News & Views

Ending on a high note: Downstream ORFs enhance mRNA translational output

Samantha Dodbele [iD] & Jeremy E Wilusz* [iD]

Eukaryotic mRNAs were long thought to only translate a single protein product, but it is now recognized that many mRNAs can also encode small open reading frames (ORFs). In this issue of *The EMBO Journal*, Wu *et al* characterized small ORFs in the 3′ untranslated regions (3′ UTRs) of human and zebrafish mRNAs and found that many are indeed translated. The peptides encoded by these downstream ORFs (dORFs) are often poorly conserved across evolution, but many dORFs are nonetheless functional, as the act of their translation can promote translation of the canonical ORF.

The dogma regarding pro- versus eukaryotic messenger RNAs (mRNAs) has long been that these are fundamentally different: Prokaryotic mRNAs often encode multiple proteins, whereas eukaryotic mRNAs typically only have a single translated open reading frame (ORF) (reviewed in Hinnebusch, 2014). Prokaryotic ribosomes can be directly recruited to the beginning of each encoded ORF through base pairing of ribosomal RNA with Shine–Dalgarno sequences. In contrast, most translation in eukaryotes occurs by a scanning mechanism, wherein the small ribosomal subunit bound to the initiator methionyl-tRNA (Met-tRNA$_i$) is recruited to the mRNA 5′ cap and then scans downstream for a sequence that is complementary to the Met-tRNA$_i$ anticodon (Fig 1A). Thus, the first AUG encountered is typically favored as the start codon and is used to generate the canonical protein product.

Nevertheless, eukaryotic mRNAs can contain short upstream ORFs (uORFs) that initiate with a canonical AUG or a non-AUG codon (e.g., CUG, GUG, or UUG) and whose translation often prevents scanning ribosomal subunits from efficiently reaching the canonical ORF (Mueller & Hinnebusch, 1986) (Fig 1B). Large-scale analyses have also revealed a number of translation events that appear to occur in mRNA 3′ UTRs (Bazzini *et al*, 2014; Ji *et al*, 2015) (Fig 1C), a location that ribosomes should rarely reach, unless there is readthrough of the stop codon (Dunn *et al*, 2013) or an internal ribosome entry site (IRES) (reviewed in Jackson, 2013). In this issue of *The EMBO Journal*, Wu *et al* comprehensively characterize these downstream ORFs (dORFs) and show that their translation enhances translation of the canonical ORF.

To explore the extent of 3′ UTR translation in human and zebrafish, Wu *et al* computationally identified all potential dORFs and found them to be very common (> 80% of mRNAs contain at least one). But are they, in fact, translated? Ribosome profiling data revealed that ~ 10% of dORFs displayed evidence of translation, including the expected three-nucleotide periodicity of ribosome footprints. Most of these translated dORFs are short (median ~ 20 amino acids) and initiate more than 100 nucleotides downstream from the stop codon of the canonical ORF (Wu *et al*, 2020). Like uORFs, dORFs often initiate at a non-AUG start codon and the peptides they encode are not well conserved across species. However, for a subset of orthologous genes, the presence of a translated dORF is evolutionarily conserved, suggesting that dORFs may function through

translation activity, rather than the polypeptide itself.

As uORFs in 5′ UTRs often repress translation of canonical ORFs (Fig 1B), the authors hypothesized that dORFs may likewise affect translational efficiency of the canonical ORF and examined ribosome profiling datasets from human cell lines and zebrafish embryonic development. This revealed that mRNAs with translated dORFs generally translate the canonical ORF at ~ 2-fold higher efficiency (Wu *et al*, 2020) (Fig 1C). Importantly, these effects were not correlated with differences in mRNA levels/stability, poly(A) tail length, or UTR lengths. To confirm that translation of a dORF can promote canonical ORF translation, reporter plasmids with distinct 3′ UTRs were assessed. Insertion of an endogenous translated dORF (including ones initiated with non-AUG codons) enhanced canonical ORF expression by ~ 2- to 3-fold, while this stimulatory effect was lost upon addition of a premature stop codon immediately downstream of the dORF start codon. Artificially designed dORFs likewise promoted reporter expression, indicating that it is the act of dORF translation itself, and not the dORF nucleotide, codon, or peptide sequence that is most critical for enhancing canonical ORF translation.

The number of uORFs in the mRNA 5′ UTR has been correlated to the strength of translational repression (Chew *et al*, 2016; Johnstone *et al*, 2016), and a similar effect was observed for dORFs. The presence of multiple translated dORFs promotes translation of the canonical ORF at even higher efficiency (Wu *et al*, 2020). It should nevertheless be noted that some differences in the degree of translation enhancement

© 2020 The Authors
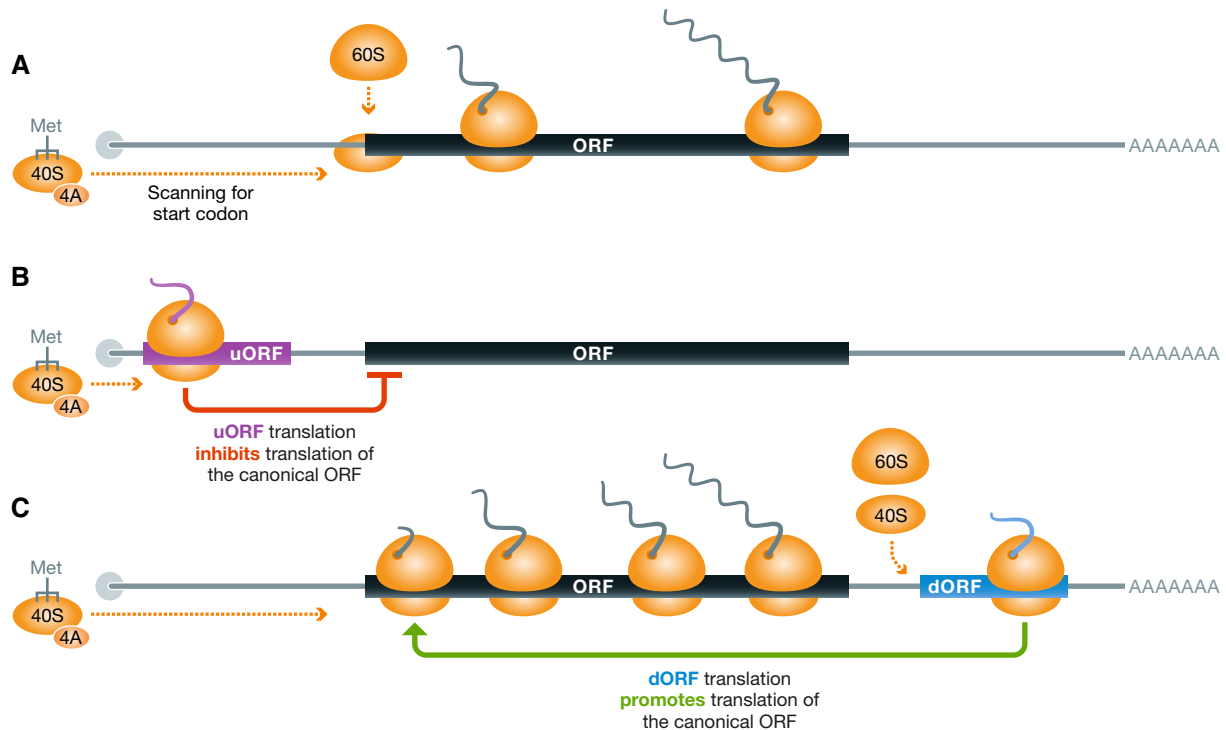
Figure 1. Translation of short ORFs controls the efficiency of canonical ORF translation.

(A) In eukaryotes, the preinitiation complex (including the 40S small ribosomal subunit, the initiator methionyl-tRNA, and several additional factors such as the helicase eIF4A) loads at the mRNA 5′ cap and then scans 5′–3′ in search of a start codon. Once the start codon is recognized, the 60S large ribosomal subunit is recruited and the complete 80S ribosome then translates the ORF. (B) Translation of a uORF can block preinitiation complexes from reaching the canonical ORF start codon, thereby downregulating the translational output of the mRNA. (C) In contrast, ribosomes can be recruited to the dORFs via cap- and scanning-independent mechanisms, and the act of dORF translation promotes the expression of the canonical ORF.

were observed, indicating that not all dORF translation events are equal and that more complicated regulation is sometimes at play.

As most eukaryotic translation occurs by a 5′ cap and scanning-dependent mechanism, how do ribosomes reach dORFs? Ribosome footprints are depleted in the region between the canonical ORF and dORF (the internal UTR—iUTR), suggesting dORFs are not translated by ribosome readthrough or re-initiation just after the canonical ORF stop codon (Wu *et al*, 2020). Instead, it appears that dORFs may be translated when new ribosomes are directly recruited to the iUTR in a manner analogous to how IRES elements function (reviewed in Jackson, 2013). Consistently, iUTRs from human *CYR61* and *CCDC167* show IRES activity when inserted into a bicistronic reporter mRNA. These iUTRs further appear to function independent of the cap as dORF translation was not affected when ribosome scanning was blocked at the mRNA 5′ end (Wu *et al*,

2020). It remains unclear how exactly iUTRs recruit ribosomal subunits, especially given that these sequences are short compared to well-characterized viral IRES elements. One possible hint is that the region near the dORF start codon is often biased for particular nucleotides, but these motifs are not similar to the Kozak consensus sequence near canonical ORF start codons.

In total, translation of dORFs appears to be a widespread post-transcriptional regulatory mechanism in eukaryotes that enhances mRNA translational output. Many of the molecular details await characterization, including how the process is regulated in a cell type or condition-specific manner. Going forward, it will be particularly critical to understand how exactly dORFs promote canonical ORF translation, for example, if dORFs help recruit initiation factors to the mRNA or promote mRNA looping. It further will be interesting to determine if/which dORF peptides accumulate and have

functional roles. Recent work has revealed that some uORF peptides are functional, e.g., by forming a stable complex with the corresponding canonical ORF product (Chen *et al*, 2020). The existence of translated dORFs further challenges the monocistronic view of eukaryotic mRNAs and suggests that there are still additional key aspects of translational control that await discovery.

References

Bazzini AA, Johnstone TG, Christiano R, Mackowiak SD, Obermayer B, Fleming ES, Vejnar CE, Lee MT, Rajewsky N, Walther TC *et al* (2014) Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation. *EMBO J* 33: 981–993

Chen J, Brunner AD, Cogan JZ, Nunez JK, Fields AP, Adamson B, Itzhak DN, Li JY, Mann M, Leonetti MD *et al* (2020) Pervasive functional translation of noncanonical human open reading frames. *Science* 367: 1140−1146

Chew GL, Pauli A, Schier AF (2016) Conservation of uORF repressiveness and sequence features in mouse, human and zebrafish. *Nature Commun* 7: 11663

Dunn JG, Foo CK, Belletier NG, Gavis ER, Weissman JS (2013) Ribosome profiling reveals pervasive and regulated stop codon readthrough in *Drosophila melanogaster. eLife* 2: e01179

Hinnebusch AG (2014) The scanning mechanism of eukaryotic translation initiation. *Ann Rev Biochem* 83: 779−812

Jackson RJ (2013) The current status of vertebrate cellular mRNA IRESs. *Cold Spring Harb Perspect Biol* 5: a011569

Ji Z, Song R, Regev A, Struhl K (2015) Many lncRNAs, 5′UTRs, and pseudogenes are translated and some are likely to express functional proteins. *eLife* 4: e08890

Johnstone TG, Bazzini AA, Giraldez AJ (2016) Upstream ORFs are prevalent translational repressors in vertebrates. *EMBO J* 35: 706−723

Mueller PP, Hinnebusch AG (1986) Multiple upstream AUG codons mediate translational control of GCN4. *Cell* 45: 201−207

Wu Q, Wright M, Gogol MM, Bradford WD, Zhang N, Bazzini A (2020) Translation of small downstream ORFs enhances translation of canonical main open reading frames. *EMBO J* 39: e104763