Stem-loop induced ribosome queuing in the uORF2/*ATF4* overlap fine-tunes stress-induced human *ATF4* translational control

Anna M. Smirnova[1], Vladislava Hronová[1], Stanislava Gunišová[1], Pasha Mahabub Mohammad[1], Denisa Petráčková[2], Petr Halada[3], Justin Rendleman[4], Štěpán Coufal[5], Anna Herrmannová[1], Petra Beznosková[1], Kristína Jendruchová[1], Maria Hatzoglou[6], Christine Vogel[4], and Leoš Shivaya Valášek[1*]

[1]Laboratory of Regulation of Gene Expression, Institute of Microbiology of the Czech Academy of Sciences, Videnska 1083, 142 20, Prague, the Czech Republic
[2]Laboratory of Post-Transcriptional Control of Gene Expression, Institute of Microbiology of the Czech Academy of Sciences, Videnska 1083, 142 20, Prague, the Czech Republic
[3]Laboratory of Structural Biology and Cell Signaling, Institute of Microbiology of the Czech Academy of Sciences, Prumyslova 595, 252 50, Vestec, Czech Republic
[4]Department of Biology, New York University; New York, NY, USA.
[5]Laboratory of Cellular and Molecular Immunology, Institute of Microbiology of the Czech Academy of Sciences, Videnska 1083, 142 20, Prague, the Czech Republic
[6]Department of Genetics and Genome Sciences, Case Western Reserve University; Cleveland, OH, USA.

*Corresponding author: *valasekl@biomed.cas.cz*
Tel: +420 724 731 661

Keywords: integrated stress response, unfolded protein response, translational control, ribosome, ATF4, ribosome queuing, translation reinitiation

SUMMARY

ATF4 is a master transcriptional regulator of the integrated stress response leading cells towards adaptation or death. Since its discovery, the delayed translation reinitiation mechanism, with reinitiation-permissive uORF1 playing a key role, was thought to underlie the ATF4 stress-induced upregulation. However, mounting evidence challenging this mechanism as the sole source of *ATF4* translation control prompted us to investigate additional means of its regulation. Contrasting the original model, we show that inhibitory uORF2 is inducible and translated even under stress; that it encompasses two post-transcriptionally modified adenines performing antagonistic roles; and that the canonical ATF4 translation start site is robustly leaky-scanned. Most importantly, we demonstrate that the presence of a stable stem-loop in the uORF2/*ATF4* overlap, immediately preceded by near-cognate CUG, introduces another layer of regulation in the form of ribosome queuing. Thus, *ATF4* translational control is much more complex than originally described underpinning its key role in diverse biological processes.

INTRODUCTION

Eukaryotic cells have evolved several mechanisms to cope with various environmental stressors. These include a complex signaling pathway referred to as the Integrated Stress Response (ISR) (Pakos-Zebrucka et al., 2016; Wek et al., 2006). Whereas its external triggers are oxygen, nutrient deprivation or viral infections, the main internal stressor is endoplasmic reticulum (ER) stress caused by the accumulation of unfolded proteins in the ER lumen. Importantly, the ISR can also be induced by activation of oncogenes (Denoyelle et al., 2006). The response is “integrated” as all stress signals are transduced by a family of four serine/threonine kinases and converge into a single event which is the phosphorylation of the α subunit of eukaryotic translation initiation factor 2 (eIF2α) (Donnelly et al., 2013).

eIF2 assembles into a ternary complex (TC) together with a GTP molecule and the initiator Met-$tRNA_i^{Met}$. This complex, together with other eIFs bound to the small ribosomal subunit, facilitates the Met-$tRNA_i^{Met}$ recruitment to the ribosomal P-site so that the pre-initiation complex (PIC) can recognize the translation start site and initiate translation (Valášek, 2012). After the AUG selection process, eIF2 with hydrolyzed GDP is ejected from the initiation complex. To participate in the next translational cycle, it must be regenerated by the eIF2B-mediated exchange of GDP for GTP (Hinnebusch, 2014). The ISR-transduced phosphorylation of the α subunit of eIF2 prevents this regeneration step, leading to a significant decrease in TC levels and a general translational shutdown (Gunisova et al., 2018).

Importantly, a handful of specific mRNAs escape this shutdown and become efficiently translated to begin adaptation, enabling cells to cope with the acute stress (Harding et al., 2000). One of these mRNAs encodes a transcription factor *ATF4,* a master regulator of the ISR (Lu et al., 2004; Vattem and Wek, 2004). Upon relief of the acute stress, eIF2α is dephosphorylated, the ISR ceases, and general protein synthesis resumes (Novoa et al., 2003). However, if the stress is too intense or persists for too long - i.e. becomes chronic (Guan et al., 2017) - the adaptive response capacity is exhausted and programmed cell death is triggered. Thus, the final outcome of the ISR, i.e. resumption of protein synthesis or programmed cell death, depends largely on the level and duration of eIF2α phosphorylation and, consequently, on the extent of upregulation of factors whose translation is paradoxically induced, such as *ATF4* (Pakos-Zebrucka et al., 2016).

Therefore, the timing and level of ATF4 expression are critical parameters determining cell fate. Even minor disturbances in the ATF4 function contribute to serious pathologies such as Parkinson's, Alzheimer's and Huntington's neurodegenerative diseases, prion diseases and various types of retinal degeneration (Pitale et al., 2017). Furthermore, persistent over-activation of ATF4, which does not induce apoptosis, is linked to many cancers because of the continuous expression of adaptive genes that sustain the stress response (Wortel et al., 2017). For example, aberrant activation of ATF4 has been associated with mutations in oncogenes from in the EGFR-PI3K pathway (Takahashi et al., 2023). Since tumor cells are under constant stress due to unregulated growth, oxygen and nutrient deprivation, ER accumulation of misfolded proteins, constitutively activated stress adaptation program maximizes

their survival. While ATF4's role in this process suggests its potential as a target of therapeutic interventions, such as through analogues designed for the ER stressor thapsigargin (Tg) (Doan et al., 2015; Lindner et al., 2020), we still do not understand molecular details of ATF4 involvement in tumorigenesis (Hetz et al., 2013).

Under normal conditions, human *ATF4* is constitutively transcribed at low levels, but the protein is undetectable. Alternative splicing produces two isoforms of *ATF4* mRNA, V1 and V2 (Chan et al., 2013). The intron-excised V2 variant, shorter in length (for simplicity hereafter referred to as the *ATF4* mRNA), represents the major form prevalent in almost all human cell types. While stress also induces *ATF4* transcription (Dey et al., 2010), increases the transcript as well as ATF4 protein stability (Gardner, 2008; Koditz et al., 2007), translation is the major contributor to ATF4 expression regulation.

This induction is mediated by the 5' leader of *ATF4* mRNA, which contains two upstream open reading frames (uORFs): a short 4-codon uORF1 followed by a longer 60-codon uORF2, which overlaps with the *ATF4* reading frame in a -1 frame. In analogy with its extensively studied yeast functional homologue *GCN4* (Gunisova et al., 2018; Gunisova and Valasek, 2014; Hinnebusch, 2005), which has a similar arrangement and nature of uORFs in its mRNA leader, it has been proposed that the primary route of *ATF4* translational control occurs via the so-called delayed translational reinitiation (REI) (Lu et al., 2004; Vattem and Wek, 2004). This model is underpinned by the stress-induced decrease in TC levels and the opposing properties of the two *ATF4* uORFs (Fig. 1A) (Gunisova et al., 2018).

Specifically, after translation of the REI-permissive uORF1, only the large ribosomal subunit is recycled, while the small subunit remains associated with the mRNA and, upon regaining scanning ability, can reinitiate downstream. This process is enabled due to the short length of uORF1, the presence of stimulatory flanking sequences, and interaction with the initiation factor eIF3 (Hronova et al., 2017). A key aspect of this delayed REI mechanism is that reacquisition of the scanning ability of the post-termination 40S is limited by its binding of a new TC, and the time required for TC acquisition increases significantly with decreasing TC levels (Lu et al., 2004; Vattem and Wek, 2004). The original model posits that, in the absence of stress, the small subunit acquires the TC relatively quickly and recognizes the AUG of uORF2 as the next initiation codon. Since uORF2 overlaps with the coding region of ATF4 out of frame, no ATF4 protein is produced. Conversely, the stress-induced decrease in TC levels delays the acquisition of a new TC by the large number of post-termination 40S subunits traversing downstream of uORF1, allowing them to skip the AUG of the inhibitory uORF2. After recruitment of the TC *en route* to the *ATF4* start codon, the ribosome reinitiates at ATF4's coding region, thereby inducing its synthesis.

Despite the depth of the original studies describing *ATF4* translational control *via* the delayed REI mechanism (Lu et al., 2004; Vattem and Wek, 2004), a growing body of evidence suggests the existence of additional regulatory modes. For example, in direct contradiction to the original model, three independent studies have reported increased translation of uORF2 under stress (Andreev et al., 2015; Sidrauski et al., 2015; Starck et al., 2016). In addition, leaky scanning over uORF2 was shown to

contribute to *ATF4* induction to some extent (Starck et al., 2016). Further, it has been proposed that a N^6-methyladenosine (m^6A) post-transcriptional modification in the non-overlapping region of uORF2 regulates translation of mouse *ATF4* mRNA (Zhou et al., 2018). Under physiological conditions, m^6A reportedly functions as a barrier to ribosomes scanning downstream of uORF1; stress triggers demethylation of m^6A and removes this block. QTI-seq performed in the same study also implicitly questioned the true initiation start site of the *ATF4* coding sequence, which begins with three nearly consecutive AUGs.

Our own work revealed that chronic stress rewires the entire control mechanism and, under the reign of the PERK eIF2-specific ER stress kinase, provokes a switch from an eIF4F/eIF2B-dependent mode of translation initiation to an eIF3-dependent mechanism (Guan et al., 2017). In addition, we have suggested that eIF3 may promote direct recruitment of selected mRNAs to 40S ribosomes during chronic ER stress. This process results in partial recovery of general translation, which, together with transcriptional reprogramming, selectively bolsters expression of ER-localized proteins. Using the Sel-TCP-Seq technique (Wagner et al., 2022), we also showed recently that uORF0 (hereafter referred to as Start-stop [St-st] because it consists only of an initiation and termination codon), which precedes uORF1 (Fig. 1A), represents an additional physical barrier that may inhibit ATF4 expression in the absence of stress (Wagner et al., 2020). Interestingly, although its existence has been known for a long time (Harding et al., 2000) and REI downstream of St-st shown to depend on non-canonical initiation factors DENR and MCTS1 (Bohlen et al., 2020b; Schleich et al., 2017), St-st's true contribution to *ATF4* translational control has not been investigated from a mechanistic point of view. Finally, other regulatory pathways unrelated to ISR have been shown to contribute significantly to ATF4 induction by unknown mechanisms (Gunisova et al., 2018). Overall, these findings paint a picture of a multilevel regulatory model that exceeds the relatively straightforward mode of delayed REI. Full understanding of this model would reveal the entire spectrum of *ATF4* regulation and further strengthen its position as a master regulator of ISR and other processes critical to human health and disease.

Here, we subjected the human *ATF4* mRNA leader to comprehensive bioinformatics analysis and uncovered several *cis*-acting features in addition to known elements. The detailed mutational analysis using our new and well-controlled reporter assay revealed that: 1) the inhibitory uORF2 is inducible and translated even under stress; 2) the first two AUGs of ATF4 are mostly leaky-scanned; 3) two adenine modification sites contribute to overall regulation in opposing ways; and, most importantly, 4) a stable stem loop in the uORF2/*ATF4* overlap causes ribosome queuing, and stalled ribosomes initiate at a near-cognate CUG codon upstream of this stem loop. Overall, our data suggest that, in addition to delayed reinitiation, translational control of ATF4 comprises a multilayered regulatory circuit of diverse sequence and structural elements that fine-tune its expression under different stimuli.

RESULTS

Revisiting translational control of human *ATF4*

Initial analysis of ATF4's translational control with respect to the role of uORFs 1 and 2 was carried out using a mouse *ATF4* mRNA reporter (mouse *ATF4* exceptionally lacks St-st, which is otherwise highly conserved among vertebrates (Fig. S1A) (Rendleman et al., 2021), where *Luciferase* was fused with *ATF4* only two codons downstream of its AUG (Vattem and Wek, 2004). The wild type (wt) and specific mutation(s)-carrying reporters were individually co-transfected with a control *Renilla* plasmid into MEF cells and Luciferase activity was measured and normalized to Renilla activity. We perceived two potential bottlenecks in this arrangement. 1) In addition to lacking St-st, the original ATF4 reporter also lacked much of the uORF2/*ATF4* overlap (as well as the entire ATF4 ORF), which limited investigation of its role. 2) Its activity was not normalized to *ATF4* mRNA levels which are known to change (Dey et al., 2010)) but to Renilla activity, which by definition drops rapidly upon stress-induced translational shutdown. Even a recent mathematical model of the *ATF4* translational control based on the original work (Vattem and Wek, 2004), which separated triplets into fast and slow TC-loading depending on their cognation to AUG, suggested several discrepancies in behavior between the natural system and the computed model, particularly regarding the role of uORF2 (Marasco et al., 2022). The authors suggested that, since the experimental results were difficult to understand, it would be useful to replicate Watt and Wek's experiments with ΔuORF2 mutants to reconcile the different findings. Therefore, to resolve these inconsistencies, we designed an entirely new reporter based on human *ATF4*.

Our newly designed CMV promoter-driven reporter preserves the entire V2 leader transcript and the entire coding region of *ATF4*, which is extended with an HA tag at its C-terminus (Fig. 1B). Twenty four hours after seeding, HEK293T cells were individually transfected with a wt reporter as a control and in every single experiment as well as its mutant derivatives; at 8 or 24 hours post-transfection, one half of the cultures was treated with thapsigargin (Tg; induces UPR) (Fig. 1C). Exactly 3 hours after Tg treatment, cells were harvested using Glo lysis buffer (Promega) and the whole cell extracts (WCE) were subjected to a fully quantitative, automated, capillary-based high throughput immunoassay ("western blotting") using JESS Protein Simple instrument as described in detail in the Methods (Fig. 1C). We monitored mRNA levels of all constructs in each individual experiment by RT-qPCR and only considered samples that showed an RNA difference of less than 1 cycle compared to the wt construct (using forward primer in the *ATF4* CDS and the reverse matching the HA tag) for further analysis. In addition, we always conducted an automated protein normalization control of the equal protein loading in all capillaries using the fluorescent channel (Protein Simple PN reagent kit) and excluded samples displaying more than 20% difference in total protein loading (Fig. 1C).

Figure 1D shows an example of the output data in the form of an electropherogram, which displays the intensity detected along the length of the capillaries; the generated peaks can be quantified by calculating the area under the curve. In case of ATF4-HA, the signal detected by the anti-HA tag antibodies is

displayed as a single peak at 53 kDa; it shows ~5.2-fold and ~1.5-fold induction (p<0.0001) upon Tg treatment either 8- or 24-hours post-transfection, respectively (Fig. 1E, see Table S1 for average values of fold-induction for wt construct measured after 8 h and 24 h and Table S2 for individual values from every experiment). For statistical analysis, see Methods.

Please note that the 24-hour transfection regime induced a stress response in all cells even without the Tg treatment, as shown by increased basal levels of ATF4. Therefore, the fold-increase upon stress induction is for this longer transfection regime relatively small, but still highly reproducible and statistically significant (p<0.05). As the 24-hour regime is much more practical with respect to the laboratory time management, we conducted our initial screening of 82 mutant reporters using this longer scheme. Then we retested all mutants with significant effects using the 8-hour transfection regime, which does not induce a stress response without the Tg treatment. For simplicity, the following text focuses on results from the 8-hour regime; data from the 24-hour regime are presented for consistency and discussed only as needed.

uORF2 has an inducible nature under stress

We began our analysis by testing the effect of the St-st and re-testing the effects of solitary uORFs on ATF4 expression by mutating the AUGs of the respective other two elements in the otherwise wt construct (Fig. 2A and B). First, we measured the control construct (d-all), where AUGs of all three elements were mutated. This construct determined the maximal expression level of ATF4, which was, in the case of the 8-hour regime, ~21.8-fold and ~3.3-fold higher in non-stress and Tg stress conditions, respectively, compared to the values of the wt construct that were set to 1 (Fig. 2A and B). The last plot in the d-all column with the d-all non-stress value set to 1 indicates that this construct is not stress-inducible compared to wt shown in Figure 2C, which was expected (see Table S1 and Table S2). Note that all Tg "fold-induction" experiments were performed separately from those carried out without and under Tg stress, which means that they were also separately controlled for equal mRNA and proteins levels (see Table S3 for average values of every construct, and Table S4 for individual values from every experiment shown in Figure 2).

Keeping St-st as the only element in the *ATF4* leader displayed ~20.6-fold increase in ATF4 expression under non-stress conditions and ~3.1-fold increase under Tg stress (Fig. 2A and B). Both values are modestly but reproducibly lower than those of d-all providing further evidence that St-st acts as a stress-independent, non-inducible repressive element, i.e. a roadblock, as observed before (Rendleman et al., 2021; Wagner et al., 2020).

The uORF1-only construct uORF1 led to 12.2-fold and 2.7-fold upregulation in non-stress and Tg stress conditions, respectively, while also showing no stress inducibility compared to wt (Fig. 2A - C). These results clearly confirm the stress-independent, REI-permissive nature of this uORF and indicate that more than 50% of ribosomes that translate uORF1 (~12-fold induction of uORF1-only over wt divided by ~22-fold induction of d-all) undergo partial recycling and reinitiate downstream upon reacquisition the TC.

According to the original delayed reinitiation model, uORF2 serves as a negative barrier that should not be inducible under stress conditions, because its sequence extends into the CDS of *ATF4.* In other words, once the ribosome initiates at uORF2, the effect on ATF4 is independent of TC levels. Accordingly, the uORF2-only construct showed an ~80% drop in reporter expression compared to the wt under both normal and TG stress conditions (Fig. 2A and B). However, comparing stress with non-stress values of uORF2-only to wt revealed an ~3.2-fold induction of the uORF2-only construct that could not be explained by the original model (see further below). The 24-hour regime confirmed this result albeit with lower values.

Next, we eliminated one upstream element at a time and compared the resulting constructs with the wt set to 1 (Fig. 2B and D). Eliminating St-st (in dSt-st) showed little to no effect under both conditions; the repeatedly observed modest decrease, which would be expected from a road-block, was not statistically significant in this experimental set-up. Nonetheless, we repeatedly observed slightly higher stress induction of dSt-st compared to wt (~5.8-fold *versus* 5.2-fold; Fig. 2D *vs.* 2C) that is consistent with its stress-independent roadblock function reducing processivity of the scanning ribosome. These results, and the fact that mouse *ATF4* lacks the Start-stop, indicate that uORFs 1 and 2 are largely sufficient to ensure stress-induced translation control of *ATF4* expression.

While removal of the REI-permissive uORF1 (in d1) significantly reduced the basal level of ATF4 expression (by ~70%), eliminating the inhibitory uORF2 (in d2) dramatically derepressed ATF4 expression (~12.1-fold) under non-stress conditions and slightly (~2.4-fold) under stress, all as expected from the original model (Fig. 2B and D). The latter result nicely corroborates that ~50% of ribosomes terminating at uORF1 reinitiate downstream, as we observed with the uORF1-only construct (Fig. 2B).

Comparison of the stress induction potentials of each construct showed that d2 is very similar to uORF1-only and d1 to uORF2-only (Fig. 2D and A, “fold-induction” plots in the respective columns), except that the presence of the St-st further magnified the uORF2 inducibility (from ~3.2-fold to ~5.8-fold; “fold-induction” plots in Fig. 2A – uORF2-only *versus* Fig. 2D – d1). This result is again in line with St-st serving as a ribosome barrier to maintain low basal levels of ATF4 expression. Taken together, our results confirmed major aspects of the original model but raised the critical question as to the mechanism behind the unexplained stress induction of uORF2. We simply asked whether there are any so-far unidentified regulatory elements in the *ATF4* mRNA leader besides St-st and uORFs?

Sequence analysis of the 5’ UTR of the human *ATF4* mRNA reveals novel elements potentially contributing to *ATF4* translational control

Therefore, next we carefully screened the *ATF4*’s 5’ UTR including the uORF2/*ATF4* overlapping coding region ('overlap’) for sequence and structure features that might suggest additional modes of regulation (Fig. 3A). We had previously predicted and experimentally validated the existence of two stem loops (SL1 and 2) preceding uORF1 (Fig. 3A), the first of which cooperates with eIF3 to unleash the full REI potential of

uORF1 (Hronova et al., 2017). SL1 also contains one CUG near-cognate start site (NC1) that is immediately followed by two overlapping GUG near-cognates (NC2/3) that are no longer part of SL1.

Our new analysis predicted another stem loop (SL3), which was relatively stable and therefore potentially inhibitory, as well as highly conserved among vertebrates (Fig. S1A). It is located roughly in the middle of the uORF2/*ATF4* overlap (Fig. 3A) and has a hairpin structure with the highest free energy predicted for this particular sequence region.

Further, our analysis predicted a total of four individual sites of m^6A methylation within well-defined motifs (Fig. 3A); DRACH1 (with A225) is located in the non-overlapping, likely unstructured region of uORF2, followed by RRACH1 (A235) in the same region, which is homologous to the predicted methylated A in mouse mRNA (Zhou et al., 2018). The DRACH2 (A286) motif occurs in the uORF2/*ATF4* overlap and extends into the *ATF4* main initiation codon with a modifiable adenine located just downstream of the AUG. Thus, its modification could interfere with AUG recognition during start site selection. Finally, RRACH2 (A326) locates to the stem of SL3; therefore, it could potentially affect its stability.

The *ATF4* gene begins with two nearly consecutive AUGs followed by a third in close proximity, all in frame (Fig. 3A, in dark green). Both the canonical AUG1 and AUG2 have a medium Kozak initiation context; AUG1 overlaps with DRACH2, and AUG2 is located two codons downstream. AUG3 has a weak Kozak initiation context and is exposed in the open loop of SL3. It represents the 17th codon downstream of AUG1. Both AUG2 and AUG3 are highly conserved among vertebrates (Fig. S1A).

The uORF2/*ATF4* overlap also contains two putative sliding sequences (FS-A and -B) and a so-called C-tract motif (Penn et al., 2020), which could potentially prompt the elongating ribosome to switch the reading frame from uORF2 to *ATF4*. We investigated this hypothesis by examining the frame of ribosome footprints from a dataset collected from HeLa cells subjected to stress (Rendleman et al., 2018). Specifically, we mapped ribosome protected fragments in the region of the uORF2/*ATF4* overlap and found that, during acute ER stress, reads continued to locate to the uORF2 reading frame (frame 0) and perhaps even increased, while there was no coverage of the *ATF4*'s frame (frame 1) (Fig. S1B). In fact, ATF4's translation appeared to initiate downstream of the predicted SL3 after 4 hours of Tunicamycin stress.

Sliding region FS-A involves 3 consecutive glycine codons, and because of its overlap with SL3, it can only function if SL3 is not formed. Switching from uORF2 to the *ATF4* reading frame could be mediated by either a +1 or -2 programmed ribosomal frameshifting (PRF). Of these two possibilities, the latter is much more likely because both types of glycine tRNAs, with GGG in the P-site and GGA in the A-site, are well positioned for re-pairing after a -2 shift (Penn et al., 2020). In support of this hypothesis, 8 nucleotides downstream from FS-A occurs the C-nucleotide-rich C-tract motif (CCCCCUUCGACC) with high similarity to the inhibitory C-motif promoting -2 PRF in arteviruses (Pakos-Zebrucka et al., 2016). Since 8 nucleotides is considered the optimal distance for the C-motif to act as the inhibitory element that can cooperate with

the upstream sliding sequence to promote -2 PRF, we designed constructs to examine this option (see below).

FS-B is located just upstream of the uORF2 stop codon, thereby partially fulfilling the +1 PRF criteria (Penn et al., 2020). Yet, neither of its glycine codons has the potential to slow or stop ribosomes, as they belong to the high codon usage category. Since FS-B also does not meet the -2 PRF criteria, it likely has no role in *ATF4* translational control.

As aforementioned, we created an array of mutants for these predicted elements and first evaluated them in a large-scale screen with the 24-hour transfection regime. This screen resulted in no observable effect for mutations in NC1 - 3 in SL1, mutations unfolding SL1 or SL2 – specifically with respect to the St-st function, and mutations of AUG3 in *ATF4* (Fig. 3A, Table S5 and Table S6). The results below focus on mutations with observable effects in the 24-hour regime, and their repeated analysis in the 8-hour transfection regime.

SL3 delays the flow of ribosomes in the uORF2/*ATF4* overlap and promotes initiation at the upstream CUG near-cognate initiation site

Next, we unfolded SL3 (in $SL3^{Mut}$) by a quadruple C to A substitution (Fig. 3A) in wt and dSt-st constructs and observed increased ATF4 expression by ~1.3- and ~1.5-fold under non-stress and Tg stress conditions, respectively (Fig. 4; Table S7 and Table S8). Accordingly, loosening SL3 also increased inducibility from ~5.2-fold to ~5.8-fold in both constructs ("fold-induction" plots in Fig. 4A – wt *versus* Fig. 4B – $SL3^{Mut}$ and dSt-st-$SL3^{Mut}$). These results suggest that SL3 forms and acts as physical barrier delaying scanning/translating ribosomes. Interestingly, eliminating uORF2 in d2-$SL3^{Mut}$ further enhanced ATF4 expression under non-stress from ~1.3- to ~1.7-fold (Fig. 4B). Eliminating SL3 in d-all and uORF2-only constructs had no such an effect (Table S5 and Table S6). Combined, these results indicate that the presence of uORF2 renders SL3 less inhibitory, in particular under non-stress conditions. In a sense, translation of uORF2 under non-stress might resemble a one-way Newton's cradle, in which a completed queue of 80S ribosomes elongating from the AUG of uORF2 is required to breach SL3 and finish uORF2 translation (see further below).

Initiating 48S PIC with the start site placed in the P-site covers ~12 nt from the mRNA entry site up to the P-site (excluding it) and ~15-25 nt from the P-site (including it) to the mRNA exit site (Fig. 3B) (Wagner et al., 2022; Wagner et al., 2020). We noticed a near cognate CUG codon in weak Kozak context positioned 20 nucleotides upstream of SL3 (Fig. 3A), which is highly conserved among vertebrates (Fig. S1A). This position is ideal to force a scanning 48S PIC stalled at SL3 to initiate at CUG (Fig. 3B). To test this hypothesis, we replaced CUG with CUA and observed that the resulting mutation significantly (~20%) reduced ATF4 expression under both conditions tested, but showing no particular effect on stress inducibility compared to wt (Fig. 5A and B, middle panel; Table S9 and Table S10).

These findings may suggest that a certain proportion of the ATF4 protein might be several amino acids shorter due to this CUG-initiated synthesis. Interestingly, combing CUG and SL3 mutations in a single construct produced a double mutation

that nullified the opposing effects of the two individual mutations, suggesting a functional interaction between the two elements (Fig. 5B, right panel; Table S9 and Table S10). Specifically, when the translation inhibitory SL3 was removed, the translation enhancing CUG, which might ensure up to 20% of ATF4 production when the SL3 was present, was no longer needed because a flux of ribosomes initiating at the *ATF4*'s AUG1 was released and produced the full-length ATF4 protein.

These findings suggest that indeed translation was initiated at the CUG and produced a shorter ATF4 variant. To investigate if the CUG-initiated ATF4 protein was produced in wt cells, we subjected HEK293 cells to Tg stress for 3 hours, extracted the proteome and enriched for ATF4 through immunoprecipitation using anti-ATF4 antibodies. We excised bands from an SDS-PAGE gel corresponding to the molecular weight of the long and short variants and digested them individually using Asp-N, chymotrypsin and CNBr. The resulting peptides were then subjected to LC-MS/MS analyses. Despite extensive efforts, the total amounts of the purified ATF4 protein were low and we only identified N-terminal fragments originating from the full-length ATF4 protein (data not shown). Since we cannot rule out that the CUG-initiated ATF4 variant was not detected because its level was below the detection limit of our approach, we cannot provide an unequivocal answer regarding the existence of a CUG-initiated ATF4 protein.

To further investigate the functional interplay between SL3 and CUG, we analyzed three publicly available ribo-seq libraries generated from cells exposed to various ER stressors. Namely, two sets of HEK293T cells treated with arsenite (Andreev et al., 2015; Ichihara et al., 2021) and HeLa cells subjected to the tunicamycin-induced stress (Rendleman et al., 2018). We compared the total number of footprints (FP) in the *ATF4* mRNA leader in stressed *versus* non-stressed cells (Fig. S2), as well as the distribution of FPs with ribosomes positioned with their P-site to a specific codon (Fig. S3). In addition to the pile-ups of FPs over St-st and AUGs of uORF1 and uORF2 (the latter two are observed in only two datasets - Fig. S2A - D), all three datasets clearly revealed a pile-up of FPs ~20 nucleotides upstream of SL3, with a slight reduction under stress. In contrast, there were virtually no FPs on the AUG1 of *ATF4*, even under stress, while the remainder of the coding region was correctly stress-induced in all datasets. A closer look at the FPs indicated a strong preference of the P-site locating to the CUG in all three datasets (Fig. S3) – regardless of the possible origin and nature of the underlying ribosomes. These analyses thus lend further support to the idea of the existence and functionality of SL3 and its associated near-cognate CUG.

Ribosome queuing contributes to the overall translational control of *ATF4*

The existence of a stable SL3 and its interplay with the upstream CUG and uORF2 intrigued us because they could trigger a) uORF2 to *ATF4* frameshifting as suggested above or b) ribosome queuing, i.e., the phenomenon implicated for example in translational control of the antizyme inhibitor mRNA (Ivanov et al., 2018). In particular, under non-stress conditions, SL3 could form a queue of 80S ribosomes elongating from the AUG of uORF2, lowering its translation rate. Such lower uORF2 translation,

while still suppressing ATF4 translation under normal conditions, is energetically desirable, as the uORF2 peptide is relatively long. Indeed, based on an 80S ribosome FP length of 30 nucleotides (Archer et al., 2016; Ingolia, 2010), exactly five 80S ribosomes could be accommodated in this queue, with the most 5' ribosome positioned with its P-site on the AUG of uORF2. The ribosome queue would not be resolved until the barrier posed by SL3 was breached (Fig. 3B).

Under stress, the first 80S ribosome initiating at the AUG of *ATF4* would soon be halted by SL3, preventing the incoming 48S PIC from moving forward and initiating at *ATF4*'s AUG due to a spatial constraint (Fig. 3B). This constraint would then be extended into a queue of 48S PICs that would span up to the AUG of uORF2. Consequently, the most 5' 48S PIC could potentially initiate at uORF2, triggering its translation under stress. Although the length of such a mixed queue of 48S PICs and 80S ribosomes is hard to predict due to variable FP length of scanning 48S PICs, this mixed queue formed under stress might be necessary to overcome the negative effect of SL3, similar to the model of a one-way Newton's cradle described above for the 80S queue under non-stress conditions. This model would also explain that, contrary to what the "delayed REI" implies, uORF2 is translated even under stress, which has been observed by us and others (Starck et al., 2016).

Next, we hypothesized that the presence of CUG would increase the rate of ATF4 production under stress as it could capture ribosomes that skipped (leaky scanned) the first two ATF4 AUGs (below we demonstrate that they are indeed rather "leaky"). Such an intricate system would ensure well-balanced induction of ATF4 which is highly desirable given its key role (Pitale et al., 2017).

To test this hypothesis, we extended the coding sequence of *ATF4* by inserting 6x c-Myc tag (30 nts each) just after AUG1 (Fig. 6, Table S11 and Table S12); i.e. two codons upstream of AUG2 (Fig. 6A). If our logic was correct, this extension would increase ATF4 expression in the otherwise wt construct under both stress and non-stress conditions, while having no effect in the absence of SL3 - which is exactly what we observed (Fig. 6B). Further supporting this model, unfolding SL3 in the absence of uORF2 (in d2_ins-SL3^{Mut}) did not eliminate the stimulatory effect of this extension (Fig. 6C and D). These results confirm the interplay between uORF2 and SL3 by showing that the SL3 effect relies largely on initiation at the AUG of uORF2 (see also below).

Due to an insufficient FP coverage in the uORF2/*ATF4* overlap upstream of SL3, the aforementioned ribo-seq datasets could not be used to further support the ribosome queuing model.

However, while the c-myc detection revealed only a single peak of 65 kDa, as expected given the 6x c-Myc tag insertion, the HA probing detected an additional peak just above, running at the original ATF4 protein size of 53 kDa (Fig. 6E). This peak could be explained by a shorter ATF4 variant(s) produced by initiation at AUG2 and/or the CUG and/or AUG3, in further support of substantial leaky scanning at AUG1, as inferred above.

uORF2 is translated even under stress

To examine the idea of uORF2 to *ATF4* frameshifting, we first mutated the first 3 AUGs of *ATF4* (Fig. S4A, in 3AUG^{Mut}), which would eliminate all ATF4 variants except for the potential frameshifted fusion protein. However, the peak corresponding to the 53 kDa wt protein did not vanish but shifted to 52 kDa (Fig. S4A) and its relative expression level decreased (by ~5.5-fold) (Table S13 and S14); note that this 52 kDa peak in 3AUG^{Mut} remained inducible under Tg stress (Fig. S4A, bottom electropherogram, Table S15 and Table S16). Further, we observed two additional, clearly discernible peaks at 44 and 34 kDa, which were also inducible upon stress. Careful inspection of the ATF4 sequence suggested that the 52 kDa peak may correspond to the almost full-length protein initiated either on the CUG or on five other nearby near-cognate codons (Fig. S5A), whereas the 44 and 34 peaks correspond to shorter ATF4 variants initiated on internal AUGs.

These results suggest that ATF4 might be translated from several alternative start sites, and the variants' inducibility is explained by the presence of uORF2. However, since none of these peaks occurred at a higher molecular weight than the original *ATF4* protein, the idea of frameshifting resulting in a longer ATF4 variant was not plausible. In agreement, shortening the *ATF4* sequence by removing 103 amino acid residues from the C-terminal segment of its CDS shifted all three peaks upward, exactly according to the corresponding loss in molecular weight (Fig. S5B – 3AUG^{Mut} *versus* Fig. S5C – 3AUG^{Mut}-short). Furthermore, inserting 2 consecutive stops immediately downstream of the uORF2/*ATF4* overlap (in the *ATF4* reading frame) eliminated only the heaviest peak but not the other two peaks, in both the full-length and shortened 3AUG^{Mut} constructs (Fig. S5D – 3AUG^{Mut}-stop, Fig. S5E – 3AUG^{Mut}-short-stop), strongly suggesting that the initiation codon for the 52-kDa protein lied within the overlap region, but in the uORF2 frame. Consistently, combining 3AUG^{Mut} with an AUG to AGG mutation of uORF2 had no effect on the size and distribution of the aforementioned 3 protein products (Fig. S6, in d2-3AUG^{Mut}). Indeed, inserting 2 consecutive stops into the wt construct at the same place as in case of the 3AUG^{Mut} construct completely eliminated the signal (data not shown).

The final piece of evidence for the absence of frameshifting arose from the an insertion of a 10x c-Myc tag into the uORF2 frame exactly 21 nt downstream of its AUG (Fig. S4B, in uORF2_ins). This mutant resulted in only one ATF4-HA peak corresponding to ~53 kDa, with inducibility unchanged compared to the wt (Fig. S4B, upper electropherogram). Note that the expected size of the fusion protein is ~62 kDa but it runs at ~57 kDa, as verified experimentally (Fig. S7, Table S17 and Table S18). Probing with c-Myc antibodies revealed a single peak at ~40 kDa corresponding to the uORF2-c-Myc fusion protein, whose presence was strictly depended on the presence of the AUG of uORF2 (Fig. S4B; bottom panel, uORF2_ins => AUG to AGG in red) and whose intensity remained strikingly constant even under Tg stress (Fig. S4B, bottom electropherogram; see Table S15 and Table S16). Consistently, combining uORF2_ins with the triple AUG mutation produced three Tg-inducible peaks of the same size as observed with the 3AUG^{Mut} alone (Fig. S8 – uORF2_ins_3AUG^{Mut} *versus* Fig. S4A – 3AUG^{Mut}). Finally, mutations of the computer-predicted slippery sequences (Fig. 3A) showed no effect (data not shown). Altogether, these findings clearly rule out

possible uORF2 to *ATF4* frameshifting and strongly support the idea that uORF2 is translated even under stress in contrast to the original model.

Strikingly, we also observed that the 10x c-Myc tag insertion in uORF2 behaved similarly to the 6x c-Myc insertion past the AUG1 of *ATF4* with respect to the role of inhibitory SL3: unfolding SL3 no longer increased the ATF4 expression levels in neither of the double mutants; compare Fig. 6B *versus* 6A (Table S11 and Table S12) and Fig. S9B *versus* S9A (Table S7 and Table S8 for wt-SL3^{Mut} values, and Table S19 and Table S20 for uORF2_ins-SL3^{Mut} values). In fact, under Tg stress the expression levels for the uORF2-only constructs without the SL3 dropped substantially (Fig. S9B). This result contrasts the well-accepted model of delayed REI, which posits that extending the distance between uORF1 and *ATF4* provides the 48S PIC with more time to pick up the TC before reaching ATF4. Also, the uORF2 levels (probed with c-myc) remained unchanged, as expected, further supporting the ribosome queuing model proposed above. Therefore, our data suggests that the exact placement of SL3 inside of uORF2 with an exactly defined length represents a precise molecular design expanding the existing delayed REI model to a more complex model that includes ribosome queuing to ensure well-tuned stimulation of ATF4 expression.

mRNA methylation further fine tunes ATF4 translation

It has been proposed that under non-stress conditions, A225 in the non-overlapping region of uORF2 in mouse *ATF4* mRNA is modified by N^6-methyladenosine (m^{6}A), which functions as a barrier blocking access of ribosomes scanning downstream of uORF1 towards the *ATF4* ORF (Zhou et al., 2018). (Please note that in the NCBI Reference Sequence NM_001287180.1 used here, A225 corresponds to A258; however, for better clarity, we refer to it as it was originally described - A225.) Upon stress, this m^{6}A is supposedly demethylated, thereby unblocking access. Of the four potential methylation sites we identified computationally, we chose to investigate the effect of A235 (NCBI Reference Sequence NM_182810.3), as it seemed to match the mouse A225 best with respect to position (no sequence), and A326, which is located within SL3 and could thus modify its function. Both sites are highly conserved among vertebrates (Fig. S1A).

Both As were substituted with unmodifiable Gs. While A235G significantly increased the basal level of ATF4 expression under non-stress, as originally reported (Zhou et al., 2018), A326G reduced expression by ~30% (Fig. 7B, Table S21 and Table S22). Under stress, the values for induction nicely mirrored these opposing effects. To support these findings, we also performed a T3-ligation assay (Liu et al., 2018b; Zhang et al., 2019) and investigated changes in the methylation status of A235 and A326 in native *ATF4* transcripts isolated from a pool of poly(A) mRNAs from HEK293T and HeLa cells under non-stress and stress conditions. The T3-ligase is sensitive to m^{6}A sites during the ligation reaction in qPCR reactions (Liu et al., 2018b; Zhang et al., 2019). We first designed specific L (left) and R (right) probes based on the sequences flanking the two prospective m^{6}A sites of interest: the *ATF4-A235* (probes L1 and R1) and *ATF4-A326* (probes L2 and R2). After ligation, the resulting fragment was amplified by qPCR using universal primers that were complementary to the same adapter

sequences attached to all L and R probes (Liu et al., 2018b). Because presence of m^6A significantly reduced the T3-ligase efficiency, less effective qPCR amplification determined under which conditions (non-stress *versus* stressed) the particular A is methylated. For control and normalization purposes, we also designed two additional probe sets complementary to two regions bearing A residues, but not present in any DRACH motifs: *ATF4-A267* (probes L3 and R3) and *ATF4-A311* (probes L4 and R4). Any changes in the amplification patterns of qPCR fragments resulting from their ligation therefore served as a quantitative measure reflecting changes in the abundance of *ATF4* transcripts in total poly(A) mRNAs isolates under either non-stress or stress conditions.

Ligation reaction time intervals of 5 and 10 minutes revealed that the *ATF4* mRNA region containing the A235 but not the A326 residue had a ~30-50 % reduction in ligation efficiency under non-stressed conditions (Figure 7C, Table S23). This result suggests that methylation of A235 might indeed contribute to inhibition of *ATF4* translation under normal conditions. Interestingly, prolonged incubation of *ATF4* mRNA from HEK293T cells with T3-ligase (60 min) led to full saturation of the ligation reaction, i.e. showing no differences under non-stress and stress conditions. This finding was consistent with observations on the time-dependent ligation efficiency of other m^6A-containing mRNAs (Zhang et al., 2019). The still-visible difference for the 60-min incubation of *ATF4* transcripts isolated from HeLa cells likely reflected the unsaturated kinetics of this reaction due to the lower amounts of poly(A) RNAs that went into in the ligation reactions.

To the best of our knowledge, the m^6A modification at A235 has not yet been described for human *ATF4* mRNA. Therefore, we cannot rule out that an adenosine modification other than m^6A, such as m^6Am or m^1A, may be present at the A235 site and inhibit the ligation reaction. Nonetheless, the direct detection of m^6A in the homologous region of the mouse *ATF4* transcript by (Zhou et al., 2018) with a similar phenotype when mutated strongly suggests that human *ATF4* A235 is also modified by m^6A and serves as a barrier to ribosome progression under non-stress conditions. In addition, our data indicate that A326 could also be methylated, which would not inhibit but stimulate basal ATF4 expression, in contrast to A235.

DISCUSSION

The work presented here reconciles the original model of delayed reinitiation to control *ATF4* translation with mounting unexplained evidence that have challenged some of its key paradigms. Our results confirmed, for human *ATF4*, a translation inhibitory role of m^6A modification of A235 observed in mouse (Zhou et al., 2018) and identified a new modification site (A326) potentially antagonizing A235's function under normal conditions. Furthermore, we uncovered an entirely new layer of regulation of ATF4 expression implemented by a near-cognate start codon (CUG) followed by a stable stem-loop (SL3) immediately downstream of a substantially leaky AUG1 of *ATF4,* all inside the region overlapping with uORF2. Detailed analysis of this overlap region provided evidence for regulatory ribosome queuing upstream of SL3, which results in initiation at the CUG and translation of uORF2 even under stress – contrasting the

original model. Therefore, we propose that the *ATF4* regulatory "armamentarium" is far more complex than previously thought.

m^6A is the most common, highly dynamic, reversible and also non-stoichiometric type of internal modification of eukaryotic mRNAs (Meyer, 2019). Almost all stages of the mRNA life cycle, including assembly, export, stability, and translation, can be regulated by m^6A (Genuth and Barna, 2018). In coding regions, it can equally slow down translation, by reducing the accumulation of tRNA to the ribosome A-site, and speed it up, either by providing protection from ribosome arrest at specific "slow" codons (i.e. GANs) or by binding specific readers that, through helicase activity, can unwind double-stranded RNA structures impeding translation progression. In 5' UTRs, m^6A can as a roadblock to slow or stop the scanning small ribosomal subunit, thereby allowing recognition of near-cognate initiation codons that would otherwise be skipped by the scanning ribosome. Indeed, global profiling of m^6A and translation initiation site occupancy during starvation showed that the occurrence of m^6A in the 5' UTR was associated with increased signal at non-AUG start codons (Zhou et al., 2018). Alternatively, it can also render the scanning process more efficient by breaking secondary structures and, through interaction with eIF3, it may even serve as an alternative, cap-independent site for ribosome-to-mRNA binding (Meyer et al., 2015). In either case, m^6A modifications are an important aspect of translational control.

Our analyses of human *ATF4* support a proposed role for the m^6A modification of A235 as a barrier to scanning ribosomes during non-stress conditions (Zhou et al., 2018). At the same time, they also pointed to a potential antagonizing modification of A326 that could promote scanning and/or elongation by destabilizing SL3 and or unfolding it with help from m^6A reader-associated helicase(s). The potential importance of the newly discovered A326 modification can be inferred from comparative analysis of the *ATF4* transcript V2 variant (analyzed here) with the transcript sequences of the two *ATF4* pseudogenes (Fig. S10). Both pseudogenes accumulated mutations in the start codon region of uORF2, so that ribosomes scanning downstream of uORF1 would automatically skip uORF2 and reinitiate directly at ATF4. Interestingly, both pseudogenes also lost the A326 through deletion, which shifts the reading frame and creates a stop codon that immediately precedes SL3 of *ATF4*; i.e. the ribosome translating ATF4 would essentially terminate shortly after the *ATF4* start site, preventing its further translation. We can only speculate whether this deletion ensured genetic tolerance for both pseudogenes or followed immediately upon their emergence in the genome to prevent dysregulated ATF4 synthesis under any circumstances.

Our results further corroborate the broad role of eIF3 in translational control, in particular in ATF4. Due to its ability to remain transiently associated with elongating ribosomes, especially on short uORFs (Bohlen et al., 2020a; Lin et al., 2020; Mohammad et al., 2017; Mohammad et al., 2021; Wagner et al., 2020), eIF3 promotes the resumption of traversing of post-termination 40S ribosomes from uORF1 downstream (Hronova et al., 2017). It has only recently been proposed that this uORF1-specific role depends on O-GlcNAcylation of eIF3's core subunit, eIF3a (Shu et al., 2022). Specifically, O-GlcNAcylated eIF3a does not interfere with the eIF3's ability to promote formation of 43-48S PICs, however, it was suggested that it

destabilizes eIF3's association with post-initiation 80S ribosomes during early rounds of elongation under non-stress conditions (Shu et al., 2022). During acute amino acid starvation, eIF3a becomes de-O-GlcNAcylated, which restores eIF3's ability to retain on elongating ribosomes translating the short uORF1 and promote reinitiation downstream of it. In addition, eIF3 also allows a switch specific to chronic stress from an eIF4F/eIF2B-dependent mode of translation initiation to an eIF3-dependent mechanism involving direct recruitment of selected mRNAs to 40S ribosomes (Guan et al., 2017). It is tempting to speculate that this switch might be enabled by one of the m^6As in ATF4. Thus, eIF3 represents yet another layer of regulation ATF4 in the integrated stress response.

uORF2 translation under stress was initially reported by Starck et al. (Starck et al., 2016), revealing a high degree of peptide expression from both uORF1 and uORF2 during both normal growth and stress conditions. The results were obtained from constructs fusing the uORF with a tracer peptide at the N-terminus of the respective sequence. The observation was supported by results from ribosome profiling experiments by different groups (Andreev et al. 2015; Sidrauski et al. 2015). To explain this discrepancy with the original model, it was suggested that either reinitiation commences after only a few 40S subunits bypass uORF2 or that translation of the *ATF4* coding region is a combination of reinitiation and leaky scanning past both uORFs. Given the complexity of the *ATF4* leader regulatory armamentarium, it is intuitive that the tracer peptide employed by Starck et al. (Starck et al., 2016) could have affected the results, e.g. by separating and/or breaking the additional *ATF4* regulatory elements we identified here. The work presented here using a more representative reporter system both confirmed high levels of uORF2 expression under normal and stress conditions, and provided a mechanistic explanation for it with respect to ribosome queuing ahead of SL3.

The mechanism of ribosome queuing is known from other systems, such as translation of the antizyme inhibitor 1 (AZIN1) responding to polyamines (Ivanov et al., 2018). The mRNA leader of *AZIN1* contains a relatively long, inhibitory, non-AUG-initiated upstream conserved coding region (uCC), with a conserved, poly-amine-dependent Pro-Pro-Trp (PPW) motif pausing elongating ribosomes. At low polyamine levels, majority of scanning ribosomes skip this uCC and translate AZIN1; those that infrequently initiate at its AUU start site will translate it, terminate, and become recycled. High polyamine levels, however, interfere with eIF5A function and cause those very few ribosomes translating uCC to stall at PPW. Subsequently scanning ribosomes, which again mostly skip the AUU start codon, as well as the occasional elongating ribosome form a queue upstream of the PPW-stalled ribosome. This queue then stimulates uCC initiation at the AUU as scanning ribosomes are jammed all the way up to this site, reinforcing the PPW-imposed elongation stall to effectively suppress AZIN1 synthesis.

The proposed new layer of *ATF4* translation control is analogous to the AZIN1 mechanism. Specifically, the scenario with 80S ribosomes elongating from the AUG of uORF2 and subsequently queuing upstream of SL3 under non-stress conditions (Fig. 3B) - which prevents translation of ATF4 and ultimately also that of uORF2 - is similar

to the AZIN1 system at high polyamine levels. The queuing re-enforces uORF2 initiation and very efficiently limits skipping of its AUG. This process minimizes the leakiness of the whole system to the lowest possible level: at zero demand, there is zero ATF4 synthesis. Upon stress, SL3 slows down 80S ribosomes elongating from the AUG of ATF4, allowing for the formation of a queue consisting of both scanning 48S PICs and elongating 80S ribosomes (Fig. 3B). The SL3-imposed queue can then (i) enhance selection of alternative, weaker start codon(s) in *ATF4*, and (ii) prevent ATF4 over-induction. The process would involve reduced but still substantial translation of uORF2 even under stress, as has been observed. The near-cognate CUG would serve as another warrant licensing ATF4 synthesis, overcoming the leakiness of the ATF4 start site(s) (Fig. 3B). Alternatively or additionally, use of this alternate (weak) start codon would expand the coding potential of *ATF4* mRNAs, the consequence of which remains to be investigated.

Our data imply that the proposed ribosome queue model depends on the distance between the elongation pause and the upstream start codon. When we increased this distance, either through the uORF2 or ATF4 reading frame, the system stopped responding. However, due to the variable length of mRNA FPs that 48S PICs protect compared to the relatively defined FP length of elongating 80S (Archer et al., 2016; Ingolia, 2010; Wagner et al., 2022; Wagner et al., 2020), the exact nature and positioning of the most 5' ribosomal species at the most likely initiation site cannot be predicted. However, this may not even be necessary because, as suggested by (Ivanov et al., 2018), ribosome queuing forces initiating ribosome to spend more time near the start codon where it can migrate back and forth by ~15 nucleotides, as previously shown (Matsuda and Dreher, 2006). Therefore, the operational space for the last 48S PIC in the queue to find the translation initiation site is relatively wide, even if the start codon is not in an optimal context.

Based on mounting evidence from the literature, it is now well-established that the precise mode of ISR induction, whether through UPR activation, starvation, or inhibition of tRNA synthetases, likely has very distinct downstream effects that manifest in highly nuanced *ATF4* translational control followed by unique modes of ATF4-mediated transcriptional responses (Neill and Masson, 2023). In light of this high degree of plasticity both in the control of ATF4 expression and its subsequent regulatory effect on its own transcriptional "regulon", it seems plausible and necessary that an extremely complex network of translation regulatory modes, as described here, is required to adjust ATF4 expression to a variety of conditions and tissues. Given the involvement of deregulated ATF4 expression in various pathologies and cancers, the complexity of its translational control should be kept in mind when considering therapies directed against this potent regulator of cell life or death.

ACKNOWLEDGEMENTS

We are grateful to Dmitry E. Andreev, Pasha Baranov and Nick Guydosh for their timely advice and all past and present members of the Valasek lab for fruitful discussions and patience with this lengthy project. This work was supported by a Grant of Excellence in Basic Research (EXPRO 2019) provided by the Czech Science Foundation (19-

25821X to L.S.V.), US National Institutes of Health 5R35GM127089 and Chan Zuckerberg Initiative (to C.V), and US National Institutes of Health DK060596 (to M.H.). We also acknowledge Structural Mass Spectrometry Core Facility of CIISB, Instruct-CZ Centre (LM2023042) and European Regional Development Fund-Project "UP CIISB" (CZ.02.1.01/0.0/0.0/18_046/0015974) supported by MEYS CR.

AUTHOR CONTRIBUTIONS

L.S.V. conceived and designed the project.
A.M.S., V.H., S.G., P.M.M., D.P., P.H., A.H., P.B., J.R., S.C. and K.J. carried out all experiments and analyzed the data.
L.S.V. interpreted results and wrote the paper with input from A.M.S., S.G., and C.V., and partly also from P.M.M., D.P., P.H., J.R., and M.H..

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at ….

DECLARATION OF INTERESTS

The authors declare no competing interests.

FIGURES AND FIGURE LEGENDS

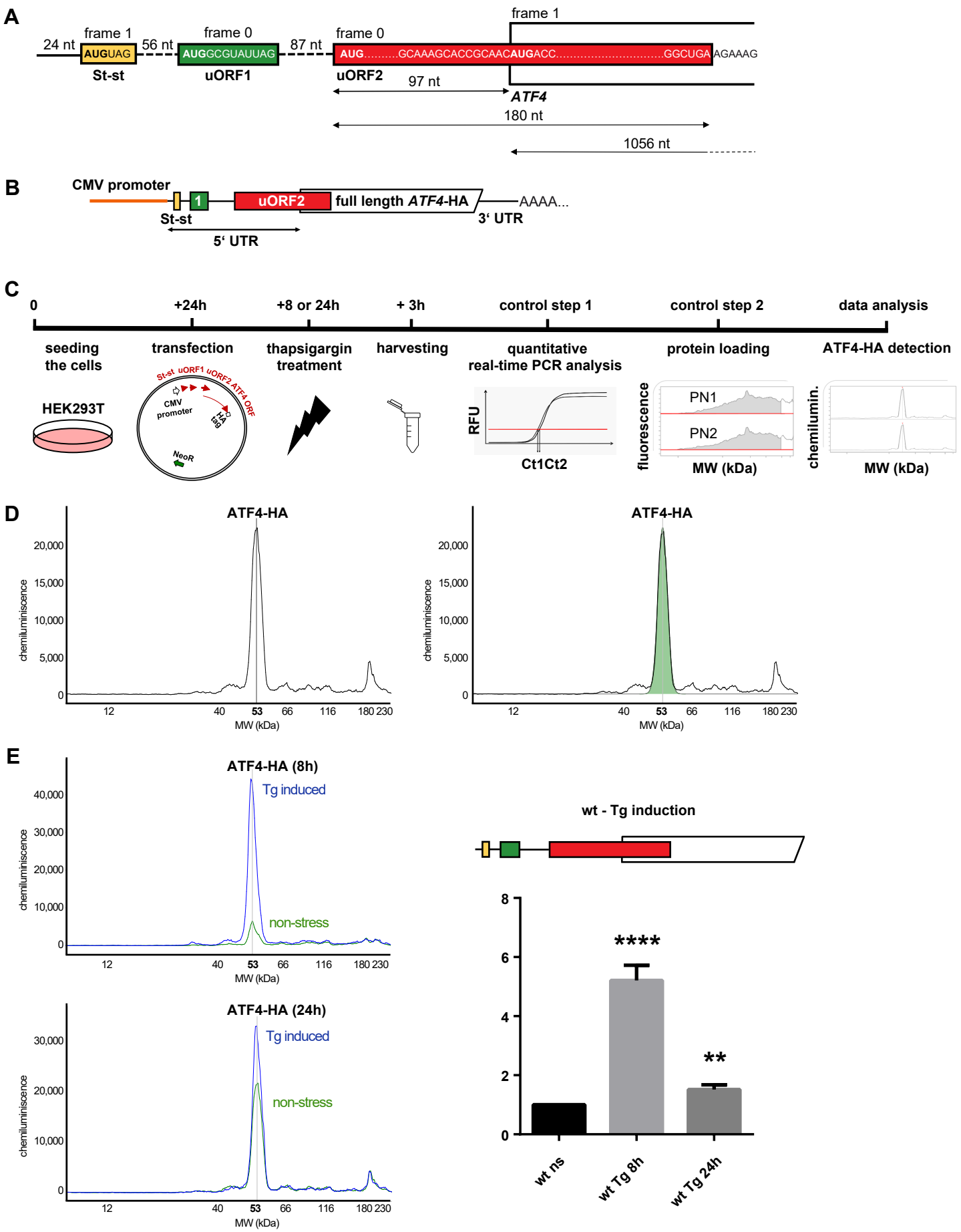

A
24 nt
frame 1
AUGUAG
St-st
56 nt
frame 0
AUGGCGUAUUAG
uORF1
87 nt
frame 0
AUG.....GCAAAGCACCGCAACAUGACC.....GGCUGA
AGAAAG
uORF2
97 nt
frame 1
ATF4
180 nt
1056 nt
B
CMV promoter
1
uORF2
full length ATF4-HA
AAAA...
St-st
5' UTR
3' UTR
C
0
+24h
+8 or 24h
+ 3h
control step 1
control step 2
data analysis
seeding the cells
transfection
thapsigargin treatment
harvesting
quantitative real-time PCR analysis
protein loading
ATF4-HA detection
HEK293T
St-st uORF1 uORF2 ATF4 ORF
CMV promoter
HA tag
NeoR
RFU
Ct1Ct2
fluorescence
PN1
PN2
MW (kDa)
chemilumin.
MW (kDa)
D
ATF4-HA
chemiluminiscence
20,000
15,000
10,000
5,000
0
12
40
53
66
116
180 230
MW (kDa)
ATF4-HA
E
ATF4-HA (8h)
Tg induced
non-stress
40,000
30,000
20,000
10,000
ATF4-HA (24h)
Tg induced
non-stress
30,000
20,000
10,000
wt - Tg induction
8
6
4
2
0
****
**
wt ns
wt Tg 8h
wt Tg 24h

Figure 1. Revisiting translational control of human *ATF4* – reporters and experimental set-up.

(A) Schematic showing the 5' end of the mRNA encoding the human activating transcription factor 4 (*ATF4*; NM_182810.2, transcript variant 2), featuring the color-coded Start-stop (St-st) element (yellow), REI-permissive (green) 3-codon uORF1, and inhibitory uORF2 (red) overlapping with the beginning of main *ATF4* ORF. Frames and distances are given in nucleotides (nt).

(B) Schematic of the wt CMV-driven ATF4-HA tagged construct; the HA tag was placed immediately upstream of the *ATF4* stop codon.

(C) Experimental workflow described in the form of a time-line diagram; PN - Jess Protein Normalization Detection Module.

(D) All HEK293T cell lysates were subjected to protein separation and immunodetection using the Jess Simple Western system (Protein Simple) that generates electropherograms. Automated detection of the ATF4-HA full-length protein forming a single peak at 53 kDa (left panel) is quantified using Compass SW Software (right panel); expression of wt construct under non-stress conditions 24 h post-transfection detected by anti-mouse HA tag antibodies is shown.

(E) Stress-induced upregulation of the ATF4-HA protein expression under Tg stress (in blue) compared to non-stress conditions (in green) either 8 (top panel) or 24 h (bottom panel) post-transfection is shown. Quantified "fold-induction" data obtained from n=17 (8 h) and n=14 (24 h) measurements were plotted (right panel). Variables were tested for normality using the D'Agostino–Pearson normality test. Based on the normality test, the differences between experimental groups (grey bars) and wt construct (black bar) were tested by the Kruskal-Wallis test accompanied by Dunn's multiple comparisons test. Variables are presented as the mean ± SD and p values of < 0.05 were considered statistically significant (i.e., $p<0.05$ *, $p<0.01$ **, $p<0.001$ ***, $p< 0.0001$ ****).

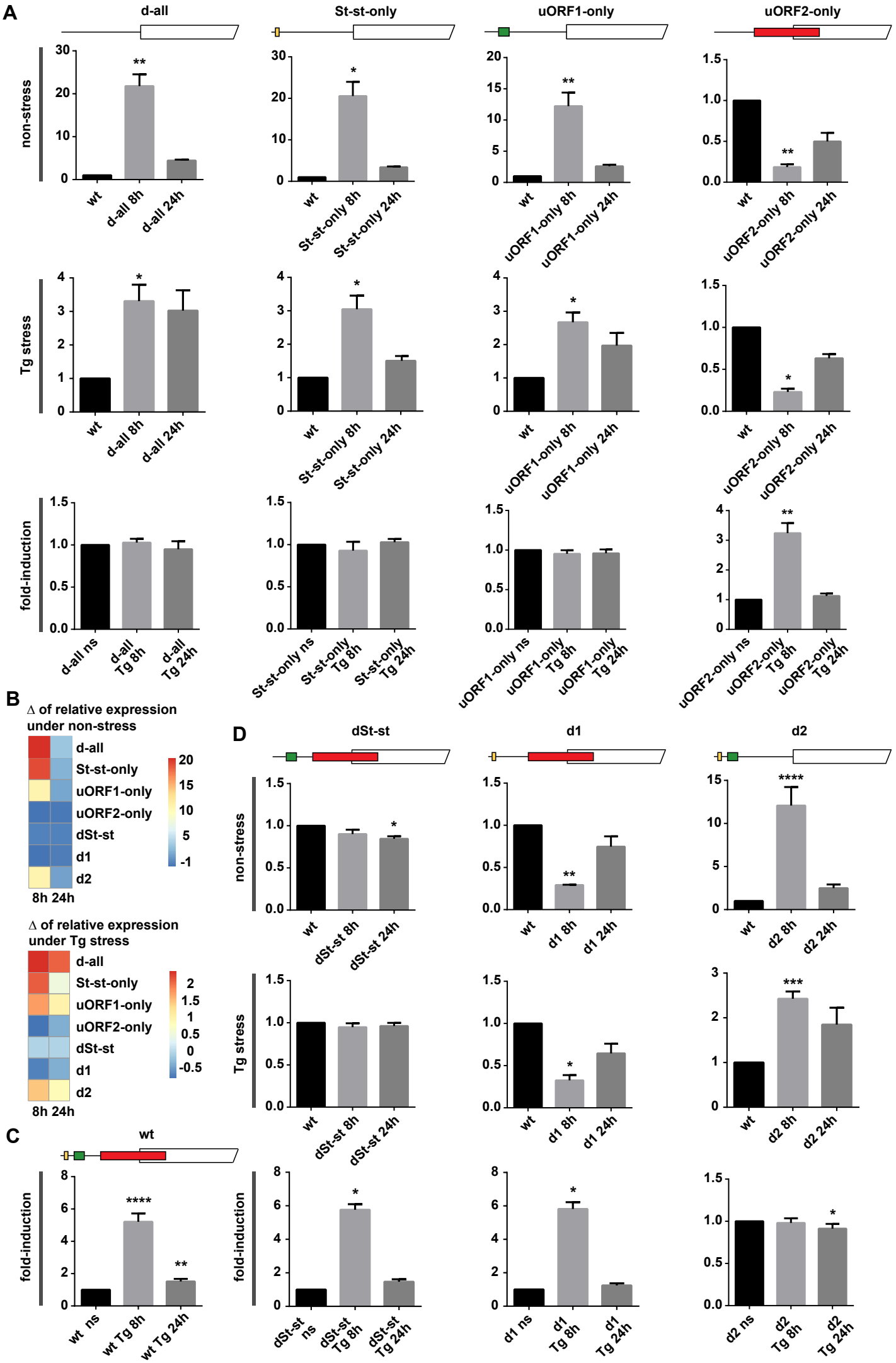

A
d-all
St-st-only
uORF1-only
uORF2-only
non-stress
Tg stress
fold-induction
wt
d-all 8h
d-all 24h
St-st-only 8h
St-st-only 24h
uORF1-only 8h
uORF1-only 24h
uORF2-only 8h
uORF2-only 24h
d-all ns
d-all Tg 8h
d-all Tg 24h
St-st-only ns
St-st-only Tg 8h
St-st-only Tg 24h
uORF1-only ns
uORF1-only Tg 8h
uORF1-only Tg 24h
uORF2-only ns
uORF2-only Tg 8h
uORF2-only Tg 24h
B
Δ of relative expression under non-stress
d-all
St-st-only
uORF1-only
uORF2-only
dSt-st
d1
d2
8h 24h
Δ of relative expression under Tg stress
C
wt
wt ns
wt Tg 8h
wt Tg 24h
D
dSt-st
d1
d2
dSt-st ns
dSt-st Tg 8h
dSt-st Tg 24h
d1 ns
d1 Tg 8h
d1 Tg 24h
d2 ns
d2 Tg 8h
d2 Tg 24h

Figure 2. Revisiting the delayed translation reinitiation mechanism of *ATF4* translational control – uORF2 shows an inducible nature under Tg stress.

(A) Schematics at the top of the corresponding panels indicate the *ATF4* mutant constructs subjected to JESS analyses. Relative ATF4-HA protein expression levels were plotted as ratios of values obtained with mutants *versus* wt set to 1 under "non-stress" and "Tg stress" conditions; "fold-induction" plots depict ratios of Tg stress *versus* non-stress values obtained with mutant constructs either 8 or 24 hours post-transfection. Statistical analyses were carried out as in Figure 1E.

(B) Heatmaps, created in R (ver. 4.2.2) using the 'pheatmap' package (ver. 1.0.12) (Kolde and Kolde, 2015), show the differences (Δ) in the relative ATF4-HA protein expression of individual mutant *versus* wt constructs that were obtained under non-stress (top panel) and Tg stress conditions (bottom panel) either 8 or 24 h post-transfection. The maximal expression levels are shown in red (for the control d-all construct), whereas the minimal levels are depicted in blue (for the uORF2-only construct).

(C) Same as Figure 1E for better comparison.

(D) Same as panel A with different constructs, depicted at the top, under study.

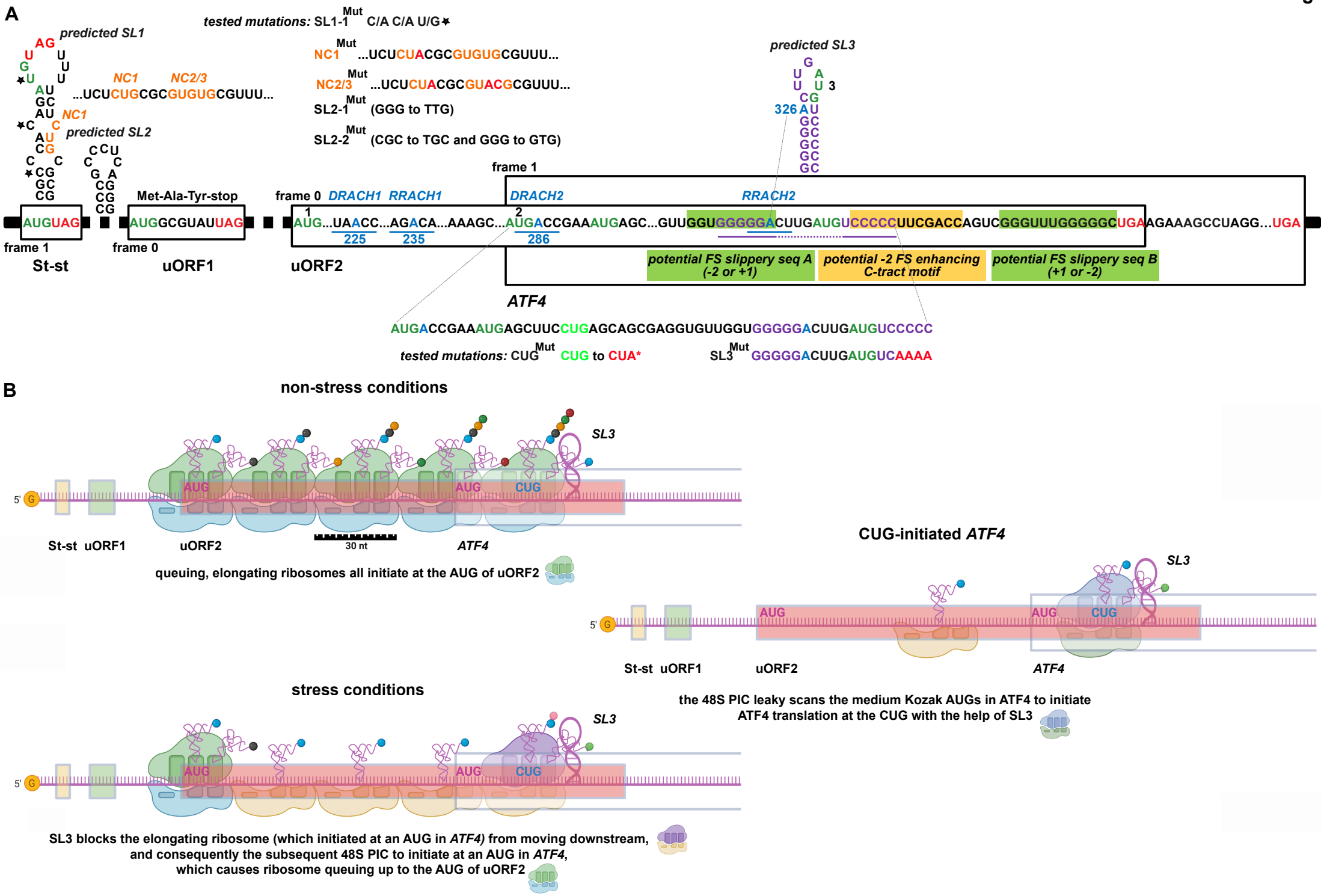

A
predicted SL1
AG
U U
G U
U U
AU
GC
AU
C C
AU
CG
C C
CG
GC
CG
NC1
NC2/3
...UCUCUGCGCGUGUGCGUUU...
NC1
predicted SL2
CCU
C C
G A
CG
CG
CG
tested mutations: SL1-1 Mut C/A C/A U/G*
NC1 Mut ...UCUCUACGCGUGUGCGUUU...
NC2/3 Mut ...UCUCUACGCGUACGCGUUU...
SL2-1 Mut (GGG to TTG)
SL2-2 Mut (CGC to TGC and GGG to GTG)
predicted SL3
G
U A
U U 3
CG
326 AU
GC
GC
GC
GC
GC
frame 1
Met-Ala-Tyr-stop
frame 0
DRACH1
RRACH1
DRACH2
RRACH2
AUGUAG
AUGGCGUAUUAG
AUG...UAACC...AGACA...AAAGC...AUGACCGAAAUGAGC...GUUGGUGGGGGACUUGAUGUCCCCCUUCGACCAGUCGGGUUUGGGGGCUGAAGAAAGCCUAGG...UGA
225
235
286
frame 1
frame 0
St-st
uORF1
uORF2
potential FS slippery seq A (-2 or +1)
potential -2 FS enhancing C-tract motif
potential FS slippery seq B (+1 or -2)
ATF4
AUGACCGAAAUGAGCUUCCUGAGCAGCGAGGUGUUGGUGGGGGACUUGAUGUCCCCC
tested mutations: CUG Mut CUG to CUA*
SL3 Mut GGGGGACUUGAUGUCAAAA
B
non-stress conditions
SL3
5' G
AUG
AUG
CUG
St-st uORF1
uORF2
30 nt
ATF4
queuing, elongating ribosomes all initiate at the AUG of uORF2
CUG-initiated ATF4
SL3
5' G
AUG
AUG
CUG
St-st uORF1
uORF2
ATF4
the 48S PIC leaky scans the medium Kozak AUGs in ATF4 to initiate ATF4 translation at the CUG with the help of SL3
stress conditions
SL3
5' G
AUG
AUG
CUG
SL3 blocks the elongating ribosome (which initiated at an AUG in ATF4) from moving downstream, and consequently the subsequent 48S PIC to initiate at an AUG in ATF4, which causes ribosome queuing up to the AUG of uORF2

Figure 3. Sequence analysis of the 5' UTR of the human *ATF4* mRNA; ribosome queuing model expanding the mode of ATF4 translational control.

(A) Schematic of novel, bioinformatically predicted, potential regulatory features within the 5' UTR of the *ATF4* mRNA and beginning of the *ATF4* main ORF; point mutations are depicted. For details, please see the main text.

(B) Model of the ribosome queuing mechanism under non-stress *versus* stress condition employing SL3 and near-cognate CUG as an additional layer of *ATF4* translational control. For details, please see the main text. Created with BioRender.com.

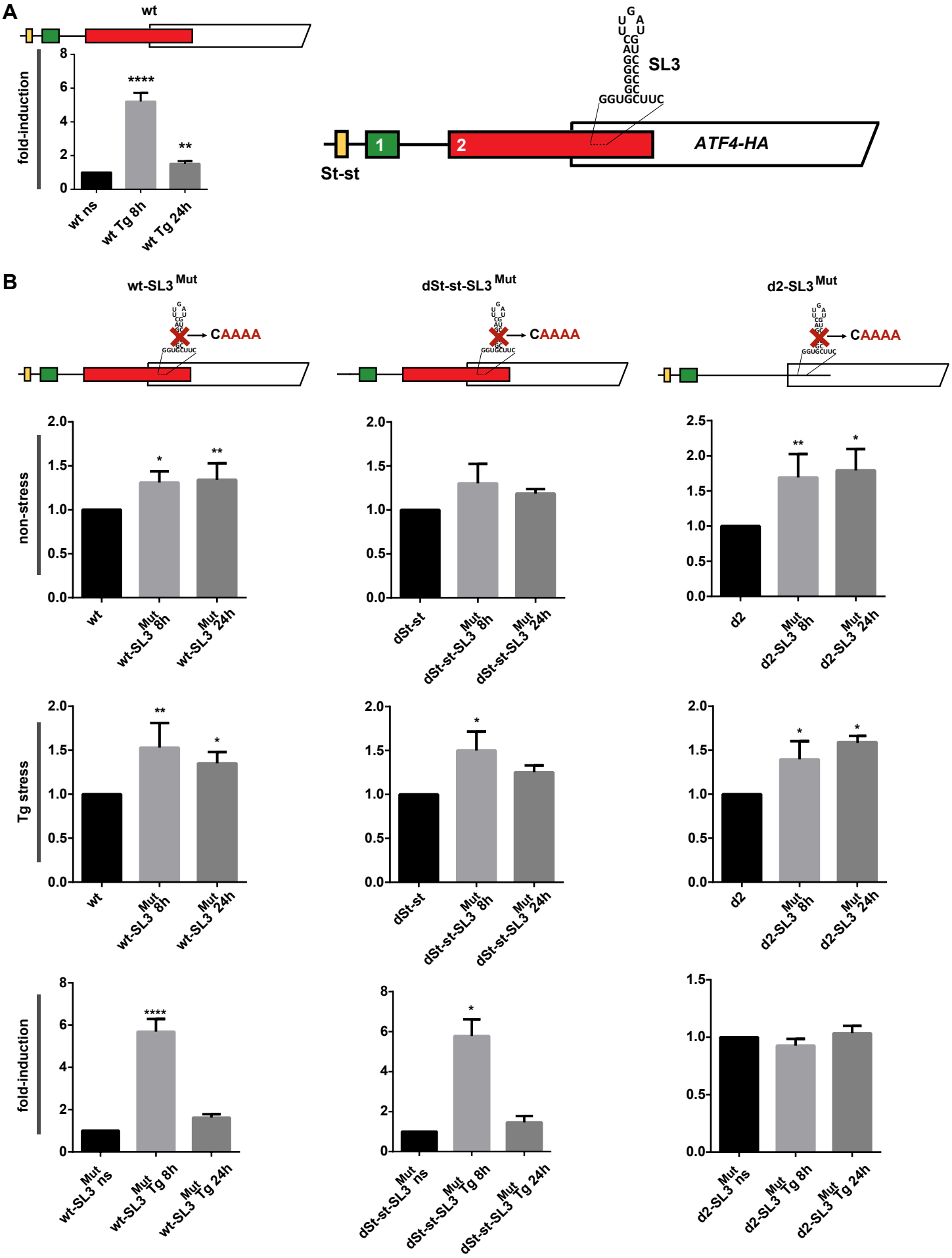

A
wt
fold-induction
****
**
wt ns
wt Tg 8h
wt Tg 24h
SL3
GGUGCUUC
St-st
1
2
ATF4-HA
B
wt-SL3 Mut
dSt-st-SL3 Mut
d2-SL3 Mut
CAAAA
non-stress
Tg stress
fold-induction
wt
wt-SL3 Mut 8h
wt-SL3 Mut 24h
dSt-st
dSt-st-SL3 Mut 8h
dSt-st-SL3 Mut 24h
d2
d2-SL3 Mut 8h
d2-SL3 Mut 24h
wt-SL3 Mut ns
wt-SL3 Mut Tg 8h
wt-SL3 Mut Tg 24h
dSt-st-SL3 Mut ns
dSt-st-SL3 Mut Tg 8h
dSt-st-SL3 Mut Tg 24h
d2-SL3 Mut ns
d2-SL3 Mut Tg 8h
d2-SL3 Mut Tg 24h

Figure 4. SL3 delays the flow of ribosomes in the uORF2/*ATF4* overlap.
(A) Same as Figure 1E for better comparison.
(B) Same as in Figure 2A except that the SL3 *ATF4* mutant constructs depicted at the top of the corresponding panels were subjected to JESS analyses.

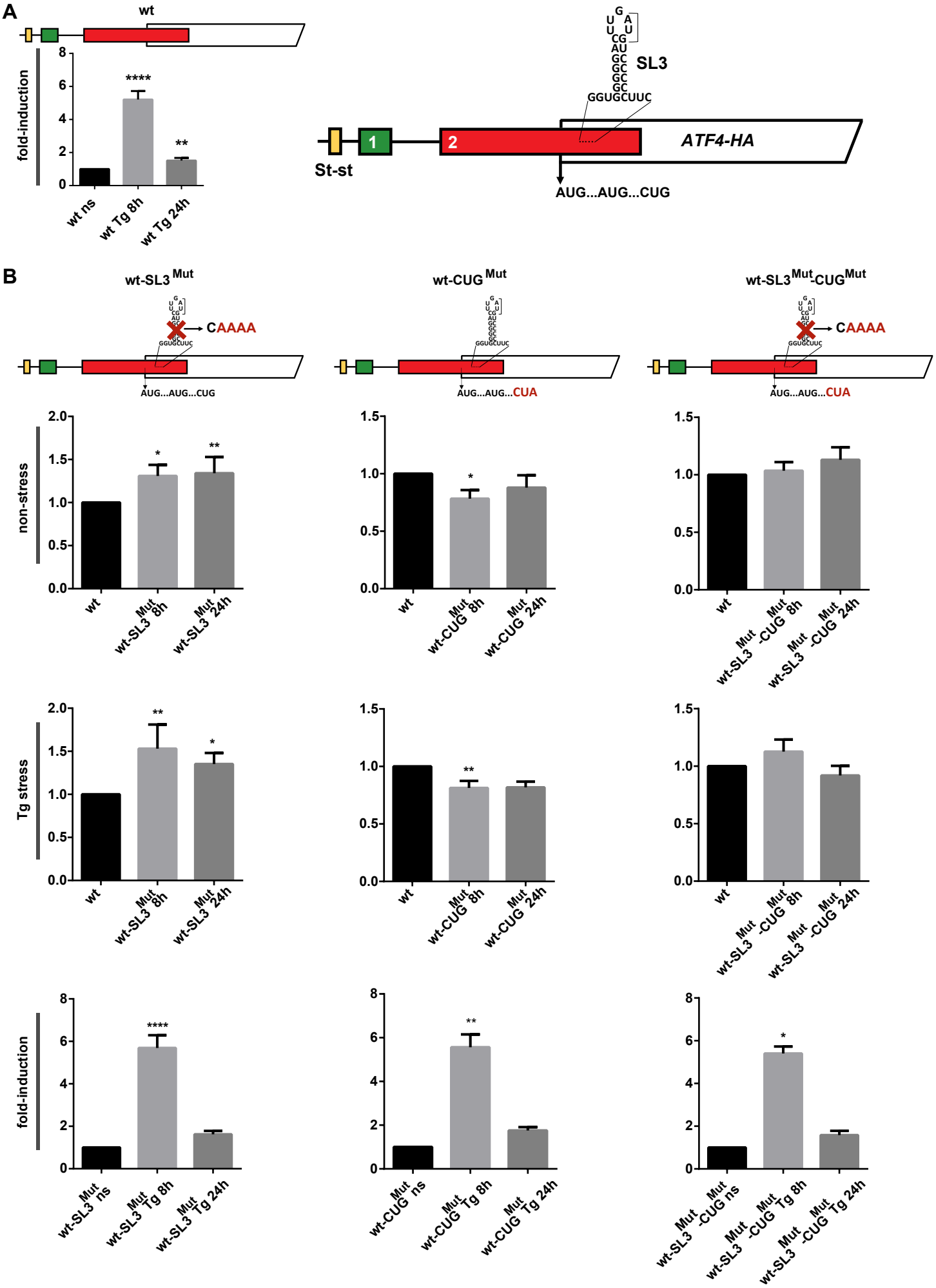

A
wt
fold-induction
****
**
wt ns
wt Tg 8h
wt Tg 24h
SL3
GGUGCUUC
St-st
1
2
ATF4-HA
AUG...AUG...CUG
B
wt-SL3 Mut
wt-CUG Mut
wt-SL3 Mut-CUG Mut
CAAAA
AUG...AUG...CUG
AUG...AUG...CUA
non-stress
Tg stress
fold-induction
wt
wt-SL3 Mut 8h
wt-SL3 Mut 24h
wt-CUG Mut 8h
wt-CUG Mut 24h
wt-SL3 Mut -CUG Mut 8h
wt-SL3 Mut -CUG Mut 24h
wt-SL3 Mut ns
wt-SL3 Mut Tg 8h
wt-SL3 Mut Tg 24h
wt-CUG Mut ns
wt-CUG Mut Tg 8h
wt-CUG Mut Tg 24h
wt-SL3 Mut -CUG Mut ns
wt-SL3 Mut -CUG Mut Tg 8h
wt-SL3 Mut -CUG Mut Tg 24h

Figure 5. SL3 promotes initiation at the upstream CUG near-cognate initiation site.

(A) Same as Figure 1E for better comparison.

(B) Same as in Figure 2A except that the CUG to CUA and SL3 *ATF4* mutant constructs depicted at the top of the corresponding panels were subjected to JESS analyses.

Smirnova Fig. 6

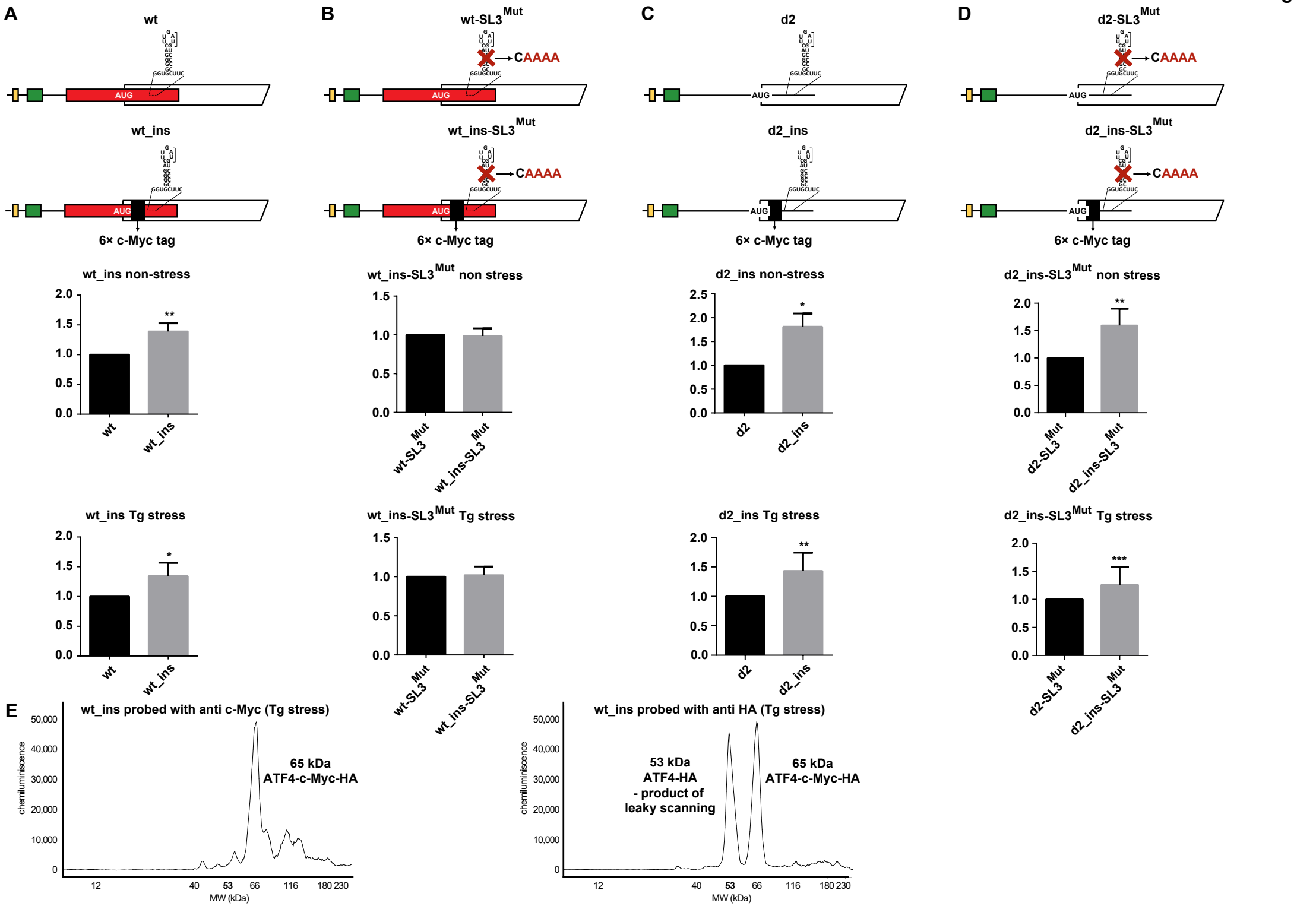


Figure 6. Ribosome queuing and a substantial leaky scanning at AUG1 of *ATF4* contributes to its overall translational control.

(A – D) Same as in Figure 2A except that the 6x c-Myc tag insertion in-frame with *ATF4* (A) combined with the SL3 mutation (B) in the otherwise wt construct (A – B) *versus* the construct lacking uORF2 (C – D), all depicted at the top of the corresponding panels, were subjected to JESS analyses.

(E) The first AUG of the *ATF4* ORF is robustly leaky scanned. The electropherograms of the *ATF4* construct bearing the 6x c-Myc tag in-frame insertion probed with the anti-c-Myc (left panel) and anti-HA (right panel) antibodies are shown. For details, see the main text.

Smirnova Fig. 7

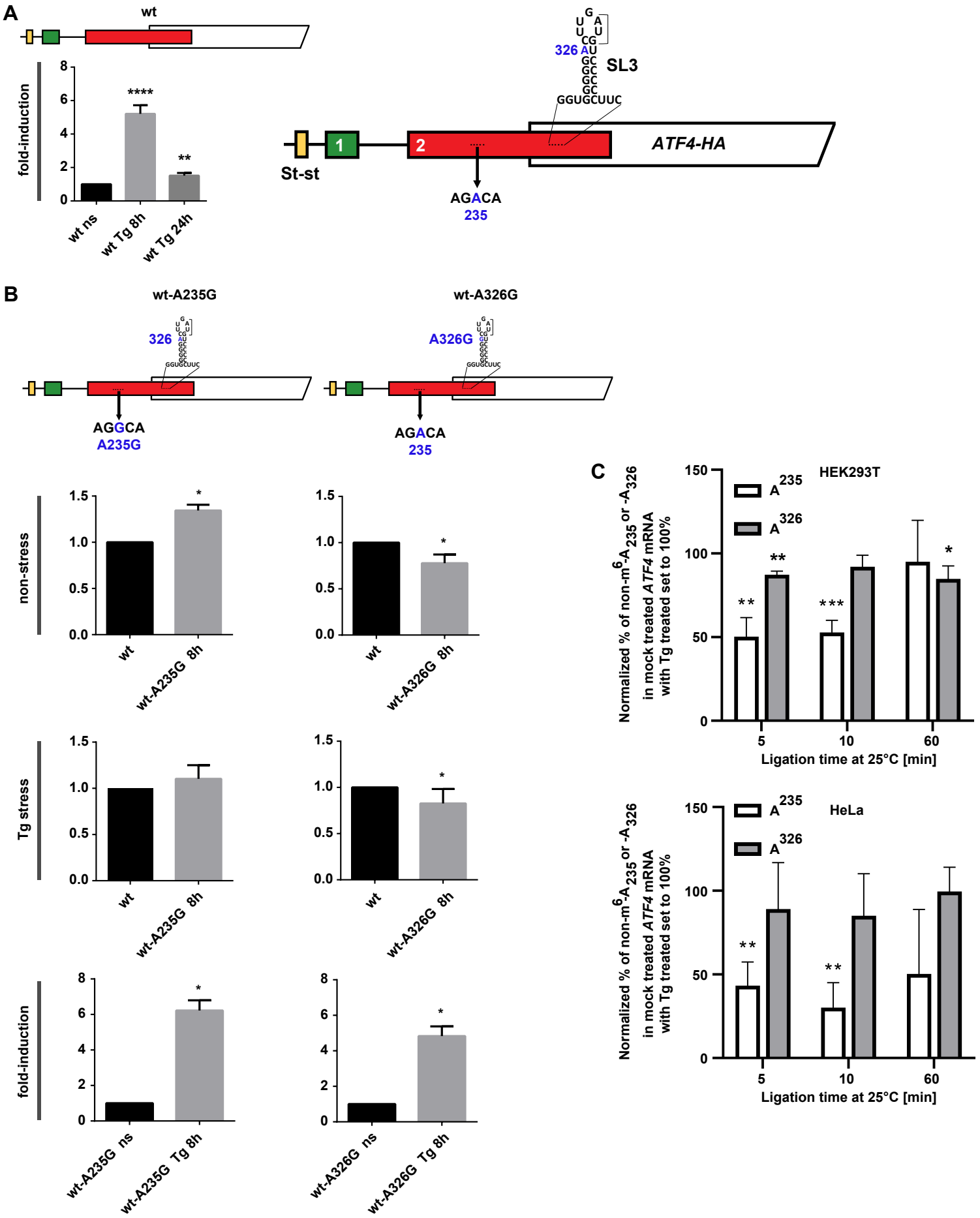


Figure 7. mRNA methylation further fine tunes ATF4 translation.

(A) Same as Figure 1E for better comparison.

(B) Same as in Figure 2A except that the A235G (right panel) and A326G (left panel) *ATF4* mutant constructs depicted at the top of the corresponding panels were subjected to JESS analyses.

(C) mRNA fragments prepared from either HEK293T (top panel) or HeLa (bottom panel) cells carrying the A235G and A326G mutations were subjected to T3-ligation assay as described in the main text. Normalized percentage of m^6A unmodified A_{235} or A_{326} bases of the *ATF4* mRNA expressed in mock treated *versus* Tg treated cells, with the latter set to 100%, was plotted shown; presence of the m^6A modification interferes with the T3-ligase reaction and thus reduces the efficiency of the subsequent qPCR amplification.

SUPPLEMENTARY FIGURES AND FIGURE LEGENDS

A

uORF2 ATF4

Human GGCTTAAGCCATGGCGCTTCTCACGGCATTCAGCAGCAGCGTTGCTGTAACCGACAAAGACACCTTCGAATTAAGCACATTCCTCGATTCCAGCAAAGCACCGCAACATGACCGAAATGA

Macaque GGCTTAAGCCATGGCGCTTCTCACGGCATTCAGCAGCAGCGTTGCTGTAACCGACAAAGACATCTTCGAATTAAGCACATTCCTCGATTCCAGCAAAGCACCGCAACATGACCGAAATGA

Chimpanzee GGCTTAAGCCATGGCGCTTCTCACGGCATTCAGCAGCAGCGTTGCTGTAACCGACAAAGACACCTTCGAATTAAGCACATTCCTCGATTCCAGCAAAGCACCACAACATGACCGAAATGA

Baboon GGCTTAAGCCATGGCGCTTCTCACGGCATTCAGCAGCAGCGTTGCTGTAACCGACAAAGACATCTTCGAATTAAGCACATTCCTCGATTCCAGCAAAGCACCGCAACATGACCGAAATGA

Marmoset GGCTTAAGCCATGGCGCTTCTCACGGCATTCAGCAGCAGCGTTGCTGTAAGCGACAAAGACACCTTCGAATTAAGCACATTCCTCGATTCCAGCGAAGCACCGCAGCATGACCGAGATGA

Mouse GGCCTAAGCCATGGCGCTCTTCACGAAATCCAGCAGCAGTGTTGCTGTAACGGACAAAGATACCTTCGAGTTAAGCACATTCCTCGAATCCAGCAAAGCCCCACAACATGACCGAGATGA

Rat GGCTTAAGCCATGGCGCTCTTCACGAAACCCAGCAGCAGCGTTGCTGTAACGGACAAAGATACCTTCGAGTTAAGCACATTCCTCGATACCAGCAAATCCCTACAACATGACCGAGATGA

Dolphin GGCTTAAGCCATGGCGCTTTTCAAGGCATTCAGCAGCAGCGTTGCTGTAACCGACAAAGACACTTTCGAGTTAAGCACATTCCTCGATTCCACCAAAACCTCGCGACATGGCCGAGAAGA

Blue whale GGCTTAAGCCATGGCGCTTTTCACGGCATTCAGCAGCAGCGTTGCTGTAACCGACAAAGACACCTTCGAATTAAGCACATTCCTCGATTTCAGCAAGGCCCCGCGTTATGGCCGAGATGA

Elephant GGCTTAAGCCATGGCGCTTCTCACGGCATTCAGCAGCAGCATTGCTGTAACCGACAGAGACACCTTCGAATTAAGCACATTCCTCGATTTCAGCAAGGCCCCGCGTTATGGCCGAGATGA

Wild boar GGCTTAGGCCATGGCGCTTCTCACGGCATTCAGCAGCAGCGTTGCTGTAACCGACAAAGACACCTTCGAATTAAGCACATTCCTCGATTCCATCAAAACTCCGCAACATGGCCGAGATGA

Human GCTTCCTGAGCAGCGAGGTGTTGGTGGGGGACTTGATGTCCCCCTTCGACCAGTCGGGTTTGGGGGCTGAAGAAAGCCTA

Macaque GCTTCCTGAGCAGCGAGGTGTTGGTGGGGGACTTGATGTCCCCCTTCGACCAGTCGGGTTTGGGGGCTGAAGAAAGCCTA

Chimpanzee GCTTCCTGAGCAGCGAGGTGTTGGTGGGGGACTTGATGTCCCCCTTCGACCAGTCGGGTTTGGGGGCTGAAGAAAGCCTA

Baboon GCTTCCTGAGCAGCGAGGTGTTGGTGGGGGACTTGATGTCCCCCTTCGACCAGTCGGGTTTGGGGGCTGAAGAAAGCCTA

Marmoset GTTTCCTGAGCAGCGAGGTGTTGGAGGGGGACTTGATGTCCCCCTTCGACCAGTCGGGTTTGGGGGCTGAGGAAAGCCTA

Mouse GCTTCCTGAACAGCGAAGTGTTGGCGGGGGACTTGATGTCCCCCTTCGACCAGTCGGGTTTGGGGGCTGAAGAAAGCCTA

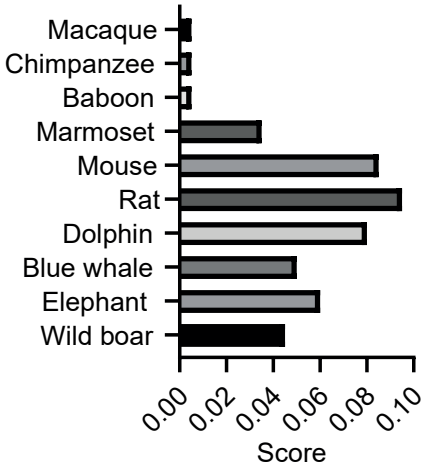
Rat GCTTCCTGAACAGCGAAGTGTTGGCGGGGGACTTAATGTCCCCCTTCGACCAGTCGGGTTTGGGGGCTGAAGAAAGCCTA

Dolphin GCTTCCTGAGCAGCGAGGTGTTGGTGGTGGACTTCATGTCCCCCTTCGACCAGTCGGGTTTGGGGGCTGAAGAGAGCCTG

Blue whale GCTTCCTGAGCAGCGAGGTGTTGGTGGGGGACTTCATGTCCCCCTTCGACCAGTCGGGTTTGGGGGCTGAAGAGAGCCTG

Elephant GCTTCCTGAGCAGCGAGGTGTTGGTGGGGGACTTGATGTCCCCCTTCGACCAGTCGGGTTTGGGGGTTGAAGAAACCCTA

Wild boar GCTTCCTGAGCAGCGAGGTGTTGGTGGGGGACTTGATGTCCCCCTTCGACCCGTCGGGTTTGGGGGCTGAAGAGAGCTTA



B

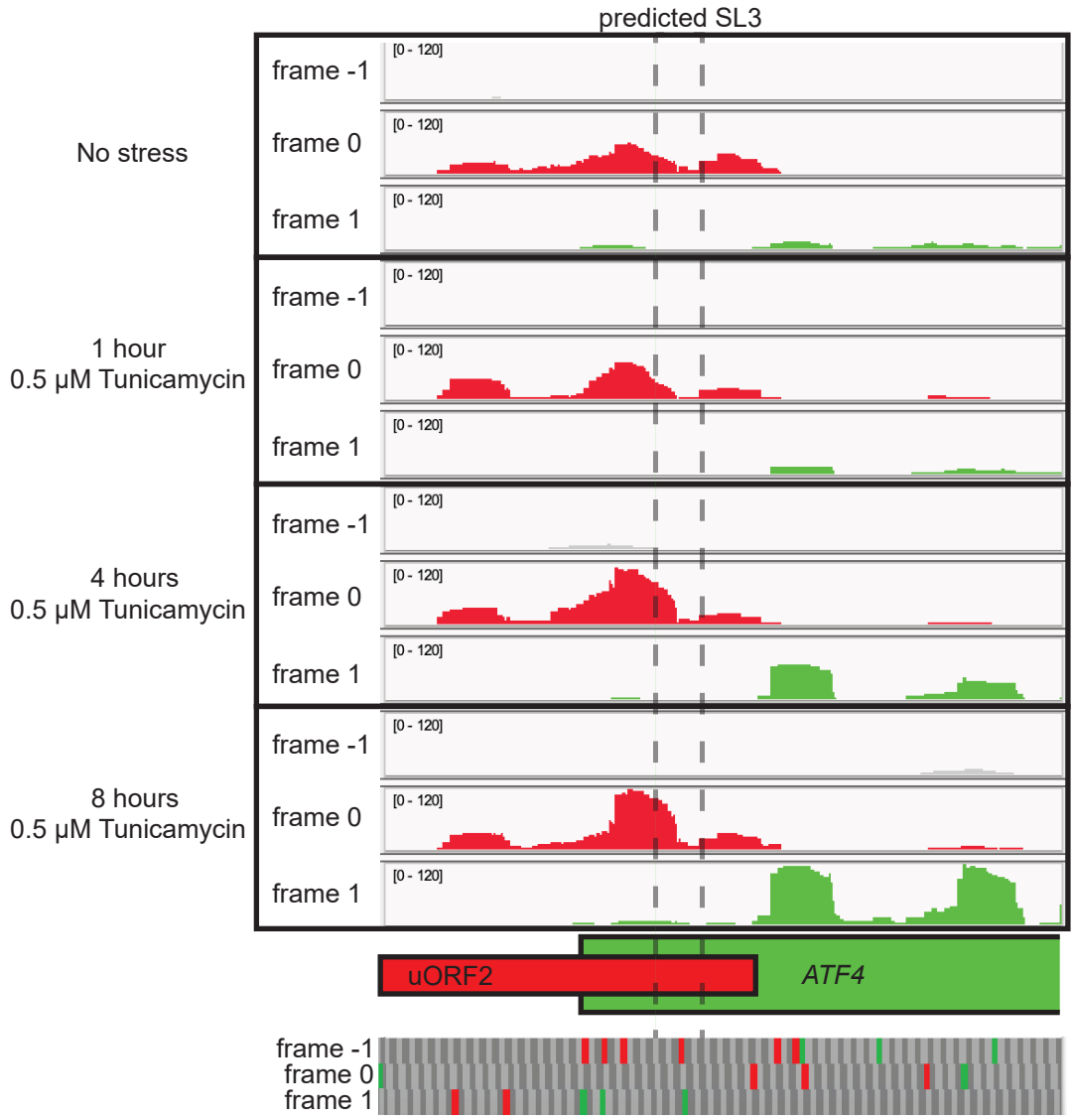


Figure S1. The SL3 and CUG are conserved in evolution; does uORF2 to *ATF4* frameshifting occurs during acute stress?

(A) Multiple Sequence Alignment of 11 mammalian species using MAFFT (Multiple Alignment using Fast Fourier Transform) is shown for the 5′ UTR of *ATF4* mRNA beginning a few nucleotides upstream of the AUG of uORF2 (in red) and ending a few nucleotides past the stop codon (in brown) of uORF2. The *ATF4*'s ATGs are shown in bold, CTG in green, SL3 in orange and A235 and A326 in blue. Nucleotides differing from the human *ATF4* mRNA sequences are underlined. The graph showing Phylogenetic Tree-based scores, that reflect sequence differences with the human *ATF4* gene.

(B) Ribosome footprints from HeLa cells undergoing acute ER stress (Rendleman et al., 2018) were mapped to the *ATF4* exon 2 based on the reading frame engaged. Reading frame assignment was determined by the +12 position in reads 28-30 bp long, corresponding to the P-site of ribosomes. ER stress was induced by 0.5 μM Tunicamycin and cells were harvested at 0, 1, 4, and 8 hours post-treatment, as described in (Rendleman et al., 2018). Predicted stem-loop 3 (SL3) start and end positions are indicated by the dashed line. Below, canonical start and stop codons within frames -1, 0, and 1 are indicated in green and red, respectively.

Smirnova Fig. S2

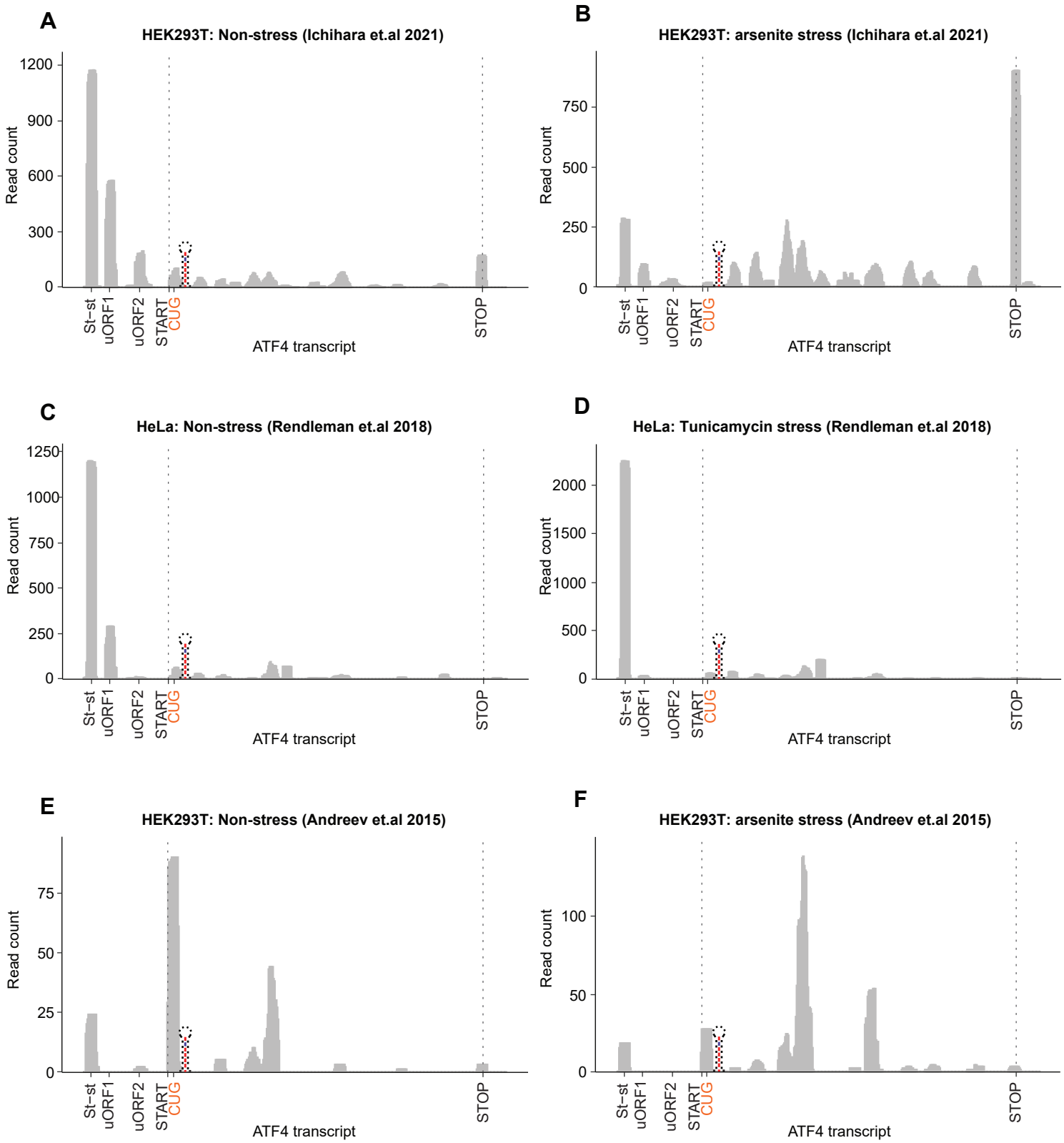


Figure S2. SL3 stalls the ribosome in the uORF2/*ATF4* overlap region. The profiles show ribosome foot-print coverage on the *ATF4* mRNA generated by RiboWaltz (Lauria et al., 2018) package using the datasets obtained from three different studies examining HEK293T cells treated with DMSO (A and E) or Sodium arsenite (B and F) or HeLa cells treated with tunicamycin for 0 (C) or 4 h (D) (Andreev et al., 2015; Ichihara et al., 2021; Rendleman et al., 2018). The *ATF4* transcript with its *cis*-acting features indicated is plotted on the X-axis. Raw read count is plotted on the Y-axis. The dotted lines indicate the 1st nucleotide of canonical start and stop codons of the *ATF4* ORF; the near cognate "CUG" codon is highlighted in red. The schematic stem loop indicates the location of the predicted RNA secondary structure designed SL3.

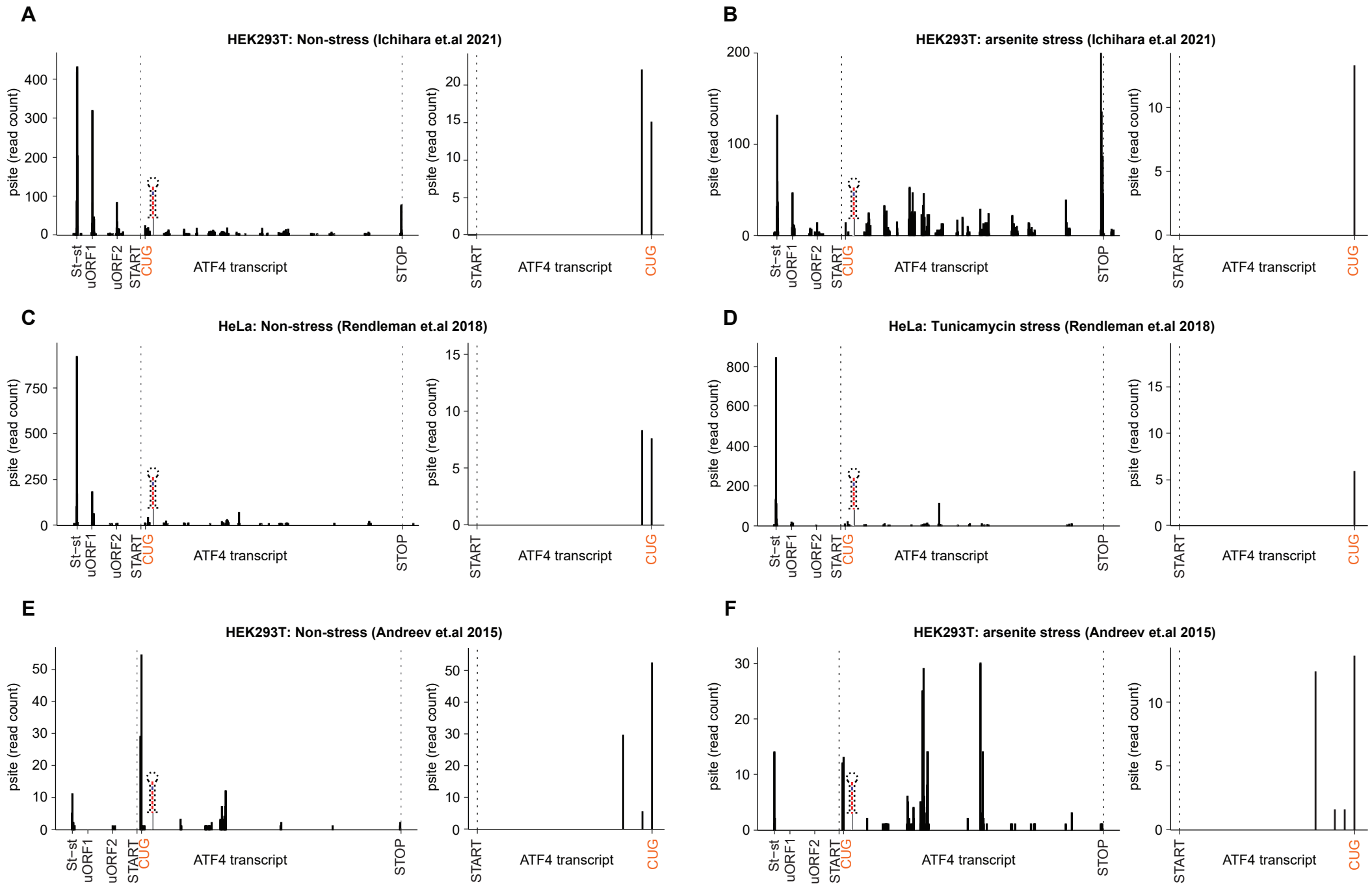

A
HEK293T: Non-stress (Ichihara et.al 2021)
B
HEK293T: arsenite stress (Ichihara et.al 2021)
C
HeLa: Non-stress (Rendleman et.al 2018)
D
HeLa: Tunicamycin stress (Rendleman et.al 2018)
E
HEK293T: Non-stress (Andreev et.al 2015)
F
HEK293T: arsenite stress (Andreev et.al 2015)
psite (read count)
St−st
uORF1
uORF2
START
CUG
ATF4 transcript
STOP

Figure S3. SL3 may promote placement of the near-cognate "CUG" in the ribosomal P-site. The profiles display the P-site occupancy of Ribosome Protected Fragments (RPFs) on the *ATF4* mRNA generated by RiboWaltz (Lauria et al., 2018) package using the datasets obtained from three different studies examining HEK293T cells treated with DMSO (A and E) or Sodium arsenite (B and F) or HeLa cells treated with tunicamycin for 0 (C) or 4 h (D) (Andreev et al., 2015; Ichihara et al., 2021; Rendleman et al., 2018). Corresponding Zoomed-in views of the P-site occupancy at the near-cognate "CUG" codon are shown to the left of main panels. The *ATF4* transcript with its *cis*-acting features indicated is plotted on the X-axis. Raw read count is plotted on the Y-axis. The dotted lines indicate the 1st nucleotide of canonical start and stop codons of the *ATF4* ORF; the near cognate "CUG" codon is highlighted in red. The schematic stem loop indicates the location of the predicted RNA secondary structure designed SL3.

Smirnova Fig. S4

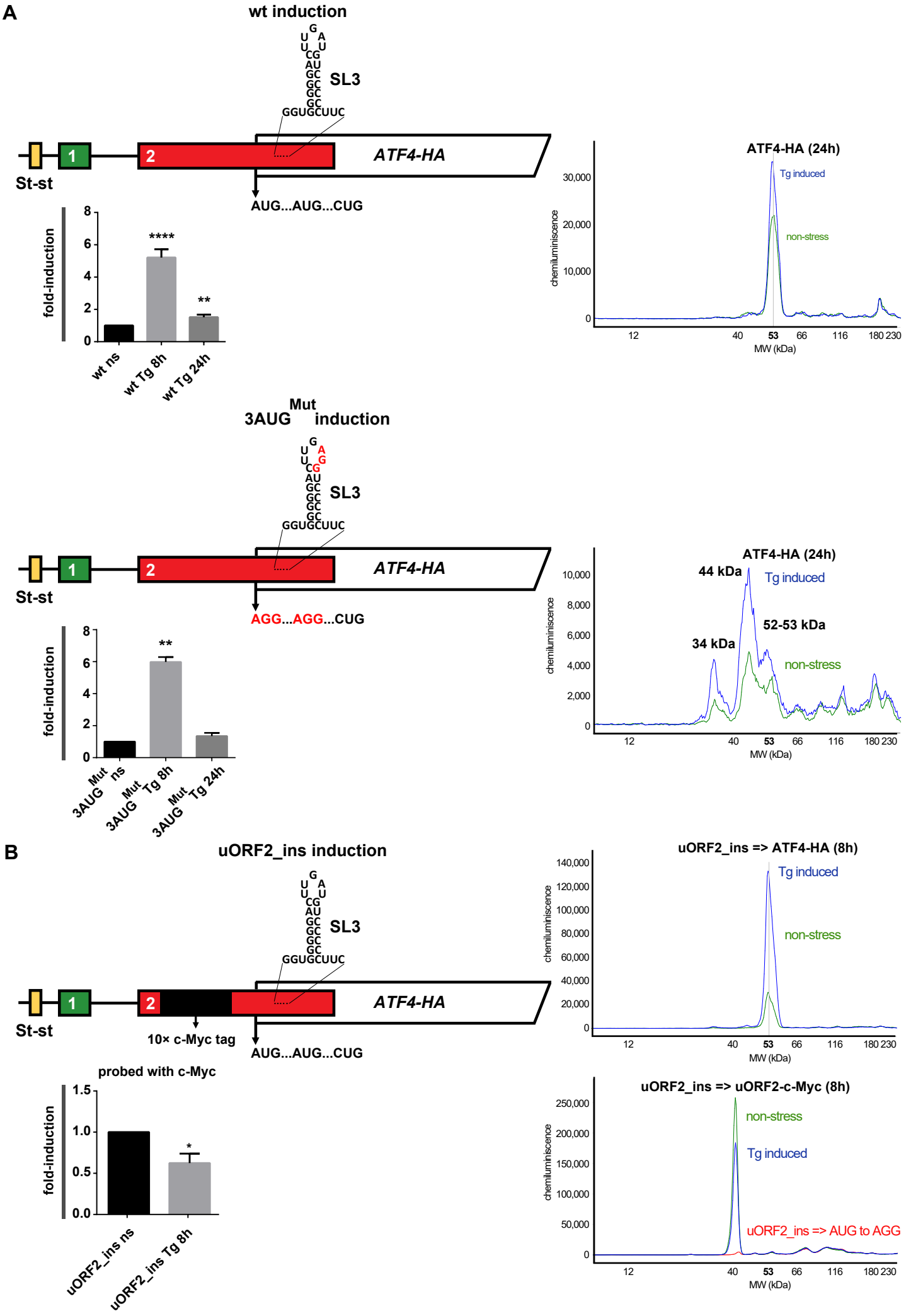


Figure S4. The *ATF4* mRNA expression is not subject to frameshifting; uORF2 is translated even under stress.

(A) Same as Figure 1E for better comparison except that the 3AUG^{Mut} construct (bottom panel; depicted at its the top) was also subjected to JESS analyses. The electropherograms of the wt (top panel) and the 3AUG^{Mut} mutant construct (bottom panel) under Tg stress (in blue) compared to non-stress conditions (in green) probed with anti-HA antibodies are shown. For details, see the main text.

(B) Same as in Figure 2A except that the 10x c-Myc tag insertion in-frame with uORF2, depicted at the top of the panel, was subjected to JESS analyses. The level of uORF2 induction under stress determined by anti-c-Myc antibodies is plotted. The electropherograms of the construct bearing the 10x c-Myc tag insertion in-frame with uORF2 under Tg stress (in blue) compared to non-stress conditions (in green) probed with the anti-HA (top panel) and anti-c-Myc (bottom panel) antibodies are shown. The electropherogram of the control construct bearing the 10x c-Myc tag insertion in-frame with uORF2, the AUG of which was mutated to AGG, is shown in red. For details, see the main text.

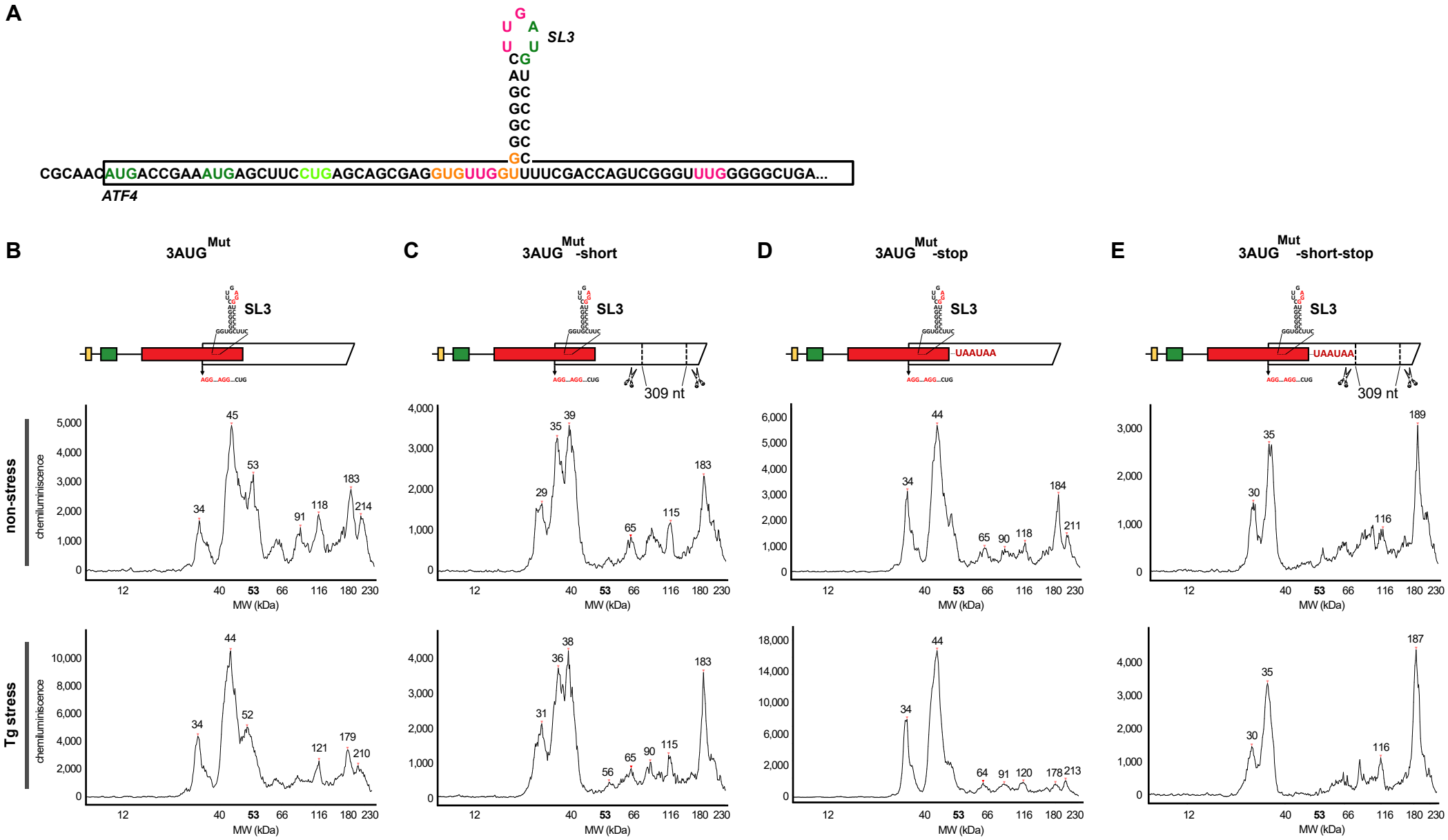

A
SL3
CGCAACAUGACCGAAAUGAGCUUCCUGAGCAGCGAGGUGUUGGUUUUCGACCAGUCGGGUUUGGGGGCUGA...
ATF4
B
3AUG Mut
C
3AUG Mut -short
D
3AUG Mut -stop
E
3AUG Mut -short-stop
SL3
GGUGCUUC
AGG...AGG...CUG
UAAUAA
309 nt
non-stress
Tg stress
chemiluminiscence
MW (kDa)

Figure S5. The *ATF4* mRNA expression is not subject to frameshifting; the canonical AUG1 translation start site of ATF4 is substantial leaky scanned.

(A) Schematic representation of the *ATF4* sequence showing the annotated AUG translation start site (TSS) of the ATF4 full-length protein, two alternative canonical TSSs, near-cognate CUG and five other near-cognate codons in uORF2/*ATF4* overlap.

(B – E) The electropherograms of the full length (B) or C-terminally shortened (C) *ATF4* constructs bearing $3AUG^{Mut}$ without (B – C) or with 2 consecutive UAA stops inserted immediately downstream of the uORF2/*ATF4* overlap (D – E), depicted at the top of the corresponding panels, are shown.

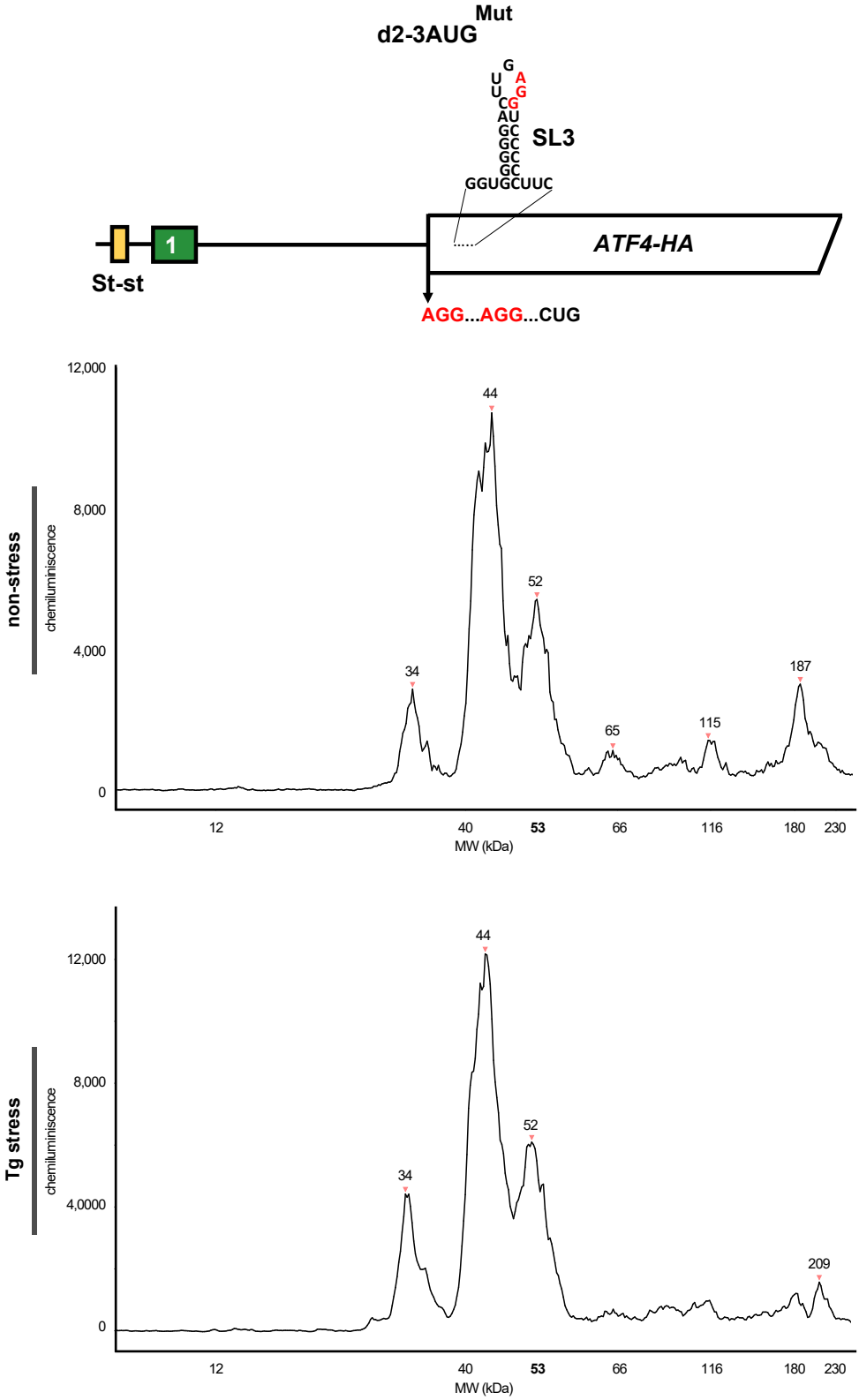
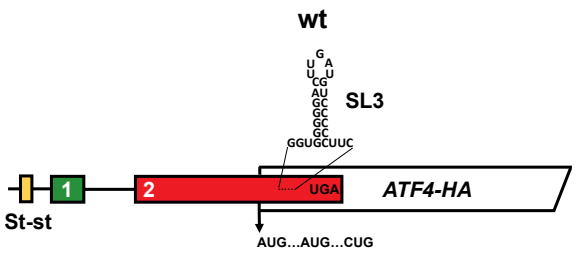
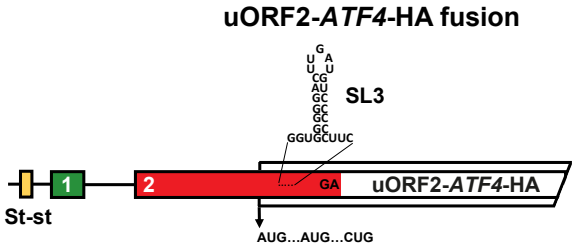
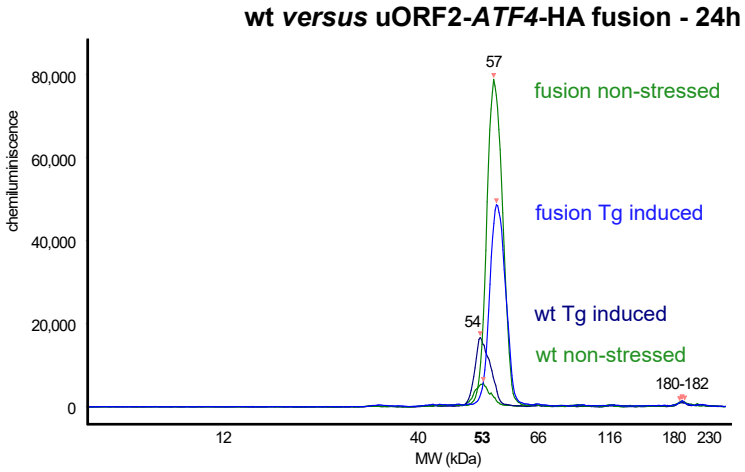

Mut
d2-3AUG
G
U A
U G
CG
AU
GC
GC
GC
GC
SL3
GGUGCUUC
1
St-st
ATF4-HA
AGG...AGG...CUG
non-stress
chemiluminiscence
12,000
8,000
4,000
0
44
52
34
65
115
187
12
40
53
66
116
180
230
MW (kDa)
Tg stress
chemiluminiscence
12,000
8,000
4,0000
0
44
52
34
209
12
40
53
66
116
180
230
MW (kDa)

Figure S6. Combining 3AUG^{Mut} with the AUG to AGG mutation of uORF2 shows no impact on the size and distribution of the ATF4 protein variants expressed with 3AUG^{Mut} alone.

The electropherogram of the 3AUG^{Mut} mutation combined with the AUG to AGG mutation of uORF2 in the same *ATF4* construct, depicted at the top, is shown.


wt
SL3
GGUGCUUC
1
2
UGA
ATF4-HA
St-st
AUG...AUG...CUG


uORF2-ATF4-HA fusion
SL3
GGUGCUUC
1
2
GA
uORF2-ATF4-HA
St-st
AUG...AUG...CUG


wt versus uORF2-ATF4-HA fusion - 24h
chemiluminiscence
80,000
60,000
40,000
20,000
0
57
54
180-182
fusion non-stressed
fusion Tg induced
wt Tg induced
wt non-stressed
12
40
53
66
116
180
230
MW (kDa)

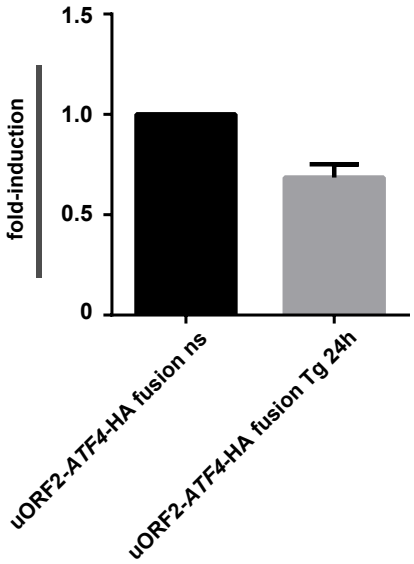

fold-induction
1.5
1.0
0.5
0
uORF2-ATF4-HA fusion ns
uORF2-ATF4-HA fusion Tg 24h

Figure S7. The artificially created uORF2-*ATF4*-HA fusion product further confirms sustained uORF2 expression even under stress conditions.

Schematics at the top depict wt and engineered uORF2-*ATF4*-HA fusion constructs with the electropherograms indicating their expression under non-stress (shades of green) *versus* stress (shades of blue) conditions. Fold-induction values of the fusion construct were plotted (bottom panel) further confirming uORF2 expression even under stress conditions.

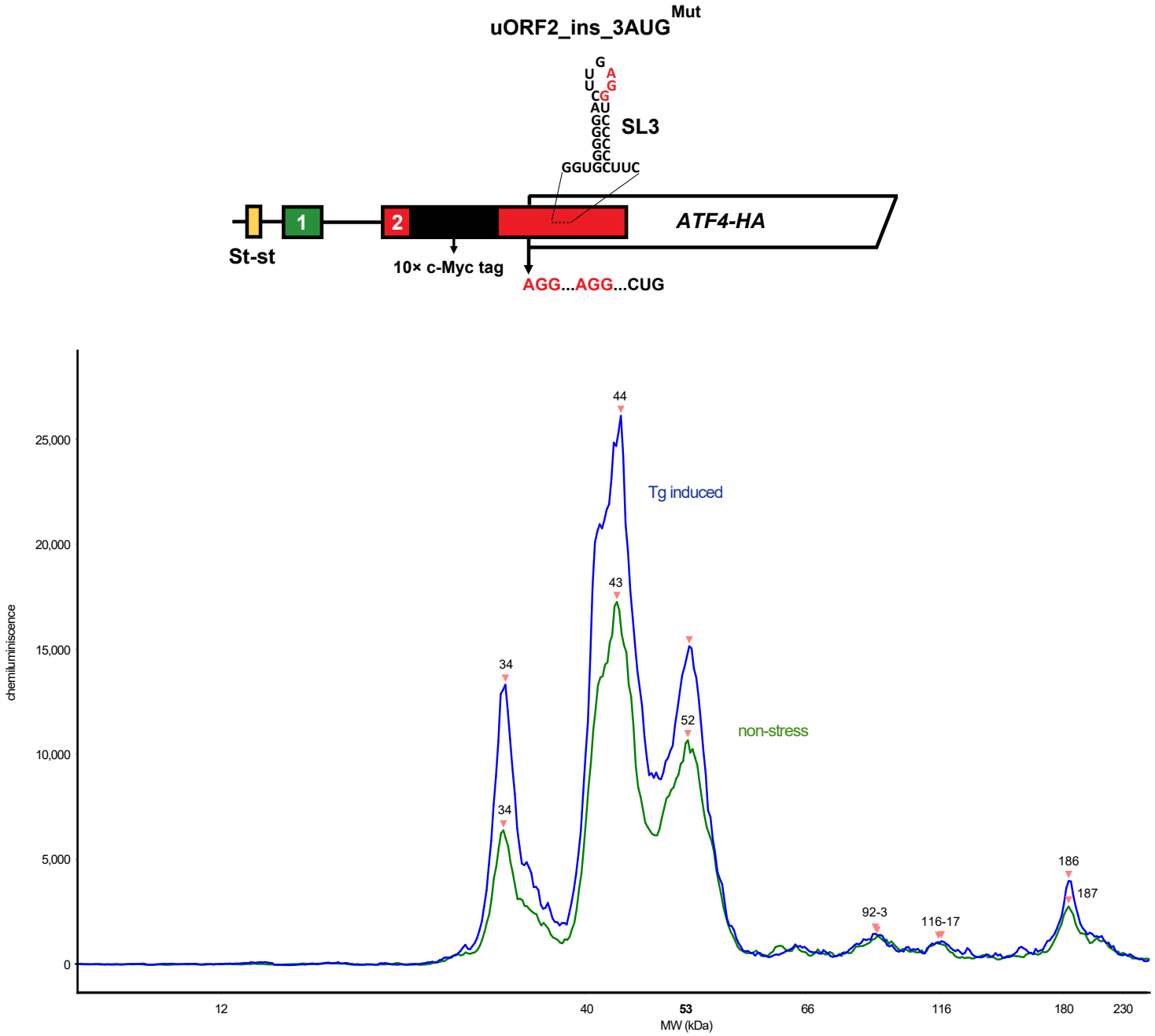

uORF2_ins_3AUG Mut
G
U A
U G
CG
AU
GC
GC
GC
GC
GGUGCUUC
SL3
1
2
ATF4-HA
St-st
10× c-Myc tag
AGG...AGG...CUG
chemiluminiscence
25,000
20,000
15,000
10,000
5,000
0
44
43
34
34
52
Tg induced
non-stress
92-3
116-17
186
187
12
40
53
66
116
180
230
MW (kDa)

Figure S8. Further evidence that the *ATF4* mRNA expression is not subject to frameshifting and that uORF2 is expressed even under stress conditions.
The schematic at the top depicts the *ATF4*-HA mutant construct where 3AUG^{Mut} was combined with 10x c-Myc tag insertion in-frame with uORF2 with the electropherograms indicating their expression under non-stress (green) *versus* stress (blue) conditions.

Smirnova Fig. S9

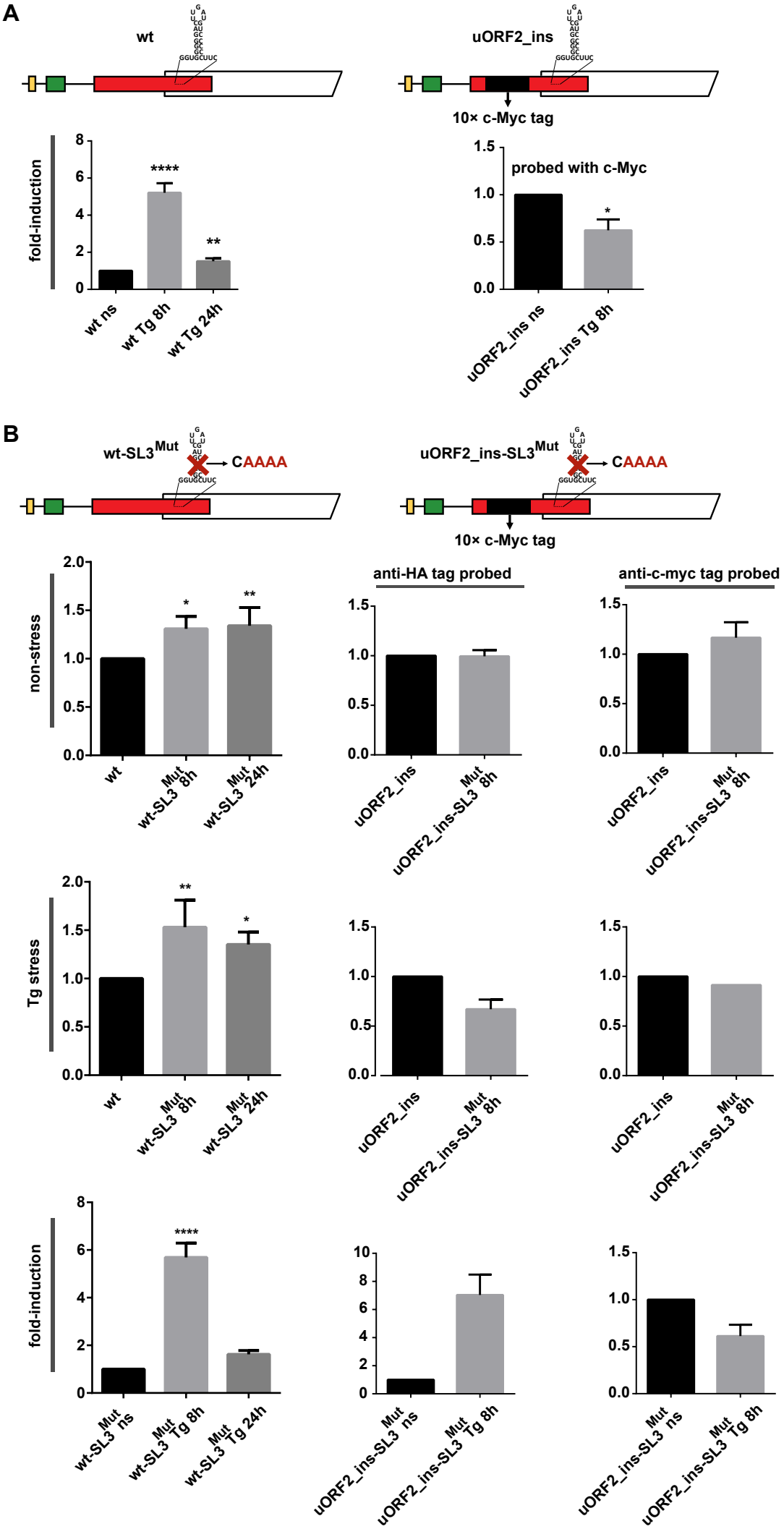


Figure S9. Precise placement of SL3 inside of uORF2 with an exactly defined length expands the existing delayed REI model to a more complex model that includes ribosome queuing.

(A) Same as in Figure S4A (top panel) and B for better comparison.

(B) Same as in Figure 5B except that the 10x c-Myc tag insertion in-frame with uORF2 was combined with the SL3 mutation, as depicted at the top of the right panel, and subjected to JESS analyses together with the SL3 mutation. Probing with both anti-HA (middle plots) and anti-c-Myc (right plots) antibodies are shown.

```
                                                St-st
human_ATF4variant2          -ATTTCTACTTTGCCCGCCCACAGATGTAGTTTTCTCTGCGCGTGTGCGTTTTCCCTCCT     59
human_ATF4pseudogene2       --TTTCTGCTCCGCCCGCCCACAGATGTAGTTTCCTCCGCACGTGCGCACCTTCCCTCCT     58
human_ATF4pseudogene1       GTTTTCTACTTTGCCCGCCCACAGATGTAGTTTTCTCTGCGCGTGTGCGTTTTCCCTCCT     60
                              ***** **  ******************** *** ** **** **   *********

                                                        uORF1
human_ATF4variant2          CCCCGCCCTCAGGGTCCACGGCCACCATGGCGTATTAGGGGCAGCAGTGCCTGCGGCAGC     119
human_ATF4pseudogene2       CCCCGCCTGCAGGGTCCATGGCCTCCATGGCGTTTTAGGGGCAGCAGTGCCTGTGGCAGC     118
human_ATF4pseudogene1       CCCCGCCCTCAGGGTCCACGGCCACCATGGCGTATTAGGGGCAGCAGTGCCTGCGGCAGC     120
                            *******  ********* **** ********* ******************* ******

human_ATF4variant2          ATTGGCCTTTGCAGCGGCGGCAGCAGCACCAGGCTCTGCAGCGGCAACCCCCAGCGGCTT     179
human_ATF4pseudogene2       CTTGGCCTTTGCAGCGGTGGCAGCAGCACCAGGCTCTGCAGTGGCATCCACCGGCGGCTT     178
human_ATF4pseudogene1       CTTGGCCTTTGCAGCGGTGGCAGCAGCACCAGGCTCTGCAGTGGCATCCACCGGCGGCTT     180
                             **************** ********************** **** ** ** *******

                                uORF2
human_ATF4variant2          AAGCCATGGCGCTTCTCACGGCATTCAGCAGCAGCGTTGCTGTAACCGACAAAGACACCT     239
human_ATF4pseudogene2       AAGCCGTAGCACTTCTCACAGCATTCAGCAGCAGCATTGCTGTAACCGACAAAGACACCT     238
human_ATF4pseudogene1       AAGCCATAGCACTTCTCACAGCATTCAGCAGCAGCATTGCTGTAACCGACAAAGACACCT     240
                            ***** * ** ******** *************** ************************
                                                                    ATF4
human_ATF4variant2          TCGAATTAAGCACATTCCTCGATTCCAGCAAAGCACCGCAACATGACCGAAATGAGCTTC     299
human_ATF4pseudogene2       TCGAATTAAGCACATTCGTTGATTCCAGCAAAGCACTGCACCATGACCGAGATGAGCTTC     298
human_ATF4pseudogene1       TCGAATTAAGCACATTCGTTGATTCCAGCAAAGCACTGCACCATGACCGAGATGAGCTTC     300
                            ***************** * **************** *** ********* *********

human_ATF4variant2          CTGAGCAGCGAGGTGTTGGTGGGGGACTTGATGTCCCCCTTCGACCAGTCGGGTTTGGGG     359
human_ATF4pseudogene2       CTGAGCAGCGAGGTGTTGGTGGGGG-CTTGATGTCCCCCCTTGACCAGTCAGGTTTGGGG     357
human_ATF4pseudogene1       CTGAGCAGCGAGGTGTTGGTGGGGG-CTTGATGTCCC-CCTTGACCAGTCAGGTTTGGGG     358
                            ************************* *********** * * ******** *********

human_ATF4variant2          GCTGAAGAAAGCCTAGGTCTCTTAGATGATTACCTGGAGGTGGCCAAGCACTTCAAACCT     419
human_ATF4pseudogene2       ACTGAAGAAAGCCTAGGTCTCTTAGATGACTACCTGGAGGTGGCCAAGCACTTCAAACCT     417
human_ATF4pseudogene1       ACTGAAGAAAGCCTAGGTCTCTTAGATGACTACCTGGAGGTGGCCAAGCACTTCAAACCT     418
                             **************************** ******************************

human_ATF4variant2          CATGGGTTCTCCAGCGACAAGGCTAAGGCGGGCTCCTCCGAATGGCTGGCTGTGGATGGG     479
human_ATF4pseudogene2       CATGGGTTCTCCAGCGACAAGGCTAAGGCAGACTCCTCCGAGTGGCTGGCTGTGGATGGG     477
human_ATF4pseudogene1       CATGGGTTCTCCAGCGACAAGGCTAAGGCAGACTCCTCCGAGTGGCTGGCTGTGGATGGG     478
                            ****************************** * ********* ******************

human_ATF4variant2          TTGGTCAGTCCCTCCAACAACAGCAAGGAGGATGCCTTCTCCGGGACAGATTGGATGTTG     539
human_ATF4pseudogene2       TTGGTCAGTGCCTCCAACGATGGCAAGGAGGATGCTTTCTTTGGGACAGATTGGATGTTG     537
human_ATF4pseudogene1       TTGGTCAGTGCCTCCAACGATGGCAAGGAGGATGCTTTCTTTGGGACAGATTGGATGTTG     538
                            ********* ******** *  ************* ****  ******************

human_ATF4variant2          GAGAAAATGGATTTGAAGGAGTTCGACTTGGATGCCCTGTTGGGTATAGATGACCTGGAA     599
human_ATF4pseudogene2       GAGAAAACTGATCTGAAGGAGTTCGACTTTGATGCCCTGTTGGGTATAGATGACCTGGAA     597
human_ATF4pseudogene1       GAGAAAACTGATCTGAAGGAGTTCGACTTTGATGCCCTGTTGGGTATAGATGACCTGGAA     598
                            *******  *** **************** ******************************
```

Figure S10. Comparative analysis of *ATF4* V2 transcript with transcripts of *ATF4* pseudogenes 1 and 2. The loss of uORF2 initiation codons in the transcripts of both pseudogenes is indicated in red and deletion of a potentially modifiable adenine nucleotide is shown in blue. Translation in the *ATF4* V2 transcript under non-stress conditions starts at uORF2 and ends at the standard stop codon of uORF2. Initiation in the transcripts of both pseudogenes starts at one of the indicated *ATF4* ORF start codons (in bold) and ends prematurely at the stop codon two nucleotides downstream of the A deletion. The uORF2 reading frame is highlighted in dark grey, the reading frame in pseudogenes in light grey.

STAR METHODS

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact Leoš Shivaya Valášek (valasekl@biomed.cas.cz).

Materials Availability

Plasmid constructs generated in this study are available upon request.

Data and Code Availability

Nothing to report.

SUBJECT DETAILS

Human cell line

The HEK293T cell line (ATCC, CRL-3216; RRID: CVCL_0063) was used for all experiments in this study; the HeLa cell line (ATCC, CCL-2; RRID: CVCL_0030) was used only for the T3-ligation assay. Cells were grown in Dulbecco's modified Eagle's Medium - high glucose medium supplemented with 10% fetal bovine serum at 37.0 ºC with 5.0% CO_2 concentration. Cells were seeded into 6-well plates to 2.5 ml of medium per well and grown 24 hours prior to transfection at approximately 40-50% confluency.

Bacterial strains

Bacterial strains DH5α and DH10B were used in this study.

METHOD DETAILS

Construction of the ATF4-HA tagged plasmids

To create the ATF4-HA tagged reporter plasmids, the h*ATF4*-wt-HA-Tag-3'UTR (NM_182810.2) ordered as GeneArt String was cloned into the empty Clontech EGFP-N2 high copy number vector, while removing the GFP insert (Key Resource Table). All hATF4-HA tagged constructs contain the CMV promoter, the full 5'UTR of *ATF4*, the full CDS of *ATF4*, and the natural 3'UTR of *ATF4*. The HA tag is placed at the very C-terminus of the *ATF4* CDS just upstream of its stop codon; all these plasmids have the kanamycin resistance. Mutant constructs were generated using either GeneArt String (Key Resource Table) or PCR using 10×ThermoPol Reaction Buffer, Vent DNA

Polymerase and specific primers; for the list of plasmids, details of their cloning, and lists of primers and GeneArt strings please see Supplemental Information (Tables S24 – S26). Plasmid isolation was performed using QIAGEN Plasmid Mini Kit and EconoSpin DNA Spin Columns. QIAGEN QIAquick Gel Extraction Kit and QIAGEN QIAquick PCR Purification Kit were used for insert isolation.

Plasmids delivery, thapsigargin treatment and preparation of whole cell extracts
HEK293T cells pre-grown for 24 hours as described in the Human cell lines section were transfected with ATF4-HA reporter constructs using TurboFect transfection reagent; 2.5 µg of plasmid DNA was used per one well in the 6-well plate in all experiments. Exactly 8 or 24 hours after transfection, cells were treated with either 1 µM thapsigargin for 3 hours or DMSO as a control. Cells were lysed directly on the plate using 200 µl Glo Lysis Buffer. One half of the lysate was saved for RNA isolation followed by a qPCR control step and the other half of the lysate was subjected to Jess SW assay according to the supplier's instructions.

RT-qPCR control step
Total RNA was isolated using RNA blue reagent according to the manufacturer's instructions. Subsequently, a TURBO DNase cleavage step was performed. For cDNA synthesis, 0.5 µg of total RNA was used for each sample using a High-Capacity cDNA Reverse Transcription kit. For qPCR reactions, HOT FIREPol EvaGreen qPCR Mix Plus 5x was mixed with 0.8 µM primers and 10-, 100-, or 1000-fold diluted cDNA and run using the following program: 95 ºC for 15 sec, followed by 43 cycles of 95 ºC for 15 sec, 60 ºC for 20 sec, and 72 ºC for 20 sec. The qPCR primers are listed in Key Resources Table. Note that mRNA levels of all constructs examined in each individual experiment were always monitored by RT-qPCR and only samples that showed a difference of less than 1 cycle compared to the wt construct were considered in our analysis.

Jess Simple Western (SW) assay (ProteinSimple)
Jess SW Assay was used according to the manufacturer's protocol. Proteins are separated by their molecular weight in a capillary pre-aspirated with separation and stacking matrixes. Once the separation is complete, UV light immobilizes the proteins to the capillary wall. After immobilization and clearing of the matrix from the capillary, the immunoprobing process is initiated, first by incubation with the primary antibody, then with the secondary HRP conjugate, and finally with the chemiluminescent substrate. The emitted chemiluminescent light is then recorded by a CCD camera and automatically quantified. Bio-Rad Protein Assay Dye Reagent Concentrate and Protein Normalization Reagents were used to control protein loading. Twenty-four samples were tested in each Jess run. In brief, the lysate was mixed with 1× Sample Buffer in accordance with the amount calculated in SW Sample Calculator for each sample, and subsequently Fluorescent 5× Master Mix containing DTT solution and 10× Sample Buffer was added. Samples were mixed by vortexing, boiled at 95 ºC for 5 min and spun down using a benchtop microcentrifuge. Together with the tested samples,

biotinylated ladder was pipetted into 12-230 kDa Pre-filled Plates from Fluorescence Separation Module. Mouse monoclonal anti-HA tag or anti-c-Myc antibodies were pipetted along with Anti-Mouse Detection Module. The plate was centrifuged for 5 minutes at 686 x g at room temperature and inserted into the Jess machine along with Fluorescent Capillary Cartridges. Data analysis and control of the results obtained by the JESS Protein Simple instrument were performed in Compass for SW software (version 6.2.0).

Sequence analysis of the 5' UTR of human ATF4 mRNA

The V2 mRNA transcript of human *ATF4* was analyzed using freely available bioinformatics tools designed to predict RNA sequence and structural motifs, such as vsfold5, Hotknot, Pknot, pKiss, RECODE, PRFdb, SCRAMP, and BERMP.

Total and poly(A) RNA isolation for T3-ligation assay

Total RNA was isolated using the RNA Blue reagent according to the manufacturer's instructions, 5 ml of the reagent was used *per* one Ø 15 cm dish containing HEK293T or HeLa cells grown to approximately 80% confluency. The resulting RNA pellets were resuspended in RNase-free water and the concentration was quantified by NanoDropOne (Thermo Scientific). Subsequently, poly(A) RNA was purified using the Poly(A)Purist™ MAG kit according to the manufacturer's instructions. The poly(A) RNA was stored in 70% ethanol at −80°C and the subsequent steps of precipitation and precipitate resuspension in RNase-free water were finished just before proceeding to the T3-ligation assay.

T3-ligation assay and qPCR

The T3-ligation assay was adopted from (Liu et al., 2018a) and (Zhang et al., 2019). Briefly, each ligation reaction mixture A consisted of 20 nM of corresponding probe SGL, 20 nM of corresponding probe SGR, 1x T3 ligation buffer, and ~200-500 ng of poly(A) RNA. The ligation reaction mixture B consisted of T3 DNA ligase diluted to 10U/µl with ligation buffer. Mixtures A were heated at 85°C for 3 min and then incubated at 35°C for 10 min, and then the ligation reaction mixtures B were added. The final volumes of the ligation reaction mixtures were 10 µl and contained 10U of T3 DNA ligase. These resultant mixtures were then incubated at 25°C for 5, 15 or 60 min and chilled on ice immediately. qPCR was carried out according to vendor's instructions (Solis BioDyne) using Bio-Rad CFX384 Real-Time PCR System. The 10 µl qPCR reactions contained 1x HOTFIREPol EvaGreen qPCR Mix Plus (no ROX), 200 nM of each adapter primer (SGadapterFOR and SGadapterREV) and 1 µl of individual ligation reactions and were run using the following program: 95°C for 15 min followed by 44 cycles of 95°C for 15 sec, 60°C for 20 sec, and 72°C for 20 sec. Melting curves were analyzed between 65 and 90°C. Results were analyzed using Bio-Rad CFX Manager. Cycle differences between thapsigargin-treated and mock-treated samples were calculated for each individually analyzed region and the resultant ratios were then used for the normalization of each of the investigated regions (containing either *ATF4* A_{235} or A_{326}) to both control regions (containing *ATF4* A_{267} and A_{311}).

Co-immunoprecipitation (CoIP) assays

HEK293T cells were pre-grown for 24 hours as described in the Human cell lines section, half of the cultures was transfected with wt-3AUG^{Mut} reporter constructs using TurboFect Transfection Reagent. After 24 hours, all the cells were treated with 1μM thapsigargin for 3 hours and harvested in ice cold PBS. The cell pellet was clarified by centrifugation (185 x g for 3 min at 4 ºC), and then washed with ice-cold PBS three times in amount of 10 ml per cell pellet from two 6-well TPP plates; the same centrifugation step was applied after each washing cycle. The pellet was collected and aliquoted, each aliquot containing $12x10^6$ of cells was stored at – 80 º C.

The aliquots were subsequently used for CoIP. The cell lysates were prepared by sonication as follows: 1 ml per an aliquot of the lysis buffer containing 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 % Igepal ® CA-630, 0.5% SDC and fresh prepared 1 mM PMSF was added to thawed cell pellets. The cells were shortly sonicated and spun down (6,000 x g for 10 min at 4 ºC) to remove cell debris. The ATF-4 (D4B8) Rabbit mAb, used in 1:50 dilution per aliquot, was added to the lysates, which were then incubated for 1 hour at 4 º C with gentle rocking to allow the complexes to form. Next, 100 µl of Protein G Mag Sepharose™ Xtra beads was added to each sample and incubated overnight at 4 ºC with gentle rotation. Magnetic beads were washed twice with 10 bed volume of the lysis buffer; twice with 10 bed volume of high salt buffer containing 50 mM Tris-HCl pH 7.5, 500 mM sodium chloride, 0.1 % Igepal and 0.05 % SDC; and finally, twice with 10 bed volume of low salt buffer containing 50 mM Tris-HCl pH 7.5. Protein complexes were eluted with 20 µl of 6x Laemmli Sample Buffer per one sample. All CoIP samples were separated by SDS-PAGE, with Criterion TGX Precast Gels 4-20 % (BioRad). The ATF4 protein enrichment was verified by western blotting using mouse monoclonal anti-ATF4 (B-3) primary antibodies. The signal was developed using SuperSignal West Femto Maximum Sensitivity Substrate and detected in a G-Box imager (Syngene) using a series of varying exposure times. According to the western blot results, the marked spots were cut from a separate SDS-PAGE blue gel, eluted, and subjected to Mass Spectrometry (MS) analysis.

Protein digestion and LC-MS/MS analysis

CBB-stained protein bands were cut out from the gel, chopped into small pieces and destained using 50 mM Tris-HCl (pH 8.1) in 50% acetonitrile (ACN). After the supernatant removal the gel was partly dried in a SpeedVac concentrator. The proteins were digested overnight at 37 °C using Asp-N (5 ng) or chymotrypsin (100 ng; both Promega) in a buffer containing 25 mM 4-ethylmorpholine acetate (pH 8.2; Sigma-Aldrich) and 5% ACN. CNBr cleavage (20 mg/ml; Sigma-Aldrich) was carried out in 70% TFA overnight. Prior to MS analysis the peptides were desalted on a peptide microtrap (Optimize Technologies, Oregon City, OR) according to the manufacturer instructions.

LC-MS/MS analyses were performed using a Vanquish Neo UHPLC system (Thermo Fisher Scientific) coupled to a timsTOF SCP mass spectrometer (Bruker Daltonics). The peptides were separated at a flow rate of 1.5 µl/min on an analytical

column (PepSep C18, 0.15 × 150 mm,1.5 µm; Bruker Daltonics) heated to 50 °C using 45 min-long linear gradient from a mobile phase A (0.1% formic acid) to a 35% of the mobile phase B (80% ACN/0.1% formic acid). MS acquisition method recorded spectra within the mass range of 100-1700 m/z with ion mobility scanning from 0.6 to 1.6 Vs/cm^2. The method consisted of a TIMS survey scan of 166 ms followed by 10 PASEF MS/MS scans. The total cycle time was 1.9 s. Precursors for data-dependent acquisition were fragmented with an ion mobility-dependent collision energy that increased linearly from 20 to 59 eV.

The raw data were processed using DataAnalysis 5.3 and the generated MGF files were searched against SwissProt 2023_02 subset of human proteins using in-house MASCOT v.2.6 search engine with the following settings: precursor mass tolerance 10 ppm, fragment ion tolerance 0.1 Da, no enzyme (for Asp-N and chymotrypsin), variable oxidation and fixed homoserine lactone modification (CNBr only) of methionine.

Ribosome foot-print coverage and P-site analysis

Riboseq fastq files of untreated or sodium arsenite (GSE17432, GSE55195) or tunicamycin (GSE113171) treated samples from previously published datasets (Andreev et al., 2015; Ichihara et al., 2021; Rendleman et al., 2018) were used for analysis in this study. Reads devoid of rRNA and tRNA sequences were aligned to the human reference genome and transcriptome (GRCh38.p13, annotation release 109) by STAR aligner (2.7.10a) (Dobin et al., 2013) using local alignment with 10% mismatches, and with indels and soft clipping allowed for transcriptomic alignment. Transcriptomic alignments were further filtered to obtain unique alignments using custom programs (Herrmannova et al., 2023). The Unique transcriptomic alignments and MANE project (v. 1.0) annotation were used to generate ribosome foot-print coverage and P-site analysis plots of ATF4 transcript by RiboWaltz package (Lauria et al., 2018).

QUANTIFICATION AND STATISTICAL ANALYSES

Variables were tested for normality using the D'Agostino–Pearson normality test. Based on the normality test, the differences between 2 experimental groups were tested by the Mann–Whitney test. Differences between more than two experimental groups were tested by the Kruskal-Wallis test supplemented by Dunn's multiple comparisons test. Variables are presented as mean ± SD and p values of < 0.05 were considered statistically significant. GraphPad Prism statistical software (ver. 8.1.1, GraphPad Software, San Diego, CA, USA and SciPy (ver. 1.10.1) (Virtanen et al., 2020), Matplotlib (ver. 3.7.0) (Hunter, 2007) libraries in Python (ver. 3.11.3) (Rossum and Drake, 2009) were used for statistical analyses and visualization. The heatmaps showing the differences in the relative expression of mutant *versus* WT constructs were created in R (ver. 4.2.2) using the 'pheatmap' package (ver. 1.0.12) (Kolde and Kolde, 2015).

REFERENCES

Andreev, D.E., O'Connor, P.B., Fahey, C., Kenny, E.M., Terenin, I.M., Dmitriev, S.E., Cormican, P., Morris, D.W., Shatsky, I.N., and Baranov, P.V. (2015). Translation of 5' leaders is pervasive in genes resistant to eIF2 repression. Elife *4*, e03971.
Archer, S.K., Shirokikh, N.E., Beilharz, T.H., and Preiss, T. (2016). Dynamics of ribosome scanning and recycling revealed by translation complex profiling. Nature *535*, 570-574.
Bohlen, J., Fenzl, K., Kramer, G., Bukau, B., and Teleman, A.A. (2020a). Selective 40S Footprinting Reveals Cap-Tethered Ribosome Scanning in Human Cells. Mol Cell *79*, 561-574 e565.
Bohlen, J., Harbrecht, L., Blanco, S., Clemm von Hohenberg, K., Fenzl, K., Kramer, G., Bukau, B., and Teleman, A.A. (2020b). DENR promotes translation reinitiation via ribosome recycling to drive expression of oncogenes including ATF4. Nat Commun *11*, 4676.
Chan, C.P., Kok, K.H., Tang, H.M., Wong, C.M., and Jin, D.Y. (2013). Internal ribosome entry site-mediated translational regulation of ATF4 splice variant in mammalian unfolded protein response. Biochim Biophys Acta *1833*, 2165-2175.
Denoyelle, C., Abou-Rjaily, G., Bezrookove, V., Verhaegen, M., Johnson, T.M., Fullen, D.R., Pointer, J.N., Gruber, S.B., Su, L.D., Nikiforov, M.A.*, et al.* (2006). Anti-oncogenic role of the endoplasmic reticulum differentially activated by mutations in the MAPK pathway. Nat Cell Biol *8*, 1053-1063.
Dey, S., Baird, T.D., Zhou, D., Palam, L.R., Spandau, D.F., and Wek, R.C. (2010). Both transcriptional regulation and translational control of ATF4 are central to the integrated stress response. J Biol Chem *285*, 33165-33174.
Doan, N.T., Paulsen, E.S., Sehgal, P., Moller, J.V., Nissen, P., Denmeade, S.R., Isaacs, J.T., Dionne, C.A., and Christensen, S.B. (2015). Targeting thapsigargin towards tumors. Steroids *97*, 2-7.
Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.
Donnelly, N., Gorman, A.M., Gupta, S., and Samali, A. (2013). The eIF2alpha kinases: their structures and functions. Cell Mol Life Sci *70*, 3493-3511.
Gardner, L.B. (2008). Hypoxic inhibition of nonsense-mediated RNA decay regulates gene expression and the integrated stress response. Mol Cell Biol *28*, 3729-3741.
Genuth, N.R., and Barna, M. (2018). Heterogeneity and specialized functions of translation machinery: from genes to organisms. Nat Rev Genet *19*, 431-452.
Guan, B.J., van Hoef, V., Jobava, R., Elroy-Stein, O., Valasek, L.S., Cargnello, M., Gao, X.H., Krokowski, D., Merrick, W.C., Kimball, S.R.*, et al.* (2017). A Unique ISR Program Determines Cellular Responses to Chronic Stress. Mol Cell *68*, 885-900 e886.
Gunisova, S., Hronova, V., Mohammad, M.P., Hinnebusch, A.G., and Valasek, L.S. (2018). Please do not recycle! Translation reinitiation in microbes and higher eukaryotes. FEMS Microbiol Rev *42*, 165-192.

Gunisova, S., and Valasek, L.S. (2014). Fail-safe mechanism of GCN4 translational control-uORF2 promotes reinitiation by analogous mechanism to uORF1 and thus secures its key role in GCN4 expression. Nucleic Acids Res *42*, 5880-5893.
Harding, H.P., Novoa, I., Zhang, Y., Zeng, H., Wek, R., Schapira, M., and Ron, D. (2000). Regulated translation initiation controls stress-induced gene expression in mammalian cells. Mol Cell *6*, 1099-1108.
Herrmannova, A., Jelinek, J., Pospisilova, K., Kerenyi, F., Vomastek, T., Brabek, J., Mohammad, M.P., Wagner, S., and Valasek, L.S. (2023). Differential expression analysis connects knock-downs of eIF3e and eIF3d with the MAP kinase signaling pathway. Nucleic Acids Res, under re-review.
Hetz, C., Chevet, E., and Harding, H.P. (2013). Targeting the unfolded protein response in disease. Nat Rev Drug Discov *12*, 703-719.
Hinnebusch, A.G. (2005). Translational regulation of GCN4 and the general amino acid control of yeast. Annu Rev Microbiol. *59*, 407-450.
Hinnebusch, A.G. (2014). The scanning mechanism of eukaryotic translation initiation. Annu Rev Biochem *83*, 779-812.
Hronova, V., Mohammad, M.P., Wagner, S., Panek, J., Gunisova, S., Zeman, J., Poncova, K., and Valasek, L.S. (2017). Does eIF3 promote reinitiation after translation of short upstream ORFs also in mammalian cells? RNA biol, doi: 10.1080/15476286.15472017.11353863.
Hunter, J.D. (2007). Matplotlib: A 2D Graphics Environment. Computing in Science & Engineering *9*, 90-95.
Ichihara, K., Matsumoto, A., Nishida, H., Kito, Y., Shimizu, H., Shichino, Y., Iwasaki, S., Imami, K., Ishihama, Y., and Nakayama, K.I. (2021). Combinatorial analysis of translation dynamics reveals eIF2 dependence of translation initiation at near-cognate codons. Nucleic Acids Res *49*, 7298-7317.
Ingolia, N.T. (2010). Genome-wide translational profiling by ribosome footprinting. Methods in enzymology *470*, 119-142.
Ivanov, I.P., Shin, B.S., Loughran, G., Tzani, I., Young-Baird, S.K., Cao, C., Atkins, J.F., and Dever, T.E. (2018). Polyamine Control of Translation Elongation Regulates Start Site Selection on Antizyme Inhibitor mRNA via Ribosome Queuing. Mol Cell *70*, 254-264 e256.
Koditz, J., Nesper, J., Wottawa, M., Stiehl, D.P., Camenisch, G., Franke, C., Myllyharju, J., Wenger, R.H., and Katschinski, D.M. (2007). Oxygen-dependent ATF-4 stability is mediated by the PHD3 oxygen sensor. Blood *110*, 3610-3617.
Kolde, R., and Kolde, M. (2015). R Package ‚pheatmap'. R package 1. web tool *790*.
Lauria, F., Tebaldi, T., Bernabo, P., Groen, E.J.N., Gillingwater, T.H., and Viero, G. (2018). riboWaltz: Optimization of ribosome P-site positioning in ribosome profiling data. PLoS Comput Biol *14*, e1006169.
Lin, Y., Li, F., Huang, L., Polte, C., Duan, H., Fang, J., Sun, L., Xing, X., Tian, G., Cheng, Y.*, et al.* (2020). eIF3 Associates with 80S Ribosomes to Promote Translation Elongation, Mitochondrial Homeostasis, and Muscle Health. Mol Cell *79*, 575-587 e577.

Lindner, P., Christensen, S.B., Nissen, P., Moller, J.V., and Engedal, N. (2020). Cell death induced by the ER stressor thapsigargin involves death receptor 5, a non-autophagic function of MAP1LC3B, and distinct contributions from unfolded protein response components. Cell Commun Signal *18*, 12.
Liu, C., Eliseeva, S.V., Luo, T.Y., Muldoon, P.F., Petoud, S., and Rosi, N.L. (2018a). Near infrared excitation and emission in rare earth MOFs via encapsulation of organic dyes. Chem Sci *9*, 8099-8102.
Liu, W., Yan, J., Zhang, Z., Pian, H., Liu, C., and Li, Z. (2018b). Identification of a selective DNA ligase for accurate recognition and ultrasensitive quantification of N(6)-methyladenosine in RNA at one-nucleotide resolution. Chem Sci *9*, 3354-3359.
Lu, P.D., Harding, H.P., and Ron, D. (2004). Translation reinitiation at alternative open reading frames regulates gene expression in an integrated stress response. J Cell Biol *167*, 27-33.
Marasco, O., Roussel, M.R., and Thakor, N. (2022). Probabilistic models of uORF-mediated ATF4 translation control. Math Biosci *343*, 108762.
Matsuda, D., and Dreher, T.W. (2006). Close spacing of AUG initiation codons confers dicistronic character on a eukaryotic mRNA. RNA *12*, 1338-1349.
Meyer, K.D. (2019). m(6)A-mediated translation regulation. Biochim Biophys Acta Gene Regul Mech *1862*, 301-309.
Meyer, K.D., Patil, D.P., Zhou, J., Zinoviev, A., Skabkin, M.A., Elemento, O., Pestova, T.V., Qian, S.B., and Jaffrey, S.R. (2015). 5 ' UTR m(6)A Promotes Cap-Independent Translation. Cell *163*, 999-1010.
Mohammad, M.P., Munzarova Pondelickova, V., Zeman, J., Gunisova, S., and Valasek, L.S. (2017). In vivo evidence that eIF3 stays bound to ribosomes elongating and terminating on short upstream ORFs to promote reinitiation. Nucleic Acids Res *45*, 2658-2674.
Mohammad, M.P., Smirnova, A., Gunisova, S., and Valasek, L.S. (2021). eIF4G is retained on ribosomes elongating and terminating on short upstream ORFs to control reinitiation in yeast. Nucleic Acids Res *49*, 8743-8756.
Neill, G., and Masson, G.R. (2023). A stay of execution: ATF4 regulation and potential outcomes for the integrated stress response. Front Mol Neurosci *16*, 1112253.
Novoa, I., Zhang, Y., Zeng, H., Jungreis, R., Harding, H.P., and Ron, D. (2003). Stress-induced gene expression requires programmed recovery from translational repression. EMBO J *22*, 1180-1187.
Pakos-Zebrucka, K., Koryga, I., Mnich, K., Ljujic, M., Samali, A., and Gorman, A.M. (2016). The integrated stress response. EMBO Rep *17*, 1374-1395.
Penn, W.D., Harrington, H.R., Schlebach, J.P., and Mukhopadhyay, S. (2020). Regulators of Viral Frameshifting: More Than RNA Influences Translation Events. Annu Rev Virol *7*, 219-238.
Pitale, P.M., Gorbatyuk, O., and Gorbatyuk, M. (2017). Neurodegeneration: Keeping ATF4 on a Tight Leash. Front Cell Neurosci *11*, 410.
Rendleman, J., Cheng, Z., Maity, S., Kastelic, N., Munschauer, M., Allgoewer, K., Teo, G., Zhang, Y.B.M., Lei, A., Parker, B.*, et al.* (2018). New insights into the cellular temporal response to proteostatic stress. Elife *7*.

Rendleman, J., Mohammad, M.P., Pressler, M., Maity, S., Hronová, V., Gao, Z., Herrmannová, A., Lei, A., Allgoewer, K., Sultanov, D.*, et al.* (2021). Regulatory start-stop elements in 5' untranslated regions pervasively modulate translation. bioRxiv, 2021.2007.2026.453809.
Rossum, G.V., and Drake, F.L. (2009). Python 3 Reference Manual (CreateSpace).
Schleich, S., Acevedo, J.M., Clemm von Hohenberg, K., and Teleman, A.A. (2017). Identification of transcripts with short stuORFs as targets for DENR*MCTS1-dependent translation in human cells. Sci Rep *7*, 3722.
Shu, X.E., Mao, Y., Jia, L., and Qian, S.B. (2022). Dynamic eIF3a O-GlcNAcylation controls translation reinitiation during nutrient stress. Nat Chem Biol *18*, 134-141.
Sidrauski, C., McGeachy, A.M., Ingolia, N.T., and Walter, P. (2015). The small molecule ISRIB reverses the effects of eIF2alpha phosphorylation on translation and stress granule assembly. Elife *4*, e05033.
Starck, S.R., Tsai, J.C., Chen, K., Shodiya, M., Wang, L., Yahiro, K., Martins-Green, M., Shastri, N., and Walter, P. (2016). Translation from the 5' untranslated region shapes the integrated stress response. Science *351*, aad3867.
Takahashi, M., Okamoto, Y., Kato, Y., Shirahama, H., Tsukahara, S., Sugimoto, Y., and Tomida, A. (2023). Activating mutations in EGFR and PI3K promote ATF4 induction for NSCLC cell survival during amino acid deprivation. Heliyon *9*, e14799.
Valášek, L.S. (2012). 'Ribozoomin' – Translation Initiation from the Perspective of the Ribosome-bound Eukaryotic Initiation Factors (eIFs). Curr Protein Pept Sci. *13*, 305-330.
Vattem, K.M., and Wek, R.C. (2004). Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. Proc Natl Acad Sci U S A. *101*, 11269-11274.
Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J.*, et al.* (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods *17*, 261-272.
Wagner, S., Bohlen, J., Herrmannova, A., Jelinek, J., Preiss, T., Valasek, L.S., and Teleman, A.A. (2022). Selective footprinting of 40S and 80S ribosome subpopulations (Sel-TCP-seq) to study translation and its control. Nat Protoc *17*, 2139-2187.
Wagner, S., Herrmannova, A., Hronova, V., Gunisova, S., Sen, N.D., Hannan, R.D., Hinnebusch, A.G., Shirokikh, N.E., Preiss, T., and Valasek, L.S. (2020). Selective Translation Complex Profiling Reveals Staged Initiation and Co-translational Assembly of Initiation Factor Complexes. Mol Cell *79*, 546-560 e547.
Wek, R.C., Jiang, H.Y., and Anthony, T.G. (2006). Coping with stress: eIF2 kinases and translational control. Biochem Soc Trans *34*, 7-11.
Wortel, I.M.N., van der Meer, L.T., Kilberg, M.S., and van Leeuwen, F.N. (2017). Surviving Stress: Modulation of ATF4-Mediated Stress Responses in Normal and Malignant Cells. Trends Endocrinol Metab *28*, 794-806.
Zhang, Z., Chen, L.Q., Zhao, Y.L., Yang, C.G., Roundtree, I.A., Zhang, Z., Ren, J., Xie, W., He, C., and Luo, G.Z. (2019). Single-base mapping of m(6)A by an antibody-independent method. Sci Adv *5*, eaax0250.

Zhou, J., Wan, J., Shu, X.E., Mao, Y., Liu, X.M., Yuan, X., Zhang, X., Hess, M.E., Bruning, J.C., and Qian, S.B. (2018). N(6)-Methyladenosine Guides mRNA Alternative Translation during Integrated Stress Response. Mol Cell *69*, 636-647 e637.