

Review Article

Before It Gets Started: Regulating Translation at the 5' UTR

**Patricia R. Araujo,¹ Kihoon Yoon,^{1,2} Daijin Ko,³ Andrew D. Smith,⁴ Mei Qiao,¹
Uthra Suresh,¹ Suzanne C. Burns,¹ and Luiz O. F. Penalva^{1,5}**

¹ Greehey Children's Cancer Research Institute, UTHSCSA, San Antonio, TX 78229-3900, USA

² Department of Epidemiology and Biostatistics, UTHSCSA, San Antonio, TX 78229-3900, USA

³ Department of Management Science and Statistics, UTSA, San Antonio, TX 78249-0631, USA

⁴ Molecular and Computational Biology, Department of Biological Sciences, USC, Los Angeles, CA 90089-2910, USA

⁵ Department of Cellular and Structural Biology, UTHSCSA, San Antonio, TX 78229-3900, USA

Correspondence should be addressed to Luiz O. F. Penalva, penalva@uthscsa.edu

Received 2 January 2012; Revised 22 February 2012; Accepted 11 March 2012

Academic Editor: Thomas Preiss

Copyright © 2012 Patricia R. Araujo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Translation regulation plays important roles in both normal physiological conditions and diseases states. This regulation requires cis-regulatory elements located mostly in 5' and 3' UTRs and trans-regulatory factors (e.g., RNA binding proteins (RBPs)) which recognize specific RNA features and interact with the translation machinery to modulate its activity. In this paper, we discuss important aspects of 5' UTR-mediated regulation by providing an overview of the characteristics and the function of the main elements present in this region, like uORF (upstream open reading frame), secondary structures, and RBPs binding motifs and different mechanisms of translation regulation and the impact they have on gene expression and human health when deregulated.

1. Translation Regulation

Gene expression can be modulated at multiple levels from chromatin modification to mRNA translation. Despite the importance of transcriptional regulation, it is clear at this point that mRNA levels cannot be used as a sole parameter to justify the protein content of a cell. In fact, in a recent study from our lab, we determined that a direct correlation between mRNA and protein exists for less than a third of analyzed genes in a human cell line. Moreover, our analysis suggested that translation regulation contributes considerably to the protein variation as several parameters related to translation like 5' UTR, 3' UTR, coding sequence length, presence of uORFs and amino acid composition, and so forth showed good correlations with the obtained mRNA/protein ratios [1]. Translation regulation functions as an important switch when rapid changes in gene expression are required in response to internal and external stimuli (*PDGF2*, *VEGF*, *TGF β* are examples of genes controlled in such way). Translation regulation also plays a significant role during development and cell differentiation by altering the levels of expression of specific mRNA subsets during

a particular time window while the majority of transcripts remain unchanged (reviewed in [2–4]).

In this paper, we will focus on the importance of 5' UTR mediated regulation and the different functional elements present in this region with the exception of IRES which is discussed in a different article of this issue. The main regulatory elements in 5' UTR are secondary structures (including IRES), binding sites for RNA binding proteins, uAUGs and uORFs (Figure 1).

2. 5' UTR

The average length of 5' UTRs is ~100 to ~220 nucleotides across species [5]. In vertebrates, 5' UTRs tend to be longer in transcripts encoding transcription factors, protooncogenes, growth factors and their receptors, and proteins that are poorly translated under normal conditions [6]. High GC content is also a conserved feature, with values surpassing 60% in the case of warm-blooded vertebrates. In the context of hairpin structures, GC content can affect protein translation efficiency independent of hairpin thermal stability and

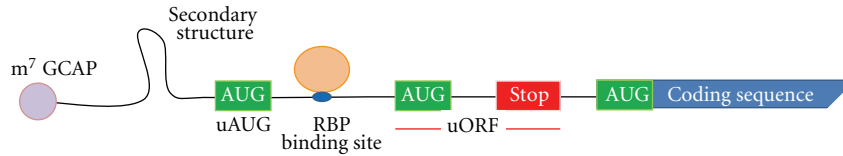


FIGURE 1: Regulatory elements present in 5' UTR.

hairpin position [7]. UTRs of eukaryotic mRNAs also display a variety of repeats that include short and long interspersed elements (SINEs and LINEs, resp.), simple sequence repeats (SSRs), minisatellites, and macrosatellites [5].

Translation initiation in eukaryotes requires the recruitment of ribosomal subunits at either the 5' m7G cap structure. The initiation codon is generally located far downstream, requiring ribosomal movement to this site. This movement appears to be nonlinear for some mRNAs (i.e., ribosomal subunits appear to bypass (shunt) segments of the 5' UTR as they move in the direction of the AUG). Shunting could allow mRNAs containing uAUGs or hairpin structures to be translated efficiently. Important examples are provided by the cauliflower mosaic virus [8] and adenovirus [9] mRNAs. The mechanism of ribosomal shunting is rather complex requiring mRNA-rRNA base pairing [10].

Genes presenting differences in the 5' UTR of their transcripts are relatively common. 10–18% of genes express alternative 5' UTR by using multiple promoters [11, 12] while alternative splicing within UTRs is estimated to affect 13% of genes in the mammalian transcriptome [13]. These variations in 5' UTR can function as important switches to regulate gene expression. Two important examples are provided by the cancer-related genes *BRCA1* (breast cancer 1) and *TGF- β* (transforming growth factor β). *BRCA1* is a tumor suppressor, frequently mutated in breast cancer with functions in cell cycle, apoptosis, and DNA damage repair. *BRCA1* produces two different transcripts that derive from two different promoters and therefore display differences in their 5' UTR. A shorter transcript is expressed in cancerous as well as noncancerous breast tissue and efficiently translated, while a longer transcript is predominantly expressed in breast cancers. The presence of several uAUGs and a more complex structure dramatically affect the translation of this longer transcript. This causes an overall decrease in *BRCA1* levels in tumor cells, leading to a relief in growth inhibition [14]. *TGF- β* is implicated in a large number of processes that include cell proliferation, migration, wound repair, development, tumorigenesis and immunosuppression. There are three known isoforms: $\beta 1$, $\beta 2$, and $\beta 3$. *TGF- $\beta 3$* produces two alternative transcripts: a 3.5 kb transcript with a very long 5' UTR (1.1 kb) and a 2.6 kb transcript with a shorter 5' UTR (0.23 kb). The presence of 11 uORFs in the longer transcript dramatically inhibits its translation while the shorter transcript is efficiently translated [15, 16].

3. Regulation by Secondary Structure

Secondary structures can function as major regulatory tools in 5' UTRs. A correlation with gene function has been

suggested; secondary structures have been determined to be particularly prevalent among mRNAs encoding transcription factors, protooncogenes, growth factors, and their receptors and proteins poorly translated under normal conditions. >90% of transcripts in these classes have 5' UTRs containing stable secondary structures with average free energies less than -50 kcal/mol. 60% of these stable secondary structures are positioned very close to the cap structure [6]. These structures are very effective in inhibiting translation. In fact, a hairpin situated close to the cap with a free energy of -30 kcal/mol would be sufficient to block the access of the preinitiation complex to the mRNA. When located further away in the 5' UTR, hairpins require a free energy stronger than -50 kcal/mol to be able to block translation [17, 18]. Stable secondary structure can resist the unwinding activity of the helicase *elf4A*. This effect can be overcome partially by the overexpression of *elf4A* in partnership with *elf4B* [19]. mRNAs with a highly structured 5' UTR like protooncogenes and other growth factors use cap-dependent translation initiation. Not surprisingly, the overexpression of components of the translation initiation machinery including *elf4E* has been linked to tumorigenesis (reviewed in [18, 20]).

The gene *TGF- $\beta 1$* provides a good example of translation inhibition mediated by secondary structure [21, 22]. An evolutionary conserved motif in the 5' UTR forms a stable stem loop. However, this structure by itself is not sufficient to block translation. Translation repression of *TGF- $\beta 1$* depends on increased binding of the RNA binding protein YB-1 to the *TGF- $\beta 1$* transcript [23]. It was then proposed that YB-1 binds the 5' UTR of *TGF- $\beta 1$* with high affinity thanks to its GC content and cooperates with the stem loop to inhibit *TGF- $\beta 1$* translation by facilitating duplex formation [24].

4. Regulation by RNA Binding Proteins

The human genome is predicted to encode circa 1,000 RNA binding proteins (RBPs) with a large percentage of them implicated in translation. They could be categorized into two main groups: RBPs that are part of the basic translation machinery and required for the translation of all expressed mRNAs (examples: PABPI, *elf4E*) and RBPs that function in a more selective way by controlling either positively or negatively the levels of translation of specific target mRNAs (examples: HuR, Musashi1). Regarding this later group, it has been observed that RBPs can use distinct mechanisms to increase or inhibit translation. Although several exceptions are known, it can be said that RBPs often recognize specific motifs in UTRs and interact with the translation machinery

to control expression. Interference with translation normally takes place during the initiation step (reviewed in [25]).

The best characterized example of RBP-mediated regulation involving 5' UTRs is provided by the iron regulatory proteins (IRP 1 and 2). These proteins recognize a highly conserved stem loop structure with circa 30 nucleotides, known as the iron response element (IRE). The most important features include a hexanucleotide loop with the sequence CAGYCX (Y = U or A; X = U, C, or A) and a 5 bp upper stem that is separated from a lower stem of variable length by an unpaired cytosine. This regulation is crucial in maintaining cellular iron homeostasis as a large number of mRNAs connected to iron storage and metabolism including ferritin, mitochondrial aconitase, succinate dehydrogenase-iron protein, erythroid 5-aminolevulinic synthetase (eALAS), and an iron-exportin molecule named ferroportin (FPN1) have their expression modulated by this system. When cellular iron levels are low, IRP1 and IRP2 bind the IRE and block translation of the downstream ORF. When intracellular iron levels are high, the RNA binding activity of both IRPs is reduced (Figure 2(a)). IREs tend to be positioned close to the cap, which causes a steric inhibition of the binding of 40S ribosomal subunits to the transcript. When located distant to the cap, rather than affecting 40S recruitment, the IRE-IRP complex blocks ribosomal scanning (reviewed in [26]). An interesting bypass of the IRE/IRP mechanism can be observed in iron-starved duodenal and erythroid precursor cells. An upstream promoter is used to generate FPN1 pre-mRNAs containing one more exon that is connected by alternative splicing to a splice acceptor in the 3' of the IRE. A mature FPN1 transcript containing the same open reading frame is generated; however, the 5' UTR does not contain the IRE [27]. Therefore, these cells express the alternative FPN1 isoform in an iron-independent manner [27, 28]. Mutations affecting IREs can lead to diseases. This is the case of hereditary hyperferritinemia-cataract syndrome (HHCS), a genetic autosomal dominant disorder in which aggregation and crystallization of ferritin in the lens leads to bilateral cataracts [29].

RBP-mediated regulation can be very elaborate and involve multiple steps. One good example showing the crosstalk between factors and distinct regulatory processes is the male-specific-lethal 2 (*msl-2*) gene in *Drosophila*, a main player in dosage compensation. The female-specific RNA binding protein sex lethal (SXL) participates in multiple aspects of *msl-2* regulation where *msl-2* expression must be prevented (Figure 2(b)). Regulation starts at the splicing level; SXL binds to two polyU stretches located in an intron that is part of the 5' UTR. This process causes intron retention and preserves critical sequences that later will be used in translation regulation [30, 31]. In the cytoplasm, the same SXL protein will function as a translation repressor of *msl-2* in two distinct mechanisms taking place at the 3' and 5' UTR [32]. SXL binds U-rich sequences in the 3' UTR and recruits the corepressor protein UNR (upstream of N-ras) and PABP blocking the recruitment of the pre-initiation complex to the 5' end of the mRNA [33–35]. To assure that *msl-2* gets fully repressed, a second regulatory step also mediated by SXL takes place at the 5' UTR. This

repression involves a novel regulatory mechanism where crosstalk between SXL and a uORF takes place to efficiently repress translation [36]. The 5' UTR of *msl-2* contains 3 uORFs but only the 3rd one is involved in the repression. Interestingly, this repression is very weak in the absence of SXL (~2-fold), but when present, SXL binds a poly U stretch a few nucleotides away from the uAUG and increases this repression to more than 14-fold. SXL acts by boosting translation initiation at the uAUG and not by acting as a simple steric arrest of scanning ribosomes. This effect may take place via an interaction between SXL and translation initiation factors; possibly members of eIF3 component as indicated by a two-hybrid screening. This mechanism potentially affects a large number of mRNAs; 268 transcripts in *Drosophila* were determined to contain SXL binding motifs associated with uAUG spaced at an appropriate distance. For instance, a reporter construct containing the 5'UTR of the gene *Irr47* was repressed ~4-fold by SXL protein [36].

RBPs can have antagonistic functions when regulating translation. An interesting example is the regulation of p21 in the context of replicative senescence, a cellular state where cells enter an irreversible growth arrest. Induction of p21 is required to initiate the process, and to inhibit cdk2-cyclin E complexes. The 5' UTR of p21 contains a GC-rich sequence that forms a stem loop. This element is recognized by two RBPs with distinct properties: CUGBP1 and calreticulin (CRT). Competition between the two proteins determines final levels of p21 expression and establishes if cells will proliferate or undergo growth arrest and senescence. Binding of CUGBP1 to p21 mRNA is dramatically increased in senescence compared to young fibroblast cells. Protein levels do not change during the process and this increase in activity is due to phosphorylation. On the other hand, CRT IPs showed a four-to-fivefold reduction of activity in senescence cells due to a decrease in expression. Both proteins were shown to affect p21 translation. However, while CUGBP1 functions as an activator, CRT acts as a repressor. Since the two proteins have opposing activity in senescent cells, they were examined to see if they compete for interaction with p21 mRNA and to control its translation. Increasing amounts of one protein were able to reverse the binding of the other protein to p21 mRNA and its effect on translation; affinity to the binding site is rather different as CUGBP1 had to be present in the binding reactions at a four-to-eightfold molar excess to CRT to antagonize its binding to p21 mRNA and impact its translation [37].

5. Regulation by uORFs and Upstream AUGs

uORFs and uAUGs are major regulatory elements in 5' UTRs. As their names suggest, uORFs are sequences defined by a start and stop codons upstream of the main coding region while uAUGs are start codons without an in-frame downstream stop codon located upstream of the main coding region. A large percentage of the human transcriptome contains uORF and/or uAUGs, with values ranging between 44 and 49% [38, 39]. Similar numbers are found in the

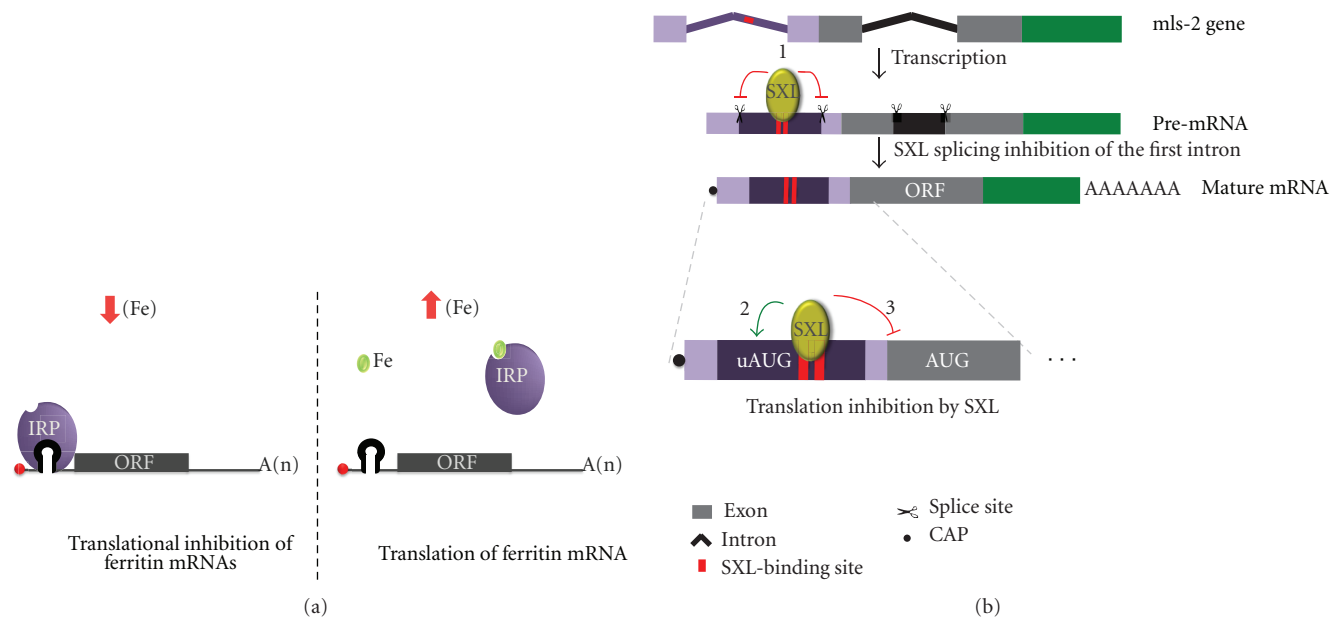


FIGURE 2: Translational regulation by RNA binding proteins. (a) In iron-deficient cells, IRPs bind to the IRE localized in the 5' UTR of ferritin mRNA, blocking its translation. Once cellular iron levels increase, a complex containing Fe binds to IRPs. Thus, these proteins are allosterically modified, which reduces IRP-IRE binding and allows the translation of ferritin mRNAs. (b) *msl-2* gene regulation in females flies. After transcription in the nucleus, SXL specifically binds to intronic U-rich regions of *msl-2* pre-mRNA and inhibits the intron removal (1). In the cytoplasm, SXL binds to the same elements localized now in the 5' UTR of mature *msl-2* mRNA, enhances the translation initiation of an upstream ORF (2), and prevents the main ORF translation (3). The regulatory elements in the 3' UTR of *msl-2* mRNA were not represented.

mouse transcriptome. Although these numbers might sound high, both uORFs and uAUGs are less frequent than expected by chance, suggesting that they are under selective pressure. uORFs and uAUG are overrepresented in particular subgroups like transcription factors, growth factors, and their receptors and proto-oncogenes [6]. Both uORFs and uAUGs are extremely diverse varying in position in relation to the cap and main AUG, number per transcript and length (in the case of uORFs) [38]. Supplementary Table 1 (in Supplementary Material available online at doi:10.1155/2012/475731) provides a comprehensive list of uORFs and uAUGs present in the human transcriptome. uORFs and uAUGs have not been extensively analyzed in terms of conservation. A pilot study done with a subset of human, mouse, and rat transcripts indicated that both elements are moderately conserved as 38% of uORFs and 24% of uAUGs were determined to be conserved among three species [39]. The modest conservation of uORFs combined with the fact that their average length (20 nucleotides) is expected by chance and uAUGs provide a stronger suppression in comparison to uORFs suggests that many uAUGs have been neutralized in the process of evolution by the acquisition of a downstream stop codon. It has been proposed then that only a few uORFs, very likely the conserved ones, have been recruited for expression regulation [39]. In yeast, it has been shown that uORFs are statistically underrepresented in 5' UTRs and were removed by selective pressure, indicating similarly that the remaining uORFs may be implicated in translation regulation [40].

Although, overall it has been suggested that uORFs are negatively correlated with protein production [1, 38, 41] until now, functional activity has been demonstrated for only a limited number of uORFs and uAUGs. In Figure 3, we show examples of the impact uAUGs can have on translation efficiency. Among the most relevant features that can contribute to functionality are long 5' cap-to-uORF distance, sequence conservation, context in which the AUG is located, strength of the initiation site for the ORF, length of the uORF, and number of AUGs in the 5' UTR [38, 42]. Different outcomes have been observed when a ribosome encounters a uAUG or uORF [43]. Since the number of characterized events is still small, it is hard to define general mechanisms; we describe then a few well-characterized and relevant events. Leaky scanning is defined when a proportion of the scanning complexes bypass the uAUG or uORF and continue scanning for the next AUG. In this case, the upstream AUG acts as a "decoy" from the ORF AUG, functioning as a negative regulator of translation at least for some fraction of ribosomes. The production of cis-acting peptides by uORFs can reduce the initiation of translation of the downstream ORF by stalling the ribosome at the end of the uORF [44]. A classical example is provided by the evolutionarily conserved eukaryotic arginine attenuator peptide (AAP), that negatively controls the translation of proteins involved in the *de novo* fungal arginine biosynthesis in high arginine concentration [45]. In this scenario, arginine changes AAP conformation and/or P site environment causing ribosomal stalling at the termination codon of AAP uORF [46, 47]. AAP also reduces translation elongation

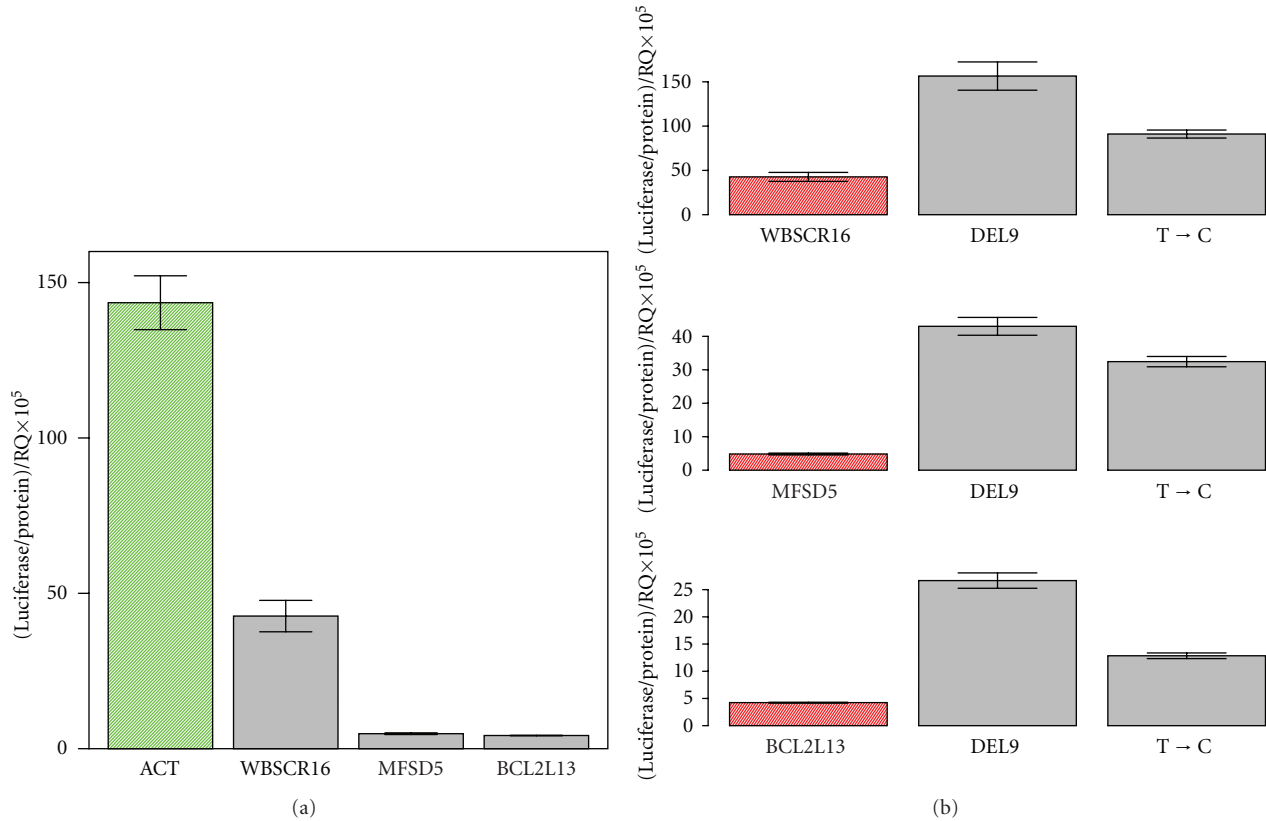


FIGURE 3: Impact of uAUG sequences on translation regulation. (a) Comparison of luciferase levels obtained for constructs having the 5' UTR of the gene ACT (control) and genes containing uAUG: WBSR16, MFSD5, and BCL2L13. (b) Deletion or mutation of uAUG sequence present in genes WBSR16, MFSD5 and BCL2L13 reverts translation repression as seen as an increase in luciferase activity.

by ribosome stalling when the uORF is inserted within an encoding sequence [48]. Another classical example of uORF-mediated regulation comes from yeast. Four uORFs are present in the 5' UTR of the transcription factor GCN4. The first of the four uORFs is always efficiently translated regardless of the nutritional conditions. In unperturbed cells, rapid reloading of ribosomes and initiation cofactors allow translation of uORFs 2–4 while inhibiting the translation of the main ORF. In situations of amino acid starvation, initiation factors are scarce, resulting in a decelerated reloading of ribosomes and scanning across the sequences containing the uORFs. A functional initiation complex is reassembled only at the main coding sequence and GCN4 expressed. This mechanism allows a fast response to nutritional stress [49, 50]. Another similar example of regulated expression via uORF is the Carnitine Palmitoyltransferase 1C (CPT1C) gene. CPT1C regulates metabolism in the brain in situations of energy surplus. The presence of uORF in the 5' UTR represses the expression of the ORF. However, this repression is relieved in response to specific stress stimuli like glucose deprivation and palmitate-BSA treatment [51]. It has been suggested that uORFs can also induce mRNA degradation. A series of 5' UTR constructs containing as a reporter the cat gene from the bacterial transposon Tn9 was tested in yeast. A single nucleotide substitution was used to create a 7-codon ORF upstream of the cat gene. The uORF was translated efficiently and caused translation inhibition of

the cat ORF and destabilization of the cat mRNA [52]. A connection between uORFs and mRNA decay was also suggested based on a comparison between average levels of expression of uORF-containing and non-uORF-containing transcripts [41].

Several mutations that eliminate or create uORFs that end up altering protein levels have been connected to human diseases. Their relevance was discussed recently [53]. Predisposition to melanoma can be caused by a mutation that introduces a uORF into the 5' UTR of the gene cyclin-dependent-kinase inhibitor protein (CDKN2A) [54]. Hereditary thrombocythemia is caused by a mutation that creates a splicing variant that eliminates a uORF, leading to an increase in protein production of the gene thrombopoietin [55]. Marie Unna hereditary hypotrichosis derives from a mutation that disrupts a uORF present in the 5' UTR of the gene hairless homolog and consequently increasing its expression [56]. A transition from G to A in one of the uORFs present in the 5' UTR of TGF- β 3 transcript was determined to be associated with arrhythmogenic right ventricular cardiomyopathy/dysplasia (ARVC) [57]. Another group of five uORFs associated with diseases have been tested recently [58] using reporter assays; they include gonadal dysgenesis (SRY) [59], Van der Woude syndrome (IRF6) [60], Carney Complex Type 1 (PRKAR1A) [61], Hereditary pancreatitis (SPINK1) [62], and Thalassemia- β (HBB) [63]. This list will certainly expand as more than

500 single-nucleotide polymorphisms (SNPs) creating or deleting uORFs have been reported.

6. Searching for Novel Regulatory Elements in the 5' UTR

Only a small fraction of the posttranscriptional regulatory elements located in human 5' UTRs have been characterized. Those identified UTR elements are catalogued in a web-resource maintained by Graziano Pesole's group called UTRdb (<http://utrdb.ba.itb.cnr.it/>) [49]. In vivo methods for the identification of posttranscriptional regulatory elements in UTRs, especially those associated with RBPs, have advanced dramatically in the last five years thanks to deep sequencing technology. CLIP and RIP-Seq are methods based on the isolation of RNA protein molecules (RNPs) via immunoprecipitation, followed by RNase digestion and precise identification of RBP binding sites with deep sequencing [64]. Although the number of RBPs analyzed so far by these methods is really small (reviewed in [65]), as deep sequencing technology becomes more accessible and the methods simplified, one could expect that very soon a large portion of the human RBP binding sites in UTRs will be mapped.

Another choice to map UTR elements regulating translation is to use purely computational methods based on analyzing the UTR sequences. These methods are based on identifying degenerate ribonucleotide patterns that have the expected properties of RBP binding sites. Similar methods have been applied for nearly 30 years to identify transcriptional regulatory in promoter sequences. These methods are reaching maturity, are very widely used, and have assisted greatly in compiling databases about transcriptional regulation (e.g., TRANSFAC) [66, 67]. Although much of the work directed towards designing and refining regulatory sequence analysis algorithms in context of transcriptional regulation can be adapted to corresponding analysis problems in the context of post-transcriptional regulatory elements, there are additional complications associated with RBP binding sites. The most obvious among these is that RBPs will have secondary structural preferences, and few existing analysis tools can incorporate information about RNA folding. Similarly, because of RNA folding regulatory elements can more easily function synergistically or display concerted binding to sequence elements that are distal in the primary sequence but very close in the folded molecule. Another difficulty is the lack of example translational regulatory elements for training the analysis. Based on a handful of well-studied examples, there is often a perception that RBP binding sites are on average shorter than transcription factors (TFs) binding sites, but this perception may be due to bias in the set of RBPs receiving the most research focus [65]. One of the most powerful methods for identifying regulatory elements is phylogenetic foot-printing, which takes advantage of locally elevated evolutionary conservation to reveal functional elements [5, 50, 51]. This logic works equally well for post-transcriptional regulatory elements. Unfortunately TF binding sites are also a major confound

to direct application of computational sequence analysis for identifying 5' UTR elements involved in translation. Elements involved in transcriptional regulation reside both up- and downstream of transcription start sites, and when 5' UTRs are sufficiently short post-transcriptional regulatory elements are likely interleaved with TF binding sites.

Ultimately the best methods for identifying post-transcriptional regulatory elements will emerge from complementary application of experimental and computational techniques.

Acknowledgments

Research in Penalva's lab is supported by the Voelcker Foundation, Children's Brain Tumor foundation, and 5R21HG004664-02 and 1R01HG006015-01A1.

References

- [1] C. Vogel, R. de Sousa Abreu, D. J. Ko et al., "Sequence signatures and mRNA concentration can explain two-thirds of protein abundance variation in a human cell line," *Molecular Systems Biology*, vol. 6, article 400, 2010.
- [2] R. J. Jackson, C. U. T. Hellen, and T. V. Pestova, "The mechanism of eukaryotic translation initiation and principles of its regulation," *Nature Reviews Molecular Cell Biology*, vol. 11, no. 2, pp. 113–127, 2010.
- [3] N. Sonenberg and A. G. Hinnebusch, "Regulation of translation initiation in eukaryotes: mechanisms and biological targets," *Cell*, vol. 136, no. 4, pp. 731–745, 2009.
- [4] F. Gebauer and M. W. Hentze, "Molecular mechanisms of translational control," *Nature Reviews Molecular Cell Biology*, vol. 5, no. 10, pp. 827–835, 2004.
- [5] G. Pesole, F. Mignone, C. Gissi, G. Grillo, F. Licciulli, and S. Liuni, "Structural and functional features of eukaryotic mRNA untranslated regions," *Gene*, vol. 276, no. 1-2, pp. 73–81, 2001.
- [6] R. V. Davuluri, Y. Suzuki, S. Sugano, and M. Q. Zhang, "CART classification of human 5' UTR sequences," *Genome Research*, vol. 10, no. 11, pp. 1807–1816, 2000.
- [7] J. R. Babendure, J. L. Babendure, J. H. Ding, and R. Y. Tsien, "Control of mammalian translation by mRNA structure near caps," *RNA*, vol. 12, no. 5, pp. 851–861, 2006.
- [8] L. A. Ryabova, M. M. Pooggin, and T. Hohn, "Viral strategies of translation initiation: ribosomal shunt and reinitiation," *Progress in Nucleic Acid Research and Molecular Biology*, vol. 72, pp. 1–39, 2002.
- [9] A. Yueh and R. J. Schneider, "Translation by ribosome shunting on adenovirus hsp70 mRNAs facilitated by complementarity to 18S rRNA," *Genes and Development*, vol. 14, no. 4, pp. 414–421, 2000.
- [10] S. A. Chappell, J. Dresios, G. M. Edelman, and V. P. Mauro, "Ribosomal shunting mediated by a translational enhancer element that base pairs to 18S rRNA," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 25, pp. 9488–9493, 2006.
- [11] N. D. Trinklein, S. J. F. Aldred, A. J. Saldanha, and R. M. Myers, "Identification and functional analysis of human transcriptional promoters," *Genome Research*, vol. 13, no. 2, pp. 308–312, 2003.
- [12] T. Zhang, P. Haws, and Q. Wu, "Multiple variable first exons: a mechanism for cell—and tissue—specific gene regulation," *Genome Research*, vol. 14, no. 1, pp. 79–89, 2004.

- [13] P. Carninci, T. Kasukawa, S. Katayama et al., "The transcriptional landscape of the mammalian genome," *Science*, vol. 309, no. 5740, pp. 1559–1563, 2005.
- [14] K. Sobczak and W. J. Krzyzosiak, "Structural determinants of BRCA1 translational regulation," *The Journal of Biological Chemistry*, vol. 277, no. 19, pp. 17349–17358, 2002.
- [15] B. A. Arrick, A. L. Lee, R. L. Grendell, and R. Derynck, "Inhibition of translation of transforming growth factor- β 3 mRNA by its 5' untranslated region," *Molecular and Cellular Biology*, vol. 11, no. 9, pp. 4306–4313, 1991.
- [16] B. A. Arrick, R. L. Grendell, and L. A. Griffin, "Enhanced translational efficiency of a novel transforming growth factor β 3 mRNA in human breast cancer cells," *Molecular and Cellular Biology*, vol. 14, no. 1, pp. 619–628, 1994.
- [17] N. K. Gray and M. W. Hentze, "Regulation of protein synthesis by mRNA structure," *Molecular Biology Reports*, vol. 19, no. 3, pp. 195–200, 1994.
- [18] B. M. Pickering and A. E. Willis, "The implications of structured 5' untranslated regions on translation and disease," *Seminars in Cell and Developmental Biology*, vol. 16, no. 1, pp. 39–47, 2005.
- [19] F. Rozen, I. Edery, K. Meerovitch, T. E. Dever, W. C. Merrick, and N. Sonenberg, "Bidirectional RNA helicase activity of eucaryotic translation initiation factors 4A and 4E," *Molecular and Cellular Biology*, vol. 10, no. 3, pp. 1134–1144, 1990.
- [20] F. Robert and J. Pelletier, "Translation initiation: a critical signalling node in cancer," *Expert Opinion on Therapeutic Targets*, vol. 13, no. 11, pp. 1279–1293, 2009.
- [21] S. J. Kim, K. Park, D. Koeller et al., "Post-transcriptional regulation of the human transforming growth factor- β 1 gene," *The Journal of Biological Chemistry*, vol. 267, no. 19, pp. 13702–13707, 1992.
- [22] D. S. Romeo, K. Park, A. B. Roberts, M. B. Sporn, and S. J. Kim, "An element of the transforming growth factor- β 1 5'-untranslated region represses translation and specifically binds a cytosolic factor," *Molecular Endocrinology*, vol. 7, no. 6, pp. 759–766, 1993.
- [23] D. J. Fraser, A. O. Phillips, X. Zhang et al., "Y-box protein-1 controls transforming growth factor- β 1 translation in proximal tubular cells," *Kidney International*, vol. 73, no. 6, pp. 724–732, 2008.
- [24] R. H. Jenkins, R. Bennagi, J. Martin, A. O. Phillips, J. E. Redman, and D. J. Fraser, "A conserved stem loop motif in the 5' untranslated region regulates transforming growth factor- β 1 translation," *PLoS one*, vol. 5, no. 8, Article ID e12283, 2010.
- [25] I. Abaza and F. Gebauer, "Functional domains of Drosophila UNR in translational control," *RNA*, vol. 14, no. 3, pp. 482–490, 2008.
- [26] D. J. Goss and E. C. Theil, "Iron responsive mRNAs: a family of Fe²⁺ sensitive riboregulators," *Accounts of Chemical Research*, vol. 44, no. 12, pp. 1320–1328, 2011.
- [27] D. L. Zhang, R. M. Hughes, H. Ollivierre-Wilson, M. C. Ghosh, and T. A. Rouault, "A ferroportin transcript that lacks an iron-responsive element enables duodenal and erythroid precursor cells to evade translational repression," *Cell Metabolism*, vol. 9, no. 5, pp. 461–473, 2009.
- [28] S. Abboud and D. J. Haile, "A novel mammalian iron-regulated protein involved in intracellular iron metabolism," *The Journal of Biological Chemistry*, vol. 275, no. 26, pp. 19906–19912, 2000.
- [29] D. Girelli, R. Corrocher, L. Bisceglia et al., "Hereditary hyperferritinemia-cataract syndrome caused by a 29-base pair deletion in the iron responsive element of ferritin L-subunit gene," *Blood*, vol. 90, no. 5, pp. 2084–2088, 1997.
- [30] L. Merendino, S. Guth, D. Bilbao, C. Martínez, and J. Valcárcel, "Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3' splice site AG," *Nature*, vol. 402, no. 6763, pp. 838–841, 1999.
- [31] P. Forch, L. Merendino, C. Martinez, and J. Valcarcel, "Modulation of msl-2 5' splice site recognition by Sex-lethal," *RNA*, vol. 7, no. 9, pp. 1185–1191, 2001.
- [32] K. Beckmann, M. Grskovic, F. Gebauer, and M. W. Hentze, "A dual inhibitory mechanism restricts msl-2 mRNA translation for dosage compensation in drosophila," *Cell*, vol. 122, no. 4, pp. 529–540, 2005.
- [33] K. Duncan, M. Grskovic, C. Strein et al., "Sex-lethal imparts a sex-specific function to UNR by recruiting it to the msl-2 mRNA 3' UTR: translational repression for dosage compensation," *Genes & Development*, vol. 20, no. 3, pp. 368–379, 2006.
- [34] I. Abaza, O. Coll, S. Patalano, and F. Gebauer, "Drosophila UNR is required for translational repression of male-specific lethal 2 mRNA during regulation of X-chromosome dosage compensation," *Genes & Development*, vol. 20, no. 3, pp. 380–389, 2006.
- [35] K. E. Duncan, C. Strein, and M. W. Hentze, "The SXL-UNR corepressor complex uses a PABP-mediated mechanism to inhibit ribosome recruitment to msl-2 mRNA," *Molecular Cell*, vol. 36, no. 4, pp. 571–582, 2009.
- [36] J. Medenbach, M. Seiler, and M. W. Hentze, "Translational control via protein-regulated upstream open reading frames," *Cell*, vol. 145, no. 6, pp. 902–913, 2011.
- [37] P. Iakova, G. L. Wang, L. Timchenko et al., "Competition of CUGBP1 and calreticulin for the regulation of p21 translation determines cell fate," *The EMBO Journal*, vol. 23, no. 2, pp. 406–417, 2004.
- [38] S. E. Calvo, D. J. Pagliarini, and V. K. Mootha, "Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 18, pp. 7507–7512, 2009.
- [39] M. Iacono, F. Mignone, and G. Pesole, "UAUG and uORFs in human and rodent 5' untranslated mRNAs," *Gene*, vol. 349, pp. 97–105, 2005.
- [40] C. Lawless, R. D. Pearson, J. N. Selley et al., "Upstream sequence elements direct post-transcriptional regulation of gene expression under stress conditions in yeast," *BMC Genomics*, vol. 10, article 7, 2009.
- [41] M. Matsui, N. Yachie, Y. Okada, R. Saito, and M. Tomita, "Bioinformatic analysis of post-transcriptional regulation by uORF in human and mouse," *The FEBS Letters*, vol. 581, no. 22, pp. 4184–4188, 2007.
- [42] M. Kozak, "Constraints on reinitiation of translation in mammals," *Nucleic Acids Research*, vol. 29, no. 24, pp. 5226–5232, 2001.
- [43] D. R. Morris and A. P. Geballe, "Upstream open reading frames as regulators of mRNA translation," *Molecular and Cellular Biology*, vol. 20, no. 23, pp. 8635–8642, 2000.
- [44] M. Oyama, C. Itagaki, H. Hata et al., "Analysis of small human proteins reveals the translation of upstream open reading frames of mRNAs," *Genome Research*, vol. 14, no. 10, pp. 2048–2052, 2004.
- [45] Z. Wang, A. Gaba, and M. S. Sachs, "A highly conserved mechanism of regulated ribosome stalling mediated by fungal arginine attenuator peptides that appears independent of the charging status of arginyl-tRNAs," *The Journal of Biological Chemistry*, vol. 274, no. 53, pp. 37565–37574, 1999.

- [46] A. Gaba, Z. Wang, T. Krishnamoorthy, A. G. Hinnebusch, and M. S. Sachs, "Physical evidence for distinct mechanisms of translational control by upstream open reading frames," *The EMBO Journal*, vol. 20, no. 22, pp. 6453–6463, 2001.
- [47] C. Wu, J. Wei, P. J. Lin et al., "Arginine changes the conformation of the arginine attenuator peptide relative to the ribosome tunnel," *Journal of Molecular Biology*, vol. 416, no. 4, pp. 518–533, 2012.
- [48] P. Fang, C. C. Spevak, C. Wu, and M. S. Sachs, "A nascent polypeptide domain that can regulate translation elongation," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 12, pp. 4059–4064, 2004.
- [49] P. P. Mueller and A. G. Hinnebusch, "Multiple upstream AUG codons mediate translational control of GCN4," *Cell*, vol. 45, no. 2, pp. 201–207, 1986.
- [50] A. G. Hinnebusch, "Translational regulation of GCN4 and the general amino acid control of yeast," *Annual Review of Microbiology*, vol. 59, no. 1, pp. 407–450, 2005.
- [51] I. Lohse, P. Reilly, and K. Zaugg, "The CPT1C 5' UTR contains a repressing upstream open reading frame that is regulated by cellular energy availability and AMPK," *PLoS one*, vol. 6, no. 9, 2011.
- [52] C. C. Oliveira and J. E. G. McCarthy, "The relationship between eukaryotic translation and mRNA stability. A short upstream open reading frame strongly inhibits translational initiation and greatly accelerates mRNA degradation in the yeast *Saccharomyces cerevisiae*," *The Journal of Biological Chemistry*, vol. 270, no. 15, pp. 8936–8943, 1995.
- [53] K. Wethmar, J. J. Smink, and A. Leutz, "Upstream open reading frames: molecular switches in (patho)physiology," *Bioessays*, vol. 32, no. 10, pp. 885–893, 2010.
- [54] L. Liu, D. Dilworth, L. Z. Gao et al., "Mutation of the CDKN2A 5' UTR creates an aberrant initiation codon and predisposes to melanoma," *Nature Genetics*, vol. 21, no. 1, pp. 128–132, 1999.
- [55] A. Wiestner, R. J. Schlemper, A. P. C. van der Maas, and R. C. Skoda, "An activating splice donor mutation in the thrombopoietin gene causes hereditary thrombocythaemia," *Nature Genetics*, vol. 18, no. 1, pp. 49–52, 1998.
- [56] Y. R. Wen, Y. Liu, Y. M. Xu et al., "Erratum: 'loss-of-function mutations of an inhibitory upstream ORF in the human hairless transcript cause Marie Unna hereditary hypotrichosis (Nature Genetics (2009) 41 (228–233))'," *Nature Genetics*, vol. 41, no. 6, p. 762, 2009.
- [57] G. Beffagna, G. Occhi, A. Nava et al., "Regulatory mutations in transforming growth factor- β 3 gene cause arrhythmogenic right ventricular cardiomyopathy type 1," *Cardiovascular Research*, vol. 65, no. 2, pp. 366–373, 2005.
- [58] S. E. Calvo, D. J. Pagliarini, and V. K. Mootha, "Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 18, pp. 7507–7512, 2009.
- [59] F. Poulat, M. Desclozeaux, S. Tuffery, P. Jay, B. Boizet, and P. Berta, "Mutation in the 5' noncoding region of the SRY gene in an XY sex-reversed patient," *Human Mutation*, vol. 10, no. 1, pp. S192–S194, 1998.
- [60] S. Kondo, B. C. Schutte, R. J. Richardson et al., "Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes," *Nature Genetics*, vol. 32, no. 2, pp. 285–289, 2002.
- [61] L. Groussin, L. S. Kirschner, C. Vincent-Dejean et al., "Molecular analysis of the cyclic AMP-dependent protein kinase A (PKA) regulatory subunit 1A (PRKAR1A) gene in patients with carney complex and primary pigmented nodular adrenocortical disease (PPNAD) reveals novel mutations and clues for pathophysiology: augmented PKA signaling is associated with adrenal tumorigenesis in PPNAD," *American Journal of Human Genetics*, vol. 71, no. 6, pp. 1433–1442, 2002.
- [62] H. Witt, W. Luck, H. C. Hennies et al., "Mutations in the gene encoding the serine protease inhibitor, Kazal type 1 are associated with chronic pancreatitis," *Nature Genetics*, vol. 25, no. 2, pp. 213–216, 2000.
- [63] R. Oner, S. Agarwal, A. J. Dimovski et al., "The G \rightarrow A mutation at position +22 3' to the cap site of the β -globin gene as a possible cause for a β -thalassemia," *Hemoglobin*, vol. 15, no. 1–2, pp. 67–76, 1991.
- [64] J. Konig, K. Zarnack, G. Rot et al., "ICLIP—transcriptome-wide mapping of protein-RNA interactions with individual nucleotide resolution," *Journal of Visualized Experiments*, no. 50, article 2638, 2011.
- [65] D. R. Morris and A. P. Geballe, "Upstream open reading frames as regulators of mRNA translation," *Molecular and Cellular Biology*, vol. 20, no. 23, pp. 8635–8642, 2000.
- [66] V. Matys, O. V. Kel-Margoulis, E. Fricke et al., "TRANSFAC and its module TRANSCOMP: transcriptional gene regulation in eukaryotes," *Nucleic Acids Research*, vol. 34, pp. D108–D110, 2006.
- [67] R. J. Keyser, L. van der Merwe, M. Venter et al., "Identification of a novel functional deletion variant in the 5'-UTR of the DJ-1 gene," *BMC Medical Genetics*, vol. 10, article 105, 2009.

