence, the module iteratively performs random synonymous codon substitutions to search sequences in order to meet user-defined criteria. The optimization process is guided by three key principles: minimizing the unpaired ratio in CDS, adjusting the GC content to a user-defined goal, and avoiding long stem structures. Minimizing unpaired ratio promotes more paired bases, reducing multi-branch RNA structures and potentially decreasing sequence degradation [17], thereby enhancing stability.

However, RNA structure formation may trigger cellular innate immune responses [21, 14]. If sequences contain stem-loop structures that are excessively long (defined as exceeding 30 bp), this module prioritizes synonymous codon substitutions in regions with stem-loops greater than 30 bp during the next iteration, disrupting excessive base pairing to mitigate potential immunogenicity. During the process of synonymous codon substitution, we also avoid using the rarest codons to prevent reducing mRNA stability [22].

To quantify the optimization, the module calculates a sequence score (seq score) using the following formula:

$$w = \lambda \cdot r_{\text{unpair}} + (1 - \lambda) \cdot \Delta\text{GC} + P(L), \quad (1)$$

where r_{unpair} represents the unpaired base ratio and ΔGC denotes the deviation of the current sequence's GC content from the target GC content. $P(L)$ serves as a penalty term for discouraging the formation of overlong stems. If stems exceed the threshold, then $P(L)$ equals 1, otherwise it equals 0. λ is a weighting factor that balances the contributions of these objectives; default is 0.5.

To efficiently search for the optimal sequence, the module leverages the Monte Carlo Tree Search (MCTS) algorithm, inspired by the RiboTree algorithm from the Eterna platform [11]. The MCTS algorithm balances exploration of new sequences and exploitation of known sequences by applying the Upper Confidence Bound for Trees loss function: $w_i/n_i + c\sqrt{\ln N_i/n_i}$, where w_i represents the sequence score after the ith step of codon substitution, n_i denotes the total visits to that sequence, and c is a constant that controls the trade-off between exploration and exploitation.

(iii) Optimal sequence selection: When the algorithm reaches the maximum number of searches, it stops. The optimized sequence is selected based on the best sequence score, reflecting optimized secondary structure, GC content, and stem length constraints, which collectively contribute to enhanced mRNA stability and expression efficiency.

Implementation of 5′ UTR optimization module

The 5′ UTR plays a crucial role in translation initiation and ribosome loading [23–25]. To optimize the 5′ UTR for enhanced translation efficiency, we collected 5′ UTR sequences from multiple databases (shown in Table 1) across various species [23, 24–28] to build a 5′ UTR library. The sequences were filtered to exclude those containing upstream AUGs (uAUGs) [25] and those exceeding 100 nucleotides in length, as extended 5′ UTRs may complicate plasmid vector construction. If user-defined Kozak sequence parameters were specified, a Kozak consensus sequence ("GCCACC"), known to enhance translation initiation, was appended to the 3′ end of each 5′ UTR. These UTR sequences were then concatenated with the first 30 nucleotides of the optimized CDS to form chimeric 5′ UTR–CDS constructs. Due to the inability to directly evaluate the relationship between 5′ UTR and translation efficiency, we use mean ribosome load (MRL), predicted by the UTR-LM model [26], to indirectly quantify translation efficiency.

We found that the secondary structure of 5′ UTR–CDS chimeric sequences, especially the folding structure of the 5′ UTR and the first 30 nt of CDS, affects translation [18], ultimately influencing protein expression levels [27]. Therefore, we predicted their secondary structures using ViennaRNA RNAfold [28, 29]. To capture structural features, the paired/unpaired states of the terminal 30 nucleotides were converted into binary matrices (1 for paired, 0 for unpaired) as structural feature vectors. Based on these structural feature vectors, all chimeric sequences were divided into different structural types using a hierarchical clustering algorithm. To balance computational complexity and structural diversity, we pruned the clustering tree to retain the five subclusters, a number determined artificially by balancing computational resources and the need to capture major structural patterns. Within each structural subcluster, we selected the 5′ UTR with the highest MRL value as a candidate sequence.

This screening, based on structure and MRL, considers the folding structural diversity of candidate UTR sequences with different CDS and ensures that each type of UTR has a relatively high MRL, facilitating wet lab scientists in testing the actual expression effects of different sequences.

Table 1. Overview of 5′ UTR databases

Database	5′ UTR counts	Source
Ensembl database	14 059	[24]
Synthetic 50 bp libraries	49 424	[23]
Cao-RF	1153	[49]
UTR-LM	78	[26]

Implementation of 3′ UTR optimization module

The 3′ UTR is also an important factor influencing mRNA expression. Although there is currently no universally accepted principle to guide the design of 3′ UTRs for improving mRNA half-life and translation efficiency, researchers have adopted strategies that utilize naturally occurring mRNA 3′ UTRs as templates for designing therapeutic mRNA 3′ UTRs [30]. Common sources of 3′ UTRs include genes such as hemoglobin subunit alpha (HBA) and beta (HBB), albumin (ALB), heat shock protein 70 (Hsp70), tyrosine hydroxylase (TH), and collagen alpha-1 (COL1A1) [31]. Additionally, some mRNA vaccines have provided beneficial attempts, such as the 3′ UTR of the AES/TLE5 gene, which is considered an ideal choice due to its low microRNA binding sites and high hybridization energy [32].

The 3′ UTR of mRNA contains various regulatory elements that significantly influence mRNA stability, such as AU-rich elements (AREs) and CU-rich elements (CUREs) [33]. AREs, in particular, are known to decrease mRNA stability by accelerating transcript degradation [34].

To assist users in selecting suitable 3′ UTR sequences, we constructed a 3′ UTR library by extracting sequences from the UTRdb database [35]. The library includes 173 979 human-derived and 69 113 mouse-derived 3′ UTRs. Sequences containing unknown nucleotides (including nonstandard nucleotides such as N, R, Y, etc.) were eliminated to ensure that all sequences in the library consist exclusively of the four clearly defined nucleotides: A, U, G, and C. Long 3′ UTR sequences were removed, filtering to retain 3′ UTR sequences ≤1000 nt in length for library construction. For sequence annotation, we utilized the Find Individual Motif Occurrences (FIMO) tool from the MEME-Suite package [36] to identify and quantify the occurrences of ARE and CURE

motifs within both user-provided 3′ UTR sequences and those in the library.

We provide three filtering options for sequence selection to accommodate diverse user requirements: (i) sequences devoid of both AREs and CUREs, (ii) sequences containing CUREs regardless of AREs presence, and (iii) sequences containing only CUREs (with AREs excluded). These options enable users to select 3′ UTR sequences based on their specific experimental preferences, thereby optimizing transcript stability and functionality. To further assist users in evaluating the potential impact of 3′ UTR selection on mRNA functionality, we also calculated the folding rate of the optimized CDS–3′ UTR. Previous studies have indicated that increased secondary structure in the CDS and the entire 3′ UTR correlated with improved protein expression [18]. We hope this predictive feature provides users with additional reference information when selecting 3′ UTR sequences.

Web interface implementation

The mRNAdesigner web interface build on Browser/Server architecture. The frontend is developed using Vue.js v3.2.36 (https://vuejs.org/), combined with Bootstrap v5.1.3 (https://getbootstrap.com/) for responsive and consistent UI design. HTTP requests and API communication are facilitated through Axios v1.6.8 (https://axios-http.com/). The frontend application is bundled into static assets using Vite v2.9.9 (https://vitejs.dev/) and deployed on an Nginx server (https://nginx.org/).

On the backend, Django v5.0.4 (https://www.djangoproject.com/) is employed to handle HTTP requests and manage application logic. Celery v5.3.6 (https://docs.celeryproject.org/) is integrated for asynchronous task processing because mRNAdesigner takes time to find the best sequence. Backend logic and sequence processing are predominantly executed using Python and shell scripts.

The mRNAdesigner does not require user registration or collect personal information. Tasks submitted by users are processed on the server, and results are temporarily stored for a period of 30 days before deletion. The platform is optimized for modern web browsers that support the HTML5, including Firefox, Google Chrome, Microsoft Edge, and Safari across various operating systems.

Cells culture

HEK293T cells were cultured in Dulbecco's modified Eagle medium (DMEM, Gibco, USA) supplemented with 10% fetal bovine serum (FBS, Sigma–Aldrich, USA), and 1% penicillin and streptomycin (P/S, Thermo Fisher Scientific, USA). All cells were cultured at 37°C, 5% CO_2 atmosphere.

Gene and plasmid preparation

The sequences of four monkeypox virus (MPV) antigens L1R, D8L, A27L, and H3L were obtained from the monkeypox virus reference sequence at NCBI (accession: GCF_014621545.1). Respiratory syncytial virus (RSV) F protein was obtained from NCBI (GenBank: MF445831.1). *Candida albicans* antigen proteins Als3 and Sap2 were obtained from *C. albicans* SC5314 at NCBI (accession: NC_032 096).

All genes used were synthesized by Tsingke Biotechnology Co., Ltd and cloned into pcDNA3.1 expression vector. The Trelief 5α Chemically Competent Cells (Tsingke Biotechnology Co., Ltd) were used for plasmid transformation, and the plasmids were extracted and purified using Endo-Free Plasmid Mini Kit (Omega Bio-tek).

Western blot analysis

Western blot was conducted to detect protein gene expression levels. HEK293T cells were seeded in six-well plates at a density of 8×10^5 cells per well. Each well was transfected with 2.5 μg of plasmid DNA using Lipofectamine 8000 (Beyotime Biotechnology) according to the manufacturer's instructions. Subsequently, the cells were lysed on ice for 20 min using RIPA lysis buffer (Beyotime, China). Cell lysates were then subjected to sodium dodecyl sulfate–polyacrylamide gel electrophoresis(SDS-PAGE) and transferred onto a polyvinylidene fluoride membrane (Millipore) for western blotting. The membrane was incubated with the anti-flag Tag mouse monoclonal Ab (Yeasen, 1:1000 dilution) for 1 h at room temperature, and subsequently with the horseradish peroxidase (HRP)-conjugated goat anti-mouse IgG (Yeasen, 1:5000 dilution) for 1 h at room temperature. Enhanced chemiluminescence substrate solution (biosharp) was utilized to detect signals.

mRNA and mRNA/LNP preparation

Linear templates of antigens containing the T7 promoter and PolyA tail were amplified from the previously constructed plasmid using PCR. The mRNA synthesis and mRNA/lipid nanoparticle(LNP) encapsulation were conducted by the protocol mentioned in our earlier work [37].

Animals and vaccination

Female BALB/c-mice (6 weeks old, female) were purchased from VitalRiver (Pinghu, China). For MPV antigen vaccination, mice were divided into three groups ($n = 7$) and immunized intramuscularly with two doses of 7.5 μg of A27L and D8L antigen-encoding mRNA and Vac-Ctrl containing noncoding mRNA as a placebo, respectively. For RSV antigen vaccination, mice were divided into two groups ($n = 7$), immunized intramuscularly with two doses of 20 μg of RSV F antigen-encoding mRNA and Vac-Ctrl. And for *C. albicans*, mice were divided into two groups ($n = 7$), immunized intramuscularly with two doses of 7.5 μg of Als3 antigen-encoding mRNA and Vac-Ctrl. The initial immunization was given on day 0, followed by a booster dose after 2 weeks. Blood samples were collected twice after the initial immunization to measure antibody levels.

Antigen purification and ELISA

The antigens used for ELISA were removed by PCR amplifications and then they were cloned into the pGEX-4T-1 vector for expression in *Escherichia coli*. The cloning construction was carried out using the in-fusion cloning protocol as performed previously [38]. These plasmids were transformed into Rosetta (DE3) strain (WEIDI, Shanghai, China). Expression of MPV-antigens was induced using 300 μmol L^{-1} isopropyl-β-D-thiogalactoside (IPTG, Beyotime) at 16°C for 18 h. The recombinant proteins were purified using the Ni-NTA agarose resin (Yeasen) or GST-tag purification resin (Beyotime) following the manufacturers' protocol.

IgG antibody titers against antigens, MPV antigens A27L and D8L, RSV antigens F protein, and *C. albicans* antigen

Als3, were determined by ELISA using the purified recombinant proteins. Ninety-six-well plates (Corning, USA) were coated with 5 μg ml^{-1} recombinant protein and incubated overnight at 4°C. Plates were washed with TBST (20 mmol L^{-1} Tris, 137 mmol L^{-1} NaCl, 0.1% Tween, pH 7.4) and blocked with 5% non-fat dry milk for 1 h at 37°C. Serial dilutions of heat-inactivated mouse serum were added to the wells and incubated for 1 h at 37°C, followed by incubation with HRP-conjugated goat anti-mouse IgG (Yeasen, diluted 1:5000 in 1% milk/TBST) for 1 h at 37°C. Next, the plates were treated with the 3,3′,5,5′-tetramethylbenzidine (MesGen, Shanghai, China) for 15 min before reactions were terminated with 2 mol L^{-1} hydrochloric acid. Absorbance at 450 nm was recorded using a Varioskan Flash microplate reader (Thermo Fisher Scientific).

Ethics statement

All animal experiments were conducted in accordance with the Guidelines for the Care and Use of Laboratory Animals and were approved by the Institutional Animal Care and Use Committee of the Shanghai Institute of Immunity and Infection, Chinese Academy of Sciences (approval number: A2023028). Mice were housed under specific pathogen-free conditions with controlled humidity (target: 55%, range: 45%–65%), temperature (target: 22°C, range: 20–24°C), and a 12-h light/12-h dark cycle, with free access to food and water.

Results

Overview of web server features and functional components

mRNAdesigner is a web server specifically developed to optimize mRNA sequences for enhanced stability and translation efficiency in eukaryotic systems. It incorporates three distinct functional modules designed to optimize the CDS, 5′ UTR, and 3′ UTR regions (Fig. 1).

The CDS optimization module is the core of mRNAdesigner. It optimizes input sequences through synonymous codon substitutions based on six parameters: target species codon usage, GC content, maximum stem length, rare codon allowance, and search iterations. The optimization process initiates with the replacement of all codons using the highest-frequency synonymous codons in the target species, generating a modified CDS with a CAI of 1. This ensures that every codon in the sequence represents the highest-frequency synonymous codon. Subsequently, the module performs random synonymous substitutions and calculates a sequence score (seq score) for the modified sequence. The seq score is determined by three factors, including the degree of base pairing, the deviation from the target GC content, and penalties for overlong stem. Lower sequence scores indicate a better fit to user-defined constraints. The process continues until reaching the specified iteration limit, ultimately outputting the CDS variant with the optimal sequence score while maintaining all specified constraints.

The 5′ UTR optimization module focuses on enhancing translation initiation by selecting the most compatible 5′ UTR for the optimized CDS. Using a pre-constructed species-specific 5′ UTR library, this module appends the first 30 nucleotides of the optimized CDS to candidate 5′ UTRs, forming chimeric sequences. We quantified the translation performance of 5′ UTRs using the MRL predicted by UTR_LM [26]. RNAfold [28, 29] was used to predict the folding structure of sequences. The folding structures were categorized by hierarchical clustering, and 5′ UTRs with high MRL were selected from different clusters as candidate 5′ UTRs.

The 3′ UTR optimization module targets regulatory elements that influence mRNA stability within the noncoding region downstream of the CDS. This module identifies and annotates stability-related elements, such as AREs and CUREs. The module scans sequences from the 3′ UTR library or user-provided 3′ UTRs for these regulatory elements by the FIMO tool [34]. Based on the count of presence of these motifs, the module filters 3′ UTR sequences to meet user-defined criteria.

The mRNAdesigner web server integrates computational outputs from three modules, presenting the synthesized data through an interactive table. This table allows evaluation of candidate sequences through advanced sorting and comparison. Additionally, researchers can export datasets, including detailed sequence annotations from all modules, for subsequent computational analysis and experimental validation.

User inputs and interface

The mRNAdesigner web server interface consists of three primary optimization modules for different regions of mRNA (Fig. 2).

The CDS optimization module requires users to input a CDS in FASTA format and select a target species codon usage table, which serves as the foundation for optimization. Users can further customize parameters, including the GC content, maximum stem-loop length, rare codon usage preferences, and the number of searching iterations. By default, the parameters are set to a GC content of 50%, rare codon exclusion, 30 bp stem-loop length limit, and 6000 iterations.

The 5′ UTR optimization module supports both a pre-constructed species-specific 5′ UTR library and user-provided sequences. The pre-constructed library includes high MRL 5′ UTRs derived from various species or artificially constructed, experimentally validated sequences. By default, the module checks the option to include a Kozak sequence, as it enhances the expression of the target protein. Users can also provide their custom 5′ UTR sequence to predict the MRL.

In the 3′ UTR optimization module, users can either select sequences from a pre-constructed 3′ UTR library or input their own 3′ UTR sequences. Due to the variable length of 3′ UTRs, users are required to specify a desired 3′ UTR length range (default 100–300 nucleotides), with a maximum allowable length of 1000 nucleotides. The module provides regulatory element filters, allowing users to prioritize specific sequence characteristics. For example, selecting the “Element-Free” option outputs sequences devoid of AREs and CUREs, and the “CU-Rich Priority” option generates sequences that maximize CUREs counts regardless of AREs, while the “CU-Rich Excluding AU” option produces sequences rich in CUREs but devoid of AREs. The default setting is “Element-Free.”

Case study 1: enhanced protein expression through multi-region mRNA design

To evaluate the effectiveness of mRNAdesigner in enhancing protein expression in eukaryotic systems, we used the tool to redesign the mRNA encoding the RSV F protein. The RSV F protein is the primary antigen for RSV vaccines [39, 40], and previous studies have demonstrated that the

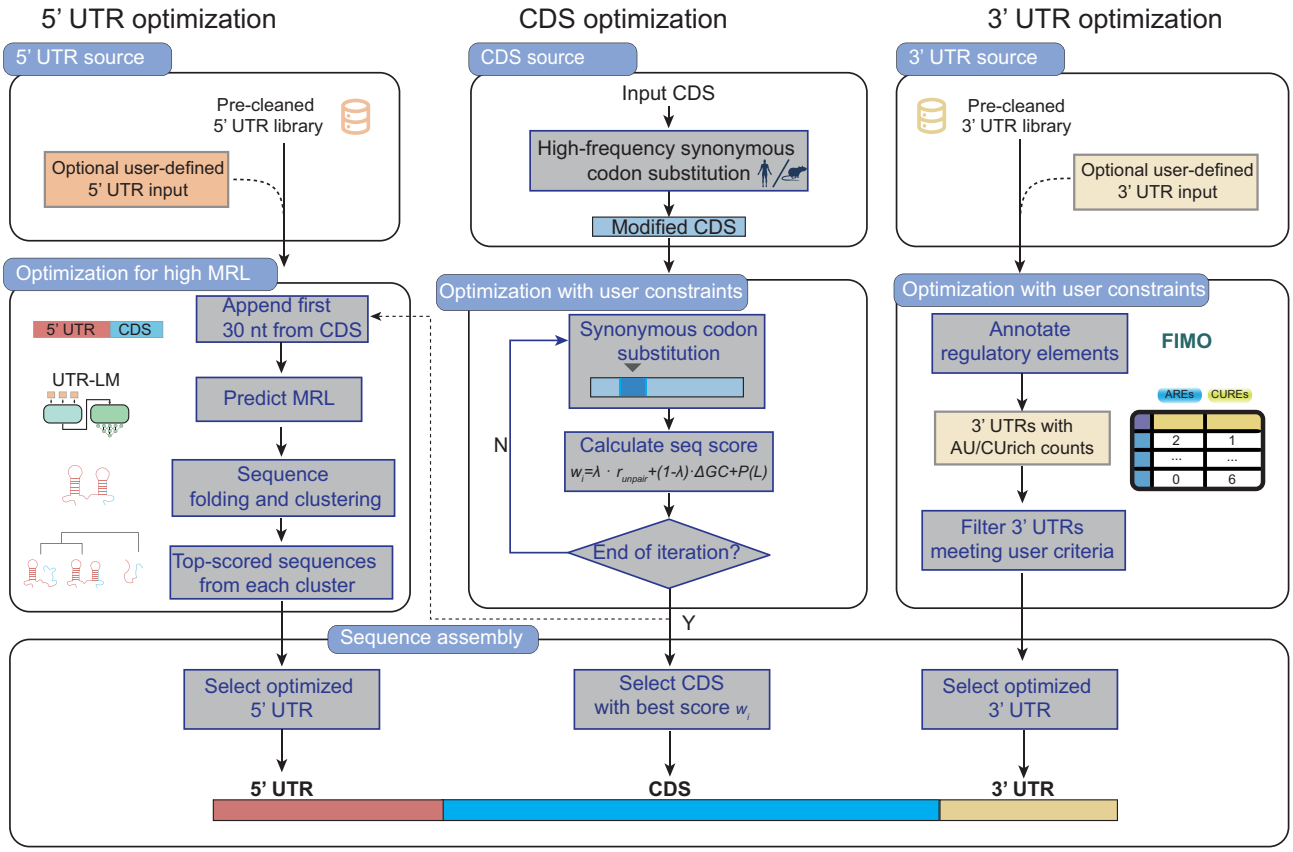


Figure 1. Overview of mRNAdesigner's functional components for mRNA sequence optimization. The web server consists of three main modules: 5′ UTR optimization, CDS optimization, and 3′ UTR optimization. 5′ UTR optimization module: starting with sequences from a pre-cleaned library or user-provided input, this module optimizes translation initiation efficiency through: (i) incorporation of the first 30 nucleotides from the optimized CDS module; (ii) prediction of MRL using UTR-LM; (iii) sequence folding and clustering; and (iv) selection of sequences with high MRL from each cluster. CDS optimization module: processes input CDSs sequences through iterative optimization: (i) performs high-frequency synonymous codon substitution; (ii) calculates sequence scores under user-defined constraints, considering base pairing, GC content deviation, and penalties for overlong stem structures; and (iii) selects the CDS with optimal sequence score. 3′ UTR optimization module: processes sequences from a pre-cleaned library or user inputs: (i) annotates regulatory elements using the FIMO tool to identify AREs and CUREs and (ii) filters 3′ UTR sequences according to user-defined criteria. Sequence assembly: combines the 5′ UTR, CDS, and 3′ UTR to generate a complete mRNA sequence.

wild-type (WT) RSV F protein exhibits inherently low expression levels [41]. The optimization parameters were configured to achieve a target GC content of 55% based on human codon usage frequencies while avoiding stem structures exceeding 30 base pairs and rare codons. The redesigned mRNA sequence showed significant improvements in multiple parameters when compared to the WT sequence (Fig. 3A). The GC content of the optimized sequence increased from 36.46% in the WT mRNA to 54.95%, closely aligning with our target. Additionally, the base-pairing probability was moderately enhanced, rising from 59.88% to 61.42%. Notably, to minimize the potential for innate immune activation, the maximum stem length was carefully controlled, increasing only slightly from 19 to 21 bp. Furthermore, the CAI demonstrated marked improvement from 0.8684 to 0.9584, indicating enhanced alignment with human codon usage preferences.

The RNA paring/folding structure of designed mRNA is one of the important outputs, which is directly correlated to the stability of mRNA [18]. The secondary structure of the full-length mRNA sequence after modification was visualized using FORGI 2.0 [42]. Structural analysis revealed that the WT mRNA exhibited a propensity to form multi-loop conformations, whereas the redesigned sequence showed a more stable secondary structure with reduced structural complexity. For instance, within the 546–591-nt region (highlighted by the red rectangle in Fig. 3A), synonymous codon substitutions resulted in a sequence that favored stable stem pairing, thereby reducing the formation of multi-branch loops. The reduction in multi-branch loops in our redesigned mRNA may contribute to enhanced stability, as these structures typically contain more single-stranded regions that can be targeted by cellular exonucleases [43, 44].

To assess the functional impact of these structural and sequence optimizations, we transfected the redesigned mRNA into 293T cells via plasmid delivery, with protein expression analyzed by western blot at 24 h post-transfection (Fig. 3B). Western blot quantitative densitometry revealed a 2.32-fold (232%) increase in RSV F protein levels compared to WT mRNA (Fig. 3C).

These findings demonstrate that mRNAdesigner serves as a practical and effective tool for mRNA sequence optimization in eukaryotes. Under user-defined constraints, the software can improve the number of base pairing and CAI, and avoid the formation of overly long stem structures that could trigger innate immune responses. The redesigned mRNA sequences

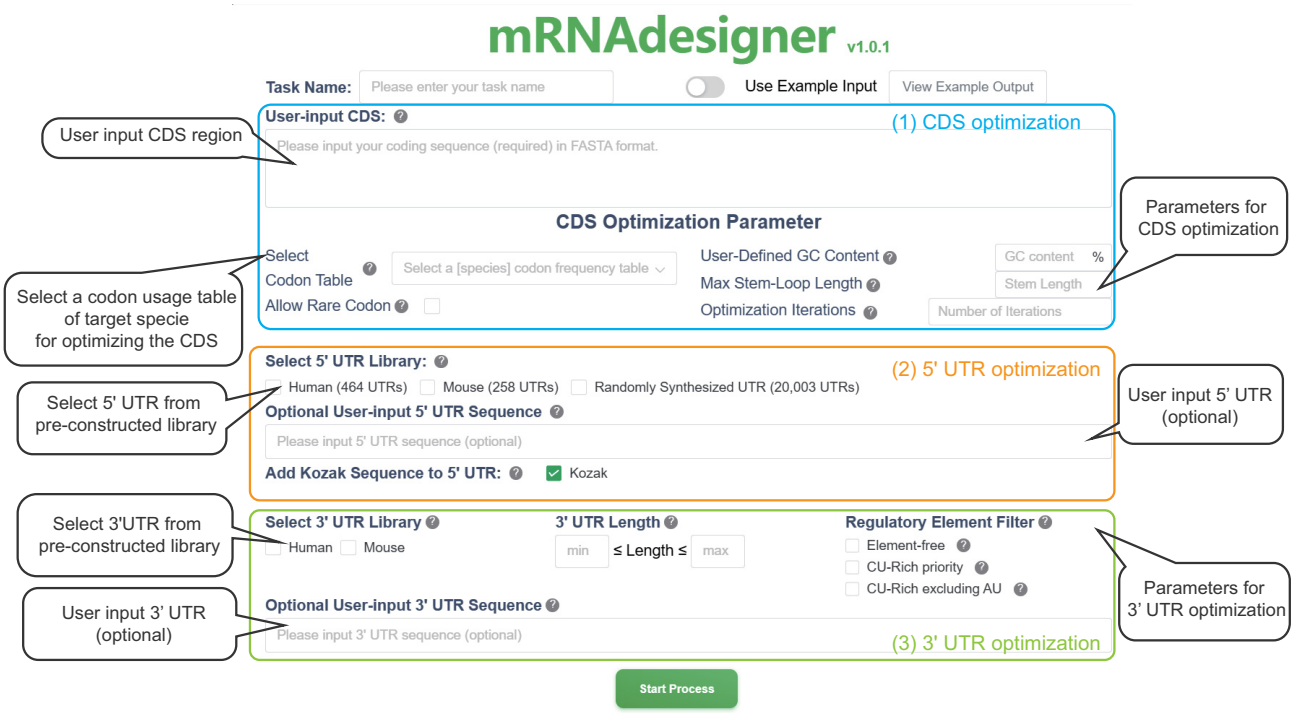


Figure 2. Modular interface of the mRNAdesigner web server. The interface consists of three main modules: (i) CDS optimization, (ii) 5′UTR optimization, and (iii) 3′ UTR optimization. The CDS module accepts FASTA input with adjustable parameters for codon usage, GC content, and structural features. The 5′ UTR module offers pre-constructed or custom sequences with Kozak sequence options. The 3′ UTR module allows sequence selection with customizable length ranges and regulatory element filters.

significantly outperformed their WT counterparts, achieving substantial improvements in protein expression.

Case study 2: broad adaptability of the mRNAdesigner across multiple species

The mRNAdesigner was developed to enhance the expression of proteins from diverse sources across eukaryotic cell lines. To evaluate its adaptability across different species, we optimized and expressed mRNA sequences derived from distinct biological origins in the 293T cell. These sequences included viral antigens, bacterial proteins, fungal antigens, and synthetically designed tumor neoantigens. The optimization was performed using default parameters (GC content: 50%, maximum stem-loop length: 30 bp, rare codon elimination, and 6000 iterations). We redesigned these sequences and transfected them into 293T cells for expression, with protein expression analyzed by western blot at 24 h post-transfection. Both the original and optimized CDSs were compared. Figure 4A and Supplementary Fig. S1 compare their secondary structures and various sequence characteristics, including CAI, base-pairing ratio, and the length of the longest stem-loop, with arrows highlighting improvements in the optimized versions. Figure 4B and C confirms the successful expression of these optimized sequences in 293T cells.

Notably, all mRNA sequences optimized by mRNAdesigner showed improvements compared to their original counterparts (Fig. 4A and Supplementary Fig. S1), with CAI values exceeding 0.9 across all optimized sequences. Structurally, mRNAdesigner-optimized sequences exhibited marked reductions in multi-branch loop structures relative to their original counterparts, resulting in enhanced pair ratio values. Additionally, the length of the longest stem-loop was reduced in the optimized sequences, thereby mitigating the risk of triggering intracellular immune responses upon delivery of these mRNAs. This balance between structural optimization and codon adaptation ensured consistently high and stable expression levels across all tested sequences in 293T cells. These results show the broad adaptability of mRNAdesigner in optimizing sequences derived from diverse species while maintaining their expression efficiency in eukaryotic cell systems.

Case study 3: immunogenicity assessment of optimized mRNA sequences in mouse models

To assess the immunogenic potential of antigens translated from mRNA sequences optimized by mRNAdesigner, we conducted immunological studies in mouse models. Mice were intramuscularly (IM) vaccinated twice with the mRNA constructs and a negative control formulation, consisting of empty LNPs without an mRNA sequence [45]. Serum samples were collected on days 0, 14, and 28 for serological analysis.

We selected antigenic targets from diverse sources, including viral and fungal proteins, to evaluate the immune responses induced by the optimized mRNA sequences. Specifically, we analyzed the immune responses against the RSV F protein, the MPXV antigens (A27L and D8L), and the fungal antigen Als3p from *C. albicans*. Serum samples were collected at two distinct time points post-immunization, and specific antibody responses were quantified using ELISA.

As shown in Fig. 5A–C, antigen-specific IgG responses were observed across all antigenic targets, with varying magnitudes that correlated with their respective antigenic properties. Significant IgG responses were detected against RSV F protein (Fig. 5A), MPXV antigens A27L and D8L (Fig. 5B),

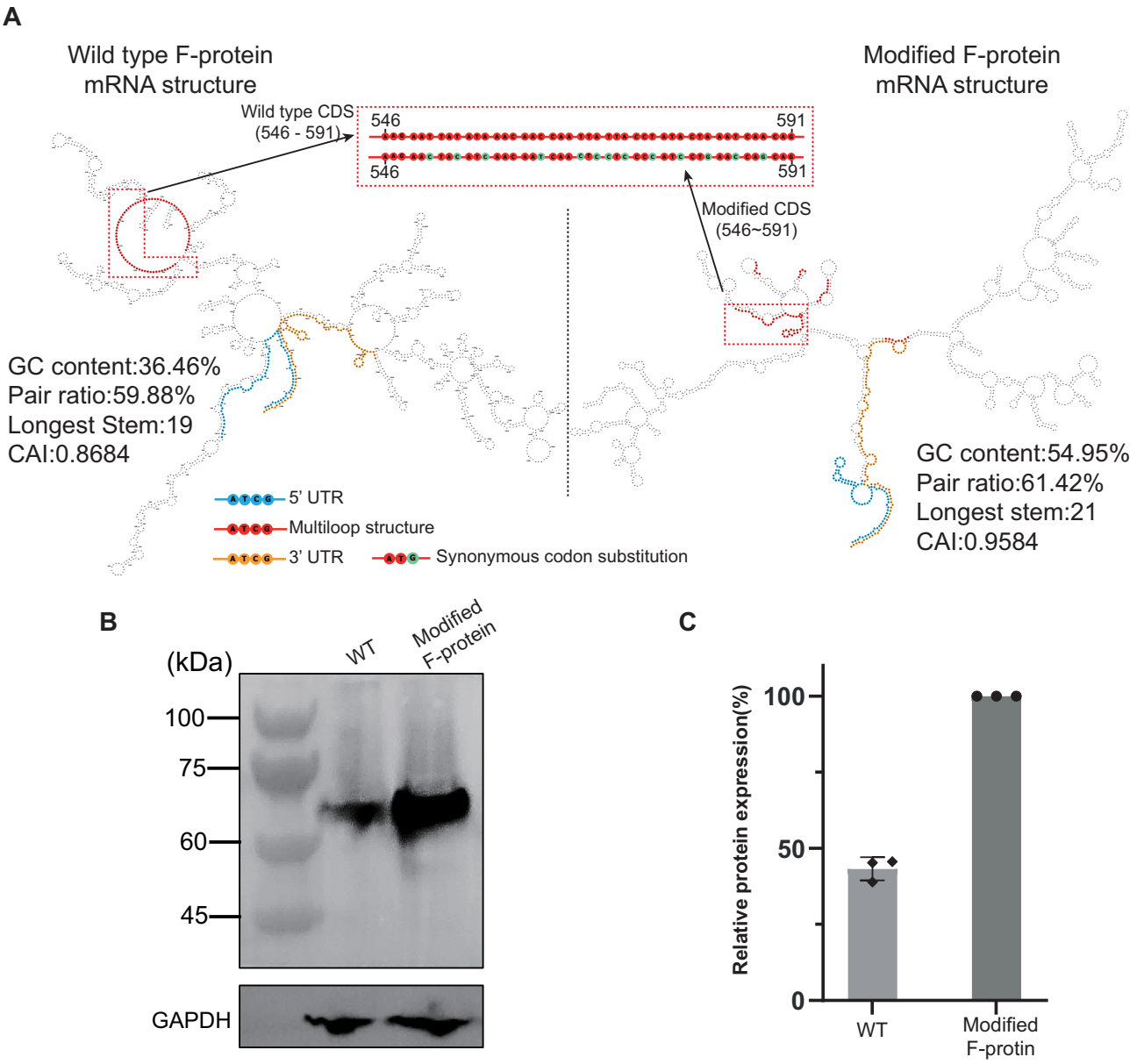


Figure 3. Optimization of RSV F-protein mRNA by multi-region design significantly enhances its protein expression. (**A**) Comparison of predicted secondary structures between WT and optimized RSV F-protein mRNA sequences. The optimized sequence exhibits improved parameters, including an increase in GC content (54.95% versus 36.46%), a higher base-pairing ratio (61.42% versus 59.88%), and an enhanced CAI (0.9584 versus 0.8684). A detailed view of nucleotides 546–591 reveals the reduction of multi-loop structures in the CDS through synonymous codon substitutions. Example changes in multi-loop structures are highlighted within the dotted rectangles. (**B**) Western blot showing F-protein expression levels in 293T cells transfected with WT or optimized mRNA sequences, with cells harvested 24 h post-transfection. (**C**) Quantitative analysis of relative F-protein expression levels based on densitometric evaluation of western blot signals. Data are presented as mean ± SD ($n = 3$). The optimized mRNA sequence exhibits a 2.32-fold increase in F-protein expression compared to the WT sequence.

and the fungal antigen Als3p (Fig. 5C). These results show the efficacy of mRNAdesigner optimized mRNA sequences in inducing specific immune responses. The consistent detection of antigen-specific antibody titers validates the capability of the optimized sequences to achieve correct protein translation and presentation to the immune system.

Conclusions and perspective

We developed a web-based server, mRNAdesigner, with an optimization framework, for the design of mRNA sequence for related applications like mRNA vaccine development and mRNA replacement therapeutics. To our knowledge, this is the first such tool that is open-access of its kind available to public. This integrated tool consists of three functional modules for CDS, 5′ UTR, and 3′ UTR optimization, simultaneously balancing the various factors, e.g. CAI, rare codon usage, GC content, unpaired ratios, and structural feature like maximum stem-loop length, which influence mRNA stability and translation efficiency. Our tool extends from traditional CDS optimizations, to cover both upstream and downstream UTRs, which are considered to play crucial roles in mRNA stability, translation efficiency, and degradation control. The

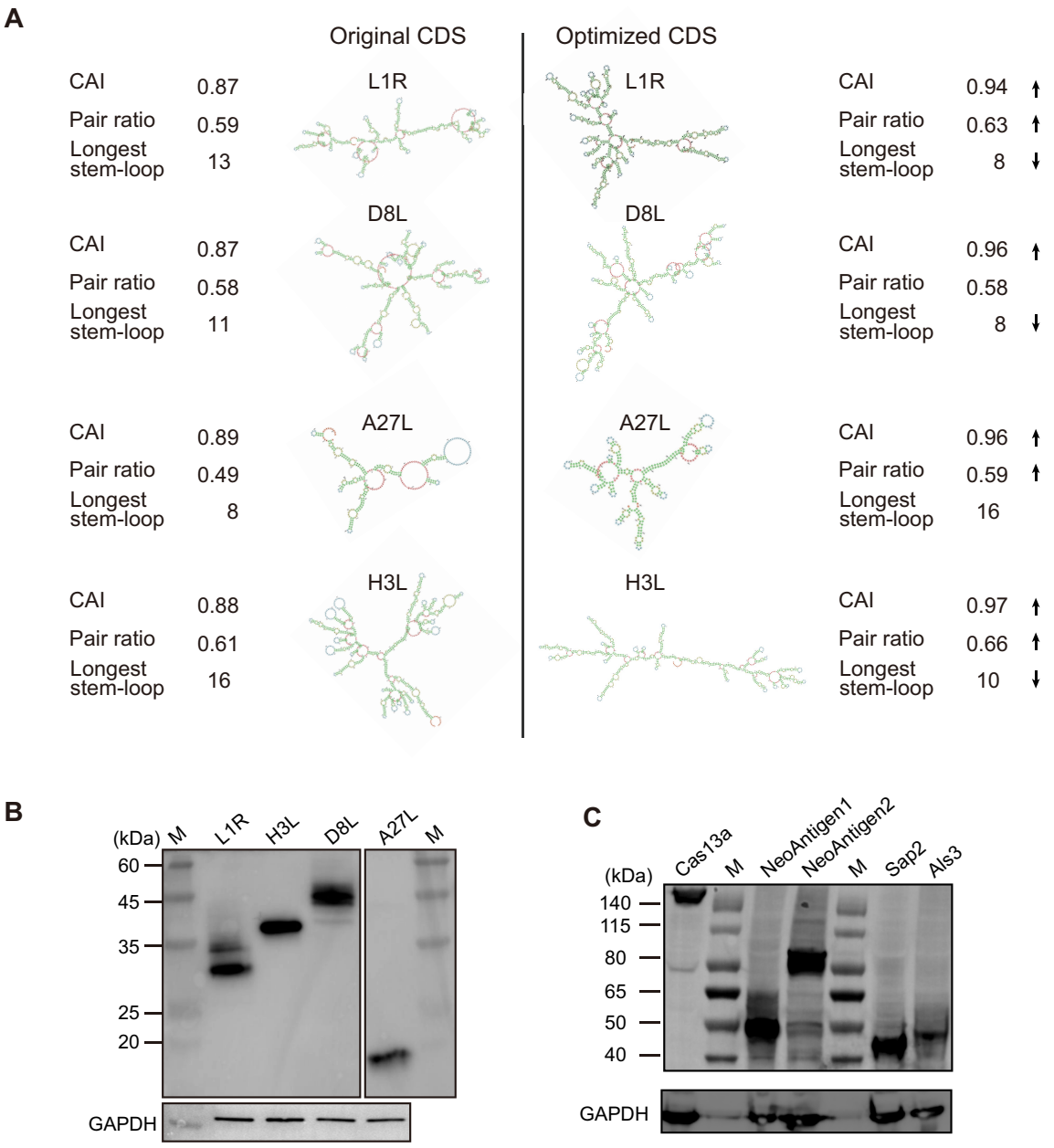


Figure 4. mRNAdesigner enhanced design and expression of proteins from diverse sources. (**A**) Secondary structures comparison between original (left) and optimized (right) mRNA sequences encoding monkeypox virus (MPXV) antigens (L1R, D8L, A27L, and H3L). Sequence metrics including CAI, base-pairing ratio, and the longest stem-loop length are shown for each pair. Arrows indicate improved parameters in optimized sequences compared to their original counterparts. (**B**) Western blotting showing expression levels of the MPXV antigens in 293T cells, with cells harvested 24 h post-transfection. GAPDH served as the internal loading control. Molecular weight markers (M) are indicated. (**C**) Western blotting showing expression levels of the optimized sequences in 293T cells. GAPDH was used as a loading control. Molecular weight markers (M) are indicated.

integrity of designed mRNA is proved by the production of proteins with good efficiency shown in three study cases. To make it easy to use for researchers, we provide a user-friendly, intuitive graphical interface, giving users the freedom to adjust inputs and parameters to meet their research needs with mRNAdesigner.

The efficient and sustained expression of proteins is critical for mRNA-based therapeutics. Achieving this goal requires carefully designing of mRNA sequences to maximize translation efficiency and protein yield. This is challenging because mRNA design is constrained by the need to balance two often competing requirements: (i) minimizing secondary structure to improve ribosome scanning speed to improve efficiency and (ii) introducing some degree of structural stability in the CDS to improve mRNA half-life and overall protein output.

Accordingly, we address these competing requirements by applying differentiated optimization strategies to distinct mRNA regions. For instance, for the 5′ UTR, we primarily use MRL to quantify the efficiency of translation. For the CDS region, we focus more on improving RNA secondary structure. Recent research suggests that when considering the trade-off between improved ribosome scanning speed (requiring less stable structures) and extended sequence half-life (requiring more stable structures), the sequence half-life often contributes more to overall protein expression [27].

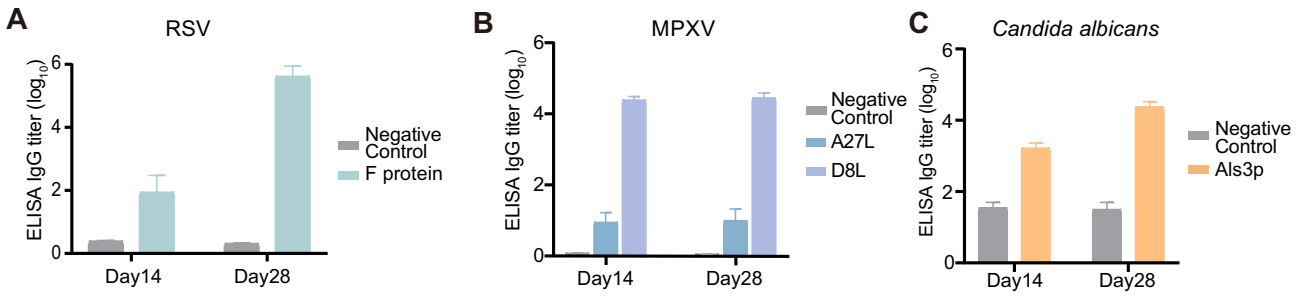


Figure 5. Evaluation of antigen-specific immune responses induced by mRNAdesigner optimized sequence. Quantification of antigen-specific IgG responses in mouse serum samples by ELISA: (**A**) RSV F protein-specific IgG responses; (**B**) MPXV A27L- and D8L-specific IgG responses; and (**C**) Als3p-specific IgG responses against the fungal antigen. Results show varying antigen-specific IgG responses. Data are presented as mean ± SEM for each antigen and the negative control. Negative controls are empty LNPs without mRNA sequence.

However, while the CDSs optimized by mRNAdesigner have high CAI values and avoid the rarest codons, it is important to note that the most frequently used codons are not always optimal. In fact, strategically using less common codons can help slow translation and facilitate proper protein folding [46]. Moreover, current codon usage frequencies are typically derived from whole-genome analyses and do not take into account tissue-specific or condition-dependent variations in tRNA abundance. Research has demonstrated that different tissues can exhibit distinct tRNA expression patterns [47], which may influence protein expression in a tissue-specific manner. Therefore, future mRNA design could benefit from incorporating tRNA transcriptomics and context-specific abundance data to improve predictive accuracy.

Despite these limitations, we have achieved a 2.3-fold increase simply by optimizing its 5′/3′ UTRs and CDSs using mRNAdesigner. The outcome from mRNA therapeutics heavily depends on the level, biological activities, and stability of the protein in expressing tissues. Specifically for mRNA vaccines, the induction of protective antibody titers is related to the protein expression levels, sustainability of the expressed protein, and also its immunogenicity. For mRNA-based protein replacement therapies, their outcomes further depend on the proper folding and stability of the expressed proteins, as well as their functionality/biological activities in target tissues. Further refinement and optimization are expected using mRNAdesigner to better meeting their needs that are often based on the therapeutics purpose of each study.

While mRNAdesigner represents an important milestone in the field, future enhancements will be critical to meet application needs and face the evolving challenges. In the model training for predicting ribosome load on 5′ UTRs, the UTR-LM model utilized high-throughput datasets that modeled the translation mechanisms of many different cell types. Therefore, the UTR-LM model should possess the best generalizability available today. However, there is certainly much room to improve mRNAdesigner on its applicability as studies are progressing with many more cell lines and tissue types, which we hope to train and update periodically. Since the impact of modified nucleotides, such as m1Ψ, on RNA folding dynamics remains unresolved [48], the current version is not applicable for structural optimization of modified nucleotides. Moreover, extending the application of mRNAdesigner from eukaryotic cells to a broader range of kingdoms holds considerable potential, but will require adaptation to account for distinct regulatory mechanisms and structural constraints present in different biological contexts. The contexts of high MRL scores that can potentially trigger translation-dependent mRNA degradation are poorly understood [27]. Furthermore, the engagement of mRNA binding proteins to various mRNA elements beyond the AREs/CUREs in the 3′ UTRs is to be defined for protein expression and mRNA stability.

In conclusion, as the first integrated mRNA designing tool, mRNAdesigner offers researchers a powerful optimization platform for mRNA application development. Future refinement will continue to expand its capabilities in both therapeutic and industrial contexts, paving the way for broader applications of mRNA technology in biomedicine.

Acknowledgements

We hope to thank Chenhui Hao and Kaituo Yan for their suggestions on carrying out the project. Gratitude is also extended to Wenya Wang (at ChemPartner PharmaTech Co.) and Dr. Guoqing Zhang (at Bio-Med Big Data Center of Shanghai Institute of Nutrition and Health, CAS) for their help in managing the joint project and for hosting the web service.

Author contributions: Ouyang Mo (Investigation [equal], Software [lead], Visualization [equal], Writing—original draft [equal]), Zhuo Zhang (Investigation [equal], Methodology [equal], Writing—original draft [equal]), Xiang Cheng (Investigation [equal], Methodology [equal], Writing—original draft [equal], Writing—review & editing [equal]), Liqi Zhu (Methodology [equal]), Kaixiang Zhang (Methodology [equal]), Niubing Zhang (Investigation [equal], Methodology [equal]), Justin Li (Data curation [equal], Investigation [equal], Software [equal]), Honglin Li (Resources [equal]), Shixin Fan (Resources [equal]), Xuan Li (Conceptualization [equal], Writing—review & editing [equal]), and Pei Hao (Conceptualization [equal], Writing—review & editing [lead])

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

None declared.

Funding

This work was supported by Noncommunicable Chronic Diseases - National Science and Technology Major Project [2023ZD0502400] and National Natural Science Founda-

tion of China [92451303, 32270695, 32270719] and The Key R&D Program of Jiangsu Province Modern Agriculture [BE2023336]. Funding to pay the Open Access publication charges for this article was provided by Noncommunicable Chronic Diseases - National Science and Technology Major Project [2023ZD0502400].

Data availability

The mRNAdesigner web server and tutorial are freely available at https://www.biosino.org/mRNAdesigner/. This resource is freely available for academic and non-commercial research purposes and does not require registration.

References

1. Kiaie SH, Majidi Zolbanin N, Ahmadi A *et al.* Recent advances in mRNA-LNP therapeutics: immunological and pharmacological aspects. *J Nanobiotechnol* 2022;**20**:276. https://doi.org/10.1186/s12951-022-01478-7
2. Liu C, Shi Q, Huang X *et al.* mRNA-based cancer therapeutics. *Nat Rev Cancer* 2023;**23**:526–43. https://doi.org/10.1038/s41568-023-00586-2
3. To KKW, Cho WCS. An overview of rational design of mRNA-based therapeutics and vaccines. *Expert Opin Drug Discov* 2021;**16**:1307–17. https://doi.org/10.1080/17460441.2021.1935859
4. Sahin U, Karikó K, Türeci Ö. mRNA-based therapeutics—developing a new class of drugs. *Nat Rev Drug Discov* 2014;**13**:759–80. https://doi.org/10.1038/nrd4278
5. Sparmann A, Vogel J. RNA-based medicine: from molecular mechanisms to therapy. *EMBO J* 2023;**42**:e114760. https://doi.org/10.15252/embj.2023114760
6. Rohner E, Yang R, Foo KS *et al.* Unlocking the promise of mRNA therapeutics. *Nat Biotechnol* 2022;**40**:1586–600. https://doi.org/10.1038/s41587-022-01491-z
7. Weng Y, Li C, Yang T *et al.* The challenge and prospect of mRNA therapeutics landscape. *Biotechnol Adv* 2020;**40**:107534. https://doi.org/10.1016/j.biotechadv.2020.107534
8. Graf M, Deml L, Wagner R. Codon-optimized genes that enable increased heterologous expression in mammalian cells and elicit efficient immune responses in mice after vaccination of naked DNA. *Methods Mol Med* 2004;**94**:197–210. https://doi.org/10.1385/1-59259-679-7:197
9. Hanson G, Coller J. Codon optimality, bias and usage in translation and mRNA decay. *Nat Rev Mol Cell Biol* 2018;**19**:20–30. https://doi.org/10.1038/nrm.2017.91
10. Terai G, Kamegai S, Asai K. CDSfold: an algorithm for designing a protein-coding sequence with the most stable secondary structure. *Bioinformatics* 2016;**32**:828–34. https://doi.org/10.1093/bioinformatics/btv678
11. Wayment-Steele HK, Kim DS, Choe CA *et al.* Theoretical basis for stabilizing messenger RNA through secondary structure design. *Nucleic Acids Res* 2021;**49**:10604–17. https://doi.org/10.1093/nar/gkab764
12. Zhang H, Zhang L, Lin A *et al.* Algorithm for optimized mRNA design improves stability and immunogenicity. *Nature* 2023;**621**:396–403. https://doi.org/10.1038/s41586-023-06127-z
13. Nieuwkoop T, Finger-Bou M, van der Oost J *et al.* The ongoing quest to crack the genetic code for protein production. *Mol Cell* 2020;**80**:193–209. https://doi.org/10.1016/j.molcel.2020.09.014
14. Metkar M, Pepin CS, Moore MJ. Tailor made: the art of therapeutic mRNA design. *Nat Rev Drug Discov* 2024;**23**:67–83. https://doi.org/10.1038/s41573-023-00827-x
15. Narula A, Ellis J, Taliaferro JM *et al.* Coding regions affect mRNA stability in human cells. *RNA* 2019;**25**:1751–64. https://doi.org/10.1261/rna.073239.119
16. Forrest ME, Pinkard O, Martin S *et al.* Codon and amino acid content are associated with mRNA stability in mammalian cells. *PLoS One* 2020;**15**:e0228730, https://doi.org/10.1371/journal.pone.0228730
17. Leppek K, Byeon GW, Kladwang W *et al.* Combinatorial optimization of mRNA structure, stability, and translation for RNA-based therapeutics. *Nat Commun* 2022;**13**:1536. https://doi.org/10.1038/s41467-022-28776-w
18. Mauger DM, Cabral BJ, Presnyak V *et al.* mRNA structure regulates protein expression through changes in functional half-life. *Proc Natl Acad Sci USA* 2019;**116**:24075–83. https://doi.org/10.1073/pnas.1908052116
19. Zhang L, More KR, Ojha A *et al.* Effect of mRNA-LNP components of two globally-marketed COVID-19 vaccines on efficacy and stability. *NPJ Vaccines* 2023;**8**:156. https://doi.org/10.1038/s41541-023-00751-6
20. Kirshina A, Vasileva O, Kunyk D *et al.* Effects of combinations of untranslated-region sequences on translation of mRNA. *Biomolecules* 2023;**13**:1677. https://doi.org/10.3390/biom13111677
21. Chen J, Chen J, Xu Q. Current developments and challenges of mRNA vaccines. *Annu Rev Biomed Eng* 2022;**24**:85–109. https://doi.org/10.1146/annurev-bioeng-110220-031722
22. Wu Q, Medina SG, Kushawah G *et al.* Translation affects mRNA stability in a codon-dependent manner in human cells. *eLife* 2019;**8**:e45396. https://doi.org/10.7554/eLife.45396
23. Sample PJ, Wang B, Reid DW *et al.* Human 5′ UTR design and variant effect prediction from a massively parallel translation assay. *Nat Biotechnol* 2019;**37**:803–9. https://doi.org/10.1038/s41587-019-0164-5
24. Cunningham F, Allen JE, Allen J *et al.* Ensembl 2022. *Nucleic Acids Res* 2022;**50**:D988–95. https://doi.org/10.1093/nar/gkab1049
25. Castillo-Hair SM, Seelig G. Machine learning for designing next-generation mRNA therapeutics. *Acc Chem Res* 2022;**55**:24–34. https://doi.org/10.1021/acs.accounts.1c00621
26. Chu Y, Yu D, Li Y *et al.* A 5′ UTR language model for decoding untranslated regions of mRNA and function predictions. *Nat Mach Intell* 2024;**6**:449–60. https://doi.org/10.1038/s42256-024-00823-9
27. Bicknell AA, Reid DW, Licata MC *et al.* Attenuating ribosome load improves protein output from mRNA by limiting translation-dependent mRNA decay. *Cell Rep* 2024;**43**:114098. https://doi.org/10.1016/j.celrep.2024.114098
28. Lorenz R, Bernhart SH, Höner Zu Siederdissen C *et al.* ViennaRNA package 2.0. Algorithms for molecular biology. *Algorithms Mol Biol* 2011;**6**:26. https://doi.org/10.1186/1748-7188-6-26
29. Hofacker IL. Vienna RNA secondary structure server. *Nucleic Acids Res* 2003;**31**:3429–31. https://doi.org/10.1093/nar/gkg599
30. Xia X. Detailed dissection and critical evaluation of the Pfizer/BioNTech and Moderna mRNA vaccines. *Vaccines* 2021;**9**:734. https://doi.org/10.3390/vaccines9070734
31. Fang E, Liu X, Li M *et al.* Advances in COVID-19 mRNA vaccine development. *Sig Transduct Target Ther* 2022;**7**:94. https://doi.org/10.1038/s41392-022-00950-y
32. Orlandini von Niessen AG, Poleganov MA, Rechner C *et al.* Improving mRNA-based therapeutic gene delivery by expression-augmenting 3′ UTRs identified by cellular library screening. *Mol Ther* 2019;**27**:824–36. https://doi.org/10.1016/j.ymthe.2018.12.011
33. Matoulkova E, Michalova E, Vojtesek B *et al.* The role of the 3′ untranslated region in post-transcriptional regulation of protein expression in mammalian cells. *RNA Biol* 2012;**9**:563–76. https://doi.org/10.4161/rna.20231
34. Koh WS, Porter JR, Batchelor E. Tuning of mRNA stability through altering 3′-UTR sequences generates distinct output expression in a synthetic circuit driven by p53 oscillations. *Sci Rep* 2019;**9**:5976. https://doi.org/10.1038/s41598-019-42509-y

35. Lo Giudice C, Zambelli F, Chiara M *et al.* UTRdb 2.0: a comprehensive, expert curated catalog of eukaryotic mRNAs untranslated regions. *Nucleic Acids Res* 2023;**51**:D337–44. https://doi.org/10.1093/nar/gkac1016
36. Bailey TL, Johnson J, Grant CE *et al.* The MEME Suite. *Nucleic Acids Res* 2015;**43**:W39–49. https://doi.org/10.1093/nar/gkv416
37. Zhang N, Cheng X, Zhu Y *et al.* Multi-valent mRNA vaccines against monkeypox enveloped or mature viron surface antigens demonstrate robust immune response and neutralizing activity. *Sci China Life Sci* 2023;**66**:2329–41. https://doi.org/10.1007/s11427-023-2378-x
38. Zhang N, Jing X, Liu Y, Chen M, Zhu X, Jiang J, Wang H, Li X, Hao P Interfering with retrotransposition by two types of CRISPR effectors: Cas12a and Cas13a. *Cell Discov* (2020) **6**30
39. McLellan JS, Chen M, Joyce MG *et al.* Structure-based design of a fusion glycoprotein vaccine for respiratory syncytial virus. *Science* 2013;**342**:592–8. https://doi.org/10.1126/science.1243283
40. Che Y, Gribenko AV, Song X *et al.* Rational design of a highly immunogenic prefusion-stabilized F glycoprotein antigen for a respiratory syncytial virus vaccine. *Sci Transl Med* 2023;**15**:eade6422. https://doi.org/10.1126/scitranslmed.ade6422
41. Huang K, Lawlor H, Tang R *et al.* Recombinant respiratory syncytial virus F protein expression is hindered by inefficient nuclear export and mRNA processing. *Virus Genes* 2010;**40**:212–21. https://doi.org/10.1007/s11262-010-0449-8
42. Thiel BC, Beckmann IK, Kerpedjiev P *et al.* 3D based on 2D: calculating helix angles and stacking patterns using forgi 2.0, an RNA Python library centered on secondary structure elements. *F1000Res* 2019;**8**:ISCB Comm J-287. https://doi.org/10.12688/f1000research.18458.2
43. Pavanello L, Hall M, Winkler GS. Regulation of eukaryotic mRNA deadenylation and degradation by the Ccr4-Not complex. *Front Cell Dev Biol* 2023;**11**:1153624. https://doi.org/10.3389/fcell.2023.1153624
44. Heinrich S, Sidler CL, Azzalin CM *et al.* Stem-loop RNA labeling can affect nuclear and cytoplasmic mRNA processing. *RNA* 2017;**23**:134–41. https://doi.org/10.1261/rna.057786.116
45. Takanashi A, Pouton CW, Al-Wassiti H. Delivery and expression of mRNA in the secondary lymphoid organs drive immune responses to lipid nanoparticle-mRNA vaccines after intramuscular injection. *Mol Pharm* 2023;**20**:3876–85. https://doi.org/10.1021/acs.molpharmaceut.2c01024
46. Komar AA, Samatova E, Rodnina MV. Translation rates and protein folding. *J Mol Biol* 2024;**436**:168384. https://doi.org/10.1016/j.jmb.2023.168384
47. Kames J, Alexaki A, Holcomb DD *et al.* TissueCoCoPUTs: novel human tissue-specific codon and codon-pair usage tables based on differential tissue gene expression. *J Mol Biol* 2020;**432**:3369–78. https://doi.org/10.1016/j.jmb.2020.01.011
48. Tang X, Huo M, Chen Y *et al.* A novel deep generative model for mRNA vaccine development: designing 5′ UTRs with N^1-methyl-pseudouridine modification. *Acta Pharm Sin B* 2024;**14**:1814–26. https://doi.org/10.1016/j.apsb.2023.11.003
49. Cao J, Novoa EM, Zhang Z *et al.* High-throughput 5′ UTR engineering for enhanced protein production in non-viral gene therapies. *Nat Commun* 2021;**12**:4138. https://doi.org/10.1038/s41467-021-24436-7

Received: March 19, 2025. **Revised:** April 19, 2025. **Editorial Decision:** April 26, 2025. **Accepted:** May 2, 2025