ACCOUNTS of chemical research

pubs.acs.org/accounts Article

Therapeutic mRNA Engineering from Head to Tail

Published as part of the Accounts of Chemical Research special issue "mRNA Therapeutics".

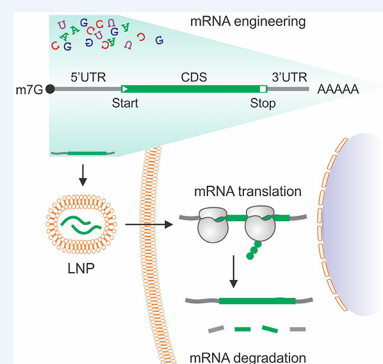
Longfei Jia and Shu-Bing Qian*

Cite This: *Acc. Chem. Res.* 2021, 54, 4272−4282

Read Online

ACCESS | Metrics & More | Article Recommendations

Downloaded via CORNELL UNIV on January 3, 2022 at 04:47:57 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.



CONSPECTUS: Synthetic messenger RNA (mRNA), once delivered into cells, can be readily translated into proteins by ribosomes, which do not distinguish exogenous mRNAs from endogenous transcripts. Until recently, the intrinsic instability and immunostimulatory property of exogenous RNAs largely hindered the therapeutic application of synthetic mRNAs. Thanks to major technological innovations, such as introduction of chemically modified nucleosides, synthetic mRNAs have become programmable therapeutic reagents. Compared to DNA or protein-based therapeutic reagents, synthetic mRNAs bear several advantages: flexible design, easy optimization, low-cost preparation, and scalable synthesis. Therapeutic mRNAs are commonly designed to encode specific antigens to elicit organismal immune response to pathogens like viruses, express functional proteins to replace defective ones inside cells, or introduce novel enzymes to achieve unique functions like genome editing. Recent years have witnessed stunning progress on the development of mRNA vaccines against SARS-Cov2. This success is built upon our fundamental understanding of mRNA metabolism and translational control, a knowledge accumulated during the past several decades. Given the astronomical number of sequence combinations of four nucleotides, sequence-dependent control of mRNA translation remains incompletely understood. Rational design of synthetic mRNAs with robust translation and optimal stability remains challenging. Massively paralleled reporter assay (MPRA) has been proven to be powerful in identifying sequence elements in controlling mRNA translatability and stability. Indeed, a completely randomized sequence in 5′ untranslated region (5′UTR) drives a wide range of translational outputs. In this Account, we will discuss general principles of mRNA translation in eukaryotic cells and elucidate the role of coding and noncoding regions in the translational regulation. From the therapeutic perspective, we will highlight the unique features of 5′ cap, 5′UTR, coding region (CDS), stop codon, 3′UTR, and poly(A) tail. By focusing on the design strategies in mRNA engineering, we hope this Account will contribute to the rational design of synthetic mRNAs with broad therapeutic potential.

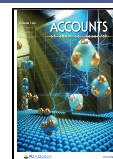
■ KEY REFERENCES

- Jia, L.; Mao, Y.; Ji, Q.; Dersh, D.; Yewdell, J.W.; Qian, S. B. Decoding mRNA translatability and stability from the 5′UTR. *Nat. Struct. Mol. Biol.* **2020**, 27(9), 814−821.[1] *This work describes a massively parallel RNA-based reporter system that reveals the function of cis-elements in 5′UTR and unravels the interplay between translational outputs and mRNA stability.*
- Liu, X. M.; Zhou, J.; Mao, Y., Ji, Q.; Qian, S. B. Programmable RNA *N6*-methyladenosine editing using CRISPR/Cas9 conjugates. *Nat. Chem. Biol.* **2019**, 15(9), 865−871.[2] *This work describes the CRISPR-based technology to achieve programmable* m^6A *editing in mRNA and the impacts on translation and degradation.*

■ INTRODUCTION

Messenger RNA (mRNA) is a single-stranded polynucleotide chain carrying protein-coding information from genome. Over the past decades, major technological innovations have enabled synthetic mRNA as a feasible therapeutic agent.[3] Compared to the conventional drugs, mRNA can be rapidly designed, cost-effectively manufactured, safely delivered, and efficiently translated into proteins inside cells. With the recent approval of mRNA-based vaccines targeting COVID-19, mRNA therapeutics has encompassed a dramatic rise in the market and triggered widespread interest in a broad range of human diseases. Great effort has been dedicated to mRNA modification to tackle the unwanted immunogenicity, mRNA production for large-scale deployment, and nanoparticle encapsulation for efficient delivery. However, much less attention has been focused toward mRNA sequence

Received: August 30, 2021
Published: November 10, 2021



ACS Publications © 2021 American Chemical Society

https://doi.org/10.1021/acs.accounts.1c00541
Acc. Chem. Res. 2021, 54, 4272−4282

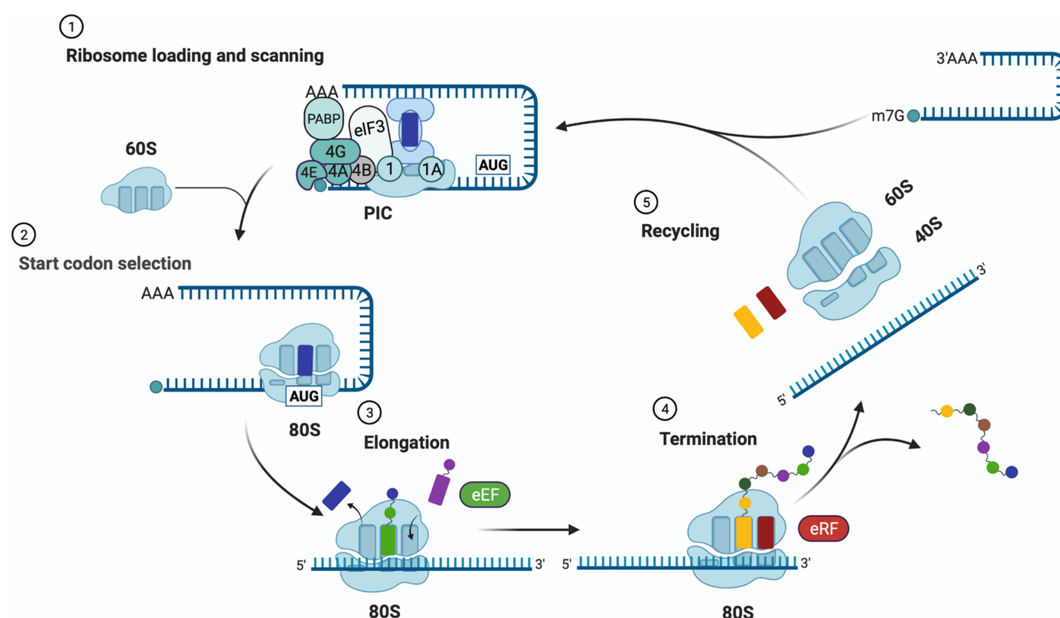


Figure 1. Schematic of the mRNA translation pathway in eukaryotic cells. PIC, preinitiation complex. 4A, eukaryotic initiation factor 4A. 4E, eukaryotic initiation factor 4E. 4G, eukaryotic initiation factor 4G. 4B, eukaryotic initiation factor 4B. eIF3, eukaryotic initiation factor 3. 1, eukaryotic initiation factor 1. 1A, eukaryotic initiation factor 1A. PABP, poly(A)-binding protein. eEF, eukaryotic elongation factor. eRF, eukaryotic release factor.

optimization. While our knowledge of mRNA translation is steadily increasing, how to design an mRNA with optimal stability and translatability remains largely unexplored.

Much like the endogenous mRNA, a therapeutic mRNA encompasses a cap structure, 5′UTR, CDS, 3′UTR, and a poly(A) tail. Such division of mRNA regions depends on the presence of start and stop codons, which are recognized by the translation machinery ribosome inside cells. The fidelity of start and stop codon recognition involves the sequence context and translation factors. Therefore, there is an urgent need to understand how mRNA sequence elements interact with ribosomes as well as other RNA-binding proteins. Another important factor in mRNA engineering is the intracellular stability that is inherently coupled with the translational output. Here, we start with an overview of eukaryotic mRNA translation followed by a general description of mRNA design principles by focusing on sequence elements influencing mRNA translatability and stability.

OVERVIEW OF EUKARYOTIC mRNA TRANSLATION

The translation of most eukaryotic mRNAs relies on a 7-Methylguanosine (m^7G) cap to load the 40S small ribosomal subunit to the mRNA, followed by a scanning process to position initiator methionine tRNA (Met-$tRNA_i^{Met}$) within the ribosome at the start codon (Figure 1).[4] This canonical cap-dependent scanning pathway commences with the assembly of the 43S preinitiation complex (PIC), which attaches to the 5′ proximal region of m^7G capped mRNAs in a manner facilitated by the eIF4F complex. eIF4F consists of a cap-binding protein eIF4E, a scaffolding protein eIF4G, and RNA helicase eIF4A. After mRNA attachment, 43S PIC scans the 5′UTR of the mRNA until the appropriate start codon appears in the P-site of 43S PIC. The start codon recognition is followed by the joining of a large ribosomal subunit (60S), forming an 80S complex competent for elongation. Notably, the stringency of start codon recognition can be influenced by the codon context as well as initiation factors, although the precise mechanism remains incompletely understood.

During translation elongation, the ribosome actively moves along the CDS of mRNA, using tRNAs to synthesize polypeptides. Translation elongation is mediated by elongation factors eEF1 and eEF2, which delivers amino acid charged tRNA to the ribosomal A site and catalyzes ribosomal translocation, respectively.[5] During elongation, the ribosome does not move at a constant speed but rather in a stop-and-go traffic manner. Both *cis* sequence elements and *trans* regulatory factors contribute to the variations of elongation speed. Although the elongation speed directly controls the translational output, our understanding of elongation control has lagged behind the knowledge of initiation regulation.

Termination is triggered when the stop codon is reached and recognized by eukaryotic release factors.[6] Subsequently, the ribosome recycling factor mediates dissociation of ribosome from mRNA and tRNA. In some cases, the 40S subunit remains associated with mRNA and could start a second round of translation from the downstream start codon, a process called reinitiation. The molecular mechanism underlying translation reinitiation remains poorly understood.

5′ END CAP

Cap Structure

The 5′ end m^7G cap structure presents in nearly all eukaryotic cellular mRNAs. The 5′-terminal G with a N^7-methyl group is connected to the first nucleotide of mRNA through a 5′-5′ linkage via two pyrophosphoryl bonds (Figure 2). This capping process is achieved by several enzymes, including guanylyltransferase and methyltransferase. In addition to the 5′ end-cap (referred to as cap-0), the first nucleotides of the mRNA can also be methylated on the 2′ hydroxyl to form cap-1 (m^7GpppN1 mN2), or the first two nucleotides are methylated to form cap-2 (m^7GpppN1 mN2m). Since the N1 nucleotide

https://doi.org/10.1021/acs.accounts.1c00541
Acc. Chem. Res. 2021, 54, 4272−4282

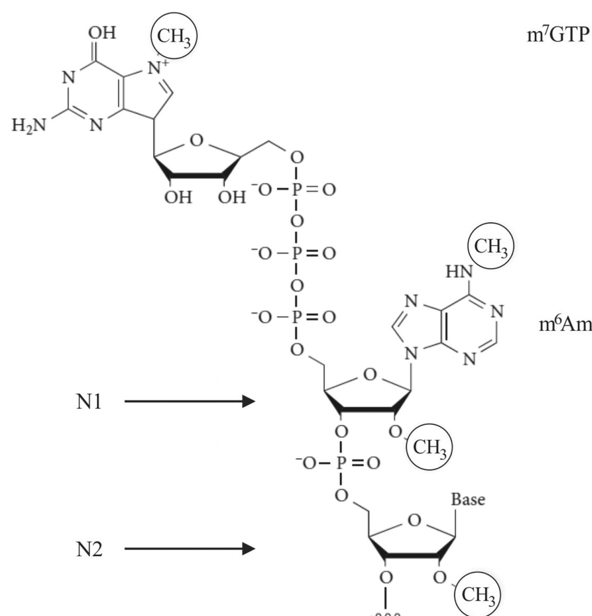


Figure 2. Cap structure of eukaryotic mRNA. The methyl group is highlighted in circle. Cap-1 and cap-2 structures are indicated as 2′-*O*-methyl groups present at N1 and N2 position, respectively.

in an mRNA is frequently an adenosine, it can be additionally methylated at the N^6 position, forming m^6Am. Recent studies demonstrated that m^6Am is mediated by phosphorylated CTD interacting factor 1 (PCIF1) in an m^7G-dependent manner. Intriguingly, m^6Am can be demethylated by FTO, suggesting that the 5′ end-cap modification is more dynamic than previously thought.[7]

Cap Function

The 5′ end-cap of mRNA is critical for splicing, poly-adenylation, mRNA stability, and translation.[8] As a hallmark of regulation of eukaryotic protein synthesis, the m^7G cap structure is recognized by different cap-binding proteins that act to impart disparate functional outcomes on gene expression. Newly synthesized mRNAs with the cap in the nucleus are bound by the cap-binding protein (CBP) heterodimer CBP80-CBP20, which regulates the mRNA transport from the nucleus to the cytoplasm. In addition to promoting pre-mRNA processing, this complex monitors mRNA quality during the pioneer round of translation. In the cytosol, the cap structure recruits the eukaryotic translation initiation factor 4E (eIF4E) to promote ribosome recognition and translation initiation. The cap-binding protein eIF4E is a component of the enzyme complex eIF4F, which contains the scaffold protein eukaryotic initiation factor 4G (eIF4G) that subsequently recruits the eIF3-associated 40S ribosomal subunit (Figure 1). The binding affinity of eIF4E for the capped mRNA is enhanced in the presence of eIF4G as part of the eIF4F complex. It is believed that the eIF4F complex assembly on the 5′ cap of mRNA is rate limiting for translation initiation. Structural studies suggest that eIF4E recognizes m^7G but not the N1 nucleotide. Other proteins may recognize the methylated states of N1 and N2 nucleotides.

2′-*O*-Methylation of the 5′penultimate and antepenultimate nucleoside is the most common modification of the first transcribed nucleotide and provides a molecular signature for distinguishing self-and nonself mRNAs.[9] Uncapped transcripts or mRNAs with the cap 0 structure can be recognized by the innate immune receptor RIG-1.[10] Therefore, capped mRNAs tend to escape the recognition by innate immune sensors. Given the diversity of cap modifications, how chemically modified cap derivatives avoid immune stimulation is of considerable interest for therapeutic mRNA engineering.

The 5′ cap structure is known to protect mRNA from 5′ → 3′ degradation mediated by the exonucleases Xrn1 in the cytoplasm and Xrn2 in the nucleus. As a result, capped transcripts show a longer half-life than uncapped counterparts. By promoting translation and increasing the stability of mRNA, the cap structure evidently enhances the yield of protein synthesis inside cells. Notably, the cap structure also involves mRNA decay owing to its capability to bind with cap-specific mRNA-decapping pyrophosphatase, Dcp1, Dcp2, and DcpS.[11] Whether the cap methylation status affects the mRNA decapping process remains to be determined.

Capping Strategy

For mRNAs synthesized by in vitro transcription (IVT), two approaches are widely used to add a 5′ cap (Figure 3). In the

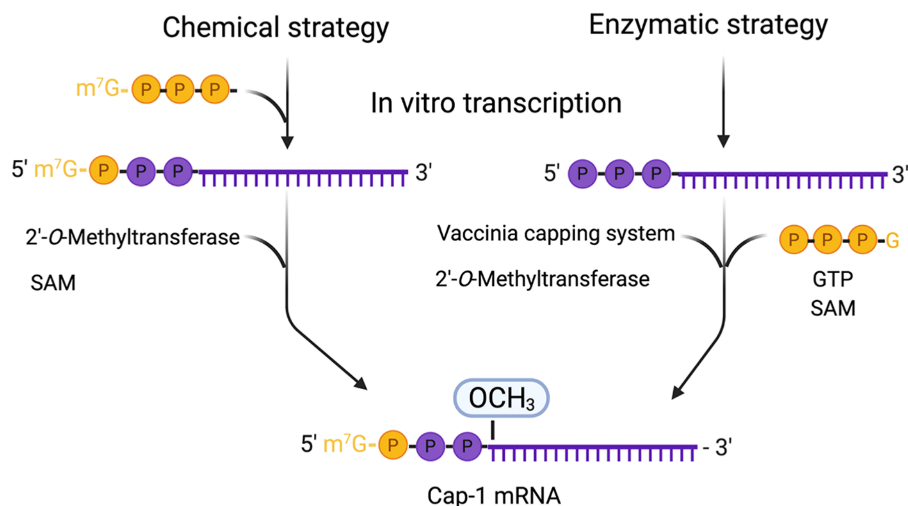


Figure 3. An illustrative graph shows the general capping approach by taking advantage of chemical strategy (left) or enzymatic strategy (right). SAM, S-adenosyl-methionine.

https://doi.org/10.1021/acs.accounts.1c00541
Acc. Chem. Res. 2021, 54, 4272−4282

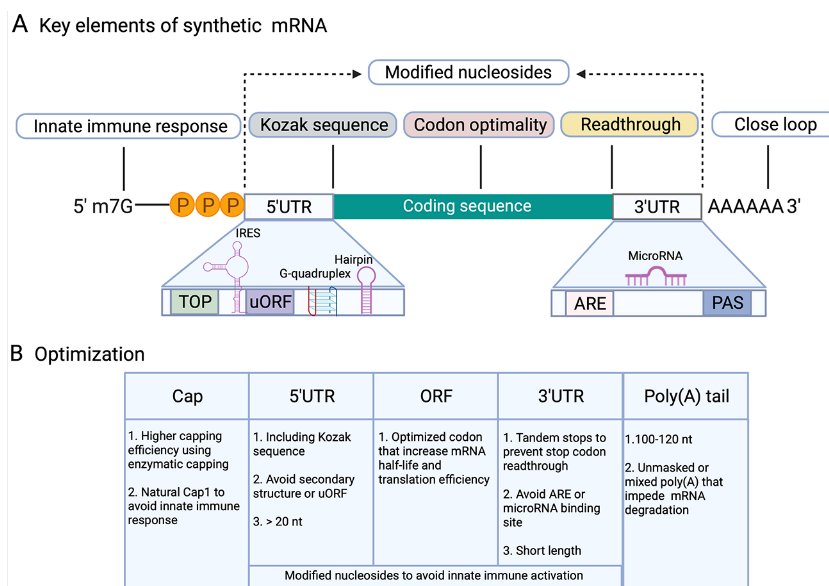


Figure 4. (A) Schematic representation of a synthetic mRNA and its key structural elements. (B) Strategies for possible optimizations by modifying these structural elements to modulate the stability, the translation efficiency as well as the intrinsic immunogenicity. 5′UTR, 5′ untranslated region. 3′UTR, 3′untranslated region. ORF, open reading frame. TOP, the terminal oligopyrimidine. IRES, internal ribosome entry sites. uORF, upstream open reading frame. ARE, adenylate uridylate-rich element. PAS, polyadenylation site.

first method, referred to as a chemical strategy, a synthetic cap analog is introduced cotranscriptionally along with the other four nucleotides during RNA synthesis. However, m^7G incorporated as cap 0 during IVT will result in two different orientations: normal $m^7GpppGpN$ and reversed $Gpppm^7GpN$. To prevent reverse incorporation, antireverse cap analogs (ARCA) have been developed. With the 3′-OH group next to m^7G replaced by $-OCH_3$, ARCA prevents m^7G in the reverse orientation by ensuring the addition of a nucleotide only at the nonmethylated guanosine during IVT.[12] ARCA capped mRNA revealed superior translation efficiency along with prolonged protein expression in a variety of cell types.[13] Due to the inherent competition with standard GTP for the RNA polymerase, ARCA incorporation is not complete (~80%). Additionally, direct incorporation of cap analogs does not give rise to cap1 or cap2 structures. The newly developed CleanCap technology appears to incorporate cap1 or cap2 during cotranscriptional capping with high efficiency.[14]

A second commonly used approach adds the m^7G cap to the 5′-triphosphate ends of RNA via vaccinia virus capping enzymes (VCE). This post-transcriptional approach is referred to as the enzymatic strategy. The VCE possesses guanylyltransferase and methyltransferase activity, thereby adding a 2′-*O*-methyl group to the penultimate nucleotide by 2′-*O*-methyltransferase in a single step. The resultant cap 1 structure mimics the naturally occurring cap structure in eukaryotic cells. Despite such an advantage, VCE-mediated large scale capped RNA production requires additional purification steps.

In addition to commonly used cap analogs, numerous chemically modified cap derivatives were developed to increase the binding affinity of cap to eIF4E and inhibit decapping to extend the RNA half-lives. These noncanonical modifications include triphosphate linkage through a bridging oxygen, O-to-CH_2 methylene-bisphosphonate, O-to-NH imidodiphosphate, nonbridging oxygen, O-to-S phosphorothioate, O-to-BH_3 boranophosphate, and lock nucleic acid (LNA).[15,16] In vivo experiments showed that mRNA with O−S substitutions induced great immune responses than those induced by mRNAs capped with standard ARCA.[17] Further studies are needed to demonstrate the feasibility of these non-natural can analogs in therapeutic mRNAs.

5′ UTR

5′UTR Features. Once the 43S PIC loads onto the eukaryotic mRNAs, it follows the scanning mechanism wherein the PIC migrates along the leader (5′UTR) of mRNAs and continuously inspects the initiation codon (AUG) using complementarity with the anticodon of Met-$tRNA_i^{Met}$. Regulation of translation initiation in eukaryotic mRNAs is primarily conferred by 5′ UTR features including secondary structures, sequence elements, and 5′UTR length (Figure 4A).[18] In eukaryotes, the average length of 5′ UTR ranges from ~100 nt to ~200 nt in mammals and ~50 nt in yeast. Although ultrashort 5′UTR is capable of translation initiation, it is generally believed that a minimal length of 20 nt is required for efficient recognition of a start codon by the ribosome.

The terminal oligopyrimidine (TOP) motif is a *cis*-regulatory RNA element located immediately downstream of the m^7G cap.[19] Comprising an invariant 5′-cytidine followed by an uninterrupted tract of 4−14 pyrimidine nucleotides, 5′ UTRs of TOP mRNAs are relatively short (40 nt on average) and free from stable secondary structures. TOP mRNAs mostly encode translation factors and nearly all ribosomal proteins. The majority of TOP mRNA are highly sensitive to the mammalian target of the rapamycin complex 1 (mTORC1) signaling pathway.

mRNA has the intrinsic propensity to fold and form a higher-order structure in a sequence-dependent manner, thereby displaying a second layer of structural information besides the primary sequence. Secondary structures in the 5′UTR affect the efficiency of translation initiation and actively participate in the regulation of protein synthesis. Complex RNA structures in 5′ UTRs, such as RNA G-quadruplexes and hairpins, serve as road blocks by impeding ribosome loading

https://doi.org/10.1021/acs.accounts.1c00541
Acc. Chem. Res. 2021, 54, 4272−4282

and the scanning process.[1] However, translation could be favored by stem-loop structures functioning as translation enhancers or activators by binding to eIF3 and the associated ribosome.[20]

The translation-promoting element in 5′UTR is reminiscent of internal ribosome entry sites (IRES) that directly recruit the 40S ribosome without recognition of the m^7G cap.[21] This cap-independent mechanism enables mRNA translation when the cap-dependent translation is suppressed. As most cellular IRES-containing mRNAs also bear the m^7G cap and can be translated by the canonical initiation mechanism, one intriguing question is how to control the switch between different modes of initiation for these IRES-containing mRNAs.

Start Codons and uORFs. Although an AUG codon is predominantly utilized to initiate mRNA translation, not all the AUG triplets are equally selected. In vertebrates, a strong start codon usually contains Kozak consensus sequence gccRcc<u>AUG</u>G, in which the start AUG is surrounded by highly conserved nucleotides. A high-throughput sequencing experiment indicated that −3 and the +4 positions are critical in determining start codon efficiency, while the −2, − 4, and +5 positions are auxiliary.[22] A large-scale analysis of human mRNAs indicates that less than half of mRNAs possess an annotated start codon in the optimal Kozak context.[23] Notably, a growing body of evidence suggests that translation initiation occurs at codons different from AUG. Among non-AUG start codons, CUG, GUG, UUG, and ACG are the most commonly used near-cognate codons, albeit at a much lower efficiency.

When a single mRNA contains multiple potential start codons, the transcript becomes polycistronic by encoding more than one open reading frames (ORFs). About 45−50% of mammalian genes encode mRNAs that have at least one short uORF upstream of the main coding region. Ribosome profiling experiments confirmed widespread translation at many of these uORFs. Generally, uORF translation functions as a physical barrier to impede ribosomal movement to the downstream CDS by blocking the scanning 43S ribosome or sequestering 80S ribosomes. However, in some cases, short uORF permits translation reinitiation, which often occurs in response to stress conditions. The best-characterized example is the activating transcription factor 4 (ATF4) in mammals.[24]

5′UTR Design

To achieve maximum expression of proteins encoded by therapeutic mRNA, various 5′UTRs have been compared. The UTRs of *β*-globin are popular sequences that have been utilized in clinical trials as well as basic research.[25] To minimize the scanning process, a shorter 5′UTR with at least 20 nt is recommended. Additionally, potential upstream start codons (especially AUG) should be avoided to eliminate the presence of uORF that often negatively regulates the translation of the main CDS. uORF translation also triggers mRNA decay, thereby further reducing the translational output. To promote the main CDS translation, a strong start codon with the consensus sequence context should be included to prevent leaky scanning. Lastly, it is important to avoid the presence of highly stable secondary structures especially near the 5′ end, which can disrupt ribosome loading as well as scanning and start codon selection (Figure 4B).

Depending on the therapeutic purposes, additional sequence elements could be introduced into 5′UTR to enable selective translation. For instance, in cancer therapy, mRNAs for intratumor injection may need special 5′UTR elements capable of translation under nutrient restriction. The 5′UTR performance is likely dependent on species, cell type, and cell state. Although bioinformatic tools can be helpful in predicting optimal 5′UTR sequences, nonbiased experimental assays like massively parallel reporter assay (MPRA) will be instrumental in uncovering 5′UTR sequences for maximal protein output.[26,27]

In an effort to dissect how the *cis* element in 5′UTR regulates mRNA translatability and stability, we employed MRPA by replacing the start codon with a 10 nt random sequence to create an mRNA reporter library with over one million variants. The result echoes the notion that the start codon sequence context determines the translation initiation efficiency. As expected, a strong uORF translation triggers mRNA decay, whereas active translation of the main ORF stabilizes the messenger. The presence of G quadraplex structures in 5′UTR hinders ribosome scanning and induces mRNA degradation in the P-body. Intriguingly, a short stretch of A residues within the 5′UTR enables cap-independent mRNA translation. By differential sequence combination, it is thus possible to fine-tune translational output.

CDS

Codon Optimality. Due to the degeneracy of the genetic code, each amino acid is often encoded by multiple synonymous codons except for methionine and tryptophan. Synonymous codons are decoded by cognate tRNAs with different efficiencies and present in unequal frequencies. It is generally believed that codon usage bias is a universal feature across species, genomes, and individual genes. Notably, the levels of tRNA gene copy numbers or cognate tRNAs are correlated with codon usage. Hence, codons can be classified into optimal and nonoptimal codons on the basis of tRNA abundance. It is commonly believed that optimal codons are decoded faster and more accurately than nonoptimal codons by the elongating ribosome. Consistent with this notion, highly expressed genes tend to contain optimal codons. By contrast, nonoptimal codons are speculated to slow down the rate of translation elongation.

In addition to affecting the translation speed, codon usage also influences protein production by regulating mRNA stability. Rare codons cause ribosome stalling, which may trigger the ribosome-mediated quality control pathway (RQC) that mediates ribosome rescue, nascent peptide degradation, and mRNA decay. Indeed, recent studies demonstrated that ribosome collision is critical in triggering RQC and mRNA cleavage.[28] Interestingly, emerging evidence has suggested that codon optimality directly impacts mRNA stability, albeit the underlying mechanism remains debatable.[29] More surprisingly, codon usage bias has been shown to influence gene expression at the transcription level, such as chromatin structure, premature transcription termination, and splicing.

The codon optimality is not limited to a single codon. In yeast cells, certain codon pairs such as CGA-CGA, CGA-CCG, and CUC-CCG inhibit translation via an interplay between tRNAs at adjacent sites in the ribosome.[30] Whether this feature applies to mammalian cells remains unknown. Apparently, our understanding of codon choices in functional consequences is far from complete.

CDS Optimization. The codon degeneracy offers multiple choices in designing different CDS sequences encoding the same desired protein. Replacing rare codons with more

https://doi.org/10.1021/acs.accounts.1c00541
Acc. Chem. Res. 2021, 54, 4272−4282

abundant and frequently used codons based on tRNA preferences is termed codon optimization. It has been shown that high guanine and cytosine (G and C) content increases mRNA stability and translation efficiency compared to their adenine and thymine (A and U) rich counterparts.[31,32] As a result, optimization of GC content in the ORF with concurrent uridine depletion is a widely applied strategy in therapeutic mRNA design.[33,34] Codon-optimized mRNAs have been successfully employed in mRNA-based therapeutics, such as vaccines against viral infections and the expression of nonviral proteins.[35,36] In addition to improving the elongation rate and translational efficiency, codon optimization may alter RNA secondary structures and other features that interfere with gene expression. Many bioinformatic tools are available to assess codon adaptation index (CAI), GC-rich property, and RNA structure prediction.[37] It is important to know that a higher elongation speed is not always beneficial. Both the quality and quantity of translational products need to be carefully considered in therapeutic mRNA engineering.

Modified nucleosides have been widely used in therapeutic mRNAs to mitigate immune response mediated by RIG-1.[38] Commonly used nucleoside derivatives include 5-methylcytidine (m^5C), pseudouridine (Ψ), and *N*(1)-methylpseudouridine ($m^1\Psi$).[39−41] Notably, many of these modified nucleosides have relatively lower incorporation efficiency by T7 RNA polymerase during IVT. Additionally, the decoding speed and fidelity could be affected when a particular codon contains these modifications.

Besides the main coding sequence, therapeutic mRNAs could contain additional sequences encoding N-terminal signal sequences/peptides. Also, desired proteins can be redirected through engineered signal sequences from a membrane-bound to secreted version.[42]

■ STOP CODON

Translation Termination at Stop Codons

Faithful mRNA translation includes ribosomal termination at the end of CDS, which requires the presence of termination codons UAA, UAG, or UGA. The absence of stop codons leads to extended translation into the poly(A) tail sequence, resulting in a poly lysine chain added to the newly synthesized protein. Translation of poly-A sequences causes ribosome stalling that initiates RQC and subsequent mRNA degradation. Thus, "non-stop" mRNAs are subjected to no-stop decay (NSD).[43] Alongside NSD, splicing errors often produce problematic mRNAs with premature stop codon (PTC) products.[44] Similar to uORF translation, the presence of PTC triggers nonsense-mediated decay (NMD), a translational surveillance pathway to mitigate deleterious products.

Stop Codon Readthrough

Translation termination in eukaryotes is an effective process, but not 100%. The potential stop codon readthrough varies in a rank of UGA > UAG > UAA.[45] Apart from the identity of a particular stop codon, both downstream and upstream sequences could influence the stop codon readthrough potential. Several studies consistently demonstrated that the nucleotide immediately following the termination codon exhibits the strongest influence on readthrough efficiency.[46,47] Similarly, the upstream sequence context surrounding the termination codon also affects the termination efficiency. For instance, the adenine adjacent to stop codons is associated with the highest readthrough, whereas uracil is associated with the lowest readthrough.[48]

Stop Codon Consideration

From the perspective of therapeutic mRNA design, it is important to prevent stop codon readthrough. Previous studies demonstrated that stop codons bearing pseudouridine tend to be skipped by ribosomes.[49] To minimize stop codon readthrough, a secondary in-frame stop codon can be included downstream of the primary stop codon. Tandem stop codons have been purported to limit the level of stop codon readthrough by providing a second chance of translation termination.[50]

3′UTR

3′UTR Features. Resembling 5′UTR, 3′UTR contains many regulatory elements that contribute profoundly to mRNA stability, subcellular localization, and translation efficiency. The 3′UTR length varies greatly, ranging from a few to thousand nucleotides. It is generally believed that a shorter 3′UTR makes the transcript more stable, presumably due to the loss of microRNA binding sites thereby escaping mRNA degradation. For endogenous transcripts, the 3′UTR length is primarily determined by the polyadenylation site (PAS). With the core sequence A(A/U)UAA, the PAS is located approximately 10−30 nt upstream of the 3′ end cleavage site. The presence of multiple PAS allows creation of various 3′UTRs by alternative polyadenylation (APA). Accumulating evidence suggests that APA occurs in a cell type-specific manner and is responsive to signaling pathways.[51,52]

3′ UTR contains many *cis*-regulatory elements that modulate mRNA localization, turnover, and translation. The most common *cis*-regulatory element is the adenylate uridylate (AU-rich) element (ARE), which is represented by the pentamer "AUUUA". The interaction with some AU-binding proteins (AUBPs) recruits the degradation machinery followed by a process termed ARE-mediated decay.[53] Apart from the ARE, there are many other *cis*-elements such as GU-rich elements (GRE), CU-rich elements (CURE), CA-rich elements (CAREs), GC-rich elements (GCREs), and iron responsive elements (IREs). Their functional consequence varies depending on the associated *trans*-acting factors.

Perhaps the best known 3′UTR *trans*-acting factor is microRNAs (miRNAs), which regulate the expression of more than 60% of human protein-coding genes via translational repression and/or mRNA degradation. miRNAs induce target RNA degradation based on interaction with complementary sequences on 3′UTRs.[54] The 5′ seed region contains 6 consecutive nucleotides in position 2−7 and can be predicted by bioinformatic tools.

3′UTR Design. For therapeutic mRNA engineering, the 3′UTR is commonly derived from *β*-globin mRNAs.[25,55] Additionally, numerous 3′UTRs have been examined for therapeutic mRNA application, such as the hepatitis B virus and bovine growth hormone, just to name a few.[56] High throughput approaches, such as massively paralleled assays and cellular library screening, have been developed for 3′UTR optimization.[57,58] Given the broad function of 3′UTR, it remains to be seen whether special *cis*-elements can be introduced to achieve localized translation and/or tunable mRNA turnover.

https://doi.org/10.1021/acs.accounts.1c00541
Acc. Chem. Res. 2021, 54, 4272−4282

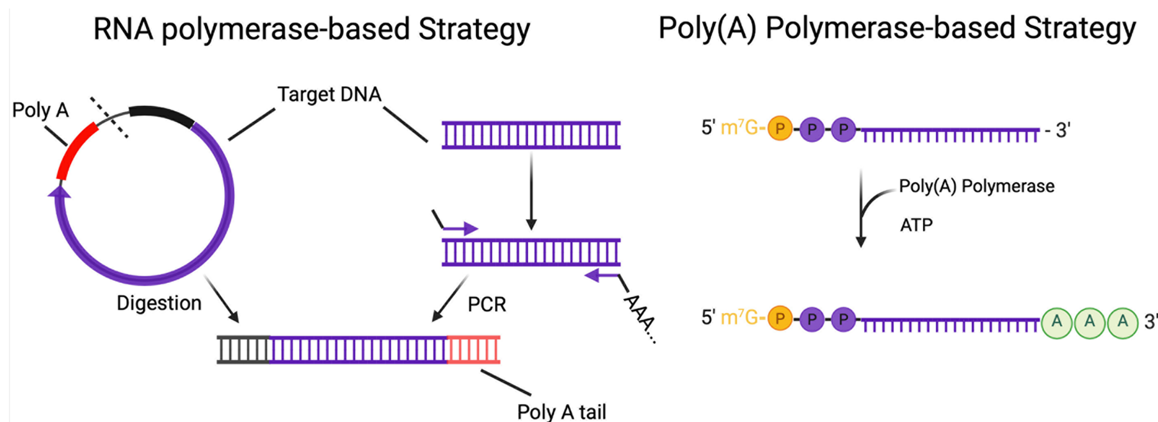


Figure 5. An illustrative graph shows the general poly(A) tailing approach using RNA polymerase-based strategy (left) and poly(A) polymerase-based strategy (right).

POLY(A) TAIL

Poly(A) Tail Features

A great portion of transcripts, except for histone mRNAs, bear poly(A) tails at the 3′ end. Followed by endonucleolytic cleavage of the pre-mRNA, the nontemplated poly(A) tail consists up to 200 to 250 adenosines. The polyadenylation of mRNA provides a binding site for multiple ribonucleoproteins (RNPs), such as poly(A) binding protein (PABP). The poly(A) tail and its associated proteins protect mRNA from enzymatic destruction in the nucleus while mediating mRNA degradation in cytosol. In addition to its function in mRNA turnover, the poly(A) tail also plays a crucial role in mRNA translation. The "close-loop" model of translation initiation in eukaryotes was first proposed based on the discovery of the association between the 5′ cap structure and 3′ poly(A) binding proteins (Figure 1). Later studies validated that this association is achieved through the interactions between PABP and eIF4G. The "close-loop" model not only explains the synergistic effect of the cap structure and the poly(A) tail in promoting translation initiation but also allows direct recycling of 40S ribosomes upon termination.

Poly(A) Tailing Strategy

Two main strategies have been developed to add the poly(A) tail to mRNAs synthesized by IVT (Figure 5). The first strategy relies on the RNA polymerase to transcribe the poly(A) sequence from the DNA template. Using a DNA template to transcribe the poly(A) tail in vitro yields mRNAs with a defined poly(A) tail length. This one-step procedure is thus preferred in clinical applications and the manufacturing industry. However, the encoded poly(A) tail length might be limited because a circular plasmid containing a long homopolymeric stretch of A is unstable during amplification in *E. coli*. It could also result in a masked poly(A) tail because cleavage by standard restriction enzymes leaves a few non-A nucleotides. To overcome this obstacle, type IIS restriction enzymes can be used because their cleavage sites and recognition sequences are separated.[59] Alternatively, an IVT template can be generated by PCR using a primer upstream of the promoter and a reverse primer containing a poly(T) sequence following the 3′ end of the inset.

The second method uses the poly(A) polymerase to add the poly(A) sequence to the 3′ end of mRNAs after IVT. The recombinant poly(A) polymerase enables the incorporation of various nucleoside analogues into the poly(A) tail, but the tailed mRNA is heterogeneous in the poly(A) length.

Poly(A) Tail Design. A suitable length of poly(A) tail is critical for its binding capacity with PABP. Exogenously delivered mRNAs with at least 16 nt of the poly(A) tail are needed for efficient translation.[60] Although the poly(A) tail length has been described to positively correlate with mRNA translation, the notion of "the longer the better" is not always true.[61] A study showed that shorter poly(A) sequence could promote the closed-loop formation of mRNA for efficient translation.[62]

Histone mRNAs lack poly(A) tails but contain a stem-loop structure close to the 3′ end. Addition of this stem-loop structure instead of the poly(A) tail appears to improve the translation potential of synthetic mRNA.[63] Since the poly(A) sequence mediates mRNA decay, a mixed poly(A) tail with intermittent non-A residues such as guanine impedes the mRNA degradation pathway.[64] The functional diversity of mixed poly(A) tails in therapeutic mRNA design remains to be explored.

CIRCULAR RNAs

CircRNAs Biogenesis and Function

Apart from linear RNAs, circular RNAs (circRNAs) are a class of covalently closed single-stranded RNA molecules without 5′ end-caps and 3′ poly(A) tails. Unlike the conventional splicing that joins an upstream 5′ splice site (splice donor) with a downstream 3′ splice site (splice acceptor), circRNAs are generated by backsplicing in which a downstream 5′ splice site is joined to an upstream 3′ splice site in reverse order across an exon or more exons. A growing body of evidence suggests that circRNAs exert a wide range of biological functions, such as transcription, alternative splicing, and chromatin looping. Additionally, circRNAs can act as miRNA sponges or sequester RNA-binding proteins. Importantly, some circRNAs can be translated into peptides via IRES or m^6A-mediated cap-independent translation.[65,66]

Therapeutic circRNAs

Exogenous circRNAs have gained considerable attention from the therapeutic perspective because of several unique features. First, circRNAs are resistant to the canonical RNA decay pathways. As a result, circRNAs exhibit longer functional persistence and improved levels of protein expression compared to the linear counterpart. Second, the circular

https://doi.org/10.1021/acs.accounts.1c00541
Acc. Chem. Res. 2021, 54, 4272−4282

structure renders circRNAs more compatible with lipid nanoparticle-mediated delivery. Third, circRNAs can be produced and purified with fewer procedures. A recent study showed that exogenous circRNAs can escape the RNA-mediated immune response, although a different observation was reported.[67,68] Future studies should clarify this important aspect in vivo.

Synthetic circRNA Strategy

RNA circularization can be achieved in vitro and in vivo (Figure 6).[69] Linear RNAs synthesized by IVT are phosphorylated at the 5′ end and can be ligated with the help of DNA or RNA ligases, such as T4 DNA ligase, T4 RNA ligase 1, and T4 RNA ligase 2. Besides enzymatic ligation, special chemical groups can be introduced to both 5′ and 3′ ends of synthetic RNAs followed by chemical ligation, such as coper-catalyzed azide alkyne cycloaddition (CuAAC). Many ribozymes are capable of catalyzing phosphodiester bond formation within the specific seqeunce context. When the sequence of interest is cloned in between the two halves of a permutated self-splicng intron, circularization occurs by the inherent ribozyme activity of the intron.[70] Although highly efficient in producing circRNAs, the ribozyme approach is not suitable for mRNAs bearing modified nucleotides.[68]

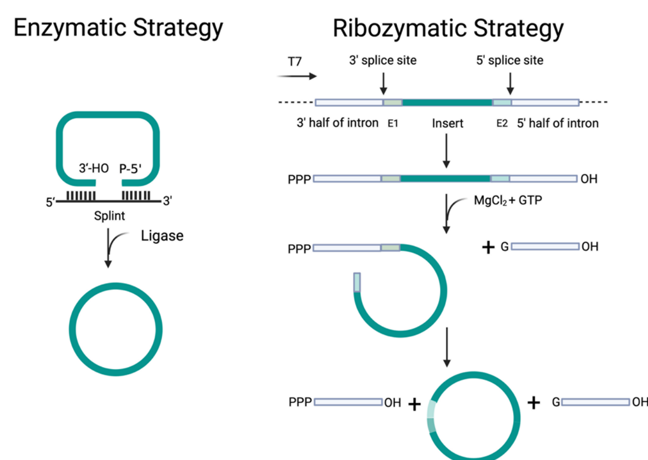


Figure 6. Schematic of engineering circular RNA using enzymatic strategy (left) and ribozymatic strategy (right).

Another limiting factor in circRNA application is the relatively low efficiency of cap-independent translation. With the identification of new *cis*-regulatory elements capable of cap-independent translation, it is possible to engineer circRNAs with a smaller size but higher translation potency.

CONCLUSIONS AND FUTURE PERSPECTIVES

The recent resurgence of interest in mRNA theraperutics has been spurred by the successful deployment of mRNA vaccines against the COVID-19 pandemic. Given the remarkable flexibility, efficient delivery, and proven safety, mRNA has become a powerful and versatile therapeutic platform. Since mRNA therapeutics rely on the translation machinery in target cells, a better understanding of the regulatory logic between *cis*-sequence elements and *trans*-acting factors will faciliate rational design of mRNAs with optimal translation potential, high stability, and low immunogenicity. Future investigations should continue focusing on understanding and utilizing sequence elements capable of selective translation in different cell types and under different growth conditions. Recent advances in genomic technologies have allowed us to shed light on new aspects of RNA life cycle. With the generation of big data on the rise, it is possible to utilize artificial intelligence/machine learning to "personalize" mRNA engineering in the future. In the foreseeable future, we envision the rapid advance of mRNA therapeutics in a broad range of human diseases.

AUTHOR INFORMATION

Corresponding Author

Shu-Bing Qian − *Division of Nutritional Sciences, Cornell University, Ithaca, New York 14853, United States;* orcid.org/0000-0002-4127-1136; Email: sq38@cornell.edu

Author

Longfei Jia − *Division of Nutritional Sciences, Cornell University, Ithaca, New York 14853, United States*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.accounts.1c00541

Funding

This work was supported by US National Institutes of Health (R01GM1222814 and DP1GM142101) and HHMI Faculty Scholar (55108556) to S.-B.Q.

Notes

The authors declare no competing financial interest.

Biographies

Longfei Jia obtained his Ph.D. degree in Animal physiology at Nanjing Agricultural University, China. He is currently a postdoctoral associate in the laboratory of Prof. Shu-Bing Qian. His research is focused on deciphering the *trans*- and *cis*-regulatory in mRNA translation, to engineer mRNA for translational output and stability in synthetic biology.

Shu-Bing Qian is a Professor of Nutritional Sciences at Cornell. His research interests span the disciplines of biochemistry and nutrition with an emphasis on studies of translational reprogramming in cell physiology and human diseases such as cancer. His lab focuses on the fundamental mechanism of translation and exploit this information for development of engineering strategies and therapeutic approaches.

ABBREVIATIONS

mRNA = messenger RNA
5′UTR = 5′untranslated region
CDS = coding region
PIC = preinitiation complex
eIF4A = eukaryotic initiation factor 4A
eIF4E = eukaryotic initiation factor 4E
eIF4G = eukaryotic initiation factor 4G
eIF4B = eukaryotic initiation factor 4B
eIF3 = eukaryotic initiation factor 3
eIF1 = eukaryotic initiation factor 1
eIF1A = eukaryotic initiation factor 1A
PABP = poly(A)-binding protein
eEF = eukaryotic elongation factor
eRF = eukaryotic release factor
m^7G = 7-Methylguanosine
Met-tRNAiMet = methionine tRNA
PIC = preinitiation complex
40S = small subunit of eukaryotic 80S ribosome

60S = large subunit of eukaryotic 80S ribosome
80S = eukaryotic 80S ribosome
P-site = peptidyl-tRNA site
cap-1 = m^7GpppN1 mN2
cap-2 = m^7GpppN1 mN2m
PCIF1 = phosphorylated CTD interacting factor 1
m^6A = *N*6-methyladenosine
FTO = fat mass- and obesity-associated protein
CBP = cap-binding protein
IVT = in vitro transcription
ARCA = antireverse cap analogs
VCE = vaccinia virus capping enzymes
LNA = lock nucleic acid
TOP = terminal oligopyrimidine motif
mTORC1 = mammalian target of rapamycin complex 1
IRES = internal ribosome entry sites
ORFs = open reading frames
ATF4 = activating transcription factor 4
MPRA = massively parallel reporter assay
RQC = ribosome-mediated quality control pathway
CAI = codon adaptation index
m^5C = 5-methylcytidine
Ψ = pseudouridine
$m^1\Psi$ = *N*(1)-methylpseudouridine
NSD = no-stop decay
PTC = premature stop codons
NMD = nonsense-mediated decay
PAS = polyadenylation site
APA = alternative polyadenylation
ARE = AU-rich element
AUBPs = AU-binding proteins
GRE = GU-rich elements
CURE = CU-rich elements
CAREs = CA-rich elements
GCREs = GC-rich elements
IREs = iron responsive elements
miRNAs = microRNAs
RNPs = ribonucleoproteins
circRNA = circular RNA
CuAAC = coper-catalyzed azide alkyne cycloaddition

REFERENCES

(1) Jia, L. F.; Mao, Y. H.; Ji, Q. Q.; Dersh, D.; Yewdell, J. W.; Qian, S. B. Decoding mRNA translatability and stability from the 5 ' UTR. *Nat. Struct. Mol. Biol.* **2020**, *27*, 814.

(2) Liu, X. M.; Zhou, J.; Mao, Y.; Ji, Q.; Qian, S. B. Programmable RNA N(6)-methyladenosine editing by CRISPR-Cas9 conjugates. *Nat. Chem. Biol.* **2019**, *15*, 865−871.

(3) Damase, T. R.; Sukhovershin, R.; Boada, C.; Taraballi, F.; Pettigrew, R. I.; Cooke, J. P. The Limitless Future of RNA Therapeutics. *Front. Bioeng. Biotechnol.* **2021**, *9*. DOI: 10.3389/fbioe.2021.628137.

(4) Shirokikh, N. E.; Preiss, T. Translation initiation by cap-dependent ribosome recruitment: Recent insights and open questions. *Wiley Interdiscip. Rev.: RNA* **2018**, *9*, No. e1473.

(5) Dever, T. E.; Dinman, J. D.; Green, R. Translation Elongation and Recoding in Eukaryotes. *Cold Spring Harbor Perspect. Biol.* **2018**, *10*, a032649.

(6) Hellen, C. U. T. Translation Termination and Ribosome Recycling in Eukaryotes. *Cold Spring Harbor Perspect. Biol.* **2018**, *10*, a032656.

(7) Wei, J.; Liu, F.; Lu, Z.; Fei, Q.; Ai, Y.; He, P. C.; Shi, H.; Cui, X.; Su, R.; Klungland, A.; Jia, G.; Chen, J.; He, C. Differential m(6)A, m(6)Am, and m(1)A Demethylation Mediated by FTO in the Cell Nucleus and Cytoplasm. *Mol. Cell* **2018**, *71*, 973−985 e975..

(8) Galloway, A.; Cowling, V. H. mRNA cap regulation in mammalian cell function and fate. *Biochim. Biophys. Acta, Gene Regul. Mech.* **2019**, *1862*, 270−279.

(9) Zust, R.; Cervantes-Barragan, L.; Habjan, M.; Maier, R.; Neuman, B. W.; Ziebuhr, J.; Szretter, K. J.; Baker, S. C.; Barchet, W.; Diamond, M. S.; Siddell, S. G.; Ludewig, B.; Thiel, V. Ribose 2′-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat. Immunol.* **2011**, *12*, 137−143.

(10) Devarkar, S. C.; Wang, C.; Miller, M. T.; Ramanathan, A.; Jiang, F.; Khan, A. G.; Patel, S. S.; Marcotrigiano, J. Structural basis for m7G recognition and 2′-O-methyl discrimination in capped RNAs by the innate immune receptor RIG-I. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 596−601.

(11) Borbolis, F.; Syntichaki, P. Biological implications of decapping: beyond bulk mRNA decay. *FEBS J.* **2021**. DOI: 10.1111/febs.15798.

(12) Stepinski, J.; Waddell, C.; Stolarski, R.; Darzynkiewicz, E.; Rhoads, R. E. Synthesis and properties of mRNAs containing the novel "anti-reverse" cap analogs 7-methyl(3 '-O-methyl)GpppG and 7-methyl(3 '-deoxy)GpppG. *Rna* **2001**, *7*, 1486−1495.

(13) Mockey, M.; Goncalves, C.; Dupuy, F. P.; Lemoine, F. M.; Pichon, C.; Midoux, P. mRNA transfection of dendritic cells: Synergistic effect of ARCA mRNA capping with Poly(A) chains in cis and in trans for a high protein expression level. *Biochem. Biophys. Res. Commun.* **2006**, *340*, 1062−1068.

(14) Vaidyanathan, S.; Azizian, K. T.; Haque, A.; Henderson, J. M.; Hendel, A.; Shore, S.; Antony, J. S.; Hogrefe, R. I.; Kormann, M. S. D.; Porteus, M. H.; McCaffrey, A. P. Uridine Depletion and Chemical Modification Increase Cas9 mRNA Activity and Reduce Immunogenicity without HPLC Purification. *Mol. Ther.–Nucleic Acids* **2018**, *12*, 530−542.

(15) Su, W.; Slepenkov, S.; Grudzien-Nogalska, E.; Kowalska, J.; Kulis, M.; Zuberek, J.; Lukaszewicz, M.; Darzynkiewicz, E.; Jemielity, J.; Rhoads, R. E. Translation, stability, and resistance to decapping of mRNAs containing caps substituted in the triphosphate chain with BH3, Se, and NH. *RNA* **2011**, *17*, 978−988.

(16) Kore, A. R.; Shanmugasundaram, M.; Charles, I.; Vlassov, A. V.; Barta, T. J. Locked Nucleic Acid (LNA)-Modified Dinucleotide mRNA Cap Analogue: Synthesis, Enzymatic Incorporation, and Utilization. *J. Am. Chem. Soc.* **2009**, *131*, 6364.

(17) Kuhn, A. N.; Diken, M.; Kreiter, S.; Selmi, A.; Kowalska, J.; Jemielity, J.; Darzynkiewicz, E.; Huber, C.; Tureci, O.; Sahin, U. Phosphorothioate cap analogs increase stability and translational efficiency of RNA vaccines in immature dendritic cells and induce superior immune responses in vivo. *Gene Ther.* **2010**, *17*, 961−971.

(18) Leppek, K.; Das, R.; Barna, M. Functional 5 ' UTR mRNA structures in eukaryotic translation regulation and how to find them (vol 19, pg 158, 2018). *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 673−673.

(19) Meyuhas, O.; Kahan, T. The race to decipher the top secrets of TOP mRNAs. *Biochim. Biophys. Acta, Gene Regul. Mech.* **2015**, *1849*, 801−811.

(20) Lee, A. S. Y.; Kranzusch, P. J.; Cate, J. H. D. eIF3 targets cell-proliferation messenger RNAs for translational activation or repression. *Nature* **2015**, *522*, 111−U292.

(21) Yang, Y.; Wang, Z. IRES-mediated cap-independent translation, a path leading to hidden proteome. *J. Mol. Cell Biol.* **2019**, *11*, 911−919.

(22) Noderer, W. L.; Flockhart, R. J.; Bhaduri, A.; de Arce, A. J. D.; Zhang, J. J.; Khavari, P. A.; Wang, C. L. Quantitative analysis of mammalian translation initiation sites by FACS-seq. *Mol. Syst. Biol.* **2014**, *10*, 748.

(23) Smith, E.; Meyerrose, T. E.; Kohler, T.; Namdar-Attar, M.; Bab, N.; Lahat, O.; Noh, T.; Li, J. J.; Karaman, M. W.; Hacia, J. G.; Chen, T. T.; Nolta, J. A.; Muller, R.; Bab, I.; Frenkel, B. Leaky ribosomal scanning in mammalian genomes: significance of histone H4 alternative translation in vivo. *Nucleic Acids Res.* **2005**, *33*, 1298−1308.

(24) Vattem, K. M.; Wek, R. C. Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 11269−11274.

(25) Kariko, K.; Kuo, A.; Barnathan, E. Overexpression of urokinase receptor in mammalian cells following administration of the in vitro transcribed encoding mRNA. *Gene Ther.* **1999**, *6*, 1092−1100.
(26) Sample, P. J.; Wang, B.; Reid, D. W.; Presnyak, V.; McFadyen, I. J.; Morris, D. R.; Seelig, G. Human 5′ UTR design and variant effect prediction from a massively parallel translation assay. *Nat. Biotechnol.* **2019**, *37*, 803−809.
(27) Cao, J. C.; Novoa, E. M.; Zhang, Z. Z.; Chen, W. C. W.; Liu, D. B.; Choi, G. C. G.; Wong, A. S. L.; Wehrspaun, C.; Kellis, M.; Lu, T. K. High-throughput 5 ʻ UTR engineering for enhanced protein production in non-viral gene therapies. *Nat. Commun.* **2021**, *12*. DOI: 10.1038/s41467-021-24436-7.
(28) Simms, C. L.; Yan, L. W. L.; Zaher, H. S. Ribosome Collision Is Critical for Quality Control during No-Go Decay. *Mol. Cell* **2017**, *68*, 361.
(29) Presnyak, V.; Alhusaini, N.; Chen, Y. H.; Martin, S.; Morris, N.; Kline, N.; Olson, S.; Weinberg, D.; Baker, K. E.; Graveley, B. R.; Coller, J. Codon optimality is a major determinant of mRNA stability. *Cell* **2015**, *160*, 1111−1124.
(30) Gamble, C. E.; Brule, C. E.; Dean, K. M.; Fields, S.; Grayhack, E. J. Adjacent Codons Act in Concert to Modulate Translation Efficiency in Yeast. *Cell* **2016**, *166*, 679−690.
(31) Mordstein, C.; Savisaar, R.; Young, R. S.; Bazile, J.; Talmane, L.; Luft, J.; Liss, M.; Taylor, M. S.; Hurst, L. D.; Kudla, G. Codon Usage and Splicing Jointly Influence mRNA Localization. *Cell Systems* **2020**, *10*, 351.
(32) Hia, F. B.; Yang, S. F.; Shichino, Y.; Yoshinaga, M.; Murakawa, Y.; Vandenbon, A.; Fukao, A.; Fujiwara, T.; Landthaler, M.; Natsume, T.; Adachi, S.; Iwasaki, S.; Takeuchi, O. Codon bias confers stability to human mRNAs. *EMBO Rep.* **2019**, *20*. DOI: 10.15252/embr.201948220.
(33) Van Gulck, E. R. A.; Ponsaerts, P.; Heyndrickx, L.; Vereecken, K.; Moerman, F.; De Roo, A.; Colebunders, R.; Van den Bosch, G.; Van Bockstaele, D. R.; Van Tendeloo, V. F. I.; Allard, S.; Verrier, B.; Maranon, C.; Hoeffel, G.; Hosmalin, A.; Berneman, Z. N.; Vanham, G. Efficient stimulation of HIV-1-specific T cells using dendritic cells electroporated with mRNA encoding autologous HIV-1 Gag and Env proteins. *Blood* **2006**, *107*, 1818−1827.
(34) Thess, A.; Grund, S.; Mui, B. L.; Hope, M. J.; Baumhof, P.; Fotin-Mleczek, M.; Schlake, T. Sequence-engineered mRNA Without Chemical Nucleoside Modifications Enables an Effective Protein Therapy in Large Animals. *Mol. Ther.* **2015**, *23*, 1456−1464.
(35) Frelin, L.; Ahlen, G.; Alheim, M.; Weiland, O.; Barnfield, C.; Liljestrom, P.; Sallberg, M. Codon optimization and mRNA amplification effectively enhances the immunogenicity of the hepatitis C virus nonstructural 3/4A gene. *Gene Ther.* **2004**, *11*, 522−533.
(36) Kim, Y. H.; Youn, H.; Na, J.; Hong, K. J.; Kang, K. W.; Lee, D. S.; Chung, J. K. Codon-optimized Human Sodium Iodide Symporter (opt-hNIS) as a Sensitive Reporter and Efficient Therapeutic Gene. *Theranostics* **2015**, *5*, 86−96.
(37) Mauger, D. M.; Cabral, B. J.; Presnyak, V.; Su, S. V.; Reid, D. W.; Goodman, B.; Link, K.; Khatwani, N.; Reynders, J.; Moore, M. J.; McFadyen, I. J. mRNA structure regulates protein expression through changes in functional half-life. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 24075−24083.
(38) Durbin, A. F.; Wang, C.; Marcotrigiano, J.; Gehrke, L. RNAs Containing Modified Nucleotides Fail To Trigger RIG-I Conformational Changes for Innate Immune Signaling. *mBio* **2016**, *7*. DOI: 10.1128/mBio.00833-16.
(39) Kormann, M. S.; Hasenpusch, G.; Aneja, M. K.; Nica, G.; Flemmer, A. W.; Herber-Jonat, S.; Huppmann, M.; Mays, L. E.; Illenyi, M.; Schams, A.; Griese, M.; Bittmann, I.; Handgretinger, R.; Hartl, D.; Rosenecker, J.; Rudolph, C. Expression of therapeutic proteins after delivery of chemically modified mRNA in mice. *Nat. Biotechnol.* **2011**, *29*, 154−157.
(40) Andries, O.; Mc Cafferty, S.; De Smedt, S. C.; Weiss, R.; Sanders, N. N.; Kitada, T. N(1)-methylpseudouridine-incorporated mRNA outperforms pseudouridine-incorporated mRNA by providing enhanced protein expression and reduced immunogenicity in mammalian cell lines and mice. *J. Controlled Release* **2015**, *217*, 337−344.
(41) Kariko, K.; Muramatsu, H.; Welsh, F. A.; Ludwig, J.; Kato, H.; Akira, S.; Weissman, D. Incorporation of Pseudouridine Into mRNA Yields Superior Nonimmunogenic Vector With Increased Translational Capacity and Biological Stability. *Mol. Ther.* **2008**, *16*, 1833−1840.
(42) Bonehill, A.; Heirman, C.; Tuyaerts, S.; Michiels, A.; Breckpot, K.; Brasseur, F.; Zhang, Y.; van der Bruggen, P.; Thielemans, K. Messenger RNA-electroporated dendritic cells presenting MAGE-A3 simultaneously in HLA class I and class II molecules. *J. Immunol.* **2004**, *172*, 6649−6657.
(43) van Hoof, A.; Frischmeyer, P. A.; Dietz, H. C.; Parker, R. Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science* **2002**, *295*, 2262−2264.
(44) Kurosaki, T.; Popp, M. W.; Maquat, L. E. Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 406−420.
(45) Wangen, J. R.; Green, R. Stop codon context influences genome-wide stimulation of termination codon readthrough by aminoglycosides. *eLife* **2020**, *9*. DOI: 10.7554/eLife.52611.
(46) Beznoskova, P.; Gunisova, S.; Valasek, L. S. Rules of UGA-N decoding by near-cognate tRNAs and analysis of readthrough on short uORFs in yeast. *RNA* **2016**, *22*, 456−466.
(47) Floquet, C.; Hatin, I.; Rousset, J. P.; Bidou, L. Statistical Analysis of Readthrough Levels for Nonsense Mutations in Mammalian Cells Reveals a Major Determinant of Response to Gentamicin. *PLoS Genet.* **2012**, *8*, e1002608.
(48) Cassan, M.; Rousset, J. P. UAG readthrough in mammalian cells: effect of upstream and downstream stop codon contexts reveal different signals. *BMC Mol. Biol.* **2001**, *2*, 3.
(49) Adachi, H.; Yu, Y. T. Pseudouridine-mediated stop codon readthrough in S. cerevisiae is sequence context-independent. *RNA* **2020**, *26*, 1247−1256.
(50) Adachi, M.; Cavalcanti, A. R. O. Tandem Stop Codons in Ciliates That Reassign Stop Codons. *J. Mol. Evol.* **2009**, *68*, 424−431.
(51) Zhang, H. B.; Lee, J. Y.; Tian, B. Biased alternative polyadenylation in human tissues. *Genome Biology* **2005**, *6*, R100.
(52) Ji, Z.; Tian, B. Reprogramming of 3 ʻ Untranslated Regions of mRNAs by Alternative Polyadenylation in Generation of Pluripotent Stem Cells from Different Cell Types. *PLoS One* **2009**, *4*, e8419.
(53) Otsuka, H.; Fukao, A.; Funakami, Y.; Duncan, K. E.; Fujiwara, T. Emerging Evidence of Translational Control by AU-Rich Element-Binding Proteins. *Front. Genet.* **2019**, *10*. DOI: 10.3389/fgene.2019.00332.
(54) Friedman, R. C.; Farh, K. K. H.; Burge, C. B.; Bartel, D. P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* **2008**, *19*, 92−105.
(55) Benteyn, D.; Anguille, S.; Van Lint, S.; Heirman, C.; Van Nuffel, A. M. T.; Corthals, J.; Ochsenreither, S.; Waelput, W.; Van Beneden, K.; Breckpot, K.; Van Tendeloo, V.; Thielemans, K.; Bonehill, A. Design of an Optimized Wilms' Tumor 1 (WT1) mRNA Construct for Enhanced WT1 Expression and Improved Immunogenicity In Vitro and In Vivo. *Mol. Ther.–Nucleic Acids* **2013**, *2*, e134.
(56) Zinckgraf, J. W.; Silbart, L. K. Modulating gene expression using DNA vaccines with different 3 ʻ-UTRs influences antibody titer, seroconversion and cytokine profiles. *Vaccine* **2003**, *21*, 1640−1649.
(57) Zhao, W. X.; Pollack, J. L.; Blagev, D. P.; Zaitlen, N.; McManus, M. T.; Erle, D. J. Massively parallel functional annotation of 3 ʻ untranslated regions. *Nat. Biotechnol.* **2014**, *32*, 387−U259.
(58) von Niessen, A. G. O.; Poleganov, M. A.; Rechner, C.; Plaschke, A.; Kranz, L. M.; Fesser, S.; Diken, M.; Lower, M.; Vallazza, B.; Beissert, T.; Bukur, V.; Kuhn, A. N.; Tureci, O.; Sahin, U. Improving mRNA-Based Therapeutic Gene Delivery by Expression-Augmenting 3 ʻ UTRs Identified by Cellular Library Screening. *Mol. Ther.* **2019**, *27*, 824−836.
(59) Grier, A. E.; Burleigh, S.; Sahni, J.; Clough, C. A.; Cardot, V.; Choe, D. C.; Krutein, M. C.; Rawlings, D. J.; Jensen, M. C.; Scharenberg, A. M.; Jacoby, K. pEVL: A Linear Plasmid for

Generating mRNA IVT Templates With Extended Encoded Poly(A) Sequences. *Mol. Ther.–Nucleic Acids* **2016**, *5*, No. e306.
(60) Holtkamp, S.; Kreiter, S.; Selmi, A.; Simon, P.; Koslowski, M.; Huber, C.; Tureci, O.; Sahin, U. Modification of antigen-encoding RNA increases stability, translational efficacy, and T-cell stimulatory capacity of dendritic cells. *Blood* **2006**, *108*, 4009−4017.
(61) Elango, N.; Elango, S.; Shivshankar, P.; Katz, M. S. Optimized transfection of mRNA transcribed from a d(A/T)100 tail-containing vector. *Biochem. Biophys. Res. Commun.* **2005**, *330*, 958−966.
(62) Lima, S. A.; Chipman, L. B.; Nicholson, A. L.; Chen, Y. H.; Yee, B. A.; Yeo, G. W.; Coller, J.; Pasquinelli, A. E. Short poly(A) tails are a conserved feature of highly expressed genes. *Nat. Struct. Mol. Biol.* **2017**, *24*, 1057.
(63) Gallie, D. R.; Lewis, N. J.; Marzluff, W. F. The histone 3′-terminal stem-loop is necessary for translation in Chinese hamster ovary cells. *Nucleic Acids Res.* **1996**, *24*, 1954−1962.
(64) Lim, J.; Kim, D.; Lee, Y. S.; Ha, M.; Lee, M.; Yeo, J.; Chang, H.; Song, J.; Ahn, K.; Kim, V. N. Mixed tailing by TENT4A and TENT4B shields mRNA from rapid deadenylation. *Science* **2018**, *361*, 701.
(65) Legnini, I.; Di Timoteo, G.; Rossi, F.; Morlando, M.; Briganti, F.; Sthandier, O.; Fatica, A.; Santini, T.; Andronache, A.; Wade, M.; Laneve, P.; Rajewsky, N.; Bozzoni, I. Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis. *Mol. Cell* **2017**, *66*, 22.
(66) Yang, Y.; Fan, X. J.; Mao, M. W.; Song, X. W.; Wu, P.; Zhang, Y.; Jin, Y. F.; Yang, Y.; Chen, L. L.; Wang, Y.; Wong, C. C. L.; Xiao, X. S.; Wang, Z. F. Extensive translation of circular RNAs driven by N-6-methyladenosine. *Cell Res.* **2017**, *27*, 626−641.
(67) Chen, Y. G.; Chen, R.; Ahmad, S.; Verma, R.; Kasturi, S. P.; Amaya, L.; Broughton, J. P.; Kim, J.; Cadena, C.; Pulendran, B.; Hur, S.; Chang, H. Y. N6-Methyladenosine Modification Controls Circular RNA Immunity. *Mol. Cell* **2019**, *76*, 96.
(68) Wesselhoeft, R. A.; Kowalski, P. S.; Parker-Hale, F. C.; Huang, Y. X.; Bisaria, N.; Anderson, D. G. RNA Circularization Diminishes Immunogenicity and Can Extend Translation Duration In Vivo. *Mol. Cell* **2019**, *74*, 508.
(69) Petkovic, S.; Muller, S. RNA circularization strategies in vivo and in vitro. *Nucleic Acids Res.* **2015**, *43*, 2454−2465.
(70) Wesselhoeft, R. A.; Kowalski, P. S.; Anderson, D. G. Engineering circular RNA for potent and stable translation in eukaryotic cells. *Nat. Commun.* **2018**, *9*. DOI: 10.1038/s41467-018-05096-6.