The genetic code as expressed through relationships between mRNA structure and protein function

David M. Mauger[#], **Nathan A. Siegfried**[#], and **Kevin M. Weeks**

Department of Chemistry, University of North Carolina Chapel Hill, North Carolina, USA 25599-3290

[#] These authors contributed equally to this work.

Abstract

Structured RNA elements within messenger RNA often direct or modulate the cellular production of active proteins. As reviewed here, RNA structures have been discovered that govern nearly every step in protein production: mRNA production and stability; translation initiation, elongation, and termination; protein folding; and cellular localization. Regulatory RNA elements are common within RNAs from every domain of life. This growing body of RNA-mediated mechanisms continues to reveal new ways in which mRNA *structure* regulates translation. We integrate examples from several different classes of RNA structure-mediated regulation to present a global perspective that suggests that the secondary and tertiary structure of RNA ultimately constitutes an additional level of the genetic code that both guides and regulates protein biosynthesis.

Introduction

RNA was long assumed to be a simple courier of the information contained within a DNA genome. This tidy, linear view of biology is quickly being replaced with models that emphasize the complex landscape of interactions among these macromolecules. At the center lies RNA. It is now well established that complex RNA structures are capable of functions previously thought to be the purview of proteins, including ligand binding and catalysis. These RNA structure-mediated functions also include regulation of nearly every step of cellular protein production. Some of the regulatory mechanisms discussed below have been thoughtfully reviewed previously. Our goal in this review is not to duplicate these reviews but to present the argument that the three-dimensional structure of messenger RNA (mRNA) constitutes an additional layer of genetic information that both guides and regulates the production of encoded proteins.

The primary sequence of an mRNA encodes the amino acid sequence of a protein, whereas structural features within mRNA molecules can determine the biological activity of the encoded protein by regulating the isoform produced, expression level, folding, localization, or stability. RNA structures that regulate biological function during translation have been identified in every kingdom of life. Consequently, developing a better understanding of how RNA structure governs protein expression and function has broad-ranging implications. These include guiding the development of novel therapeutics for combating bacterial [1] and viral pathogens [2,3] and extend to understanding and mitigating diverse human genetic

diseases such as Huntington's disease [4], myotonic dystrophy type 1 [5], and cystic fibrosis [6].

mRNA as a sensor

An mRNA can govern its own translation and transcription using ligand-binding structural elements called riboswitches. The best characterized riboswitches are located in the 5′ untranslated regions (UTRs) of bacterial mRNAs. Upon ligand binding, the RNA undergoes allosteric rearrangement that regulates transcription or translation initiation, elongation efficiency, mRNA stability, or splicing [7-9]. Riboswitches contain two domains: a metabolite-binding region known as the "aptamer domain" and an allosteric domain termed the "expression platform" (Figure 1A). The expression platform enacts the regulatory function signaled by the aptamer domain. Typically one structure of the riboswitch occludes an important regulatory element such as the ribosome-binding Shine-Dalgarno sequence. Riboswitch aptamer domains have evolved to bind diverse small molecules, and those domains that bind the same ligand tend to be highly conserved. In contrast, the expression domains vary in both sequence and function in different organisms. Thus, riboswitches are modular; a given aptamer domain has a specific target metabolite but the ultimate function depends on the linked expression platform. Riboswitches that bind ions (Mg^{2+}, F^-), carbohydrates, metabolites, proteins, and co-enzymes have been characterized. Protein expression requires mRNA to be (*i*) transcribed, (*ii*) processed (for example, 5′ capped, spliced, polyadenylated), and (*iii*) translated by the ribosome. Interfering with any of these processes reduces the number of available mRNA transcripts or decreases protein production. Examples of ligand-sensing riboswitches that influence each of these steps are discussed below.

Control of transcription elongation

In an mRNA with a riboswitch, RNA polymerase synthesizes the aptamer domain before the expression domain. This allows the nascent aptamer domain to sample the cellular environment and potentially bind a cognate ligand before the expression domain is transcribed and folded, as recently observed in a co-transcriptional folding study of the adenine riboswitch [10]. Critically, this timing allows folding of one of the mutually exclusive structures that can be adopted by the riboswitch. For example, in many strains of bacteria, ligand binding to an aptamer domain causes an expression platform to form an intrinsic terminator stem, an RNA hairpin followed by a stretch of six or more uracil (U) residues, that inhibits RNA polymerase extension (Figure 1A) [11,12]. In the absence of ligand, the unbound aptamer domain structure allows the anti-terminator stem-loop to form, sequestering the poly-U stretch and allowing mRNA synthesis to proceed. This type of regulation is exemplified by the T-box mechanism: if an uncharged tRNA binds to an element in the mRNA encoding the synthetase responsible for charging the particular tRNA, an anti-terminator helix forms that permits transcription of the message encoding the aminoacyl-tRNA synthetase [13]. In contrast, the charged tRNA does not bind strongly to the terminator helix. In most cases, as in this example, the riboswitch mechanisms are negative feedback loops wherein a given ligand down regulates enzymes involved in ligand production. Although uncommon, there are examples of riboswitches that turn protein

expression on rather than off. For example, ligand binding to the adenine riboswitch in *Bacillus subtilis* results in increased expression of proteins involved in adenine export [14].

Regulation of splicing

Although most riboswitches identified to date exist in simple prokaryotes, a thiamine pyrophosphate (TPP) sensing riboswitch is a widely distributed element [15,16] found in bacteria, archaea [17], and eukaryotes [16,18,19] including both simple [20] and complex plants [19,21]. While bacterial TTP riboswitches typically exert control at the level of transcription [22], these elements regulate alternative splicing in eukaryotes. Eukaryotic TPP riboswitches are typically located in intronic regions of genes associated with thiamine metabolism. Differences in secondary structure between ligand-bound and unbound forms of the precursor mRNA (pre-mRNA) sequester, expose, or relocate splice sites resulting in alternatively spliced mRNAs. Inclusion or exclusion of upstream open reading frames (ORFs) in mRNAs affects the identity of the synthesized protein. *Neurospora crassa* contains several TPP riboswitches and provides the best understood examples of eukaryotic riboswitches. In the *NMT1* riboswitch ligand binding causes formation of a structure that exposes an alternative splice site that prevents production of the major ORF product (Figure 1B) [19]. A recent study, also in *N. crassa*, suggests that a TPP riboswitch in the *NCU01977* gene controls formation of long-distance base pairs [23]. These base pairs do not cover or reveal splice sites, but rather place a 5′ splice donor and 3′ acceptor in favorable proximity [23]. These different forms of splicing control within a single organism highlight the functional diversity of riboswitch elements.

Regulation of mRNA stability

The *glmS* ribozyme, a self-cleaving RNA that requires ligand binding for activation, is an example of an RNA structure that regulates mRNA stability. The ribozyme is located in the 5′ UTR of the mRNA encoding the glucosamine-6-phosphate synthetase (GlmS) enzyme, which is widely distributed in Gram positive bacteria [24]. When glucosamine-6-phosphate levels are high, the molecule binds to the aptamer domain, activating the ribozyme. When the *glmS* ribozyme self-cleaves, a terminus recognized by RNase J1 is produced, and RNase J degrades the remaining mRNA. Another example of riboswitch modulation of mRNA stability is found in the *p27* tumor suppressor mRNA [25]. A riboswitch in this mRNA binds the Pumilio RNA-binding protein (PUM1) and causes an allosteric RNA rearrangement that reveals an miRNA target. Subsequent miRNA binding results in miRNA-mediated gene silencing.

Translation

Translation can be regulated by two mutually exclusive RNA structures. In this case, one hairpin – the 'anti-RBS' stem – sequesters the ribosome binding site (RBS) and diminishes translation efficiency, whereas the second structure leaves the RBS site accessible. In this case, a region containing a lysine-responsive riboswitch in *E. coli* also contains two RNase E cleavage sites [26]. Upon ligand binding, a structure forms that inhibits translation by sequestering the RBS and that also exposes the enzymatic cleavage sites, leading to subsequent mRNA degradation (Figure 1C).

Expression of some proteins is regulated by mRNA structure at multiple levels. For example, cellular tryptophan levels control expression of the *trp* operon in *B. subtilis* at the levels of transcription and translation through binding of a protein complex called tryptophan-activated RNA-binding attenuation protein (TRAP) [27,28]. The combination of these two strategies highlights the highly sensitive regulation of gene expression achievable via RNA-based mechanisms. In the event that metabolite concentrations do not exceed the threshold required to halt mRNA transcription, these same concentrations may trigger translational inhibition. Such dual control may also allow rapid responses to changes in metabolite concentrations. Similarly, examples of cooperative riboswitches have been characterized [29]. These include the glycine and tetrahydrofolate riboswitches that contain tandem aptamer domains or bind multiple ligand molecules, respectively.

Other mechanisms

Although most riboswitches function in *cis*, a *trans*-acting SAM riboswitch was discovered in *Listeria monocytogenes* [30]. Two of the seven SAM riboswitches in *L. monocytogenes* (SreA through SreG) [30] (SreA and B) act in *trans*: These RNAs bind to the 5′ UTR of the *prfA* virulence factor mRNA and sequester the Shine-Dalgarno sequence to inhibit *prfA* translation. This *trans* activity does not require SAM binding and therefore can be considered similar to antisense mechanisms employed by other non-coding RNAs. A riboswitch has also been identified that binds two glycine ligands in a cooperative process that likely evolved to allow precise response to cellular concentrations of glycine [31].

Novel categories of riboswitch-like mechanisms also exist. For example, *L. monocytogenes* contains an RNA thermosensor, a 5′ UTR structure that inhibits translation of *prfA* virulence factor at temperatures below 37 °C [32]. Infection of warm host cells results in an RNA rearrangement that allows translation. This use of temperature as a 'ligand' demonstrates the wide range of regulation mechanisms achievable by RNA structure.

Hijacking the ribosome: IRES elements

Cap-dependent translation

Under most conditions, translation initiation in eukaryotes is dependent on a 5′ 7-methylguanylate cap on a mRNA. Initiation is mediated by eukaryotic initiation factors (eIFs) that are recruited through a series of recognition events (reviewed in [33]). This process begins with recognition and binding of the 5′ cap by the eIF4 cap-binding complex. This event leads to formation of the 48S initiation complex, a ribonucleoprotein complex containing the 40S ribosomal subunit and several eIFs. This initiation complex scans the mRNA from 5′ to 3′ for an AUG start codon [34,35]. At the start codon, the complex stalls, the eIFs dissociate, and the remainder of the ribosome is recruited to form an 80S unit. It is only at this point that the elongation stage of translation begins. Initiation is the rate-limiting step of translation and regulation of this event allows tight control of protein expression. RNA structures play regulatory roles in canonical translation initiation; for example, extensive, stable mRNA structure can interfere with a scanning ribosomal complex [36].

Cap-independent translation

Canonical translation initiation is attenuated by cellular mechanisms in response to environmental stimuli such as nutrient stress, mitosis, apoptosis, and viral infection. Even during times of stress, however, synthesis of certain proteins is required, and specific eukaryotic mRNAs use a structure-based strategy to 'hijack' the ribosome and initiate translation through an internal ribosomal entry site (IRES). Viral mRNAs also co-opt host protein synthesis machinery through cap-independent binding. At these internal entry sites, typically in the 5′ UTR of a message, RNA structure replaces many or all of the eIF proteins required to recruit the ribosome (Figure 2).

Eukaryotic and viral RNAs use IRES elements in strategically different ways. Eukaryotic mRNAs with IRES elements encode proteins required for apoptosis, cell development, oncogenesis and for survival of nutrient stress, hypoxia, heat shock, and viral infection (reviewed in [37]). These conditions coincide with a decline in cap-dependent translation initiation, and IRES mechanisms allow the cell to respond with synthesis of appropriate proteins. For example, during apoptosis, the mRNA encoding c-myc, a transcription enhancer important to apoptosis and cancer mechanisms, switches from canonical to IRES-mediated translation initiation and is actively produced despite inactivation of cap-dependent translation [38]. This may be a widespread cellular strategy. *In silico* analyses estimate that nearly 10% of cellular mRNAs have IRES activity [39]. In contrast, viruses often contain highly structured 5′ regions incompatible with traditional translation initiation, and may lack the 5′ cap or poly(A) tail required for canonical translation [40]. RNA viruses use IRES elements to proliferate in the cellular environment.

IRES structures

The first IRES elements were discovered in the positive sense RNAs of picornaviruses and encephalomyocarditis viruses [41,42]. IRES elements in closely-related viral species exhibit a high degree of structural conservation, despite large differences in primary sequence, emphasizing an important role for RNA structure in directing protein expression. In contrast, cellular IRES elements show much less sequence and structural conservation, suggesting these elements are more highly specialized.

IRES elements may require eIFs and/or RNA-binding proteins called IRES trans-acting factors (ITAFs), whereas others directly bind the ribosome to initiate translation. The number of eIF and ITAF factors required to initiate translation is roughly inversely related to the amount of compactly folded RNA structure in an IRES[43]. Viral IRES elements tend to be more structured than their cellular counterparts and bind fewer co-factors [44]. Four classes of IRES elements have been defined for *Picornaviridae* (Figure 2A). Class I viral IRES elements are self-sufficient, RNA-driven systems that do not require any factors for initiation. At the opposite end of the spectrum, Class IV elements require canonical eIFs as well as ITAFs [45].

Perhaps the best characterized Class I viral IRES is from a *Dicistroviridae* intergenic region. An RNA pseudoknot in this IRES occupies the ribosomal P-site and mimics the structure of a bound peptidyl tRNA, allowing complete ribosomal assembly [46,47]. A Class I IRES in

Israeli acute bee paralysis virus controls levels of two proteins, one expressed from a +1 reading frame by forming an extra GU base pair in the pseudoknot occupying the P-site, thereby shifting the frame of the codon:anticodon mimic by one nucleotide [40].

Class II viral IRES domains are less structured. The Class II HCV IRES binds directly to the ribosome [48] but requires canonical elements such as initiator tRNA^{Met}, eIF2, and GTP for translation initiation [49]. The HCV IRES contains several major domains (II, III, and IV) that are conserved among related flaviviruses and some picornaviruses (Figure 2B). Domain III binds eIF3 and contains a distinct 40S binding site (reviewed in [50]). A pseudoknotted element at the base of domain III is located near the AUG start codon in domain IV (Figure 2B), and computational modeling suggests that it forms a three-dimensional structure similar to that of tRNA [51]. This tRNA-like structure likely interacts with the ribosome, ultimately aligning the start codon in the ribosomal P-site [48,49,51-54]. The use of mimicry, especially of tRNAs, appears to be a general strategy in IRES function [55].

Class III and Class IV viral IRES elements require more initiation factors than do Class II IRESs and do not directly bind the 40S ribosome. Examples of Group III IRESs are found in foot and mouth disease virus and encephalomyocarditis virus RNAs. Group IV IRES elements include those from poliovirus and hepatitis A virus. An important distinction between these two IRES classes is in their proximity to an AUG start codon. Class III IRESs immediately precede the AUG, whereas Class IV IRESs can be located at significant distance from their target ORFs.

The mRNA encoding the transcription factor c-myc contains perhaps the best characterized cellular IRES element. Normally, proteins from the myc family control cell proliferation and apoptosis; however, a mutant form is associated with carcinogenesis [56,57]. As discussed, *c-myc* mRNA can use either canonical or IRES-mediated translation initiation [58]. Use of the IRES element occurs under specific cellular conditions when one of four possible promoter sequences is used to produce a long carcinogenic mRNA [58] (Figure 2C). RNA pseudoknots, similar to those important to tRNA mimicry in HCV and *Dicistroviridae* Class I IRESs, may form near the *c-myc* ribosome landing sites in the IRES elements [59]. Understanding these RNA structures will be key to determining how these complex protein regulatory pathways function.

Derailing translation with programed ribosomal frameshifting

Numerous cellular and viral RNAs undergo non-canonical translation where the genetic information within the message is "recoded" [60] by signals within the RNA (reviewed [40,61]). We will highlight the role that RNA structure plays in inducing and regulating one example of non-canonical translation: programed ribosomal frameshifting.

Ribosomal frameshifting occurs when elements within the RNA cause the ribosome to switch registers to either the −1 or +1 reading frame. Additionally, some studies suggest that −2 and +2 frameshifting events are possible [62,63]. Frameshifting sites are minimally composed of two distinct parts: a "slippery sequence" and a *cis*-acting element. Various *cis*-acting elements, such as Shine-Dalgarno-like sequences and stop codons [64,65], are capable of inducing frameshifting, but most frameshifting events are mediated by highly

structured RNA elements [66-68]. The *cis*-element within the mRNA pauses a translating ribosome's active site over the slippery sequence (Figure 3A). The codon-anticodon interactions are momentarily disrupted, and the ribosome advances two or four nucleotides in the case of −1 and +1 frameshifting, respectively, before translation continues. The efficiencies of frameshifting sites vary depending on the features of the elements involved [66,69,70]. Even low levels of programmed frameshifting, less than 10%, can have large functional consequences.

RNA pseudoknots are structural motifs that are often found at both −1 and +1 ribosomal frameshifting sites, presumably due to their ability to resist the helicase activity of the ribosome [71]. Pseudoknots contain a minimum of two base-paired helices; nucleotides within the loop of one helix base pair with nucleotides outside the stem region to form the "pseudoknotted" structure [72]. An example of pseudoknot within a −1 frameshifting element in mammalian cells is found in the mouse *embryonal carcinoma differentiation regulated* (*Edr*) gene [73] (Figure 3B). *Edr* and the human ortholog *PEG10* are highly expressed during development and encode proteins with two overlapping reading frames [74,75]. A pseudoknot structure also can be found in the +1 frameshifting sites within eukaryotic antizyme genes, from yeast to mammals [76 ,77], which encode inhibitors of ornithine decarboxylase[78]. Translation of functional protein requires a +1 frameshift at the end of the first reading frame [79]. The AZ1 frameshifting site is complex and is composed of a (UGC UCC UGA) slippery sequence flanked by a 5′ regulatory element and pseudoknot (Figure 3C) [76,79]. The slippery sequence contains a UGA stop-codon that will terminate protein translation unless a frameshift occurs. Increased levels of polyamines, the downstream products of ornithine decarboxylase, up regulate levels of frameshifting through action of the 5′ regulatory element [80].

One of the first −1 frameshifting elements identified occurs at the end of the Gag reading frame in HIV-1 [70] (Figure 3D). This frameshifting element induces a −1 frameshift for 5-10% of translating ribosomes, allowing for the translation of a gag-pol fusion protein that contains gag protein fused to the viral protease and reverse transcriptase enzymes [70]. The frameshifting element was initially suggested to consist of the slippery sequence (UUUUUUA) and a simple stem-loop structure [70]; however, this simple model has been revised as studies of this element suggested that additional sequences are required for biological activity [81]. Studies of the frameshifting element in the context of the intact HIV-1 genomic RNA suggest that the stem-loop is part of a larger structured RNA domain that regulates frameshifting *in vivo* presumably by affecting the translating ribosome (Figure 3D) [82]. The surrounding structural domain functions to attenuate the frameshifting efficiency [83].

A complex −1 frameshifting element exists within the genome of the SARS corona-virus (SARS-CoV) at the end of the first ORF. Frameshifting results in the translation of the SARS-CoV the replicase-transcriptase peptide is cleaved by proteases to generate up to 16 individual proteins [68]. The programed frameshift in SARS-CoV appears to regulate the ratios of SARS viral proteins. The frameshifting element in SARS-CoV contains a pseudoknot with a stem-loop in one of the pseudoknot loops (Figure 3F) [84]. The end of this internal stem-loop structure contains a palindromic sequence that allows the element to

dimerize (Figure 3F) [85]. Mutations that abolished the dimerization of this domain, both reduced frameshifting efficiency and inhibited viral replication [85]. These data suggest that the frameshifting efficiency within SARS-CoV is regulated by the ability of the genome to dimerize, connecting two critical steps in the replication cycle of the virus.

The frameshifting elements described here are just a few examples of a rapidly expanding list of frameshifting elements that function in diverse organisms and viruses. RNA structures within these frameshifting elements often have modular properties, and stable stem-loop structures can substitute for pseudoknots in some frameshifting elements [86]. The frameshifting efficiency of the core elements can be regulated by the surrounding RNA sequence and structure [83], tertiary interactions [85], and ligands [80] that ultimately create dynamic elements capable of both precisely controlling frameshifting efficiency and response to conditions within the cellular environment.

RNA structure and setting ribosomal speed limits

Folding of a nascent chain of amino acids into functional protein domains begins immediately after the peptide emerges from the ribosome (reviewed in [87,88]). The rate of translation can significantly impact protein folding [89]. The most extensively studied examples of regulated translational pausing in cells are at rare codons. Rare codons, those codons with low abundance tRNA partners, often occur at the junctions between independently folding protein domains [90]. Ribosomal pausing at these rare codons is thought to facilitate the folding of native proteins in cells [91]. The correlation between translational pausing and rare codon frequency is imperfect, and there is a growing body of evidence that mRNA structure might also regulate translation rates. For example, when individual ribosomes translating an RNA hairpin were monitored using optical tweezers, the length of time that the ribosome paused during translation was dependent on the stability of the RNA structure [92]; stable structures are eventually unwound by the mechanical forces exerted by the ribosome[93].

An extensive body of evidence demonstrating the effect of mRNA structure on translation comes from studies of the yeast *ASH1* gene. Ash1p is a transcription factor that is localized to newly formed daughter cells during mitosis and regulates mating-type switching. During mitosis, four structured elements within the coding sequence of the *ASH1* mRNA function both to localize the mRNA to newly forming cells and also to translationally silence the mRNA in the mother cell [94,95]. These two activities are separable: When the structured RNA elements are moved from the coding sequence to the 3′ UTR, the mRNA is correctly localized, but localization of the Ash1p protein is impaired [96]. The structured elements within the *ASH1* mRNA appear to function as ribosome stop-lights, preventing protein translation until proper localization of the message. Another example of the effect of mRNA structure on protein translation comes from the analysis of single nucleotide polymorphisms in the human pain regulator gene *catechol-O-methyltransferase* (*COMT*) [97]. Polymorphisms within the *COMT* gene are associated with increased pain sensitivity in humans [98]. The increased pain sensitivity phenotypes are caused by dramatic increases in COMT protein levels due to changes in the global structure of the mRNA [97,99]. These

differences in COMT protein expression are attributed, in part, to structural changes near the start codon that affect translational efficiency of the mRNA [99].

Recent innovations in RNA structure probing technologies have facilitated systematic analyses of RNA structure on genome-wide scales and examination of relationships between RNA structure and the organization of the encoded protein. The RNA structure probing technology SHAPE was used to analyze the structure of an HIV-1 genomic RNA [82]. The regions of the genomic RNA that encode polyprotein linkers and interdomain loops within HIV-1 proteins are highly structured (Figure 4A-B). This direct correlation between the structure of the mRNA and the organization of protein domains suggests that these structured regions of the mRNA function as ribosomal pause sites and facilitate the folding of active viral proteins. The hypothesis that structures within coding regions of mRNAs play an important role in protein folding is also supported by analysis of the mRNA transcriptome in yeast [100]. A technology termed PARS was used to probe the structures of over 3000 yeast mRNAs. There is a correlation between mRNA structure and the organization of coding sequences such that both the start stop codons generally occur in regions that lack stable structures (Figure 4C). The internal coding regions of most mRNAs appear to have a significantly higher level of structure than the 5′ and 3′ UTRs. Taken together, these recent studies suggest that RNA structure represents a second layer of information, an RNA structurome [101], that regulates the kinetics of protein production to ensure proper expression and folding

Protein coding regions within mRNAs are often depicted as long rectangles, as if they are translation highways. The underlying assumption is that once the ribosome clears the obstacles imposed by initiation, translation proceeds uninterrupted until a stop codon is encountered. Current evidence indicates, however, that coding regions within mRNAs are complex landscapes defined by both codon usage and the physical structure of the mRNA.

Conclusions

We have presented an overview of strategies by which RNA structure controls and modulates the synthesis and function of proteins. These examples emphasize that structured elements in RNA function as an additional level of information to govern protein expression. Conversely, in the absence of structured RNA regulatory elements, many transcriptional and translational control functions would be lost. Proteins carry out the vast majority of cellular catalytic, signaling, regulatory, and replicative functions. Ultimately, the three-dimensional structure of an mRNA exerts a great deal of influence over the function of its protein product. The influence and complexity of RNA structure in governing protein expression, synthesis and function can appropriately be considered another level of the genetic code, one that we are likely to have only glimpsed thus far.

Acknowledgements

References

[1]. Deigan KE, Ferre-D'Amare AR. Riboswitches: discovery of drugs that target bacterial gene-regulatory RNAs. Acc. Chem. Res. 2011; 44:1329–38. [PubMed: 21615107]

[2]. Ahn DG, et al. Interference of ribosomal frameshifting by antisense peptide nucleic acids suppresses SARS coronavirus replication. Antiviral Res. 2011; 91:1–10. [PubMed: 21549154]

[3]. Marcheschi RJ, Tonelli M, Kumar A, Butcher SE. Structure of the HIV-1 frameshift site RNA bound to a small molecule inhibitor of viral replication. ACS Chem. Biol. 2011; 6:857–64. [PubMed: 21648432]

[4]. Yu D, et al. Single-stranded RNAs use RNAi to potently and allele-selectively inhibit mutant huntingtin expression. Cell. 2012; 150:895–908. [PubMed: 22939619]

[5]. Parkesh R, et al. Design of a bioactive small molecule that targets the myotonic dystrophy type 1 RNA via an RNA motif-ligand database and chemical similarity searching. J. Am. Chem. Soc. 2012; 134:4731–42. [PubMed: 22300544]

[6]. Liu X, et al. Partial correction of endogenous DeltaF508 CFTR in human cystic fibrosis airway epithelia by spliceosome-mediated RNA trans-splicing. Nat. Biotechnol. 2002; 20:47–52. [PubMed: 11753361]

[7]. Mandal M, Breaker RR. Gene regulation by riboswitches. Nat. Rev. Mol. Cell Biol. 2004; 5:451–63. [PubMed: 15173824]

[8]. Smith AM, Fuchs RT, Grundy FJ, Henkin TM. Riboswitch RNAs: regulation of gene expression by direct monitoring of a physiological signal. RNA Biol. 2010; 7:104–10. [PubMed: 20061810]

[9]. Breaker RR. Ancient, giant riboswitches at atomic resolution. Nat. Struct. Mol. Biol. 2012; 19:1208–9. [PubMed: 23142981]

[10]. Frieda KL, Block SM. Direct observation of cotranscriptional folding in an adenine riboswitch. Science. 2012; 338:397–400. [PubMed: 23087247]

[11]. Gusarov I, Nudler E. The mechanism of intrinsic transcription termination. Mol. Cell. 1999; 3:495–504. [PubMed: 10230402]

[12]. Yarnell WS, Roberts JW. Mechanism of intrinsic transcription termination and antitermination. Science. 1999; 284:611–5. [PubMed: 10213678]

[13]. Grundy FJ, Winkler WC, Henkin TM. tRNA-mediated transcription antitermination in vitro: codon-anticodon pairing independent of the ribosome. Proc. Natl. Acad. Sci. U. S. A. 2002; 99:11121–6. [PubMed: 12165569]

[14]. Mandal M, Breaker RR. Adenine riboswitches and gene activation by disruption of a transcription terminator. Nat. Struct. Mol. Biol. 2004; 11:29–35. [PubMed: 14718920]

[15]. Rodionov DA, Vitreschak AG, Mironov AA, Gelfand MS. Comparative genomics of thiamin biosynthesis in procaryotes. New genes and regulatory mechanisms. J Biol Chem. 2002; 277:48949–59. [PubMed: 12376536]

[16]. Sudarsan N, Barrick JE, Breaker RR. Metabolite-binding RNA domains are present in the genes of eukaryotes. RNA. 2003; 9:644–7. [PubMed: 12756322]

[17]. Miranda-Rios J, Navarro M, Soberon M. A conserved RNA structure (thi box) is involved in regulation of thiamin biosynthetic gene expression in bacteria. Proc. Natl. Acad. Sci. U. S. A. 2001; 98:9736–41. [PubMed: 11470904]

[18]. Kubodera T, Watanabe M, Yoshiuchi K, Yamashita N, Nishimura A, Nakai S, Gomi K, Hanamoto H. Thiamine-regulated gene expression of Aspergillus oryzae thiA requires splicing of the intron containing a riboswitch-like domain in the 5′-UTR. FEBS Lett. 2003; 555:516–20. [PubMed: 14675766]

[19]. Cheah MT, Wachter A, Sudarsan N, Breaker RR. Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. Nature. 2007; 447:497–500. [PubMed: 17468745]

[20]. Croft MT, Moulin M, Webb ME, Smith AG. Thiamine biosynthesis in algae is regulated by riboswitches. Proc. Natl. Acad. Sci. U. S. A. 2007; 104:20770–5. [PubMed: 18093957]

[21]. Bocobza S, Adato A, Mandel T, Shapira M, Nudler E, Aharoni A. Riboswitch-dependent gene regulation and its evolution in the plant kingdom. Genes Dev. 2007; 21:2874–9. [PubMed: 18006684]

[22]. Barrick JE, Breaker RR. The distributions, mechanisms, and structures of metabolite-binding riboswitches. Genome Biol. 2007; 8:R239. [PubMed: 17997835]

[23]. Li S, Breaker RR. Eukaryotic TPP riboswitch regulation of alternative splicing involving long-distance base pairing. Nucleic Acids Res. 2013

[24]. Ferre-D'Amare AR. The glmS ribozyme: use of a small molecule coenzyme by a gene-regulatory RNA. Q. Rev. Biophys. 2010; 43:423–47. [PubMed: 20822574]

[25]. Kedde M, van Kouwenhove M, Zwart W, Oude Vrielink JA, Elkon R, Agami R. A Pumilio-induced RNA structure switch in p27-3′ UTR controls miR-221 and miR-222 accessibility. Nat. Cell Biol. 2010; 12:1014–20. [PubMed: 20818387]

[26]. Caron MP, Bastet L, Lussier A, Simoneau-Roy M, Masse E, Lafontaine DA. Dual-acting riboswitch control of translation initiation and mRNA decay. Proc. Natl. Acad. Sci. U. S. A. 2012; 109:E3444–53. [PubMed: 23169642]

[27]. Babitzke P, Stults JT, Shire SJ, Yanofsky C. TRAP, the trp RNA-binding attenuation protein of Bacillus subtilis, is a multisubunit complex that appears to recognize G/UAG repeats in the trpEDCFBA and trpG transcripts. J Biol Chem. 1994; 269:16597–604. [PubMed: 7515880]

[28]. Babitzke P, Yanofsky C. Reconstitution of Bacillus subtilis trp attenuation in vitro with TRAP, the trp RNA-binding attenuation protein. Proc. Natl. Acad. Sci. U. S. A. 1993; 90:133–7. [PubMed: 7678334]

[29]. Serganov A, Patel DJ. Molecular recognition and function of riboswitches. Curr. Opin. Struct. Biol. 2012; 22:279–86. [PubMed: 22579413]

[30]. Loh E, et al. A trans-acting riboswitch controls expression of the virulence regulator PrfA in Listeria monocytogenes. Cell. 2009; 139:770–9. [PubMed: 19914169]

[31]. Mandal M, Lee M, Barrick JE, Weinberg Z, Emilsson GM, Ruzzo WL, Breaker RR. A glycine-dependent riboswitch that uses cooperative binding to control gene expression. Science. 2004; 306:275–9. [PubMed: 15472076]

[32]. Johansson J, Mandin P, Renzoni A, Chiaruttini C, Springer M, Cossart P. An RNA thermosensor controls expression of virulence genes in Listeria monocytogenes. Cell. 2002; 110:551–61. [PubMed: 12230973]

[33]. Jackson RJ, Hellen CU, Pestova TV. The mechanism of eukaryotic translation initiation and principles of its regulation. Nat. Rev. Mol. Cell Biol. 2010; 11:113–27. [PubMed: 20094052]

[34]. Hinnebusch AG. Molecular mechanism of scanning and start codon selection in eukaryotes. Microbiol. Mol. Biol. Rev. 2011; 75:434–67. first page of table of contents. [PubMed: 21885680]

[35]. Kozak M. An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. 1987; 15:8125–48. [PubMed: 3313277]

[36]. Kozak M. Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs. Mol. Cell. Biol. 1989; 9:5134–42. [PubMed: 2601712]

[37]. Fitzgerald KD, Semler BL. Bridging IRES elements in mRNAs to the eukaryotic translation apparatus. Biochim. Biophys. Acta. 2009; 1789:518–28. [PubMed: 19631772]

[38]. Stoneley M, Paulin FE, Le Quesne JP, Chappell SA, Willis AE. C-Myc 5′ untranslated region contains an internal ribosome entry segment. Oncogene. 1998; 16:423–8. [PubMed: 9467968]

[39]. Spriggs KA, Stoneley M, Bushell M, Willis AE. Re-programming of translation following cell stress allows IRES-mediated translation to predominate. Biol. Cell. 2008; 100:27–38. [PubMed: 18072942]

[40]. Firth AE, Brierley I. Non-canonical translation in RNA viruses. J. Gen. Virol. 2012; 93:1385–409. [PubMed: 22535777]

[41]. Pelletier J, Sonenberg N. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. Nature. 1988; 334:320–5. [PubMed: 2839775]

[42]. Jang SK, Krausslich HG, Nicklin MJ, Duke GM, Palmenberg AC, Wimmer E. A segment of the 5′ nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. J. Virol. 1988; 62:2636–43. [PubMed: 2839690]

[43]. Filbin ME, Kieft JS. Toward a structural understanding of IRES RNA function. Curr. Opin. Struct. Biol. 2009; 19:267–76. [PubMed: 19362464]

[44]. Hellen CU, Sarnow P. Internal ribosome entry sites in eukaryotic mRNA molecules. Genes Dev. 2001; 15:1593–612. [PubMed: 11445534]

[45]. Kieft JS. Viral IRES RNA structures and ribosome interactions. Trends Biochem. Sci. 2008; 33:274–83. [PubMed: 18468443]

[46]. Schuler M, et al. Structure of the ribosome-bound cricket paralysis virus IRES RNA. Nat. Struct. Mol. Biol. 2006; 13:1092–6. [PubMed: 17115051]

[47]. Spahn CM, Jan E, Mulder A, Grassucci RA, Sarnow P, Frank J. Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor. Cell. 2004; 118:465–75. [PubMed: 15315759]

[48]. Spahn CM, Kieft JS, Grassucci RA, Penczek PA, Zhou K, Doudna JA, Frank J. Hepatitis C virus IRES RNA-induced changes in the conformation of the 40s ribosomal subunit. Science. 2001; 291:1959–62. [PubMed: 11239155]

[49]. Pestova TV, Shatsky IN, Fletcher SP, Jackson RJ, Hellen CU. A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. Genes Dev. 1998; 12:67–83. [PubMed: 9420332]

[50]. Lukavsky PJ. Structure and function of HCV IRES domains. Virus Res. 2009; 139:166–71. [PubMed: 18638512]

[51]. Lavender CA, Ding F, Dokholyan NV, Weeks KM. Robust and generic RNA modeling using inferred constraints: a structure for the hepatitis C virus IRES pseudoknot domain. Biochemistry. 2010; 49:4931–3. [PubMed: 20545364]

[52]. Boehringer D, Thermann R, Ostareck-Lederer A, Lewis JD, Stark H. Structure of the hepatitis C virus IRES bound to the human 80S ribosome: remodeling of the HCV IRES. Structure. 2005; 13:1695–706. [PubMed: 16271893]

[53]. Otto GA, Puglisi JD. The pathway of HCV IRES-mediated translation initiation. Cell. 2004; 119:369–80. [PubMed: 15507208]

[54]. Fraser CS, Hershey JW, Doudna JA. The pathway of hepatitis C virus mRNA recruitment to the human ribosome. Nat. Struct. Mol. Biol. 2009; 16:397–404. [PubMed: 19287397]

[55]. Hammond JA, Rambo RP, Filbin ME, Kieft JS. Comparison and functional implications of the 3D architectures of viral tRNA-like structures. RNA. 2009; 15:294–307. [PubMed: 19144910]

[56]. Paulin FE, West MJ, Sullivan NF, Whitney RL, Lyne L, Willis AE. Aberrant translational control of the c-myc gene in multiple myeloma. Oncogene. 1996; 13:505–13. [PubMed: 8760292]

[57]. Pelengaris S, Khan M, Evan G. c-MYC: more than just a matter of life and death. Nat. Rev. Cancer. 2002; 2:764–76. [PubMed: 12360279]

[58]. Nanbru C, Lafon I, Audigier S, Gensac MC, Vagner S, Huez G, Prats AC. Alternative translation of the proto-oncogene c-myc by an internal ribosome entry site. J. Biol. Chem. 1997; 272:32061–6. [PubMed: 9405401]

[59]. Jopling CL, Spriggs KA, Mitchell SA, Stoneley M, Willis AE. L-Myc protein synthesis is initiated by internal ribosome entry. RNA. 2004; 10:287–98. [PubMed: 14730027]

[60]. Gesteland RF, Weiss RB, Atkins JF. Recoding: reprogrammed genetic decoding. Science. 1992; 257:1640–1. [PubMed: 1529352]

[61]. Namy O, Rousset JP, Napthine S, Brierley I. Reprogrammed genetic decoding in cellular gene expression. Mol. Cell. 2004; 13:157–68. [PubMed: 14759362]

[62]. Fang Y, et al. Efficient −2 frameshifting by mammalian ribosomes to synthesize an additional arterivirus protein. Proc. Natl. Acad. Sci. U. S. A. 2012; 109:E2920–8. [PubMed: 23043113]

[63]. Chung BY, Miller WA, Atkins JF, Firth AE. An overlapping essential gene in the Potyviridae. Proc. Natl. Acad. Sci. U. S. A. 2008; 105:5897–902. [PubMed: 18408156]

[64]. Craigen WJ, Caskey CT. Expression of peptide chain release factor 2 requires high-efficiency frameshift. Nature. 1986; 322:273–5. [PubMed: 3736654]

[65]. Mejlhede N, Atkins JF, Neuhard J. Ribosomal −1 frameshifting during decoding of Bacillus subtilis cdd occurs at the sequence CGA AAG. J. Bacteriol. 1999; 181:2930–7. [PubMed: 10217788]

[66]. Jacks T, Varmus HE. Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. Science. 1985; 230:1237–42. [PubMed: 2416054]

[67]. Jacks T, Madhani HD, Masiarz FR, Varmus HE. Signals for ribosomal frameshifting in the Rous sarcoma virus gag-pol region. Cell. 1988; 55:447–58. [PubMed: 2846182]

[68]. Marra MA, et al. The Genome sequence of the SARS-associated coronavirus. Science. 2003; 300:1399–404. [PubMed: 12730501]

[69]. Cao S, Chen SJ. Predicting ribosomal frameshifting efficiency. Phys. Biol. 2008; 5:016002. [PubMed: 18367782]

[70]. Jacks T, Power MD, Masiarz FR, Luciw PA, Barr PJ, Varmus HE. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. Nature. 1988; 331:280–3. [PubMed: 2447506]

[71]. Tholstrup J, Oddershede LB, Sorensen MA. mRNA pseudoknot structures can act as ribosomal roadblocks. Nucleic Acids Res. 2012; 40:303–13. [PubMed: 21908395]

[72]. Pleij CW, Rietveld K, Bosch L. A new principle of RNA folding based on pseudoknotting. Nucleic Acids Res. 1985; 13:1717–31. [PubMed: 4000943]

[73]. Manktelow E, Shigemoto K, Brierley I. Characterization of the frameshift signal of Edr, a mammalian example of programmed −1 ribosomal frameshifting. Nucleic Acids Res. 2005; 33:1553–63. [PubMed: 15767280]

[74]. Shigemoto K, Brennan J, Walls E, Watson CJ, Stott D, Rigby PW, Reith AD. Identification and characterisation of a developmentally regulated mammalian gene that utilises −1 programmed ribosomal frameshifting. Nucleic Acids Res. 2001; 29:4079–88. [PubMed: 11574691]

[75]. Ono R, Kobayashi S, Wagatsuma H, Aisaka K, Kohda T, Kaneko-Ishino T, Ishino F. A retrotransposon-derived gene, PEG10, is a novel imprinted gene located on human chromosome 7q21. Genomics. 2001; 73:232–7. [PubMed: 11318613]

[76]. Ivanov IP, Gesteland RF, Atkins JF. Antizyme expression: a subversion of triplet decoding, which is remarkably conserved by evolution, is a sensor for an autoregulatory circuit. Nucleic Acids Res. 2000; 28:3185–96. [PubMed: 10954585]

[77]. Bekaert M, Ivanov IP, Atkins JF, Baranov PV. Ornithine decarboxylase antizyme finder (OAF): fast and reliable detection of antizymes with frameshifts in mRNAs. BMC Bioinformatics. 2008; 9:178. [PubMed: 18384676]

[78]. Murakami Y, Matsufuji S, Kameji T, Hayashi S, Igarashi K, Tamura T, Tanaka K, Ichihara A. Ornithine decarboxylase is degraded by the 26S proteasome without ubiquitination. Nature. 1992; 360:597–9. [PubMed: 1334232]

[79]. Matsufuji S, Matsufuji T, Miyazaki Y, Murakami Y, Atkins JF, Gesteland RF, Hayashi S. Autoregulatory frameshifting in decoding mammalian ornithine decarboxylase antizyme. Cell. 1995; 80:51–60. [PubMed: 7813017]

[80]. Petros LM, Howard MT, Gesteland RF, Atkins JF. Polyamine sensing during antizyme mRNA programmed frameshifting. Biochem. Biophys. Res. Commun. 2005; 338:1478–89. [PubMed: 16269132]

[81]. Dulude D, Baril M, Brakier-Gingras L. Characterization of the frameshift stimulatory signal controlling a programmed −1 ribosomal frameshift in the human immunodeficiency virus type 1. Nucleic Acids Res. 2002; 30:5094–102. [PubMed: 12466532]

[82]. Watts JM, Dang KK, Gorelick RJ, Leonard CW, Bess JW Jr. Swanstrom R, Burch CL, Weeks KM. Architecture and secondary structure of an entire HIV-1 RNA genome. Nature. 2009; 460:711–6. [PubMed: 19661910]

[83]. Mouzakis KD, Lang AL, Vander Meulen KA, Easterday PD, Butcher SE. HIV-1 frameshift efficiency is primarily determined by the stability of base pairs positioned at the mRNA entrance channel of the ribosome. Nucleic Acids Res. 2013; 41:1901–13. [PubMed: 23248007]

[84]. Baranov PV, Henderson CM, Anderson CB, Gesteland RF, Atkins JF, Howard MT. Programmed ribosomal frameshifting in decoding the SARS-CoV genome. Virology. 2005; 332:498–510. [PubMed: 15680415]

[85]. Ishimaru D, et al. RNA dimerization plays a role in ribosomal frameshifting of the SARS coronavirus. Nucleic Acids Res. 2012; 41:2594–2608. [PubMed: 23275571]

[86]. Yu CH, Noteborn MH, Pleij CW, Olsthoorn RC. Stem-loop structures can effectively substitute for an RNA pseudoknot in −1 ribosomal frameshifting. Nucleic Acids Res. 2011; 39:8952–9. [PubMed: 21803791]

[87]. Cabrita LD, Dobson CM, Christodoulou J. Protein folding on the ribosome. Curr. Opin. Struct. Biol. 2010; 20:33–45. [PubMed: 20149635]

[88]. Komar AA. A pause for thought along the co-translational folding pathway. Trends Biochem. Sci. 2009; 34:16–24. [PubMed: 18996013]

[89]. Huard FP, Deane CM, Wood GR. Modelling sequential protein folding under kinetic control. Bioinformatics. 2006; 22:e203–10. [PubMed: 16873473]

[90]. Purvis IJ, Bettany AJ, Santiago TC, Coggins JR, Duncan K, Eason R, Brown AJ. The efficiency of folding of some proteins is increased by controlled rates of translation in vivo. A hypothesis. J Mol Biol. 1987; 193:413–7. [PubMed: 3298659]

[91]. Kimchi-Sarfaty C, Oh JM, Kim IW, Sauna ZE, Calcagno AM, Ambudkar SV, Gottesman MM. A "silent" polymorphism in the MDR1 gene changes substrate specificity. Science. 2007; 315:525–8. [PubMed: 17185560]

[92]. Wen JD, Lancaster L, Hodges C, Zeri AC, Yoshimura SH, Noller HF, Bustamante C, Tinoco I. Following translation by single ribosomes one codon at a time. Nature. 2008; 452:598–603. [PubMed: 18327250]

[93]. Qu X, Wen JD, Lancaster L, Noller HF, Bustamante C, Tinoco I Jr. The ribosome uses two active mechanisms to unwind messenger RNA during translation. Nature. 2011; 475:118–21. [PubMed: 21734708]

[94]. Chartrand P, Meng XH, Singer RH, Long RM. Structural elements required for the localization of ASH1 mRNA and of a green fluorescent protein reporter particle in vivo. Curr. Biol. 1999; 9:333–6. [PubMed: 10209102]

[95]. Gonzalez I, Buonomo SB, Nasmyth K, von Ahsen U. ASH1 mRNA localization in yeast involves multiple secondary structural elements and Ash1 protein translation. Curr. Biol. 1999; 9:337–40. [PubMed: 10209099]

[96]. Chartrand P, Meng XH, Huttelmaier S, Donato D, Singer RH. Asymmetric sorting of ash1p in yeast results from inhibition of translation by localization elements in the mRNA. Mol. Cell. 2002; 10:1319–30. [PubMed: 12504008]

[97]. Nackley AG, Shabalina SA, Tchivileva IE, Satterfield K, Korchynskyi O, Makarov SS, Maixner W, Diatchenko L. Human catechol-O-methyltransferase haplotypes modulate protein expression by altering mRNA secondary structure. Science. 2006; 314:1930–3. [PubMed: 17185601]

[98]. Diatchenko L, et al. Genetic basis for individual variations in pain perception and the development of a chronic pain condition. Hum. Mol. Genet. 2005; 14:135–43. [PubMed: 15537663]

[99]. Tsao D, Shabalina SA, Gauthier J, Dokholyan NV, Diatchenko L. Disruptive mRNA folding increases translational efficiency of catechol-O-methyltransferase variant. Nucleic Acids Res. 2011; 39:6201–12. [PubMed: 21486747]

[100]. Kertesz M, Wan Y, Mazor E, Rinn JL, Nutter RC, Chang HY, Segal E. Genome-wide measurement of RNA secondary structure in yeast. Nature. 2010; 467:103–7. [PubMed: 20811459]

[101]. Westhof E, Romby P. The RNA structurome: high-throughput probing. Nat. Methods. 2010; 7:965–7. [PubMed: 21116245]

[102]. Le Quesne JP, Stoneley M, Fraser GA, Willis AE. Derivation of a structural model for the c-myc IRES. J. Mol. Biol. 2001; 310:111–26. [PubMed: 11419940]
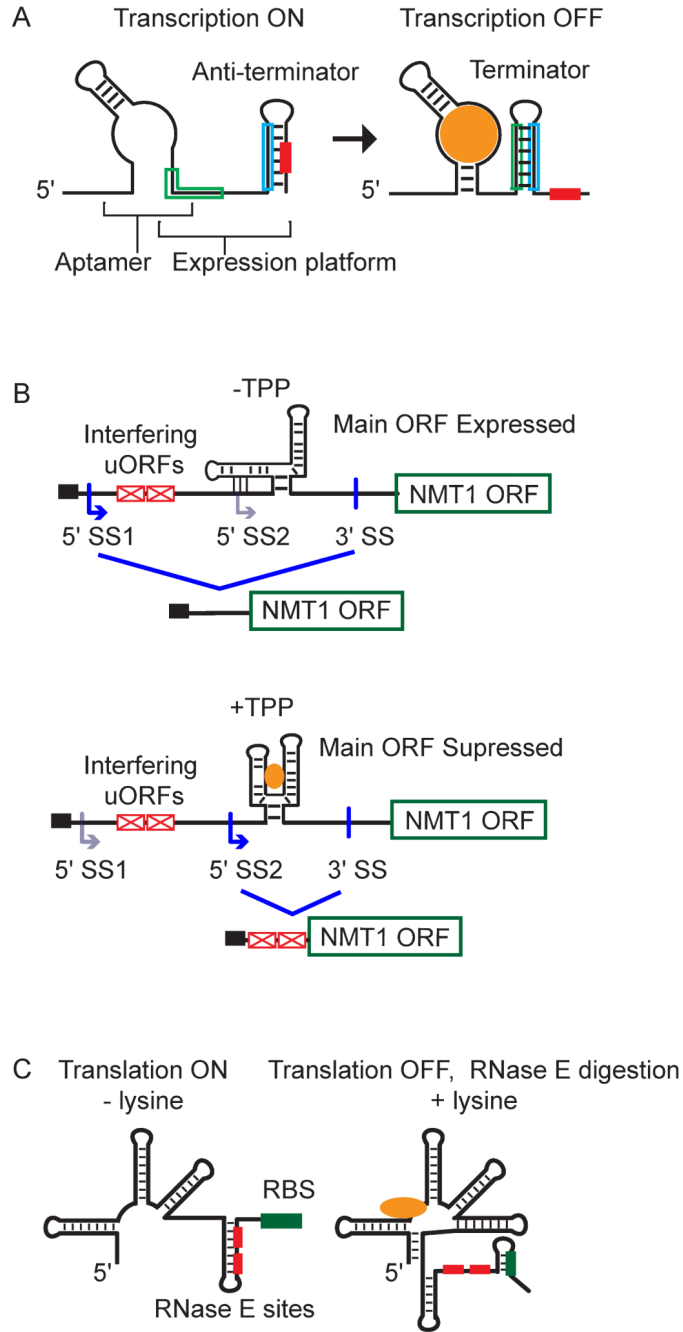
Figure 1. Mechanisms of riboswitch regulation

(A) Schematic example of transcription termination resulting from ligand binding to the aptamer domain of a riboswitch and subsequent stabilization of a terminator stem-loop and exposure of a poly-uracil stretch (red box). (B) Splicing regulation by a TPP riboswitch in the *N. crassa NMT1* gene [19]. The preferred 5′ splice site (SS2) is obscured in the absence of ligand. Ligand binding attenuates ORF protein expression by allowing use of SS2. (C) Combined translational and stability control by a lysine riboswitch. Ligand binding

sequesters the RBS (green) and exposes two RNase E cleavage sites (red), leading to mRNA degradation.
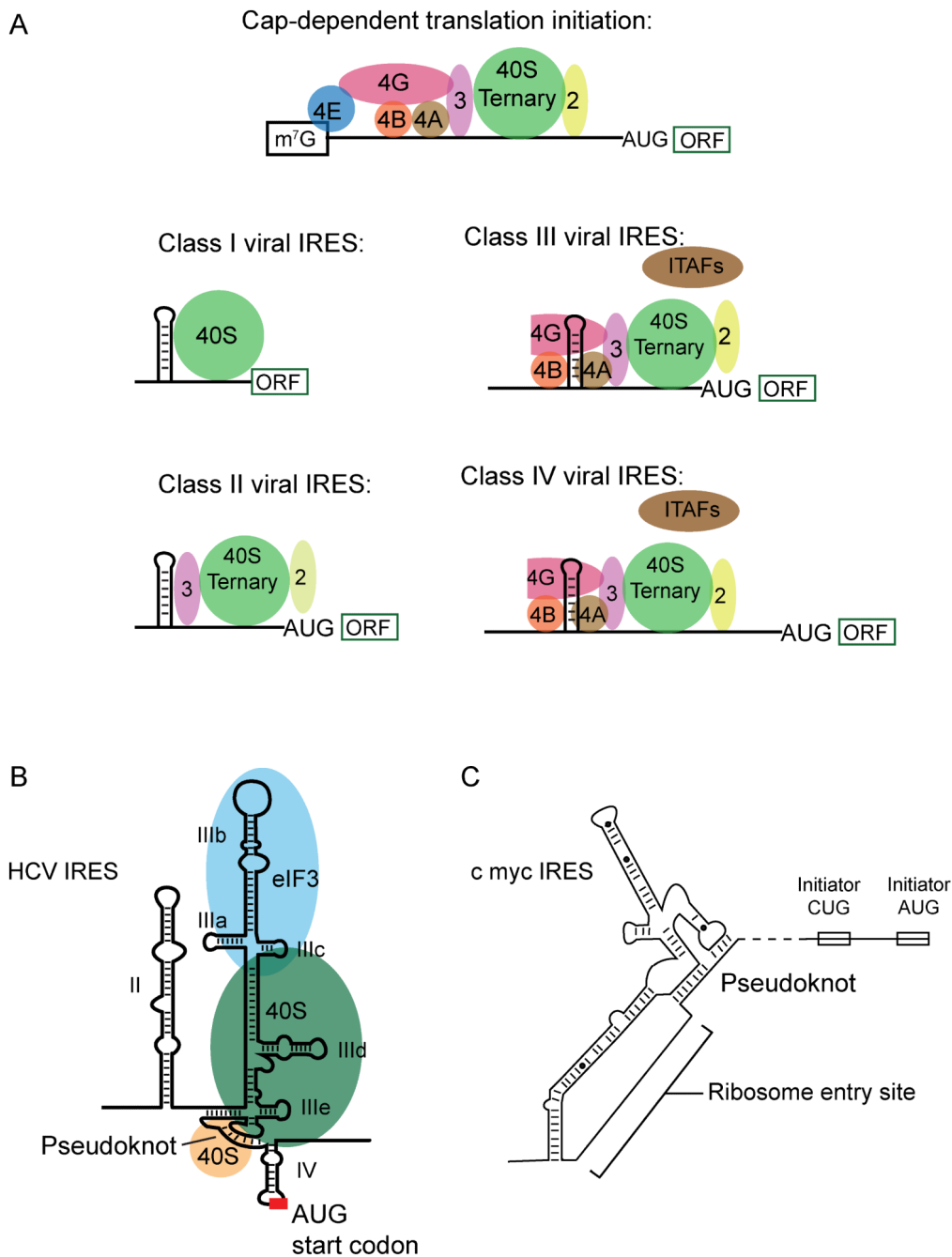
Figure 2. Cap-dependent translation initiation and classes of IRES-mediated initiation

(A) Schematic of canonical cap-dependent initiation [102] and viral IRES classes. (B) Model of HCV IRES initiation. The pseudoknot at the base of domain III mimics a tRNA [51] and may position the AUG in the ribosome active site. (C) Simplified model of the *c-myc* IRES element. The ribosome entry site is adjacent to an RNA pseudoknot [102].
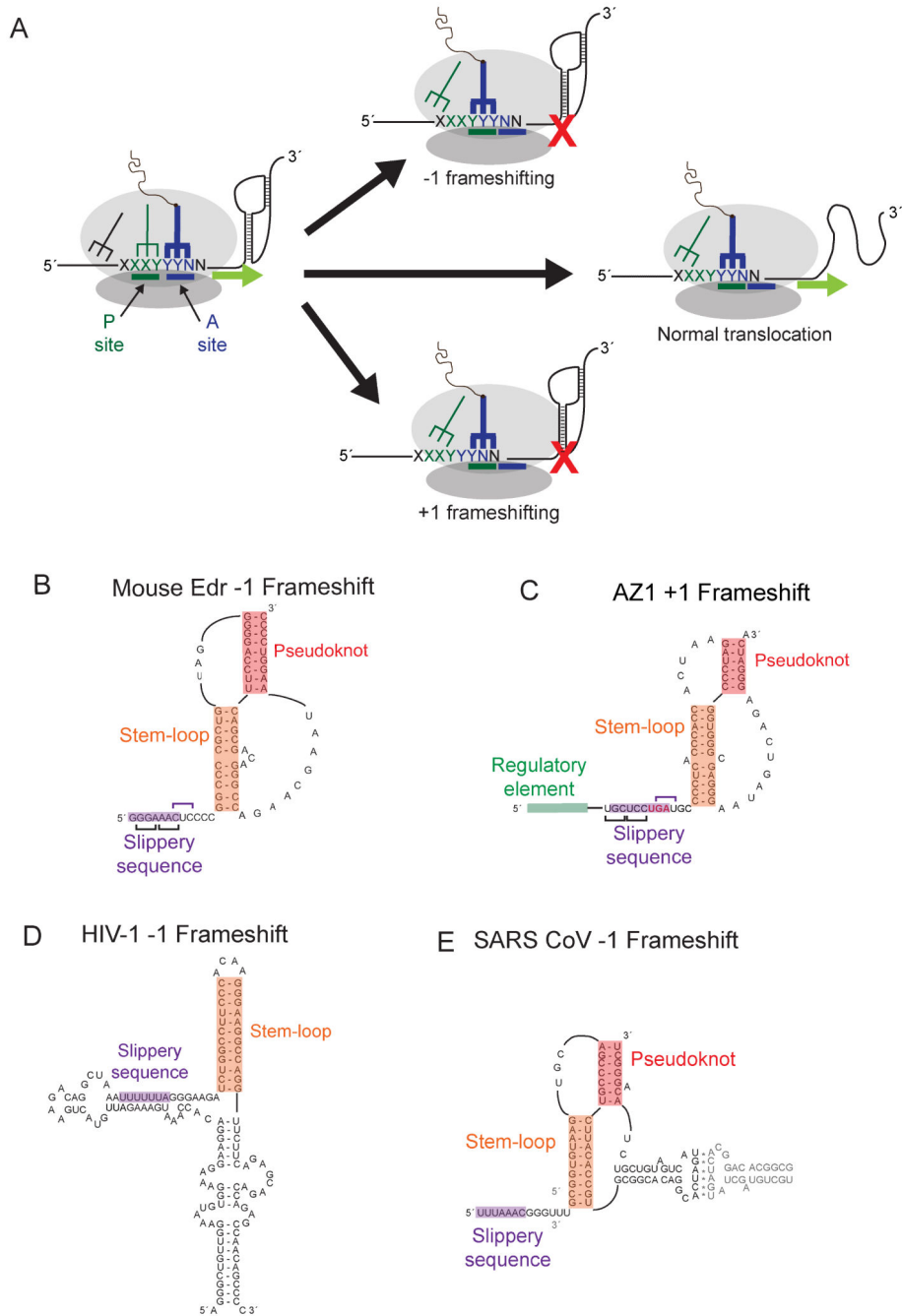
Figure 3. Structural diversity within programmed ribosomal frameshifting elements
(A) Schematic diagram of −1 and +1 programmed frameshifting events. A translating ribosome encounters a slippery sequence upstream of a stable RNA structure. The post-translocation position of the peptide-linked tRNA (blue structure) is shown for normal translocation (middle) and −1 and +1 frameshifting events. Proposed secondary structures for (B) the −1 frameshift element for the mouse *Edr* gene [73], (C) the +1 frameshifting region of the human AZ1 gene [76,79] (D) the gag-pol frameshift in HIV-1 [82], and (E) the

replicase-transcriptase frameshift of SARS-CoV [84]. In each example, the "slippery sequence" (purple box) is located upstream of a stable structure (orange and red boxes).
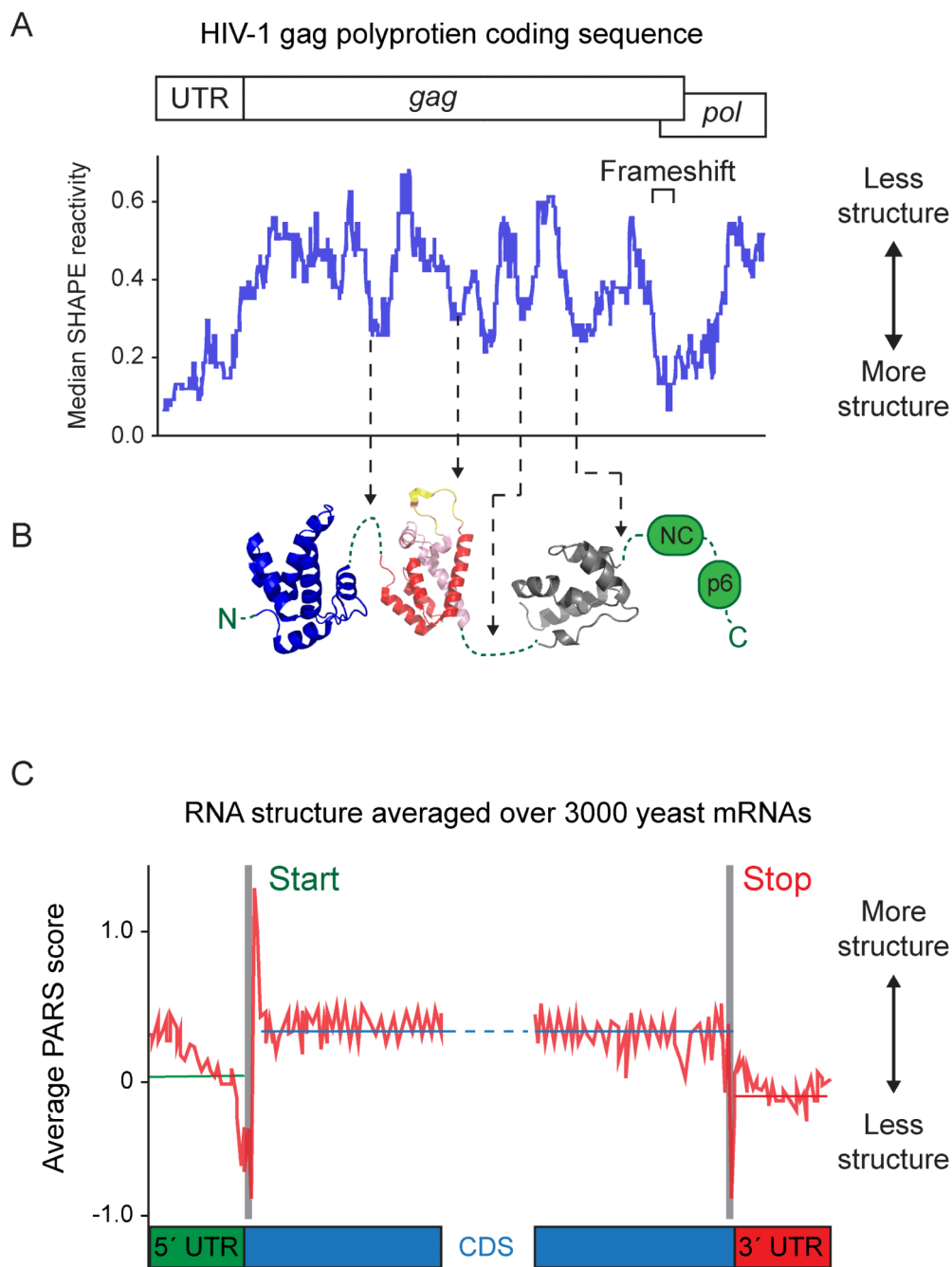
Figure 4. Relationships between RNA structure and protein domain organization
(A) Median SHAPE reactivities over 75-nucleotide windows, which provide a metric for
overall RNA structure, for the 5′ end of the HIV-1 genome [82]. (B) Low SHAPE
reactivities, reflecting higher levels of RNA structure, correlate with unstructured protein
linkers in the gag polyprotein. Interdomain linkers (green) and the intra-domain loops
(yellow) are illustrated between the independently folding domains of *gag* (blue, red, pink,

and grey structures). (C) PARS scores averaged over 3000 yeast mRNAs for 5′ UTR (green), coding region (blue), and 3′ UTR (red) regions [100].