HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2024 April 30.

Published in final edited form as:

Cell Rep. 2024 April 23; 43(4): 113976. doi:10.1016/j.celrep.2024.113976.

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], **Mahabub Pasha Mohammad**[1], **Anna Herrmannová**[1], **Stanislava Gunišová**[1], **Denisa Petráčková**[2], **Petr Halada**[3], **Štěpán Coufal**[4], **Michał Świrski**[5], **Justin Rendleman**[6], **Kristína Jendruchová**[1], **Maria Hatzoglou**[7], **Petra Beznosková**[1], **Christine Vogel**[6,*], **Leoš Shivaya Valášek**[1,8,*]

[1]Laboratory of Regulation of Gene Expression, Institute of Microbiology of the Czech Academy of Sciences, Videnska 1083, 142 20 Prague, 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, 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]Laboratory of Cellular and Molecular Immunology, Institute of Microbiology of the Czech Academy of Sciences, Videnska 1083, 142 20 Prague, Czech Republic

[5]Institute of Genetics and Biotechnology, Faculty of Biology, University of Warsaw, Warsaw, Poland

[6]Department of Biology, New York University, New York, NY, USA

[7]Department of Genetics and Genome Sciences, Case Western Reserve University, Cleveland, OH, USA

[8]Lead contact

SUMMARY

Activating transcription factor 4 (ATF4) is a master transcriptional regulator of the integrated stress response, leading cells toward adaptation or death. ATF4's induction under stress was thought to be due to delayed translation reinitiation, where the reinitiation-permissive upstream open reading frame 1 (uORF1) plays a key role. Accumulating evidence challenging this mechanism as the sole source of ATF4 translation control prompted us to investigate additional

*Correspondence: cvogel@nyu.edu (C.V.), valasekl@biomed.cas.cz (L.S.V.).

AUTHOR CONTRIBUTIONS

L.S.V. conceived and designed the project. A.M.S., V.H., M.P.M., S.G., D.P., P.H., M. Ś.,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 from M.P.M., D.P., P.H., J.R., M. Ś., and M.H.

SUPPLEMENTAL INFORMATION

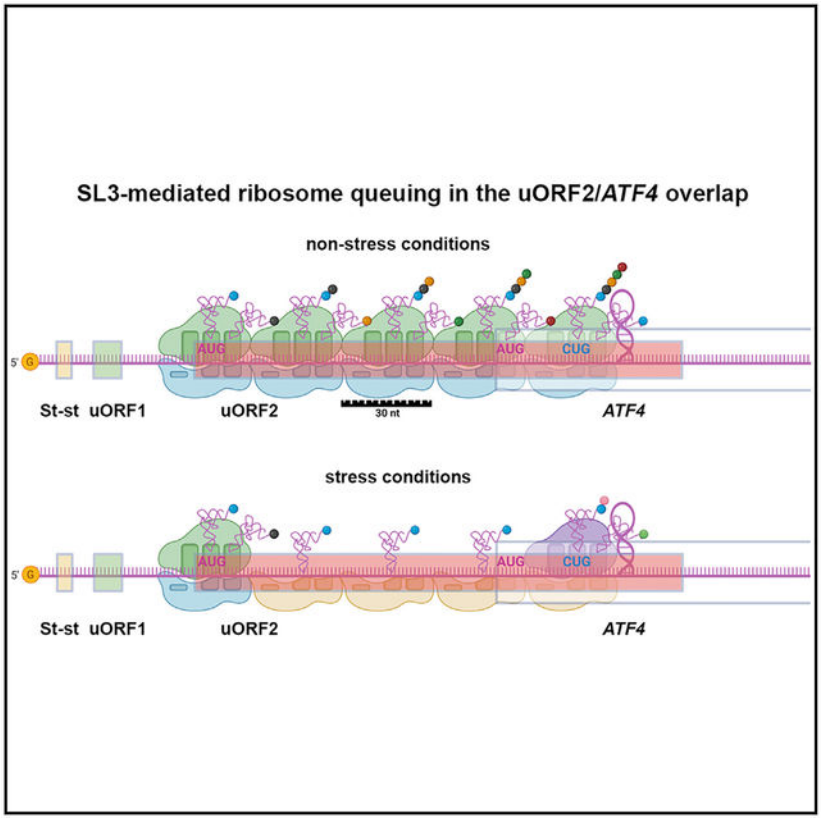
Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2024.113976.

DECLARATION OF INTERESTS

The authors declare no competing interests.

regulatory routes. We identified a highly conserved stem-loop in the uORF2/*ATF4* overlap, immediately preceded by a near-cognate CUG, which introduces another layer of regulation in the form of ribosome queuing. These elements explain how the inhibitory uORF2 can be translated under stress, confirming prior observations but contradicting the original regulatory model. We also identified two highly conserved, potentially modified adenines performing antagonistic roles. Finally, we demonstrated that the canonical ATF4 translation start site is substantially leaky scanned. Thus, ATF4's translational control is more complex than originally described, underpinning its key role in diverse biological processes.

Graphical Abstract



In brief

Traditionally, ATF4 translational control was thought to comprise only delayed reinitiation. Smirnova et al. describe additional layers of ATF4 translational regulation, including substantial leaky scanning of *ATF4*'s AUG1 and ribosome queuing implemented by a stem-loop (SL3). SL3 co-operates with an upstream near-cognate start codon and localizes to the uORF2/*ATF4* overlap.

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).[1,2] Whereas its external triggers are oxygen or nutrient deprivation or viral

infection, the main internal stressor is the accumulation of unfolded proteins in the lumen of the endoplasmic reticulum (ER). Importantly, the ISR can also be induced by activation of oncogenes.[3] 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).[4]

eIF2 assembles into a ternary complex (TC) together with guanosine triphosphate (GTP) and the initiator Met-tRNA$_i^{Met}$.This complex, together with other eIFs, facilitates 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.[5] After the AUG selection process, eIF2 with hydrolyzed guanosine diphosphate (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.[6] 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 (reviewed in Gunišová et al.[7] and Dever et al.[8]).

A handful of specific mRNAs escape this shutdown and become efficiently translated to begin adaptation to the acute stress.[9] One of them encodes the transcription factor ATF4, a master regulator of the ISR.[10,11] Upon relief from the acute stress, eIF2α is dephosphorylated, the ISR ceases, and general protein synthesis resumes.[12] However, if the stress is too intense or persists for too long, then it becomes chronic,[13] the adaptive response capacity is exhausted, and programmed cell death is triggered. Thus, the final outcome of the ISR depends largely on the level and duration of eIF2α phosphorylation and, consequently, on the extent of upregulation of factors escaping the shutdown.[2]

Therefore, the timing and level of ATF4 expression are critical parameters determining cell fate. Even minor disturbances of ATF4 function contribute to serious pathologies, such as the neurodegenerative Parkinson’s, Alzheimer’s, and Huntington’s diseases; prion diseases, and various types of retinal degeneration.[14] Furthermore, persistent overactivation 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.[15]

Under normal conditions, human *ATF4* mRNA is constitutively transcribed at low levels, but the protein is undetectable. Induction of its expression is mediated by the *ATF4* 5′ mRNA leader containing 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 homolog *GCN4*,[7,8,16–18] it has been proposed that the primary route of ATF4 translational control occurs via the so-called delayed translational reinitiation (REI).[10,11] This model is underpinned by the stress-induced decrease in TC levels and the opposing properties of the two *ATF4* uORFs (Figure 1A).[7,8]

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 by the short length of uORF1 and the presence of stimulatory flanking sequences interacting with the initiation

factor eIF3.[19] This interaction allows eIF3 to remain transiently associated with elongating and terminating ribosomes on short uORFs to allow REI.[20–22]

A key aspect of this delayed REI mechanism is that reacquisition of the scanning competence of the post-termination 40S is limited by its binding of a new TC and that the time required for TC acquisition increases significantly with decreasing TC levels.[10,11] The original model posits that, in the absence of stress, the 40S acquires the TC relatively quickly, recognizes the uORF2 AUG, and, due to its overlap with *ATF4*, no ATF4 protein is produced. Conversely, the stress-induced decrease in TC levels delays the TC acquisition by the majority of 40Ss traversing downstream of uORF1. This allows them to skip the AUG of uORF2 and acquire the TC while *en route* to the *ATF4* start codon to induce its synthesis.

Despite the depth of the original studies,[10,11] a growing body of evidence suggests the existence of additional regulatory modes beyond the delayed REI mechanism. For example, three independent studies have reported substantial translation of uORF2 even under stress.[23–25] In addition, leaky scanning over uORF2 has been shown to contribute to ATF4 induction.[23] It has been proposed that an N^6-methyladenosine (m^6A) post-transcriptional modification in the non-overlapping region of uORF2 of mouse *ATF4* mRNA functions as a barrier to ribosomes scanning downstream of uORF1 and that stress triggers demethylation of m^6A, removing this block.[26] Quantitative translation initiation sequencing (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. (Note that the *ATF4* transcript in the mouse is largely conserved with the human transcript [see also Figure S1A) and was examined in the original studies.[10]) Employing the Selective Translation Complex Profiling (Sel-TCP-seq) technique,[27] we showed that the 5′ most proximal and so far the least studied uORF0 (hereafter referred to as Start-stop [St-st] because it consists only of an initiation and termination codon), which precedes uORF1 (Figure 1A), represents an additional barrier that may inhibit ATF4 expression in the absence of stress.[22] All of these findings suggest a more complex regulatory system exceeding the relatively straightforward mode of delayed REI, whose full understanding may help to explain all aspects of the ATF4 contribution to ISR as its master regulator.

Here, we subjected the human *ATF4* mRNA leader to complex bioinformatics analysis and uncovered several previously unknown *cis*-acting features. Following detailed mutational analysis using our newly developed and well-controlled reporter assay revealed that (1) a highly conserved stem-loop in the uORF2/*ATF4* overlap causes ribosome queuing, and stalled or slowed-down ribosomes may initiate at a near-cognate CUG codon upstream of this stem-loop; (2) the inhibitory uORF2 has an unexpected stress-inducible character and is relatively highly translated even under stress, which could be explained by the ribosome queuing; (3) the first two AUGs of ATF4 are substantially leaky scanned; and(4) two prospective, highly conserved adenine modification sites contribute to overall regulation in opposing ways, one of them antagonizing the role of the stem-loop. Overall, our data suggest that 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

Published analyses of ATF4's translational control were carried out using a mouse *ATF4* mRNA reporter (mouse and possum *ATF4* lack St-st, which is otherwise highly conserved among vertebrates; Figure S1A),[28] and luciferase was fused with *ATF4* only two codons downstream of its AUG.[10] The wild type (WT) and specific mutation-carrying reporters were individually co-transfected with a control *Renilla* plasmid into mouse embryonic fibroblast (MEF) cells and luciferase activity was measured. 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,[29] but to *Renilla* activity, which, by definition, drops rapidly upon stress-induced translational shutdown. To resolve these discrepancies, we designed an entirely new ATF4-based reporter, this time using the human sequence, and an experimental workflow that avoided normalizing ATF4 expression levels to any reference genes whose expression would be translationally shut down under stress conditions.

This CMV promoter-driven reporter preserves the entire V2 transcript leader sequence and the entire coding region of *ATF4*, which is extended with a hemagglutinin (HA) tag at its C terminus (Figure 1B). Twenty-four hours after seeding, HEK293T cells were individually transfected with the same amount of a control WT reporter or its mutant derivatives, and after 8 h, one-half of the cultures was treated with thapsigargin (Tg; induces the unfolded protein response (UPR)) and the other half with DMSO (control) (Figure 1C). Therefore, the WT control was used as a reference in every single experiment. Exactly 3 h after Tg treatment, cells were harvested using a standard Glo lysis buffer supplemented with protease inhibitors by the vendor (Promega), and the lysates were subjected to a fully quantitative, automated, and capillary-based high-throughput immunoassay ("western blotting") using the Jess™ Protein Simple instrument, as described in detail in the STAR Methods (Figure 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 highly specific signal detected by the anti-HA tag antibodies is displayed as a single peak at 53 kDa and shows ~5.2-fold induction ($p < 0.0001$) upon Tg treatment 8 h post-transfection (Figure 1E; see Table S1 for average values of fold induction for the WT construct and Table S2 for individual values from each experiment). Tunicamycin-stressed cells (treated the same way only exposed to stress for 4 h) generated nearly 3-fold induction (Figure 1F; Tables S1 and S2). For statistical analysis, see STAR Methods. Please note that we routinely recorded a slightly increased activity of our reporter even under non-stress conditions (Figure 1E, green peak), the source of which is explained in STAR Methods. This feature of our reporter system proved to be very useful throughout the study because it served as a "background" value used to compare the effects of the mutations on both basal ATF4 expression and its inducibility under Tg stress; this would have been impossible with a background value of

zero. For experiments validating that our newly developed reporter system faithfully mimics endogenous *ATF4* regulation, please see STAR Methods and Figures S1B–S1D and S2A.

St-st modestly inhibits ATF4 expression, uORF1 allows for 50% downstream REI, and a solitary, inhibitory uORF2 allows for high ATF4 stress inducibility

We began our analysis by testing the effect of the St-st and retesting the effects of uORFs (on their own) on ATF4 expression by mutating the AUGs of the respective other two elements in the otherwise WT construct (Figures S3A and S3B). 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 ~22.7-fold and ~3.3-fold higher under "non-stress" and "Tg stress" conditions, respectively, compared with the values of the WT construct that were set to 1 for each of these two conditions individually (Figures S3A and S3B; Tables S3, S4). The last plot in the d-all column ("fold induction") with the d-all non-stress value set to 1 indicates that this construct is expectedly not stress inducible, in contrast to the WT (Figures 1E and S3C).

Note that all "fold induction" experiments throughout the study (e.g., row 3 of Figure S3A) were conducted separately from the "non-stress" and "Tg stress" experiments (e.g., rows 1 and 2, Figure S3A). To further clarify our setup, in the "non-stress" and "Tg stress" experiments, we compared individual mutants with the WT separately under the respective conditions, whereas in the "fold induction" experiments, we compared expression of a given mutant under stress with that of the same mutant under no stress (DMSO) (Figure 1G). This means that both sets of experiments were individually controlled for the equal amount of reporter mRNA and protein levels, as described above and in STAR Methods.

Our results, presented in detail in STAR Methods and Figure S3, confirmed major aspects of the original model and supported the idea that St-st acts as a general repressive element (i.e., a roadblock), as observed before.[22,28] Unexpectedly, they also revealed a stress induction capability of the uORF2-only reporter of an unknown mechanism (see below).

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

Next, we carefully screened the *ATF4*'s 5′ UTR and the uORF2/*ATF4* overlap in the coding region for sequence and structure features that might suggest additional modes of regulation (Figure 2A). Their detailed description is provided in STAR Methods. The most notable element, predicted by our analysis, is a stem-loop (SL3), with $\Delta G = -15.40$ kcal/mol, that is highly conserved among vertebrates (Figure S1A) and could be inhibitory. It is located roughly in the middle of the uORF2/*ATF4* overlap (Figure 2A) and has a hairpin structure with the highest free energy predicted for this particular sequence region. Our analysis also predicted a total of four potential sites of m^6A methylation within well-defined motifs (Figure 2A). Interestingly, the *ATF4* gene begins with two nearly consecutive AUGs followed by a third in proximity, all in frame (Figure 2A, 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 (Figure S1A).

Given this knowledge, we created an array of mutants for all of these predicted elements (Figure 2A) and first evaluated them in a large-scale screen. Subsequently, we focused our analysis on mutations in elements that had observable effects, like SL3.

SL3 delays the flow of ribosomes in the uORF2/*ATF4* overlap and genetically interacts with the upstream near-cognate CUG codon

Next, we unfolded SL3 by either quadruple C-to-A (in $SL3^{Mut-1}$) or GGG to AUC (in $SL3^{Mut-2}$) substitutions (Figure 2A) and observed ATF4 expression to increase by ~1.3- and ~1.7-fold under non-stress and by and ~1.5 and 1.8-fold under Tg stress conditions, respectively, over the WT construct set to 1 (Figure 3B, rows 1 and 2; Tables S5 and S6). This means that SL3 mutations add a further ~50%–80% increase (Figure 3B, row 2) to the ~5.2-fold induction of the WT reporter (Figure 3A) to reach the robust ~7.8- to 9.4-fold induction (5.2 multiplied by 1.5 or 1.8) under stress conditions. Figure 1G illustrates how these fold induction changes were calculated. Accordingly, loosening SL3 also increased the inducibility of both of these mutant constructs, measured and compared separately under stress vs. non-stress conditions, from ~5.2-fold to ~5.7- to 5.9-fold (Figure 3A vs. Figure 3B). These results suggest that SL3 forms and acts as physical barrier, delaying scanning/translating ribosomes under both conditions.

Eliminating SL3 in the d2 construct (in d2-$SL3^{Mut-1}$) showed a similar increase over the d2 construct alone as in the case of the wt-$SL3^{Mut-1}$ mutant over the WT (Figure 3B, rows 1 and 2; Tables S5 and S6). Note that the overall induction (d2-$SL3^{Mut-1}$ over the WT) under Tg stress was robustly increased by ~17.5-fold (Figure 3B, row 2). These results strongly suggest that both uORF2 and SL3 act together to reduce ATF4 translation.

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 downstream of the P site (including it) to the mRNA exit site (Figure 2B).[22,27] We noticed a near-cognate CUG codon in weak Kozak context positioned 20 nt upstream of SL3 (Figure 2A), which is highly conserved among vertebrates (Figure S1A). This position is ideal to force a scanning 48S PIC stalled at or slowed by SL3 to initiate at CUG (Figure 2B). To test this hypothesis, we replaced CUG with CUA (in CUG^{Mut}) and observed that the mutation significantly reduced ATF4 expression by ~20% under both conditions tested, which corresponds to a 4.2-fold reduction (5.2 of the WT construct multiplied by 0.8) of “fold induction” under stress (Figure 4B, center, row 2; Tables S7 and S8).

These findings suggest that a small proportion of the ATF4 protein might be several amino acids shorter due to this CUG-initiated synthesis downstream of the main initiation site. Interestingly, combining CUG and SL3 mutations in a single $SL3^{Mut-1}$-CUG^{Mut} construct nullified the opposing effects of the two individual mutations, suggesting a functional interaction between the two elements (Figure 4B, right; Tables S7 and S8). Specifically, when the translation-inhibitory SL3 was removed, the translation-enhancing CUG was no longer needed, perhaps because a flux of ribosomes initiating at the *ATF4* was released

to maximize the synthesis of the full-length ATF4 protein. It remains to be investigated whether the CUG-initiated ATF4 protein is stably produced in WT cells.

To further investigate the functional interplay between SL3 and CUG, we analyzed three publicly available ribosome sequencing (ribo-seq) libraries generated from cells exposed to various ER stressors; namely, two sets of HEK293T cells treated with arsenite[24,30] and HeLa cells subjected to tunicamycin-induced stress.[31] We compared the total number of footprints (FPs) in the *ATF4* mRNA leader under stressed vs. non-stressed cells (Figure S4) as well as the distribution of FPs with ribosomes positioned with their P site to a specific codon (Figure S5). In addition to the pile-ups of FPs on the St-st and AUGs of uORF1 and uORF2 (the latter two are observed in only two datasets; Figures S4A–S4D), all three datasets clearly revealed a pile-up of FPs ~20 nt upstream of SL3, with a visible reduction under stress (Figure S5). 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 (Figure S5). We argue that, even though the FP coverage in the uORF2/*ATF4* overlap upstream of SL3 was relatively low, the repeated occurrence of this prominent “CUG” peak in 3 independent datasets did not occur by chance. Therefore, the existence and functionality of SL3 and its associated near-cognate CUG are a distinct possibility.

Ribosome queuing contributes to the overall translational control of ATF4

The existence of SL3 and its interplay with the upstream CUG intrigued us because it could trigger (1) uORF2-to-*ATF4* frameshifting (STAR Methods) or (2) ribosome queuing; i.e., the phenomenon implicated, for example, in translational control of the antizyme inhibitor mRNA.[32] Specifically, under non-stress conditions, SL3 could form a queue of 80S ribosomes elongating from the AUG of uORF2, reducing its translation rate and, thus, completely eliminating any leaky scanning. This, in turn, would result in even tighter suppression of ATF4 translation under normal conditions. Indeed, based on an 80S ribosome’s average FP length of 30 nt,[33,34] altogether 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 (Figure 2B).

According to this model, under stress, the first 80S ribosome initiating at the AUG of *ATF4* would soon be halted or at least slowed down by SL3, preventing the incoming 48S PIC from moving forward to reach the *ATF4*’s AUG due to a spatial constraint (Figure 2B). 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 the variable FP length of scanning 48S PICs. In any case, this model could explain that, contrary to what the “delayed REI” implies, uORF2 is translated even under stress at substantial levels, which has been observed by us (see below) and others.[23–25] Our model would also imply that, thanks to SL3, CUG could be utilized as an alternative start site for all 40S ribosomes that leaky scan the *ATF4*’s AUG

(Figure 2B), and below we demonstrate that this AUG is indeed rather "leaky." Such an intricate system would ensure well-balanced induction of ATF4, which is highly desirable given its key role.[14]

To test this model, we extended the coding sequence of *ATF4* by inserting a 63 c-Myc tag (30 nt each) just after AUG1 (Figure 5; Tables S9 and S10); i.e., two codons upstream of AUG2 (Figure 5A). If our logic was correct, then 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 because extending the spacing between SL3 and the AUGs of uORF2 and *ATF4* should eliminate SL3's negative effect. This is exactly what we observed (Figure 5B). Further supporting this model, unfolding SL3 in the absence of uORF2 (d2_ins-SL3^{Mut-1}) did not eliminate the stimulatory effect of the 63 c-Myc tag extension (Figures 5C and 5D). These results further underscore the observed interplay between uORF2 and SL3 by showing that the SL3 effect relies largely on initiation at the AUG of uORF2 (see also below).

Importantly, while the c-myc detection revealed only a single peak of 65 kDa generated by the wt_ins construct, as expected given the 63 c-Myc tag insertion, the HA probing detected an additional peak just above, running at the size of the original ATF4 protein (53 kDa; Figure 5E). This peak could only 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.

To further support the ribosome queuing hypothesis, we treated the non-stressed and stressed HEK293T cells (without any reporter) with the cross-linking agent formaldehyde (HCHO) or control (non-crosslinking) cycloheximide, lysed the cells, applied RNase I, isolated total RNA, and subjected the resulting RNA sample to RT-qPCR. For this ribosome protection assay, three sets of primers were used: A1 (for amplicon 1), covering almost the entire putative queuing region (132 nt of a total of ~150 nt, beginning with uORF2's AUG and ending right in front of SL3); A2 of a similar length, covering the region immediately downstream of SL3; and A3 of a similar length, covering the region in the middle of the *ATF4* CDS (Figure 6A). Primer sets A2 and A3 were used independently for normalization and were designed based on the publicly available ribo-seq data used in this study (Figures S6B, S4, and S5) in the regions with detectable but low ribosome coverage. All 3 primer pairs had comparable efficiency ranging from 1.95 to 2.05 and generated fragments of the expected length upon RNase I treatment, as also verified by sequencing (data not shown). We reasoned that, if ribosomes were queued, then they should protect the putative queuing region (A1) from the RNase I digestion more so than the other two control regions (A2 and A3). Accordingly, we found a robust enrichment of A1 over A2 in stressed (~42-fold) and also to a lesser but still very high extent (~11-fold) in non-stressed cells in the RNase I-digested HCHO samples compared with undigested samples (Figure 6B; for raw data, see Data S1). Consistently, the cycloheximide non-cross-linking control, which, by definition, should protect queued 80S ribosomes with lower efficiency, showed a much smaller difference (~3-fold); moreover, regardless of stress (Figure 6C). Importantly, given that only HCHO, but not cycloheximide, can protect 40S-bound mRNAs, the clear difference in fragment protection under stress vs. non-stress conditions observed in RNase I-digested

HCHO-cross-linked samples suggests the existence of a mixed queue of 48S PICs and 80S ribosomes under stress. Collectively, these data further support our model proposing the formation of the 80S queue under non-stress conditions and a mixed queue under stress conditions (Figure 2B), the latter of which seems to be more prevalent (Figure 6B). Note that similar results were also obtained when A1 was compared with A3 (Figures S6A and S6B; Data S1).

Importantly, to demonstrate that it is the SL3 that prompts ribosome queuing, we repeated the ribosome-protection assay, but instead of non-transfected cells, we employed HEK293T cells transiently transfected with plasmids carrying either WT or SL3-mutated ($SL3^{Mut-1}$) *ATF4* reporters. As shown in Figures 6D and S6C (Data S1), the SL3 elimination significantly reduced the difference in fragment protection efficiency of A1 over A2 and A3 under both non-stress (by >2-fold) and stress conditions (by >3-fold). The fact that the fragment protection was not abolished completely by the $SL3^{Mut-1}$ mutation is naturally caused by the presence of the endogenous *ATF4* mRNA with the fully preserved SL3.

Finally, we employed the recently developed RiboCrypt tool (https://ribocrypt.org) to analyze the disome sequencing (disome-seq) data (mRNA fragments protected by two stacked ribosomes) in a study where eIF5A, generally recognized as a ribosome rescue factor at polyproline or proline-proline-glycine stretches,[35,36] was knocked down.[37] We observed by far the most prominent disome peak downstream of the uORF2 start codon (5′ ends of the reads were mapped), followed by a much smaller and less sharp peak about 60 nt downstream and another smaller peak upstream of SL3; i.e., another 60 nt downstream, in WT cells (Figure S6E).

Thus, compared with the rest of the leader sequence, except for the extreme 5′ disome peak of unknown origin, the uORF2/*ATF4* overlap region showed considerable disome coverage, with the two peaks set apart by 60 nt (the expected size of a disome). This fits well with our model predicting a queue of five ribosomes, with the first one in the queue located with its P site on the AUG of uORF2, together covering the region of roughly 150 nt. Interestingly, in the eIF5A knockdown cells, the major peak downstream of uORF2 practically disappeared, as did the extreme 5′ disome peak, while the coverage of the entire stretch of the following 60 nt downstream of AUG of uORF2 increased substantially. Although we cannot explain the loss of the two major peaks, we propose that the increased coverage of this relatively broad region suggests defective collision clearance due to the lack of eIF5A, further supporting our queuing hypothesis.

uORF2 translation under stress is more prevalent than expected

To further investigate the characteristics of uORF2 translation under stress, we next tested for evidence of uORF2-to-*ATF4* frameshifting. As presented in detail in STAR Methods, multiple approaches completely ruled out this possibility but clearly suggested that ATF4 might be translated from several alternative start sites, and all of the resulting variants are induced by uORF2 (Figures S7–S9; Tables S11–S14).

Noteworthy, the final piece of evidence for the absence of frameshifting arose from an insertion of a 10× c-Myc tag into the uORF2 frame exactly 21 nt downstream of its AUG

(Figure S7B, uORF2_ins). Strikingly, using this construct, we also observed that the 10× c-Myc tag insertion in uORF2 behaved similarly as the 6× c-Myc insertion past the AUG1 of *ATF4* with respect to the role of SL3. In contrast to the WT, where mutating inhibitory SL3 increased the ATF4 expression (Figure 4A), unfolding SL3 in the 10× c-Myc tag insertion in uORF2 had no effect (Figure S10A, left column; Tables S15 and S16) under non-stress conditions, the same as in the case of the 6× c-Myc insertion past the AUG1 of *ATF4* (Figure S10B; Tables S17 and S18), and even decreased ATF4 expression under stress. The uORF2 levels (probed with c-myc) seemed to be unchanged with or without SL3 under both conditions (Figure S10A, right column; Tables S15 and S16), as would be expected. Importantly, even with this SL3 mutant and uORF2-extended construct, we reproduced the unexpectedly small reduction (by ~40%) in the uORF2-c=Myc expression level under Tg stress (compare Figure S10A, right column, “fold induction”, and Figure S7B), further strengthening the interpretation that uORF2 is translated at a much higher level even under stress than could be expected based on the original model. Collectively, our data suggests that the exact placement of SL3 within the uORF2 coding sequence with a defined length represents a precise molecular design that expands 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 m^6A, which functions as a barrier, blocking access of ribosomes scanning downstream of uORF1 toward the *ATF4* ORF.[26] (Please note that in the NCBI reference sequence NM_ 001287180.1 used here, mouse A225 corresponds to mouse A258; however, for clarity, we refer to it as it was originally described: A225.) Upon stress, this m^6A is supposedly demethylated, thereby unblocking access. Of the four potential methylation sites we identified computationally, we chose to investigate the effect of human A235 (NCBI reference sequence NM_182810.2), as it seemed to match the mouse A225 best with respect to position (not sequence), and human A326, which is located within SL3 and could thus modify its function.

Both sites, highly conserved among vertebrates (Figure S1A), were substituted with unmodifiable Gs. While A235G significantly increased (by 30% compared with the WT), the basal level of ATF4 expression under non-stress conditions, as originally reported,[26] A326G reduced expression by ~20% under both conditions (Figure 7B, left and center; Table S19; Table S20), indicating that they may have opposing effects. To support these findings, we performed a T3 ligation assay[38,39] and investigated changes in the modification (possibly 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 modification sites like m^6A during the ligation reaction in qPCR reactions.[38,39] We first designed specific L (left) and R (right) probes based on the sequences flanking the two prospective m^6A sites of interest: *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.[38] Because the presence of m^6A

significantly reduces the T3 ligase efficiency, less effective qPCR amplification determines under which conditions (non-stress vs. stressed) the particular A is modified (possibly 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 min revealed that the *ATF4* mRNA region containing the A235 but not the A326 residue had an ~30%–50% reduction in ligation efficiency under non-stressed conditions (Figures 7C; Table S21). This result suggests that modification (possibly methylation) of A235 might indeed contribute to inhibition of ATF4 translation under normal conditions. To the best of our knowledge, modification at A235 has not yet been described for human *ATF4* mRNA. Therefore, it is difficult to predict whether A235 is modified with m^6A, m^6Am, m^1A or something else. Nonetheless, the direct detection of m^6A in the homologous region of the mouse *ATF4* transcript by Zhou et al.[26] with a similar phenotype when mutated strongly suggests that human *ATF4* A235 is also m^6A-modified and serves as a barrier to ribosome progression, like SL3, but only under non-stress conditions.

Although our T3 ligation assay did not confirm that A326 is also modified, the fact that it may stimulate basal ATF4 expression, at least based on our reporter assays (Figure 7B), and that it is situated in the open loop of inhibitory SL3, prompted us to test the effect of the A326G $SL3^{Mut-1}$ and A235G $SL3^{Mut-1}$ (as a control) double mutations. Whereas the latter phenocopied the effect of $SL3^{Mut-1}$ alone under both conditions tested (data not shown), the former double mutation abolished the stimulatory effect of the SL3 elimination on ATF4 expression, the same as the CUG^{Mut}, but only under non-stress conditions (Figure 7B, right; Tables S22 and S23).

These findings indicate that the residue A326 plays a stimulatory role under non-stress conditions and acts in concert with SL3, of which it is a part. Overall, our data suggest the importance of the A326 residue specifically under non-stress conditions, as its mutation to G renders SL3 even more inhibitory. At the sametime, it neutralizes a stimulatory effect of the SL3 elimination. To reconcile these ostensibly contradictory observations, we propose that, under normal conditions, potentially modified A326 destabilizes the inhibitory SL3 to balance its effect but requires SL3 to be unfolded to fully unleash its stimulatory potential. The precise molecular mechanism is unknown, nor is it not known why it does not work the same way under stress conditions.

DISCUSSION

The work presented capitalizes on our newly developed human ATF4-based reporter using an experimental workflow that avoided normalizing ATF4 expression levels to any reference genes whose expression would be translationally shut down under stress conditions.
It extends the original model of delayed reinitiation to control ATF4 translation and

explains some newer evidence that has challenged some of its paradigms. In particular, we uncovered an additional layer of regulation of ATF4 expression implemented by stable SL3 immediately downstream of a substantially leaky AUG1 of *ATF4*, co-operating with a near-cognate start codon (CUG), all inside the region overlapping with uORF2. Detailed analysis of this uORF2/*ATF4* overlap region suggests ribosome queuing upstream of SL3, which may result in initiation at CUG and unexpectedly high translation of uORF2 even under stress, which the original model did not consider. Furthermore, our results confirmed, for human *ATF4*, a translation-inhibitory role of m^6A modification of A235 observed in the mouse[26] and identified a previously unknown, potential modification site (A326) antagonizing the SL3 function under normal conditions. Therefore, we propose that the *ATF4* regulatory "armamentarium" is much more complex than previously thought.

Unexpectedly high uORF2 translation under stress was initially reported by Starck et al.,[23] revealing a high degree of peptide expression from both uORF1 and uORF2 during both normal growth and under stress conditions. The results were obtained from constructs fusing the uORF with a tracer peptide to the respective sequence. The observation was supported by results from ribosome profiling experiments by different groups.[24,25] To explain this discrepancy with the original model, it was suggested that translation of the *ATF4* coding region under stress is a combination of reinitiation and leaky scanning past both uORFs. Given the complexity of the *ATF4* leader, it is conceivable that the tracer peptide employed by Starck et al.[23] could have unintentionally affected the results; e.g., by separating and/or breaking the additional *ATF4* regulatory elements we identified here. However, our reporter system not only confirmed the reported high levels of uORF2 expression under stress conditions, it additionally proposed a mechanistic explanation for this phenomenon based on ribosome queuing ahead of SL3.

Ribosome queuing in the uORF2/ATF4 overlap region

The mechanism of ribosome queuing is known from other systems, such as translation of antizyme inhibitor 1 (AZIN1).[32] The mRNA leader of *AZIN1* contains a relatively long, inhibitory, and non-AUG-initiated upstream conserved coding (uCC) region, with a conserved, polyamine-dependent Pro-Pro-Trp (PPW) motif pausing elongating ribosomes. At low polyamine levels, the majority of scanning ribosomes skip this uCC and translate AZIN1; those that infrequently initiate at its AUU start site will translate the uCC, terminate, and become recycled. High polyamine levels, however, interfere with eIF5A function and cause those very few ribosomes translating the uCC to stall at PPW. Subsequently scanning ribosomes, which again mostly skip the AUU start codon, as well as the occasional elongating ribosomes 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 efficiently suppress AZIN1 synthesis.

The additional layer of ATF4 translation control that we propose is at least in part 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 (Figure 2B), which prevents translation of ATF4 and ultimately also that of uORF2, is

similar to the AZIN1 system at high polyamine levels. This way, the queuing would reinforce uORF2 initiation and very efficiently limit skipping of its AUG. It would also minimize the leakiness of the whole system to the lowest possible level: at zero demand, there is zero ATF4 synthesis. Upon stress, SL3 could slow 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 (Figure 2B).

Interestingly, and somewhat paradoxically, we also found that removal of a potential modification of A326 upon stress would further stabilize SL3. We suggest that the SL3-imposed queue could then (1) enhance selection of alternative start codon(s) in *ATF4* and (2) prevent ATF4 overinduction. This process would involve reduced but still substantial translation of uORF2 even under stress, as has been observed. We further speculate that the near-cognate CUG could serve as another warrant licensing ATF4 synthesis, overcoming the leakiness of the *ATF4* start site(s) (Figure 2B). It seems that the *ATF4* queuing system is well optimized for its complex function, as our data imply that it depends strongly on the precise spacing between SL3 and upstream start codons (either uORF2 or *ATF4*). When we increased this spacing, either through the uORF2 or *ATF4* reading frame, the tight ATF4 control was relaxed.

Importantly, in addition to the reporter analysis, several other observations supported the ribosome queuing hypothesis: (1) the published disome-seq data suggested formation of roughly two disomes (80S couples) in the region spanning the uORF2 AUG and SL3 whose distribution changes upon eIF5A knockdown (Figure S6E); (2) our ribosome protection assays indicated a specific protection of a 132-nt-long fragment in the same region depending on intact SL3 (Figures 6 and S6A–S6C); and (3) analysis of publicly available ribo-seq libraries generated from cells exposed to various ER stressors revealed prominent FP peaks with a strong preference for the CUG placement in the P site (Figures S4 and S5). All of these results are consistent with a queue of 4–5 ribosomal species (either 80S couples or a mixture of 80Ss and 48S PICs) accumulated between uORF2 AUG and SL3.

It is understandable that, given the variable length of mRNA FPs of 48S PICs compared with the relatively well-defined FP length of elongating 80S,[22,27,33,34] it is difficult to predict the exact position of the most 5′ ribosomal species. This is made even more difficult because, as suggested by Ivanov et al.,[32] ribosome queuing forces an initiating ribosome to spend more time near the start codon, where it can migrate back and forth by ~15 nt, as shown previously.[40] Therefore, the operational space for the last 48S PIC in the queue to find the translation initiation site is relatively wide, increasing the probability for initiation even if the start codon is not in an optimal context. Concordantly, altering the spacing between uORF2 and SL3 cancels their functional interaction, indicating that the defined length of the queue for exerting its role is critical.

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.[41] 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 that a 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.

Limitations of the study

One of the limitations of this study is that data obtained with our newly developed reporter system, which we consider to be otherwise robust, are not supported by insertion of key mutations directly into the endogenous *ATF4* locus by, for example, CRISPR-based gene editing. Although they were not feasible given the scale of this study, this will be addressed by the ongoing experiments. Another issue to consider is that, owing to the fact that our reporter mRNAs were highly overexpressed in cells (STAR Methods), we recorded slightly increased ATF4 expression under non-stress conditions even with the WT reporter. This was, however, highly desirable because at least minimal ATF4 expression under non-stress conditions was instrumental for monitoring and comparing effects of our mutations also on the basal level of ATF4 expression. Finally, more experiments are needed to demonstrate that A235 and A326 are indeed modified, possibly by m^6A, in human cells and that ATF4 can be stably expressed from the near-cognate CUG codon upstream of SL3.

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

- The data generated in this paper and used for the preparation of main Figures is available in Tables S1–S23, as well as from the lead contact upon request. Additionally, this paper analyzes existing, publicly available data. These accession numbers for the datasets are listed in the key resources table. Data have been deposited at Mendeley and are publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Human cell line—The HEK293T cell line (ATCC) was used for all experiments in this study; the HeLa cell line (ATCC) was used only for the T3 ligation assay. The source for commercial HEK293T as well as HeLa cell line is reported to be female. Cells were grown in Dulbecco's modified Eagle's Medium (DMEM) high glucose medium supplemented with 10% fetal bovine serum (FBS) at 37.0°C with 5.0% CO_2 concentration. Cells were seeded into 6-well plates (Techno Plastic Products) to 2.5 mL of medium per well and grown 24 h prior to transfection at approximately 40–50% confluency.

Bacterial strains—Bacterial strains DH5α and DH10B (Invitrogen) were used in this study.

METHOD DETAILS

Validation of the newly developed ATF4-HA reporter system—To understand what lies behind the "background" ATF4 expression under non-stress conditions and to validate that our newly developed reporter system faithfully mimics endogenous ATF4 translational regulation, we used traditional western blot analysis. We reasoned that the non-stress peak could originate from the DNA transfection which is known to be stressful to cells[48] and/or from CMV promoter-driven transcription generating a high level of the *ATF4-HA* mRNA, which could magnify the normally negligible leakiness of ATF4 repression under non-stress conditions. As shown in Figure S1B, transfection of our WT reporter using turbofect, but not turbofect or DMEM alone, resulted in a mild ATF4 induction even without Tg treatment (lanes 2 and 8 vs. lanes 3, 4 and 9, 10). This result could suggest that DNA transfections trigger mild stress as reported before, however, control transfections of an empty vector did not trigger such a response (Figure S1C). Therefore, we concluded that a high level of our CMV-driven reporter mRNA (>100 times higher than endogenous *ATF4* mRNA levels, as demonstrated below) resulted in detectable basal levels of ATF4 protein in otherwise unstressed cells, as also observed by others using the CMV-driven ATF4 expression.[49,50] Nonetheless, clear Tg-mediated induction of the endogenous ATF4 in lanes 6 and 7 in the "ATF4 (anti-ATF4)" panel of Figure S1B (compared to lanes 3 and 4), and of our ATF4-HA reporter in lane 5 in the "ATF4-HA (anti-HA)" panel (compared to lane2), demonstrate that both ATF4 variantsrespond to Tg treatment in the same manner. Due to the absence of signal in lanes 3 and 4 we could not quantify the level of induction of the endogenous ATF4 alone. Nonetheless, the level of induction of ATF4-HA calculated from Jess (5.2 ± 0.5) or western blots (3.9 ± 1.3) and the level of induction of combined endogenous ATF4 and ATF4-HA (lanes 5 versus 2 in the "ATF4 (anti-ATF4)" panel showing the cumulative amounts of both HA-tagged and endogenous proteins that anti-ATF4 antibodies recognize) calculated from Jess (5.5 ± 0.6) or western blots (3.4 ± 0.2) correspond nicely to what we expected (Figure S1D). Given the fully quantitative nature of Jess measurements, we consider them to be the most reliable.

Importantly, the observed induction pattern is solely attributable to ATF4 translational control, as the mRNA levels of both *ATF4-HA* and endogenous *ATF4* (measured separately using highly specific reverse primers matching the *ATF4* stop codon region, by which these two alleles differ due to the HA tag sequence) remained virtually unchanged under

non-stress vs. stress conditions (Figure S2A). Collectively, these results document that our reporter mirrors the behavior of endogenous ATF4 regulation.

As aforementioned, when we compared *ATF4* mRNA levels under non-stress conditions using primer sets A1 and A2 downstream of SL3 (Amplicons 2 and 3 defined in the results chapter "Ribosome queuing contributes to the overall translational control of ATF4" and Figure S6), we revealed over 100-fold more *ATF4* mRNA in transfected (expressing both endo and exogenous *ATF4* mRNA) than in non-transfected (expressing endogenous *ATF4* mRNA alone) (Figure S6D; Data S1).

St-st modestly inhibits ATF4 expression, uORF1 allows for 50% downstream REI, and a solitary, inhibitory uORF2 allows for high ATF4 stress-inducibility—Keeping St-st as the only element in the *ATF4* leader resulted in ~20.6-fold increase in ATF4 expression under non-stress conditions and ~3.1-fold increase under Tg stress over WT (Figures S3A and S3B; Tables S3 and S4). The former value is only modestly but reproducibly lower than that of d-all (reduced by ~10%) supporting the idea that St-st acts as a general repressive element, i.e., a roadblock, as observed before.[22,28]

The uORF1-only construct, where only the AUG of uORF1 is preserved, 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 (Figures S3A–S3; Tables S3 and S4). 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 [Figure S3A - row 1] divided by ~23-fold induction of d-all [Figure S3A - row 1]) undergo partial recycling and reinitiate downstream upon reacquisition the TC.

According to the original delayed reinitiation model, uORF2 serves as a barrier capturing most, if of not all, 40S ribosomes that resumed scanning after termination at uORF1 under non-stress conditions, thus preventing ATF4 expression because its sequence extends into the *ATF4* CDS out of frame. Under stress conditions, characterized by low TC levels, its expression should be dramatically reduced, allowing ATF4 expression. In the WT set-up, uORF2 might still be expressed to some degree because not 100% of ribosomes skip its AUG even under decreased TC levels, but the amount of those that do should be substantially higher when compared to non-stress conditions. In any case, in the uORF2-only set-up, this strong barrier should theoretically not allow any significant inducibility of the reporter under stress conditions, because once the ribosome initiates at uORF2 as the first and only upstream uORF, the effect on ATF4 is independent of the TC levels.

Accordingly, we found that the uORF2-only construct showed an ~80% drop in reporter expression compared to the WT under both normal and Tg stress conditions indicating dramatically reduced initiation at *ATF4* (Figures S3A and S3B – rows 1 and 2; Tables S3 and S4). However, in stark contrast to the above logic, comparing stress with non-stress values of uORF2-only ("fold-induction" – row 3) revealed an unexpected ~3.2-fold increase in ATF4 translation under stress that could not be explained by the original model. We argue

that this large increase cannot be explained purely by the proposed increased leaky scanning over uORF2 under stress.[23]

Next, we constructed mutants in which we eliminated one upstream element at a time and compared the resulting constructs with the WT set to 1 (Figures S3B and S3D; Tables S3 and S4). Eliminating St-st (in dSt-st) showed no significant effect under both conditions. Nonetheless, we repeatedly observed higher, but not significant, stress induction of dSt-st compared to WT (increased by ~10% from WT ~5.2-fold–~5.8-fold; Figures S3C vs. S3D – row3; Tables S3 and S4), which argues for the St-st's role as a stress-independent roadblock reducing processivity of the scanning ribosome. These results, and the fact that mouse and possum *ATF4* lacks the St-st, indicate that uORFs 1 and 2 are largely sufficient to ensure stress-induced translation of ATF4.

While removal of the REI-permissive uORF1 (in d1) significantly reduced ATF4 expression by ~70% under both non-stress and Tg stress, eliminating the inhibitory uORF2 (in d2) dramatically increased ATF4 expression (~12.1-fold) under non-stress conditions and slightly (~2.4-fold) under stress (Figures S3B and S3D rows 1 and 2; Tables S3 and S4), all as expected from the original model. The latter result nicely corroborates that ~50% of ribosomes terminating at uORF1 reinitiate downstream (~12-fold induction of d2 over WT [Figure S3D - row 1] divided by ~23-fold induction of d-all [Figure S3A - row 1]), as we observed with the uORF1-only construct (Figure S3A).

Comparison of the stress induction potentials of each construct showed that d2 is very similar to uORF1-only and d1 to uORF2-only (Figures S3D and S3A, "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 Figure S3A – uORF2-only versus Figure S3D – 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 question as to the mechanism behind the unexpected stress induction of the uORF2-only reporter.

Sequence analysis of the 5′ UTR of the human ATF4 mRNA reveals additional elements potentially contributing to ATF4 translational control—Here we asked if there are some so-far unidentified regulatory elements in the *ATF4* mRNA leader besides St-st and both uORFs. We had previously predicted and experimentally validated the existence of two stem loops (SL1 and 2) preceding uORF1 (Figure 2A), the first of which cooperates with eIF3 to unleash the full REI potential of uORF1.[19] SL1 also contains one near-cognate CUG start site (NC1) that is immediately followed by two overlapping GUG near-cognates (NC2/3) that are no longer part of SL1.

Notably, the SL1 and its NC1, and NC2/3 are not universally conserved in eukaryotes (Figure S1A). The St-st is conserved in most mammals examined, except mouse a few others. The SL2 is conserved in primates and, interestingly, its formation is even slightly

more favorable in mouse and rat than in human. Indeed, the presence of uORF1 and uORF2 is well conserved.

The next element, predicted by our analysis, is another stem-loop (SL3), with ΔG = −15.40 kcal/mol, which could potentially be inhibitory and is highly conserved among vertebrates (Figure S1A). It is located roughly in the middle of the uORF2/*ATF4* overlap (Figure 2A) 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 m6A methylation within well-defined motifs (Figure 2A); 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[26]; RRACH (R = G or A; H = A, C or U), DRACH (D = A, G or U). The DRACH2 (A286) motif occurs in the uORF2/*ATF4* overlap and covers 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. Although the DRACH1 and 2 motifs are not universally conserved in eukaryotes, RRACH1 and 2 motifs are highly conserved in mammals (Figure S1A) and could therefore play an important role in ATF4 regulation.

The *ATF4* gene begins with two nearly consecutive AUGs followed by a third in close proximity, all in frame (Figure 2A, 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 (Figure S1A).

The uORF2/*ATF4* overlap also contains two putative conserved sliding sequences (FS-A and -B) and a so-called C-tract motif,[51] 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.[31] 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) (Figure S2B). In fact, ATF4's translation appeared to initiate downstream of the predicted SL3 after 4 h 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.[51] In support of this hypothesis, 8 nucleotides downstream of FS-A occurs the C-nucleotide-rich C-tract motif (CCCCCUUCGACC) with high similarity to the inhibitory C-motif promoting −2 PRF in arteviruses.[2] 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.[51] Yet, neither of its glycine codons had the potential to slow or stop ribosomes, as they belong to the high codon usage category. Since FS-B also did not meet the −2 PRF criteria, it likely has no role in ATF4 translational control.

Given this knowledge, we created an array of mutants for all these predicted elements (Figure 2A), and first evaluated them in a large-scale screen. It resulted in no observable effect for mutations in NC1 to NC3 in SL1, both DRACH motifs, mutations unfolding SL1 or SL2 (specifically with respect to the St-st function), and mutations of AUG3 inside the *ATF4* coding region (data not shown). Therefore, we focused our analysis on mutations in other elements that had observable effects.

uORF2 translation under stress is more prevalent than expected—To investigate characteristics of uORF2 translation further, we next tested for evidence of uORF2 to *ATF4* frameshifting. To do so, we mutated the first 3 AUGs of *ATF4* (Figure S7A, in 3AUG^{Mut}), which would eliminate all ATF4 variants except for the potential frameshifted fusion protein. However, we observed that the peak corresponding to the 53 kDa WT protein did not vanish but shifted to 52 kDa (Figure S7A) and its relative expression level decreased by ~5.5-fold under both conditions when the 3 AUGs were mutated (Figure S7A, bottom electropherogram; Tables S11 and S12. This 52 kDa peak in 3AUG^{Mut} remained inducible under Tg stress, indicating that despite the lack of these three *ATF4* AUGs, the regulatory system remained fully responsive. Further, we observed two additional, clearly discernible peaks at 44 and 34 kDa, which were consistently 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 (Figure S8A), whereas the 44 and 34 peaks corresponded to shorter ATF4 variants initiated on internal AUGs further downstream.

These results clearly suggest that ATF4 might be translated from several alternative start sites, and the variants are induced by 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 supported. Concordantly, 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 (Figure S8B – 3AUG^{Mut} versus Figure S8C – 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 (Figure S8D – 3AUG^{Mut}-stop; Figure S8E – 3AUG^{Mut}-short-stop), strongly suggesting that the initiation codon for the 52-kDa protein lied within the overlap region, but not 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 (Figure S9A, 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 an insertion of a 10x c-Myc tag into the uORF2 frame exactly 21 nt downstream of its AUG (Figure S7B, in uORF2_ins). This mutant resulted in only one ATF4-HA peak corresponding to ~53 kDa, with inducibility unchanged compared to the WT (Figure S7B, upper electropherogram). Note that the expected size of the fusion protein is ~62 kDa but it runs at ~57 kDa, as verified experimentally (Figure S9B; Tables S13 and S14). 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 (Figure S7B; bottom panel, uORF2_ins => AUG to AGG in red) and whose intensity remained unexpectedly high (reduced only by ~40%) even under Tg stress (Figure S7B, bottom electropherogram; see Tables S11 and S12). Consistently, combining uORF2_ins with the triple AUG mutation of *ATF4* produced three Tg-inducible peaks of the same size as observed with the 3AUG^{Mut} alone (Figure S9C [uORF2_ins_3AUG^{Mut}] versus Figure S8B [3AUG^{Mut}]). Finally, mutations of the computer-predicted slippery sequences (Figure 2A) showed no effect (data not shown). Altogether, these findings clearly ruled out possible uORF2 to *ATF4* frameshifting and strongly supported the idea that uORF2 was translated even under stress and at a much higher level than could be expected based on the original model.

Construction of the ATF4-HA tagged plasmids—To create the ATF4-HA tagged reporter plasmids, the hATF4-wt-HA-Tag-3′UTR (NM_182810.2) ordered as GeneArt Strings DNA Fragment (Invitrogen) was cloned into the empty Clontech EGFP-N2 high copy number vector,[42] while removing the GFP insert. 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 Strings (Invitrogen) or PCR using 10×ThermoPol Reaction Buffer (NEB), Vent DNA Polymerase (NEB) and specific primers. Cloning details of all plasmids are described below; for the list of all plasmids, primers and GeneArt strings, please see Data S2–S4. Plasmid isolation was performed using QIAGEN Plasmid Mini Kit and EconoSpin DNA Spin Columns (Epoch Life Science). QIAGEN QIAquick Gel Extraction Kit and QIAGEN QIAquick PCR Purification Kit (QIAGEN) were used for insert isolation.

WT ATF4-HA Tag was created by inserting *Hind*III-*Not*I digested h*ATF4*-wt-HA-Tag (GeneArt String DNA Fragment; Invitrogen) into *Hind*III-*Not*I digested *pCMV-EGFP-N2* high copy vector.

dSt-st was created by inserting the *Hind*III-*Pst*I digested fusion PCR product obtained with primers h*ATF4*-St-st_AGG-F and h*ATF4*-St-st_AGG-R using wt ATF4-HA Tag as a template into *Hind*III-*Pst*I digested wt ATF4-HA Tag.

d1 was created by inserting the *Hind*III-*Pst*I digested fusion PCR product obtained with primers h*ATF4*-uORF1_AGG-F and h*ATF4*-uORF1_AGG-R primers using wt ATF4-HA Tag as a template into *Hind*III-*Pst*I digested wt ATF4-HA Tag.

d2 was created by inserting the *Pst*I-*EcoR*V digested fusion PCR product obtained with primers h*ATF4*-uORF2_AGG and SW *ATF4* d120 SphI R using wt ATF4-HA Tag as a template into *Pst*I-*EcoR*V digested wt ATF4-HA Tag.

d-all was created by inserting *Hind*III-*Not*I digested h*ATF4-d*-all-HA-Tag (GeneArt String DNA Fragment; Invitrogen) into *Hind*III-*Not*I digested *pCMV-EGFP-N2* high copy vector.

St-st-only was created using fusion PCR with wt ATF4-HA Tag serving as a template and the following combination of primers: (i) h*ATF4*-uORF1_AGG-F + h*ATF4*-uORF1_AGG-R and (ii) h*ATF4*-uORF2_AGG and SW *ATF4* d120 SphI R. The resulting PCR product was digested with *Hind*III and *EcoR*V and inserted into *Hind*III-*EcoR*V digested wt ATF4-HA Tag.

uORF1-only was created using fusion PCR with wt ATF4-HA Tag serving as a template and the following combination of primers: (i) h*ATF4*-St-st_AGG-F + h*ATF4*-St-st_AGG-R and (ii) h*ATF4*-uORF2_AGG and SW *ATF4* d120 SphI R. The resulting PCR product was digested with *Hind*III and *EcoR*V and inserted into *Hind*III-*EcoR*V digested wt ATF4-HA Tag.

uORF2-only was created using fusion PCR with wt ATF4-HA Tag serving as a template and the following combination of primers:(i) h*ATF4*-St-st_AGG-F + h*ATF4*-St-st_AGG-R and (ii) h*ATF4*-uORF1_AGG-F + h*ATF4*-uORF1_AGG-R. The resulting PCR product was digested with *Hind*III and *Pst*I and inserted into *Hind*III-*Pst*I digested wt ATF4-HA Tag.

dSt-st-NC1 was created by inserting the *Hind*III-*Pst*I digested PCR product obtained with primers h*ATF4*-dSt-st_NC1 and h*ATF4*-PstI-R using dSt-st as a template into *Hind*III-*Pst*I digested dSt-st.

St-st-only-NC1 was created by inserting the *Hind*III-*Pst*I digested PCR product obtained with primers h*ATF4*-St-st-only_NC1 and h*ATF4*-PstI R using St-st-only as a template into *Hind*III-*Pst*I digested St-st-only.

St-st-only-NC2/3 was created by inserting the *Hind*III-*Pst*I digested PCR product obtained with primers h*ATF4*-St-st-only_NC2/3 and h*ATF4*-PstI-R using St-st-only as a template into *Hind*III-*Pst*I digested St-st-only.

dSt-st-SL1 was created by inserting the *Hind*III-*Pst*I digested PCR product obtained with primers h*ATF4*-dSt-st-SL1 and h*ATF4*-PstI-R using dSt-st as a template into *Hind*III-*Pst*I digested dSt-st.

St-st-only-SL2–1 was created by inserting the *Hind*III-*Pst*I digested fusion PCR product obtained with primers h*ATF4*-SL2–1-F and h*ATF4*-SL2–1-R using St-st-only as a template into *Hind*III-*Pst*I digested St-st-only.

St-st-only-SL2–2 was created by inserting the *Hind*III-*Pst*I digested fusion PCR product obtained with primers h*ATF4*-SL2–2-F and h*ATF4*-SL2–2-R using St-st-only as a template into *Hind*III-*Pst*I digested St-st-only.

dSt-st-SL2–1 was created by inserting the *Hind*III-*Pst*I digested fusion PCR product obtained with primers h*ATF4*-SL2–1-F and h*ATF4*-SL2–1-R using dSt-st as a template into *Hind*III-*Pst*I digested dSt-st.

dSt-st-SL2–2 was created by inserting the *Hind*III-*Pst*I digested fusion PCR product obtained with primers h*ATF4*-SL2–2-F and h*ATF4*-SL2–2-R using dSt-st as a template into *Hind*III-*Pst*I digested dSt-st.

d-all-SL2–2 was created by inserting the *Hind*III-*Pst*I digested fusion PCR product obtained with primers h*ATF4*-SL2–2-F and h*ATF4*-SL2–2-R using d-all as a template into *Hind*III-*Pst*I digested d-all.

wt-A235G was created by inserting the *Pst*I-*EcoR*V digested fusion PCR product obtained with primers h*ATF4*-A235G-F and h*ATF4*-A235G-R using wt ATF4-HA Tag as a template into *Pst*I-*EcoR*V digested wt ATF4-HA Tag.

wt-SL3$^{Mut\ 1}$ was created by inserting the *Pst*I-*EcoR*V digested fusion PCR product obtained with primers h*ATF4*-SL3_CAAAA-F and h*ATF4*-SL3_CAAAA-R using wt ATF4-HA Tag as a template into *Pst*I-*EcoR*V digested wt ATF4-HA Tag.

wt-SL3$^{Mut\ 2}$ was created by *Pst*I-*EcoR*V digested h*ATF4*-wt-SL3$^{Mut\ 2}$ (GeneArt String DNA Fragment; Invitrogen) into *Pst*I-*EcoR*V digested wt ATF4-HA Tag.

dSt-st-SL3$^{Mut\ 1}$ was created by inserting the *Pst*I-*EcoR*V digested fusion PCR product obtained with primers h*ATF4*-SL3_CAAAAF and h*ATF4*-SL3_CAAAA-R using dSt-st as a template into *Pst*I-*EcoR*V digested dSt-st.

d-all-SL3$^{Mut\ 1}$ was created by inserting the *Pst*I-*EcoR*V digested fusion PCR product obtained with primers h*ATF4*-SL3_CAAAA-F and h*ATF4*-SL3_CAAAA-R using d-all as a template into *Pst*I-*EcoR*V digested d-all.

uORF2-only-SL3$^{Mut\ 1}$ was created by inserting the *Pst*I-*EcoR*V digested fusion PCR product obtained with primers h*ATF4*-SL3_CAAAA-F and h*ATF4*-SL3_CAAAA-R using uORF2-only as a template into *Pst*I-*EcoR*V digested uORF2-only.

wt-AUG3^{Mut} (3rd *ATF4* AUG) was created by inserting the *Pst*I-*EcoR*V digested fusion PCR product obtained with primers h*ATF4*-3rdAGG-F and h*ATF4*-3rdAGG-R wt ATF4-HA Tag as a template into *Pst*I-*EcoR*V digested wt ATF4-HA Tag.

dSt-st-AUG3^{Mut} (3rd *ATF4* AUG) was created by inserting the *Pst*I-*EcoR*V digested fusion PCR product obtained with primers h*ATF4*-3rdAGG-F and h*ATF4*-3rdAGG-R dSt-st as a template into *Pst*I-*EcoR*V digested dSt-st.

wt-UUG^{Mut}-AUG3 was created by inserting *Pst*I-*EcoR*V digested h*ATF4*-wt-UUG^{Mut}_AUG3 (GeneArt String DNA Fragment; Invitrogen) into *Pst*I-*EcoR*V digested wt ATF4-HA Tag.

wt-SL3$^{Mut\ 1}$-AUG3^{Mut} was created by inserting *Pst*I-*EcoR*V digested h*ATF4*-wt-SL3$^{Mut\ 1}$-AUG3^{Mut} (GeneArt String DNA Fragment; Invitrogen) into *Pst*I-*EcoR*V digested wt ATF4-HA Tag.

wt-A326G was created by inserting *Pst*I-*EcoR*V digested h*ATF4*-wt-A326G (GeneArt String DNA Fragment; Invitrogen) into *Pst*I-*EcoR*V digested wt ATF4-HA Tag.

wt-SL3$^{Mut\ 1}$-A326G was created by *Pst*I-*EcoR*V digested h*ATF4*-wt-SL3$^{Mut\ 1}$-A326G (GeneArt String DNA Fragment; Invitrogen) into *Pst*I-*EcoR*V digested wt ATF4-HA Tag.

wt-3AUG^{Mut} was created by inserting *Pst*I-*EcoR*V digested h*ATF4*-wt-3AUG^{Mut} of *ATF4* (GeneArt String DNA Fragment; Invitrogen) into *Pst*I-*EcoR*V digested wt ATF4-HA Tag.

d2-3AUG^{Mut} was created by inserting *Pst*I-*EcoR*V digested h*ATF4*-d2-3AUG^{Mut} of *ATF4* (GeneArt String DNA Fragment; Invitrogen) into *Pst*I-*EcoR*V digested wt ATF4-HA Tag.

3AUG^{Mut}-stop was created by inserting *Pst*I-*EcoR*V digested h*ATF4*-3AUG^{Mut}-stop (GeneArt String DNA Fragment; Invitrogen) into *Pst*I-*EcoR*V digested wt ATF4-HA Tag.

3AUG^{Mut}-short was created by inserting *EcoR*V-*Not*I digested h*ATF4*-3AUG^{Mut}-short (GeneArt String DNA Fragment; Invitrogen) into *EcoR*V-*Not*I digested wt-3AUG^{Mut}.

3AUG^{Mut}-short-stop was created by inserting *EcoR*V-*Not*I digested h*ATF4*-3AUG^{Mut}-short (GeneArt String DNA Fragment; Invitrogen) into *EcoR*V-*Not*I digested 3AUG^{Mut}-stop.

uORF2_ins was created by inserting *Pst*I-*EcoR*V digested h*ATF4*-uORF2_ins (GeneArt String DNA Fragment; Invitrogen) into PstI-EcoRV digested wt ATF4-HA Tag.

d2_ins (uORF2_ins => AUG to AGG) was created by inserting *Pst*I-*EcoR*V digested h*ATF4*-d2_ins (GeneArt String DNA Fragment; Invitrogen) into *Pst*I-*EcoR*V digested wt ATF4-HA Tag.

uORF2_ins_3AUG^{Mut} was created by inserting *Pst*I-*EcoR*V digested h*ATF4*-uORF2_ins_3AUG^{Mut} (GeneArt String DNA Fragment; Invitrogen) into *Pst*I-*EcoR*V digested wt ATF4-HA Tag.

wt-CUG^{Mut} was created by inserting *Pst*I-*EcoR*V digested h*ATF4*-wt-CUG^{Mut} (GeneArt String DNA Fragment; Invitrogen) into *Pst*I-*EcoR*V digested wt ATF4-HA Tag.

wt-SL3$^{Mut\ 1}$-CUG^{Mut} was created by inserting *Pst*I-*EcoR*V digested h*ATF4*-wt—SL3$^{Mut\ 1}$-CUG^{Mut} (GeneArt String DNA Fragment; Invitrogen) into *Pst*I-*EcoR*V digested wt ATF4-HA Tag.

wt_ins was created by inserting *Pst*I-*EcoR*V digested h*ATF4*-wt_ins (GeneArt String DNA Fragment; Invitrogen) into *Pst*I-*EcoR*V digested wt ATF4-HA Tag.

wt_ins-SL3$^{Mut\ 1}$ was created by inserting *Pst*I-*EcoR*V digested h*ATF4*-wt_ins-SL3$^{Mut\ 1}$ (GeneArt String DNA Fragment; Invitrogen) into *Pst*I-*EcoR*V digested wt ATF4-HA Tag.

d2_ins was created by inserting *BstB*I digested d2 into *BstB*I digested wt_ins.

d2_ins-SL3$^{Mut\ 1}$ was created by inserting *BstB*I digested d2 into *BstB*I digested wt_ins-SL3$^{Mut\ 1}$.

uORF2_ins-SL3$^{Mut\ 1}$ was created by inserting *BstB*I digested uORF2_ins into *BstB*I digested wt-SL3$^{Mut\ 1}$.

uORF2-*ATF4*-HA fusion was created by inserting *Pst*I-*EcoR*V digested h*ATF4*- uORF2-*ATF4*-HA fusion (GeneArt String DNA Fragment; Invitrogen) into *Pst*I-*EcoR*V digested wt ATF4-HA Tag.

Preparation of whole cell extracts—HEK293T cells pre-grown for 24 h as described in the Human cell lines section were transfected with ATF4-HA reporter constructs using TurboFect transfection reagent (Thermo Fisher); 2.5 μg of plasmid DNA was used per one well in the 6-well plate in all experiments. Exactly 8 h after transfection, cells were treated with either 1 μM thapsigargin (Invitrogen) for 3 h or Dimethyl sulfoxide (DMSO) (Sigma) as a control. Cells were lysed directly on the plate using 200 μL Glo Lysis Buffer (Promega), which is recommended to be used with the Jess™ Automated Western Blot System (Protein Simple Bio-Techne). 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. In the experiments with tunicamycin (Sigma-Aldrich), 8 h after transfection, HEK293T cells were treated with either 0.5 μM tunicamycin for 4 h or DMSO as a control. Harvesting of the cell lysates was performed essentially the same as for thapsigargin.

RT-qPCR control step—Total RNA was isolated using RNA blue reagent (Top-Bio) according to the manufacturer's instructions. Subsequently, a TURBO DNase (Invitrogen) cleavage step was performed. For cDNA synthesis, 0.5 mg of total RNA was used for each sample using a High-Capacity cDNA Reverse Transcription kit (Applied Biosystems). As described before,[52,53] at least three 10-fold serial dilutions of cDNA for each mutant with the WT pair, examined in each individual experiment, were compared using RT-qPCR, and the maximum difference criterion of 1 cycle had to be met for all sample dilutions in all relevant controls (including control for the transfection efficiency) at the time of lysis, otherwise samples were discarded. To compare the reporter levels for all mutants, the reverse primer was designed to be in match with HA tag (to avoid interference with endogenous *ATF4*). Primers matching the plasmid region carrying neomycin/kanamycin resistance were created (to ensure the same level of transfection efficiency for each pair of samples). Internal controls for RNA isolation efficiency and cDNA synthesis were used as well (SPIKE RNA and the corresponding primers). The qPCR primers are listed in Data S3. For qPCR reactions, HOT FIREPol EvaGreen qPCR Mix Plus 5x (Solis BioDyne) was mixed with0.8 μM primers and cDNA and run using the following program: 95°C for 15 s, followed by 43 cycles of 95°C for 15 s, 60°C for 20 s, and 72°C for 20 s.

Jess™ Simple western (SW) assay (Protein Simple)—Jess™ SW Assay was used according to the manufacturer's protocol. In brief, the entire process in the Jess machine can be described as follows: 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. For the experiments conducted as part of this study, 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 (Protein Simple bio-techne). Twenty-four samples were tested in each Jess run. Samples were mixed by vortexing, boiled at 95°C 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 (Protein Simple bio-techne). Mouse monoclonal anti-HA tag (abcam) or anti-*c*-Myc antibodies (provided by Dr. Tomáš Vomastek) were pipetted along with Anti-Mouse Detection Module (Protein Simple bio-techne). The plate was centrifuged for 5 min at 686 × 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 instrument were performed in Compass for SW software (version 6.3.0) (ProteinSimple).

The key point to mention is that the same amount of total protein loaded in all capillaries was pre-estimated using Bio-Rad Protein Assay Dye Reagent Concentrate (BIO-RAD) and BSA titration curve was performed before each Jess run with serial dilutions of BSA in PBS. Afterward, Protein Normalization Module (Protein Simple bio-techne) was applied for a thorough comparison and finally only up to 20% difference between samples was tolerated using PN Module, otherwise samples were discarded. Please note that the 180 kDa peak, which occurs occasionally in electropherograms, represents (based on the manufacturer's instructions) a non-specific peak arising from cross-reactions between fluorescence (PN kit) and chemiluminescence (target antibody) reagents during capillary detection. According to the manufacturer, the only occasionally nature of this peak is not known.

Human ATF4 mRNA transcript sequence analysis—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.

Western blot—All samples were resolved using Criterion TGX Precast Gels 4–20% (BIO-RAD) followed by western blotting. All primary antibodies used in this study are listed in key resources table. The signal was developed using SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Fisher) and detected in a G-Box imager (Syngene) using series of varying exposure time. Data were analyzed using Quantity One Software (Bio-Rad) and deposited at Mendeley; accession numbers are listed in the key resources table.

Ribosome protection assay—HEK293T cells were seeded in 20 mL of media in the 15 cm dish and grown to a confluency of 60–70% before being treated with either DMSO (control) or thapsigargin (1μM) for 3h. For experiments involving either WT or $SL3^{Mut}$ 1*ATF4* reporters (Figures 7D and S5C), cells were independently transfected with either reporter plasmid at approximately 40–50% cell confluency and – 8 h after transfection – treated with DMSO (Control) or thapsigargin (1μM) for additional 3h. Cells were then treated with formaldehyde ("cross-linked", incubated with HCHO 0.8% for 5 min at 4°C and the reactions were quenched with 75 mM glycine for 5 min, as described[22,27]) or cycloheximide ("non-crosslinked"; incubated with 100 μg/mL cycloheximide for 1 min at 37°C, as shown before,[54] following which cell lysates were prepared. Concentration of cell lysates were measured on NanoDrop at OD_{260} and all samples were divided into two-halves. One-half (5 AU OD_{260} per 1 mL of a lysate) was treated with Ambion RNase I (4U) (Invitrogen) at 30°C for 10 min at 500 rpm to digest all enzyme-accessible RNA. The other half was processed the same way but without the RNase I treatment. Nuclease reactions were stopped with SUPERaseIn RNase inhibitor (8U) (Invitrogen). All samples (RNase I treated and non-treated) were supplemented with 1 mL of RNA blue (Top Bio) and mixed vigorously by inverting the tubes. To reverse crosslinking, formaldehyde-crosslinked samples were incubated at 65°C for 20 min with an intermittent shaking by inverting the tubes. Subsequently, all the samples were cooled down at RT for 5 min and 5μL of SPIKE RNA (*in vitro* transcribed yeast RPL41a) was added as an internal control of RNA isolation to all the samples and RNA was isolated according to the manufacturer's (Top Bio s.r.o) instructions. The resulting RNA pellets were dissolved in 40 μL of nuclease free water. RNA concentrations were measured using NanoDrop. To remove any DNA contaminants, 2.5 μg of RNA samples were treated with TURBO DNase (Invitrogen) according to the manufacturer's instructions. For cDNA synthesis, 1 μg of DNase-treated 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 0.4 μM primers (for the SPIKE primer set) and 0.25 μM (for all Amplicon 1–3 primer sets) primers and cDNA, and run using the following program: 95°C for 15 s, followed by 40 cycles of 95°C for 15 s, 57.4°C (for SPIKE primer set) or 62°C (for all Amplicon 1–3 primer sets) for 20 s, and 72°C for 20 s. Melting curve was measured between 60°C and 90°C to assess amplification of the single specific product. qPCR data was analyzed as described in the main text and corresponding figure legends.

Total and poly(A) RNA isolation—Total RNA for T3 ligation assay was isolated using the RNA Blue reagent (Top-Bio) according to the manufacturer's instructions, 5 mL of the reagent was used *per* one 15 cm dish containing HEK293T or HeLa cells grown to approximately 80% confluency. The resulting RNA pellets were resuspended in RNase-free water (Thermo-Fisher) and the concentration was quantified by NanoDrop. Subsequently, poly(A) RNA for T3-ligation assay was purified using the Poly(A)Purist MAG kit (Thermo Fisher) 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.[38] Briefly, each ligation reaction mixture A consisted of 20 nM of corresponding probe SGL (Data S3), 20 nM of corresponding probe SGR, 1x T3 ligation buffer (NEB), and ~200–500 ng of poly(A) RNA. The ligation reaction mixture B consisted of T3 DNA ligase (NEB) 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, 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 s, 60°C for 20 s, and 72°C for 20 s. Melting curves were analyzed between 65°C 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}).

Ribosome profiling datasets analysis—Riboseq fastq files of untreated or sodium arsenite (GSE17432, GSE55195) or tunicamycin (GSE113171) treated samples from previously published datasets[24,30,31] 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)[55] 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.[56] 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.[47]

QUANTIFICATION AND STATISTICAL ANALYSIS

Variables were tested for normality using the Shapiro-Wilk normality test. Based on the normality test, the differences between experimental groups were tested by the t test or by Mann–Whitney test (stated in the Figure's legend). Variables are presented as mean ± SD and p values of <0.05 were considered statistically significant. GraphPad Prism statistical software (ver. 10.1.2, GraphPad Software, San Diego, CA, USA, RRID:SCR_002798) and SciPy (ver. 1.10.1)[43], Matplotlib (ver. 3.7.0)[45] libraries in Python (ver. 3.11.3)[44] 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).[46] For the ribosome protection assay, statistical significance was assessed using unpaired, two-sided, *t* test with Bonferroni correction.

ADDITIONAL RESOURCES

This work does not use additional resources.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We are grateful to Dmitry E. Andreev, Pavel Baranov, and Nick Guydosh for 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) and CZ.02.01.01/00/22_008/0004575 RNA for Therapy by ERDF and MEYS CR (both to L.S.V.), National Institutes of Health 5R35GM127089 and the Chan Zuckerberg Initiative (to C.V.), and National Institutes of Health DK060596 (to M.H.). We also acknowledge the Structural Mass Spectrometry Core Facility of CIISB, Instruct-CZ Center (LM2023042), and European Regional Development Fund Project “UP CIISB” (CZ.02.1.01/0.0/0.0/18_046/0015974) supported by MEYS CR.

REFERENCES

1. Wek RC, Jiang HY, and Anthony TG (2006). Coping with stress: eIF2 kinases and translational control. Biochem. Soc. Trans 34, 7–11. 10.1042/BST20060007. [PubMed: 16246168]
2. Pakos-Zebrucka K, Koryga I, Mnich K, Ljujic M, Samali A, and Gorman AM (2016). The integrated stress response. EMBO Rep. 17, 1374–1395. 10.15252/embr.201642195. [PubMed: 27629041]
3. Denoyelle C, Abou-Rjaily G, Bezrookove V, Verhaegen. M, Johnson TM, Fullen DR, Pointer JN, Gruber SB, Su LD, Nikiforov MA, et al. (2006). Anti-oncogenic role of the endoplasmic reticulum differentially activated by mutations in the MAPK pathway. Nat. Cell Biol 8, 1053–1063. 10.1038/ncb1471. [PubMed: 16964246]
4. Donnelly N, Gorman AM, Gupta S, and Samali A (2013). The eIF2αlpha kinases: their structures and functions. Cell. Mol. Life Sci 70, 3493–3511. 10.1007/s00018-012-1252-6. [PubMed: 23354059]
5. Valášek LS (2012). Ribozoomin’ – Translation Initiation from the Perspective of the Ribosome-bound Eukaryotic Initiation Factors (eIFs). Curr. Protein Pept. Sci 13, 305–330. [PubMed: 22708493]
6. Hinnebusch AG (2014). The scanning mechanism of eukaryotic translation initiation. Annu. Rev. Biochem 83, 779–812. [PubMed: 24499181]
7. Gunišová S, Hronová V, Mohammad MP, Hinnebusch AG, and Valášek LS (2018). Please do not recycle! Translation reinitiation in microbes and higher eukaryotes. FEMS Microbiol. Rev 42, 165–192. 10.1093/femsre/fux059. [PubMed: 29281028]
8. Dever TE, Ivanov IP, and Hinnebusch AG (2023). Translational regulation by uORFs and start codon selection stringency. Genes Dev. 37, 474–489. 10.1101/gad.350752.123. [PubMed: 37433636]
9. Harding HP, 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. [PubMed: 11106749]
10. Vattem KM, and Wek RC (2004). Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. Proc. Natl. Acad Sci. USA 101, 11269–11274. [PubMed: 15277680]
11. Lu PD, Harding HP, 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. 10.1083/jcb.200408003. [PubMed: 15479734]
12. Novoa I, Zhang Y, Zeng H, Jungreis R, Harding HP, and Ron D (2003). Stress-induced gene expression requires programmed recovery from translational repression. EMBO J. 22, 1180–1187. 10.1093/emboj/cdg112. [PubMed: 12606582]
13. Guan BJ, van Hoef V, Jobava R, Elroy-Stein O, Valasek LS, Cargnello M, Gao XH, Krokowski D, Merrick WC, Kimball SR, et al. (2017). A Unique ISR Program Determines Cellular Responses to Chronic Stress. Mol. Cell 68, 885–900.e886. 10.1016/j.molcel.2017.11.007. [PubMed: 29220654]

14. Pitale PM, Gorbatyuk O, and Gorbatyuk M (2017). Neurodegeneration: Keeping ATF4 on a Tight Leash. Front. Cell. Neurosci 11, 410. 10.3389/fncel.2017.00410.
15. Wortel IMN, van der Meer LT, Kilberg MS, and van Leeuwen FN (2017). Surviving Stress: Modulation of ATF4-Mediated Stress Responses in Normal and Malignant Cells. Trends Endocrinol. Metab 28, 794–806. 10.1016/j.tem.2017.07.003. [PubMed: 28797581]
16. Hinnebusch AG (2005). Translational regulation of GCN4 and the general amino acid control of yeast. Annu. Rev. Microbiol 59, 407–450. [PubMed: 16153175]
17. Gunišová S, and Valášek LS (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. [PubMed: 24623812]
18. Gunišová S, Beznosková P, Mohammad MP, Vlcková V, and Valášek LS (2016). In-depth analysis of cis-determinants that either promote or inhibit reinitiation on GCN4 mRNA after translation of its four short uORFs. RNA 22, 542–558. 10.1261/rna.055046.115. [PubMed: 26822200]
19. Hronova V, Mohammad MP, Wagner S, Panek J, Gunisova S, Zeman J, Poncova K, Valasek LS, et al. (2017). Does eIF3 promote reinitiation after translation of short upstream ORFs also in mammalian cells? RNA Biol. 10.1080/15476286.15472017.11353863.
20. Mohammad MP, Munzarová Pondelícková V, Zeman J, Gunišová S, and Valášek LS (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. 10.1093/nar/gkx049. [PubMed: 28119417]
21. Mohammad MP, Smirnova A, Gunišová S, and Valášek LS (2021). eIF4G is retained on ribosomes elongating and terminating on short upstream ORFs to control reinitiation in yeast. Nucleic Acids Res. 49, 8743–8756. 10.1093/nar/gkab652. [PubMed: 34352092]
22. Wagner S, Herrmannová A, Hronová V, Gunišová S, Sen ND, Hannan RD, Hinnebusch AG, Shirokikh NE, Preiss T, and Valášek LS (2020). Selective Translation Complex Profiling Reveals Staged Initiation and Co-translational Assembly of Initiation Factor Complexes. Mol. Cell 79, 546–560.e547. 10.1016/j.molcel.2020.06.004. [PubMed: 32589964]
23. Starck SR, Tsai JC, 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. 10.1126/science.aad3867. [PubMed: 26823435]
24. Andreev DE, O'Connor PBF, Fahey C, Kenny EM, Terenin IM, Dmitriev SE, Cormican P, Morris DW, Shatsky IN, and Baranov PV (2015). Translation of 5' leaders is pervasive in genes resistant to eIF2 repression. Elife 4, e03971. 10.7554/eLife.03971. [PubMed: 25621764]
25. Sidrauski C, McGeachy AM, Ingolia NT, and Walter P (2015). The small molecule ISRIB reverses the effects of eIF2alpha phosphorylation on translation and stress granule assembly. Elife 4, e05033. 10.7554/eLife.05033. [PubMed: 25719440]
26. Zhou J, Wan J, Shu XE, Mao Y, Liu XM, Yuan X, Zhang X, Hess ME, Bruning JC, and Qian SB (2018). N(6)-Methyladenosine Guides mRNA Alternative Translation during Integrated Stress Response. Mol. Cell 69, 636–647.e637. 10.1016/j.molcel.2018.01.019. [PubMed: 29429926]
27. Wagner S, Bohlen J, Herrmannová A, Jelínek J, Preiss T, Valášek LS, and Teleman AA (2022). Selective footprinting of 40S and 80S ribosome subpopulations (Sel-TCP-seq) to study translation and its control. Nat. Protoc 17, 2139–2187. 10.1038/s41596-022-00708-4. [PubMed: 35869369]
28. Rendleman J, Haizel S, Wu S, Liu J, Ge X, Zou H, Mohammad MP, Pressler M, Maity S, Hronová V, et al. (2023). Regulatory start-stop elements in 5' untranslated regions pervasively modulate translation. Preprint at bioRxiv. 10.1101/2021.07.26.453809.
29. Dey S, Baird TD, Zhou D, Palam LR, Spandau DF, and Wek RC (2010). Both transcriptional regulation and translational control of ATF4 are central to the integrated stress response. J. Biol. Chem 285, 33165–33174. 10.1074/jbc.M110.167213. [PubMed: 20732869]
30. Ichihara K, Matsumoto A, Nishida H, Kito Y, Shimizu H, Shichino Y, Iwasaki S, Imami K, Ishihama Y, and Nakayama KI (2021). Combinatorial analysis of translation dynamics reveals eIF2 dependence of translation initiation at near-cognate codons. Nucleic Acids Res. 49, 7298–7317. 10.1093/nar/gkab549. [PubMed: 34226921]

31. Rendleman J, Cheng Z, Maity S, Kastelic N, Munschauer M, Allgoewer K, Teo G, Zhang YBM, Lei A, Parker B, et al. (2018). New insights into the cellular temporal response to proteostatic stress. Elife 7, e39054. 10.7554/eLife.39054. [PubMed: 30272558]
32. Ivanov IP, Shin BS, Loughran G, Tzani I, Young-Baird SK, Cao C, Atkins JF, and Dever TE (2018). Polyamine Control of Translation Elongation Regulates Start Site Selection on Antizyme Inhibitor mRNA via Ribosome Queuing. Mol. Cell 70, 254–264.e256. 10.1016/j.molcel.2018.03.015. [PubMed: 29677493]
33. Archer SK, Shirokikh NE, Beilharz TH, and Preiss T (2016). Dynamics of ribosome scanning and recycling revealed by translation complex profiling. Nature 535, 570–574. 10.1038/nature18647. [PubMed: 27437580]
34. Ingolia NT (2010). Genome-wide translational profiling by ribosome footprinting. Methods Enzymol. 470, 119–142. [PubMed: 20946809]
35. Sfakianos AP, Raven RM, and Willis AE (2022). The pleiotropic roles of eIF5A in cellular life and its therapeutic potential in cancer. Biochem. Soc. Trans 50, 1885–1895. 10.1042/BST20221035. [PubMed: 36511302]
36. Schuller AP, Wu CC, Dever TE, Buskirk AR, and Green R (2017). eIF5A Functions Globally in Translation Elongation and Termination. Mol. Cell 66, 194–205.e195. 10.1016/j.molcel.2017.03.003. [PubMed: 28392174]
37. Han P, Shichino Y, Schneider-Poetsch T, Mito M, Hashimoto S, Udagawa T, Kohno K, Yoshida M, Mishima Y, Inada T, and Iwasaki S (2020). Genome-wide Survey of Ribosome Collision. Cell Rep. 31, 107610. 10.1016/j.celrep.2020.107610. [PubMed: 32375038]
38. Liu W, Yan J, Zhang Z, Pian H, Liu C, and Li Z (2018). 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. 10.1039/c7sc05233b. [PubMed: 29780465]
39. Zhang Z, Chen LQ, Zhao YL, Yang CG, Roundtree IA, Zhang Z, Ren J, Xie W, He C, and Luo GZ (2019). Single-base mapping of m(6) A by an antibody-independent method. Sci. Adv 5, eaax0250. 10.1126/sciadv.aax0250. [PubMed: 31281898]
40. Matsuda D, and Dreher TW (2006). Close spacing of AUG initiation codons confers dicistronic character on a eukaryotic mRNA. RNA 12, 1338–1349. 10.1261/rna.67906. [PubMed: 16682564]
41. Neill G, and Masson GR (2023). A stay of execution: ATF4 regulation and potential outcomes for the integrated stress response. Front. Mol. Neurosci 16, 1112253. 10.3389/fnmol.2023.1112253. [PubMed: 36825279]
42. Roithová A, Feketová Z, Vanáčová Š, and Staněk D (2020). DIS3L2 and LSm proteins are involved in the surveillance of Sm ring-deficient snRNAs. Nucleic Acids Res. 48, 6184–6197. 10.1093/nar/gkaa301. [PubMed: 32374871]
43. Virtanen P, Gommers R, Oliphant TE, 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. 10.1038/s41592-019-0686-2. [PubMed: 32015543]
44. Rossum GV, and Drake FL (2009). Python 3 Reference Manual (CreateSpace).
45. Hunter JD (2007). Matplotlib: A 2D Graphics Environment. Comput. Sci. Eng 9, 90–95. 10.1109/MCSE.2007.55.
46. Kolde R, and Kolde M (2015). R Package 'pheatmap'. R Package 1. Web Tool 790.
47. Lauria F, Tebaldi T, Bernabo P, Groen EJN, Gillingwater TH, and Viero G (2018). riboWaltz: Optimization of ribosome P-site positioning in ribosome profiling data. PLoS Comput. Biol 14, e1006169. 10.1371/journal.pcbi.1006169. [PubMed: 30102689]
48. Andreev DE, Terenin IM, Dmitriev SE, and Shatsky IN (2016). Pros and cons of pDNA and mRNA transfection to study mRNA translation in mammalian cells. Gene 578, 1–6. 10.1016/j.gene.2015.12.008. [PubMed: 26680098]
49. Iwawaki T, Akai R, Toyoshima T, Takeda N, Ishikawa TO, and Yamamura KI (2017). Transgenic mouse model for imaging of ATF4 translational activation-related cellular stress responses in vivo. Sci. Rep 7, 46230. 10.1038/srep46230. [PubMed: 28387317]

50. Wang L, Zhao W, Xia C, Li Z, Zhao W, Xu K, Wang N, Lian H, Rosas IO, and Yu G (2022). TRIB3 Mediates Fibroblast Activation and Fibrosis though Interaction with ATF4 in IPF. Int. J. Mol. Sci 23, 15705. 10.3390/ijms232415705. [PubMed: 36555349]

51. Penn WD, Harrington HR, Schlebach JP, and Mukhopadhyay S (2020). Regulators of Viral Frameshifting: More Than RNA Influences Translation Events. Annu Rev Virol 7, 219–238. 10.1146/annurev-virology-012120-101548. [PubMed: 32600156]

52. Chiu W-L, Wagner S, Herrmannová A, Burela L, Zhang F, Saini AK, Valásek L, and Hinnebusch AG (2010). The C-Terminal Region of Eukaryotic Translation Initiation Factor 3a (eIF3a) Promotes mRNA Recruitment, Scanning, and, Together with eIF3j and the eIF3b RNA Recognition Motif, Selection of AUG Start Codons. Mol. Cell. Biol 30, 4415–4434. [PubMed: 20584985]

53. Khoshnevis S, Gunišová S, Vlčková V, Kouba T, Neumann P, Beznosková P, Ficner R, and Valášek LS (2014). Structural integrity of the PCI domain of eIF3a/TIF32 is required for mRNA recruitment to the 43S pre-initiation complexes. Nucleic Acids Res. 42, 4123–4139. [PubMed: 24423867]

54. Herrmannová A, Prilepskaja T, Wagner S, Šikrová D, Zeman J, Poncová K, and Valášek LS (2020). Adapted formaldehyde gradient cross-linking protocol implicates human eIF3d and eIF3c, k and l subunits in the 43S and 48S pre-initiation complex assembly, respectively. Nucleic Acids Res. 48, 1969–1984. 10.1093/nar/gkz1185. [PubMed: 31863585]

55. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. 10.1093/bioinformatics/bts635. [PubMed: 23104886]

56. Herrmannová A, Jelínek J, Pospíšilová K, Kerényi F, Vomastek T, Watt K, Brábek J, Mohammad MP, Wagner S, Topisirovic I, and Valášek LS (2023). Perturbations in eIF3 subunit stoichiometry alter expression of ribosomal proteins and key components of the MAPK signaling pathway. Elife, in press. 10.1101/2023.06.29.547003.

Highlights

- A highly conserved stem-loop (SL3) occurs in the uORF2/*ATF4* overlap
- SL3 enables ribosome queuing to provide another layer ofATF4 translational control
- SL3 is preceded by a near-cognate CUG in the *ATF4* frame with which it genetically interacts
- The canonical *ATF4* translation start site is a subject of substantial leaky scanning

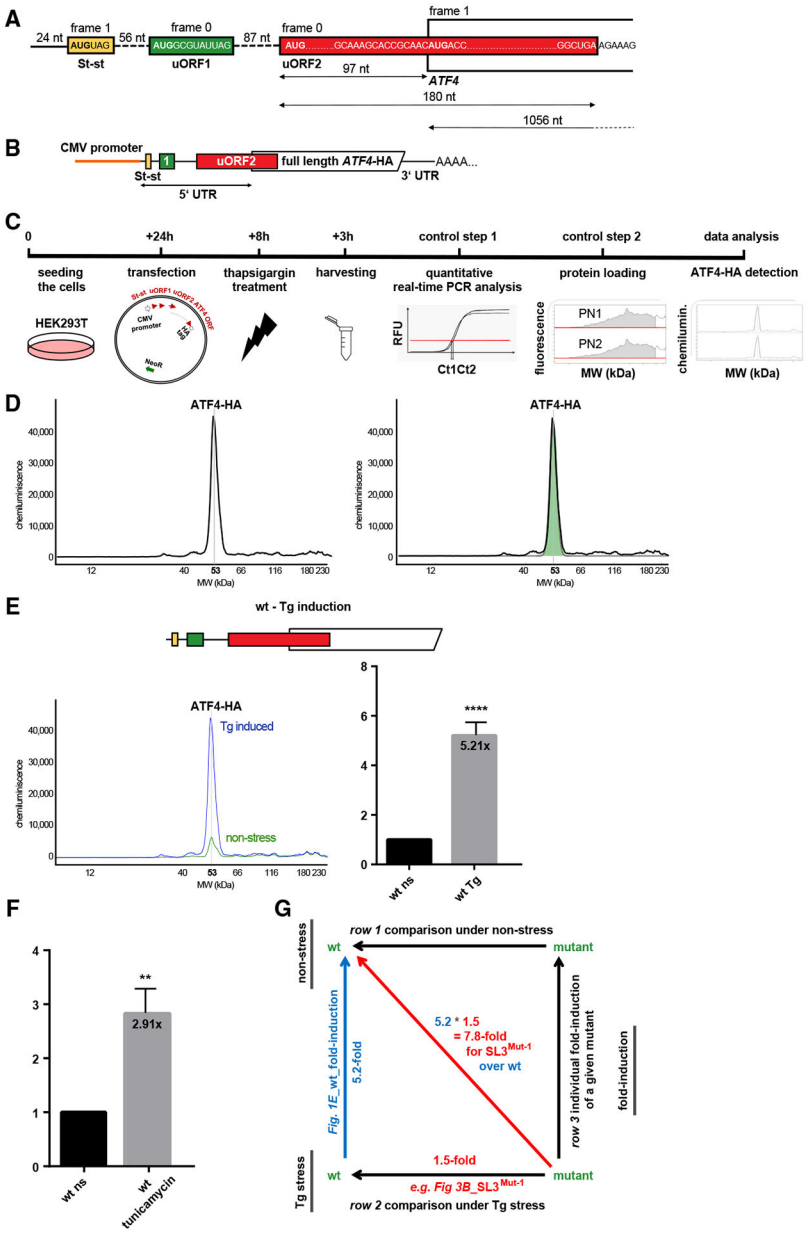


Figure 1.

Revisiting translational control of human ATF4: Reporters and experimental setup

(A) Schematic showing the 5′ end of the mRNA encoding 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 the main *ATF4* ORF, frames 0 and 1. Distances are given in nucleotides.

(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 timeline diagram. PN, Jess protein normalization detection module.

(D) All HEK293T cell lysates were subjected to protein separation and immunodetection using the Jess system. The signal, detected in the capillary, is represented as an electropherogram (a single peak of the ATF4-HA tag full-length protein, size of 53 kDa, left) and was automatically quantified (right). Expression of the WT construct under 3 h of Tg stress conditions 8 h post transfection, detected by anti-mouse HA tag antibodies, is shown.

(E) Stress-induced upregulation of ATF4-HA protein expression under Tg stress (blue) compared with non-stress conditions (green). Quantified “fold induction” data were plotted ($n = 17$, right). The differences between experimental groups were tested by a t test. Variables are presented as mean ± SD, and $p < 0.05$ was considered statistically significant (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$).

(F) Stress-induced upregulation of the ATF4-HA protein expression from the WT reporter under tunicamycin stress compared with non-stress conditions (set to 1). Quantified “fold induction” data were plotted ($n = 3$) and analyzed as in (E).

(G) Experimental setup illustrating determination of fold change values when comparing (1) each mutant reporter construct vs. the WT construct separately under non-stress (top horizontal arrow) and Tg stress conditions (bottom horizontal arrow) and (2) the fold induction expression of a mutant under stress with the same mutant under non-stress (right vertical arrow) or the WT under stress vs. no stress (blue left vertical arrow). The red diagonal arrow indicates calculated fold induction changes of a given mutant under stress vs. WT under non-stress conditions; i.e., how much each mutant increases or decreases the ~5.2-fold induction of the WT reporter.

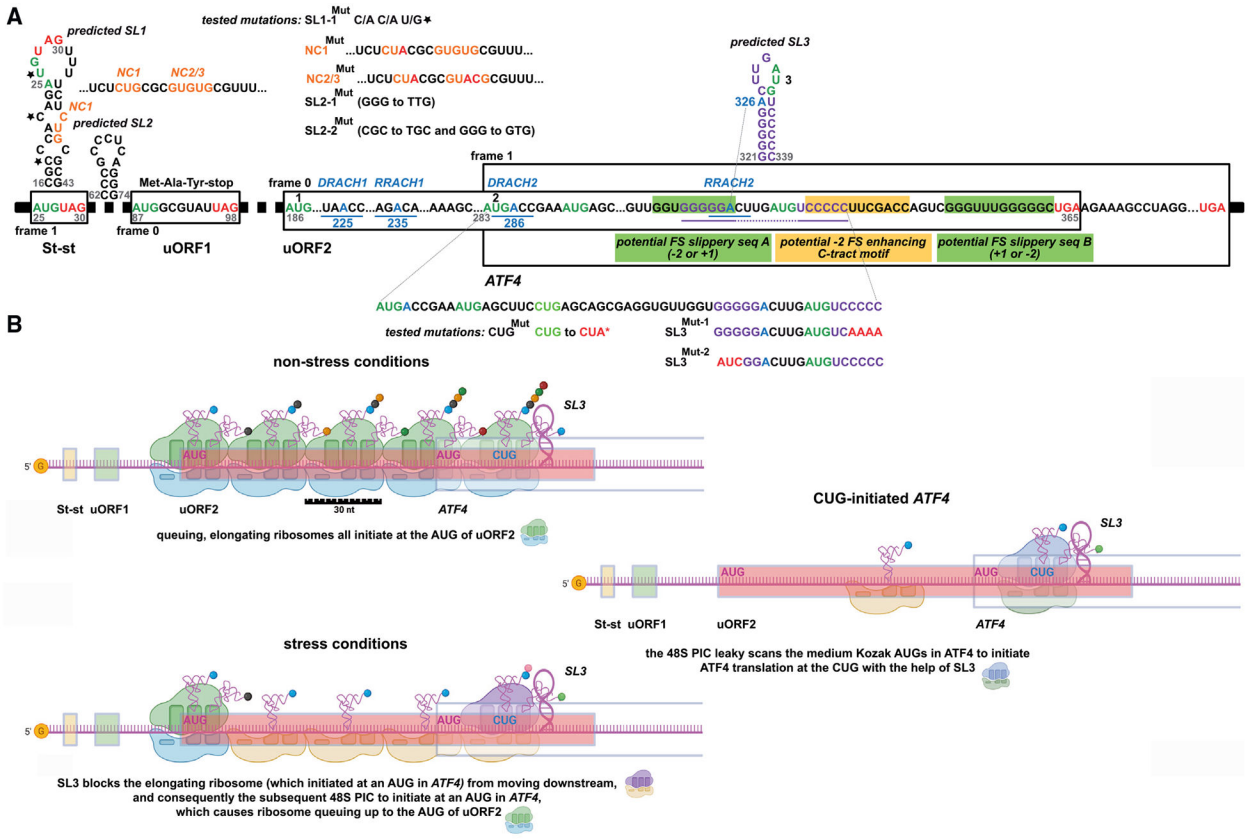


Figure 2.

Sequence analysis of the 5′ UTR of the human ATF4 mRNA and ribosome queuing model expanding the mode of ATF4 translational control

(A) Schematic of previously unknown, 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 vs. 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.

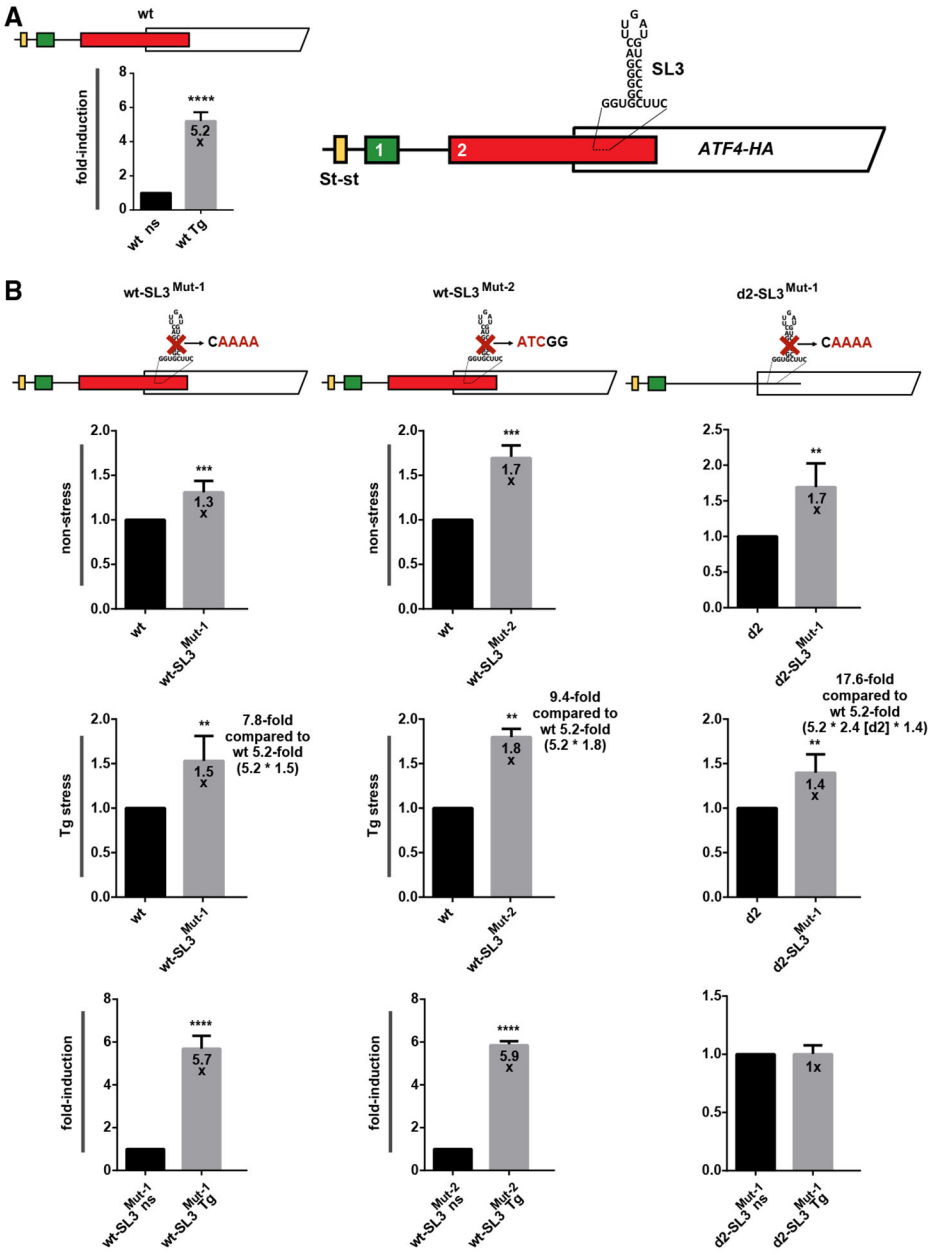


Figure 3.

SL3 delays the flow of ribosomes in the uORF2/ATF4 overlap

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

(B) Same as Figure 1E except that the SL3 *ATF4* mutant constructs depicted at the top of the corresponding panels were subjected to Jess analyses. Relative ATF4-HA protein expression levels were plotted as ratios of values obtained with an indicated mutant construct vs. the WT set to 1 under “non-stress” (row 1) and “Tg stress” (row 2) conditions; “fold-induction” plots (row 3) depict ratios of Tg stress vs. non-stress values obtained with a given mutant construct. The differences between experimental groups were tested by the t test, except wt-SL3^{Mut-1} Tg stress and d2-SL3^{Mut-1} non-stress, where a Mann-Whitney test was used ($n \geq 3$).

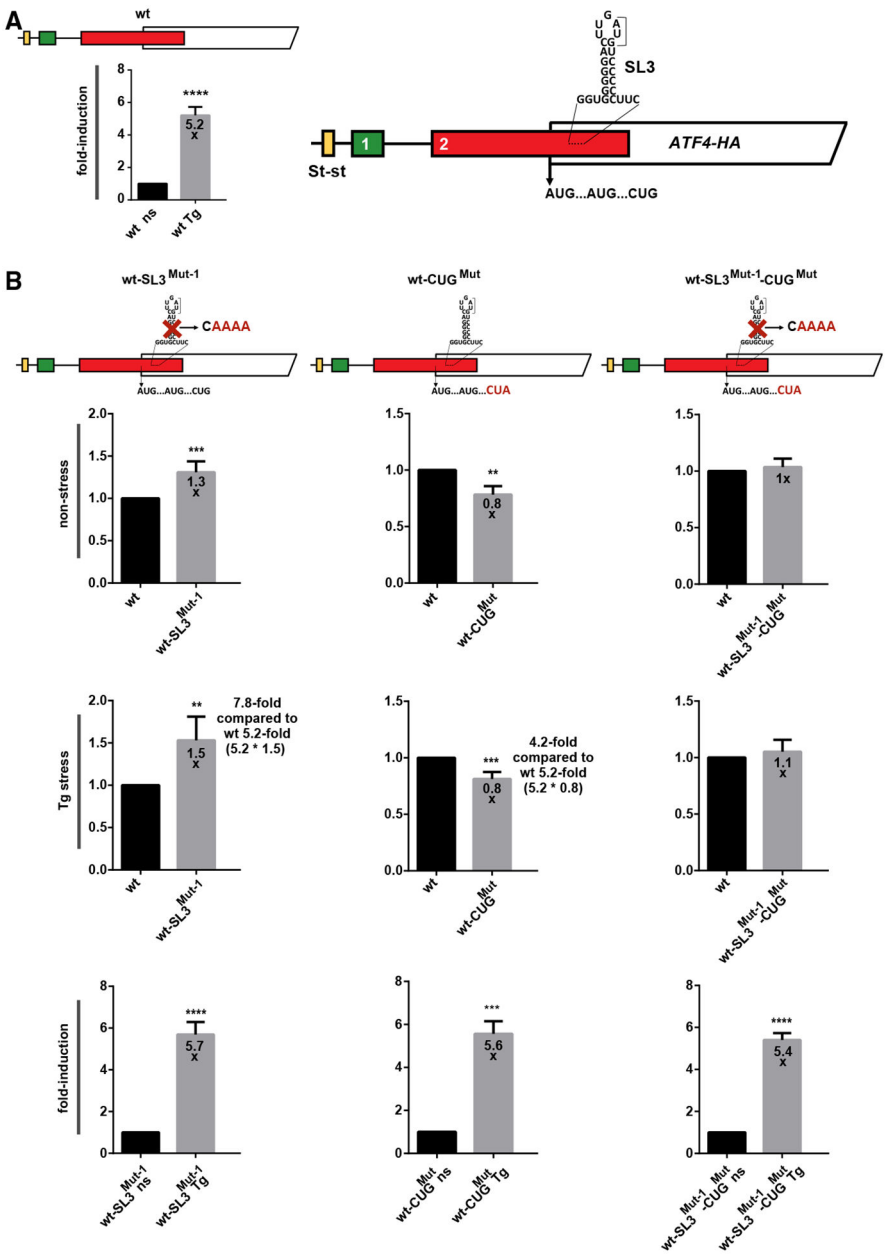


Figure 4.

SL3 genetically interacts with the upstream CUG near-cognate codon

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

(B) Same as Figure 3B except that the CUG to CUA and SL3^{Mut-1} *ATF4* mutant constructs depicted at the top of the corresponding panels were subjected to Jess analyses (n ≥ 3).

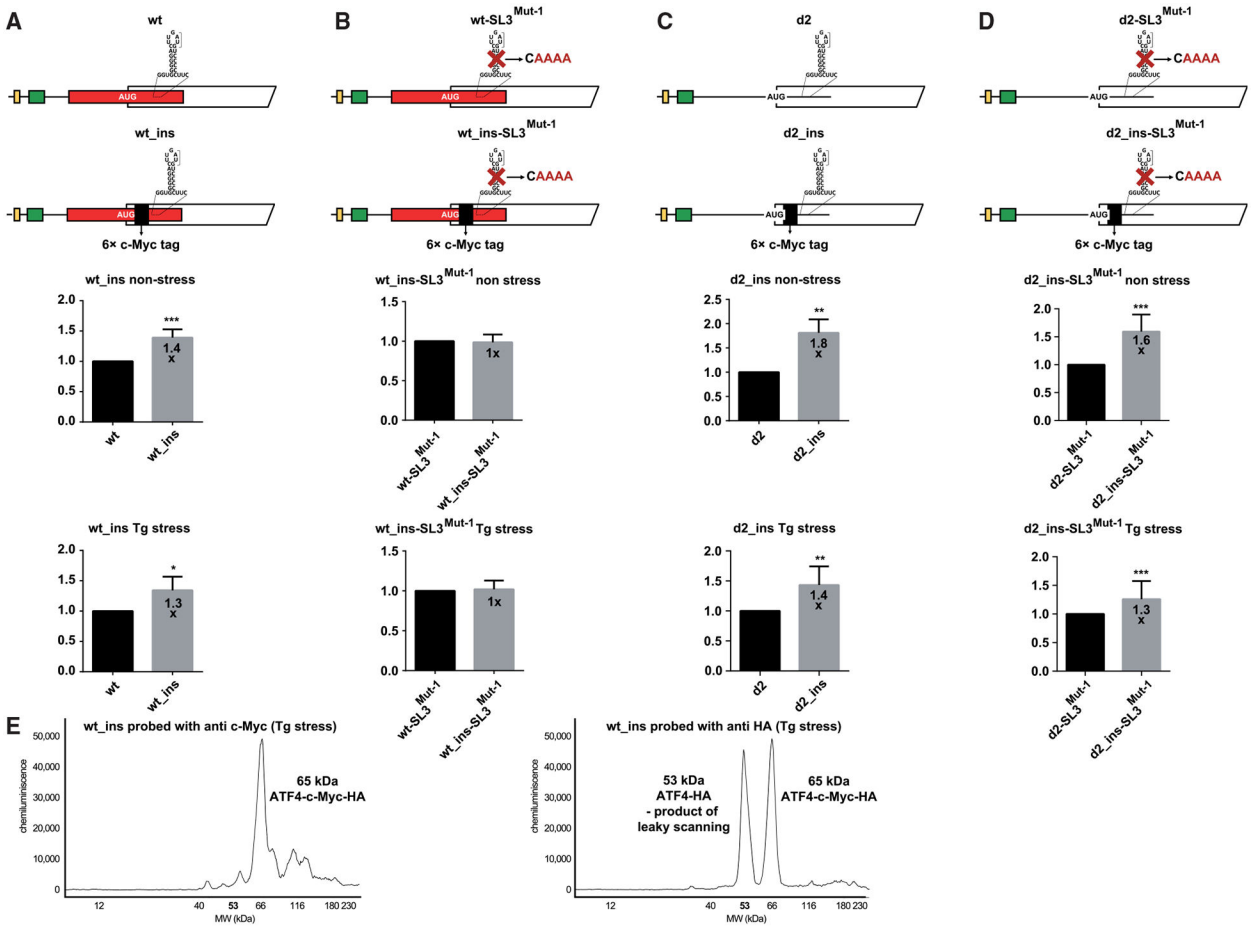


Figure 5.

Ribosome queuing and substantial leaky scanning at AUG1 of ATF4 contributes to its overall translational control

(A–D) Same as in Figure 3B except that the 63 c-Myc tag insertion in frame with *ATF4* (A), combined with the SL3^{Mut-1} mutation (B), in the otherwise WT construct (A and B) vs. the construct lacking uORF2 (C and D), all depicted at the top of the corresponding panels, were subjected to Jess analyses. The differences between experimental groups were tested by a t test, except d2_ins_SL3^{Mut-1} Tg stress, where a Mann-Whitney test was used ($n \geq 4$).

(E) The first AUG of the *ATF4* ORF is robustly leaky scanned. The electropherograms of the *ATF4* construct bearing the 63 c-Myc tag in-frame insertion probed with the anti-c-Myc (left) and anti-HA (right) antibodies are shown. For details, see the main text.

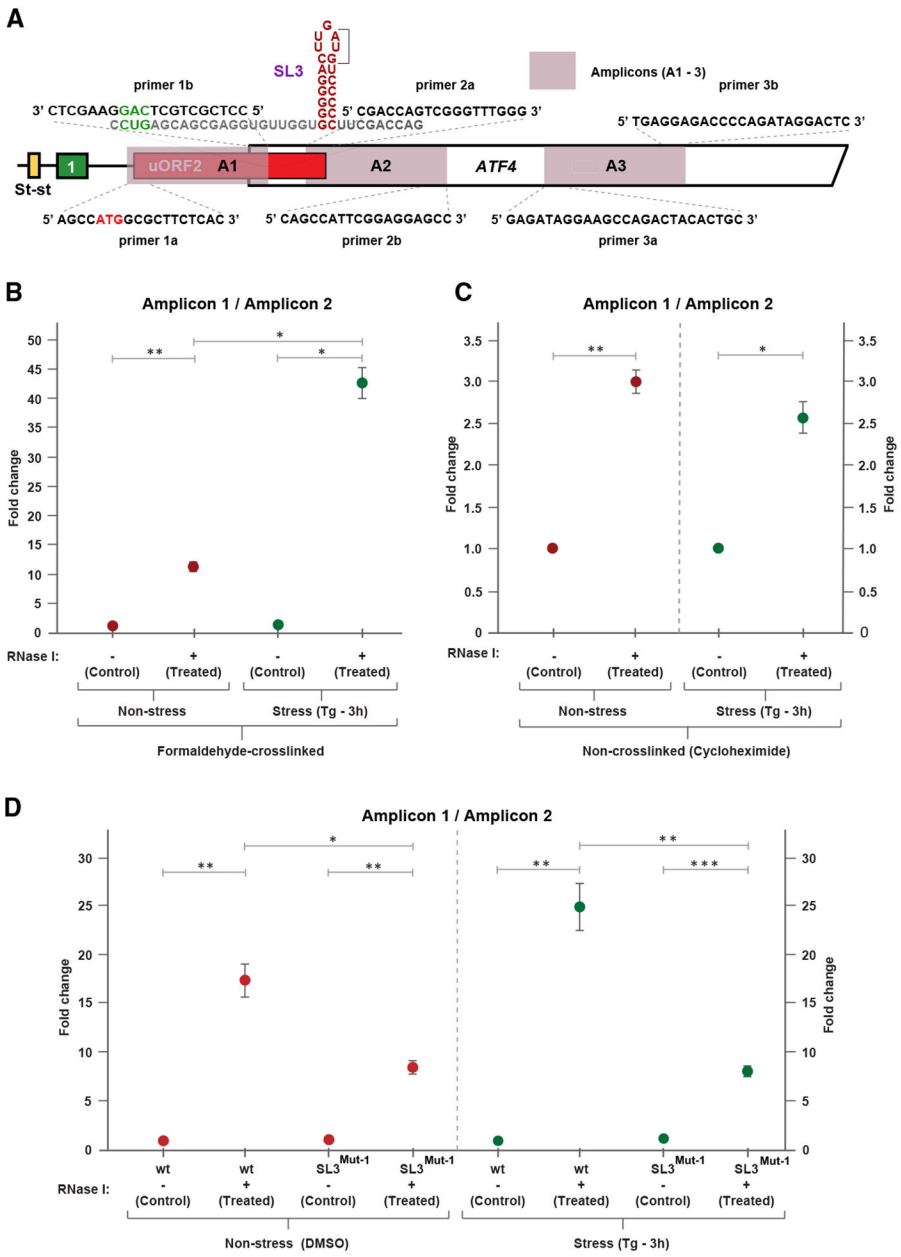


Figure 6.

Ribosome protection assay demonstrating that SL3 pauses ribosomes and prompts their queuing under both non-stress and stress conditions

(A) Schematic showing the *ATF4* mRNA with the sequences of three primer pairs amplifying three different amplicons (A1–A3; also indicated in Figure S6E): 1a-1b (the latter shown in the 3′ to 5′ direction for better illustrative purposes) for the putative queuing fragment and 2a-2b and 3a-3b for two control fragments downstream of SL3 used in the ribosome protection assay.

(B) HEK293T cells were cross-linked with formaldehyde (HCHO) and then subjected to the ribosome protection assay as described in STAR Methods. qPCR product levels of the recovered putative queuing region (A1) are normalized to the region immediately

downstream of SL3 (A2) as well as to the internal RNA isolation control (SPIKE) with the non-stress values set to 1. Results are representative of three independent replicates, and values are expressed as mean ± SD. Statistical significance was assessed using unpaired two-sided t test (*p < 0.01, **p < 0.001) with Bonferroni correction.

(C) HEK293T cells were treated with cycloheximide (a non-cross-linking agent) and then subjected to the ribosome-protection assay as described in STAR Methods. Results from three independent replicates were analyzed as described in (B) with non-stress values set to 1 (*p < 0.01, **p < 0.001).

(D) HEK293T cells were transiently transfected with plasmids carrying either WT or SL3-mutated (in $SL3^{Mut-1}$) *ATF4* reporters and treated as described in (B). Results from three independent replicates were analyzed as described in (B) with the WT values set to 1 (*p < 0.01, **p < 0.001, ***p < 0.0001).

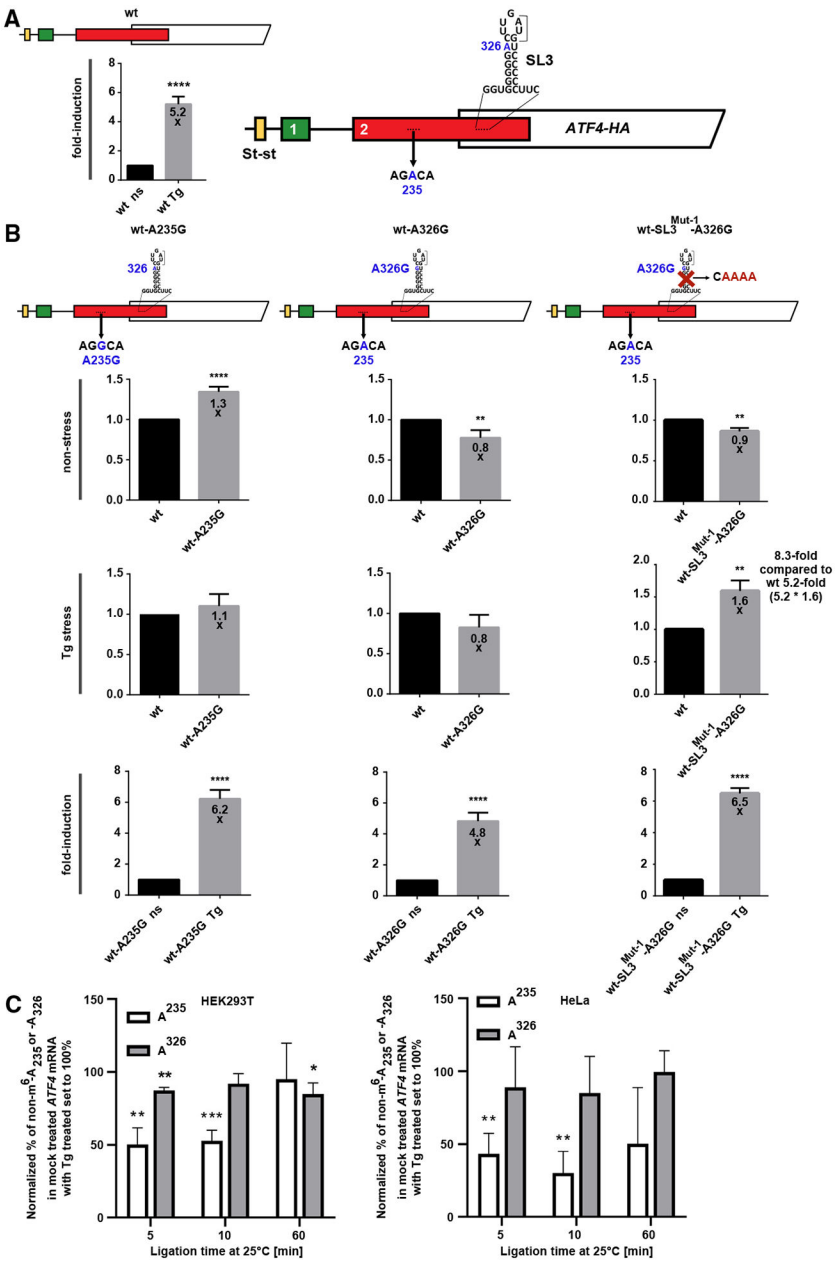


Figure 7.

mRNA methylation further fine tunes ATF4 translation

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

(B) Same as Figure 3B except that A235G (left) and A326G either alone (center) or in combination with SL3^{Mut-1} (right) *ATF4* mutant constructs depicted at the top of the corresponding panels were subjected to Jess analyses (n ≥ 3).

(C) mRNA fragments prepared from either HEK293T (left) or HeLa (right) cells carrying the A235G and A326G mutations were subjected to the T3 ligation assay as described in the main text. The normalized percentage of unmodified A_{235} or A_{326} bases of the *ATF4* mRNA

expressed in mock-treated vs. Tg-treated cells, with the latter set to 100%, was plotted as shown ($n \geq 4$).

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti-HA tag [HA.C5]	abcam	Cat# ab18181; RRID:AB_444303
Rabbit monoclonal anti-ATF-4 (D4B8)	Cell Signaling	Cat# 11815S; RRID:AB_2616025
Mouse monoclonal anti-ATF4 (B-3)	Santa Cruz	Cat# sc-390063; RRID:AB_2810998
Mouse monoclonal anti-*c*-Myc tag [9E10]	Dr. Tomáš Vomastek	N/A
Bacterial and virus strains		
DH5alpha	Invitrogen	Cat# EC0112
DH10B	Invitrogen	Cat# 18290015
Chemicals, peptides, and recombinant proteins		
Dimethyl sulfoxide (DMSO)	Sigma	Cat# D2438-5X
Thapsigargin	Invitrogen	Cat# T7458
Tunicamycin	Sigma-Aldrich	Cat# T7765-1MG
Glo Lysis Buffer	Promega	Cat# E2661
TurboFect	Thermo Fisher	Cat# R0532
RNA blue	Top-Bio	Cat# R013
UltraPure Distilled Water	Thermo Fisher	Cat# 109977-035
Chloroform	VWR BDH Chemicals	Cat# 22711.290
2-Propanol	VWR BDH Chemicals	Cat# 20842.312
Ethanol 96% vol	VWR BDH Chemicals	Cat# 20822.290
Glycogen	Thermo Fisher	Cat# AM5910
High-capacity cDNA reverse Transcription Kit	Applied Biosystems	Cat# 4368814
RNase Inhibitor	Applied Biosystems	Cat# N8080119
TURBO DNase	Invitrogen	Cat# 2238G
10×TURBO DNase Buffer	Invitrogen	Cat# 8167G
DNase Inactivation Reagent	Invitrogen	Cat# 8174G
HOT FIREPol EvaGreen qPCR Mix Plus 5×	Solis BioDyne	Cat# 08-25-00020
10×ThermoPol Reaction Buffer	NEB	Cat# B9004S
Vent DNA Polymerase	NEB	Cat# M0254L
1x T3 ligation buffer	NEB	Cat# B0535S

REAGENT or RESOURCE	SOURCE	IDENTIFIER
T3 DNA ligase	NEB	Cat# M0317L
Plasmid Buffer P1	QIAGEN	Cat# 19051
RNase A	QIAGEN	Cat# 19101
Plasmid Buffer P2	QIAGEN	Cat# 19052
Plasmid Buffer N3	QIAGEN	Cat# 19064
Plasmid Buffer PE	QIAGEN	Cat# 19065
Plasmid Buffer PB	QIAGEN	Cat# 19066
Buffer QG Solubilization Buffer	QIAGEN	Cat# 1014876
T4 DNA Ligase	Roche	Cat# 10716359001
Ligation Buffer, 10×	Roche	Cat# 11243292001
Quick CIP	NEB	Cat# M0525S
Midori Green Direct	NIPPON Genetics EUROPE	Cat# MG06
BstBI	NEB	Cat# R0519S
HindIII-HF	NEB	Cat# R3104S
HpaI	NEB	Cat# R0105S
EcoRV-HF	NEB	Cat# R3195S
PstI-HF	NEB	Cat# R3140S
rCutSmart Buffer	NEB	Cat# B6004S
Criterion TGX Precast Gels 4–20%	BIO-RAD	Cat# 5671093
SuperSignal West Femto Maximum Sensitivity Substrate	Thermo Fisher	Cat# 34096
Bio-Rad Protein Assay Dye Reagent Concentrate	BIO-RAD	Cat# 5000006
GeneArt Strings DNA Fragments	Invitrogen	N/A
EcoSpin Spin Colums for DNA	Epoch Life Science	Cat# 1910-25
SUPERaseIn RNase inhibitor	Invitrogen	Cat# AM2694
Ambion™ RNase I	Invitrogen	Cat# AM2294
Tissue culture test plate, 6 wells	Techno Plastic Products	Cat# 92406
Critical commercial assays		
12-230 kDa Fluorescence Separation Module	Protein Simple bio-techne	SM-FL004-1
8×EZ Standard Pack 1	Protein Simple bio-techne	PS-ST01EZ-8
Anti-Mouse Detection Module	Protein Simple bio-techne	DM-002
Protein Normalization Module	Protein Simple bio-techne	DM-PN02

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Analyzed data	Ichihara, et al., 2021[30]	GSE174329
Analyzed data	Rendleman et al., 2018[31]	GSE113171
Analyzed data	Andreev, et al., 2015[24]	GSE55195
Experimental models: Cell lines		
HEK293T	ATCC	CRL-3216; RRID: CVCL_0063
HeLa	ATCC	CCL-2; RRID: CVCL_0030
Oligonucleotides		
See Data S3 for a list of oligonucleotides		N/A
Recombinant DNA		
pCMV-EGFP-N2	Roithova et al., 2020[42]	N/A
WT *ATF4-HA* Tag	This paper	N/A
dSt-st	This paper	N/A
d1	This paper	N/A
d2	This paper	N/A
d-all	This paper	N/A
St-st-only	This paper	N/A
uORF1-only	This paper	N/A
uORF2-only	This paper	N/A
dSt-st-NC1	This paper	N/A
St-st-only-NC1	This paper	N/A
St-st-only-NC2/3	This paper	N/A
dSt-st-SL1	This paper	N/A
St-st-only-SL2-1	This paper	N/A
St-st-only-SL2-2	This paper	N/A
dSt-st-SL2-1	This paper	N/A
dSt-st-SL2-2	This paper	N/A
d-all-SL2-2	This paper	N/A
wt-A235G	This paper	N/A
wt-SL3^{Mut-1}	This paper	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
wt-SL3^{Mut-2}	This paper	N/A
dSt-st-SL3^{Mut-1}	This paper	N/A
d-all-SL3^{Mut-1}	This paper	N/A
uORF2-only-SL3^{Mut-1}	This paper	N/A
wt-AUG3^{Mut} (3rd *ATF4* AUG)	This paper	N/A
dSt-st-AUG3^{Mut} (3rd *ATF4* AUG)	This paper	N/A
wt-UUG^{Mut}-AUG3	This paper	N/A
wt-SL3^{Mut-1}-AUG3^{Mut}	This paper	N/A
wt-A326G	This paper	N/A
wt-SL3^{Mut-1}-A326G^{Mut}	This paper	N/A
wt-3AUG^{Mut}	This paper	N/A
d2-3AUG^{Mut}	This paper	N/A
3AUG^{Mut}-stop	This paper	N/A
3AUG^{Mut}-short	This paper	N/A
3AUG^{Mut}-short-stop	This paper	N/A
uORF2_ins	This paper	N/A
d2_ins (uORF2_ins => AUG to AGG)	This paper	N/A
uORF2_ins_3AUG^{Mut}	This paper	N/A
wt-CUG^{Mut}	This paper	N/A
wt-SL3^{Mut-1}-CUG^{Mut}	This paper	N/A
wt_ins	This paper	N/A
wt_ins-SL3^{Mut-1}	This paper	N/A
d2_ins	This paper	N/A
d2_ins-SL3^{Mut-1}	This paper	N/A
uORF2_ins-SL3^{Mut-1}	This paper	N/A
uORF2-*ATF4*-HA fusion	This paper	N/A
Software and algorithms		
Compass for SW ver. 6.3.0	ProteinSimple	https://www.bio-techne.com/resources/instrument-software-download-center

REAGENT or RESOURCE	SOURCE	IDENTIFIER
GraphPad Prism ver. 10.1.2	GraphPad Software, San Diego, CA, USA	RRID:SCR_002798
SciPy ver. 1.10.1	Virtanen et al., 2020[43]	RRID:SCR_008058
Python 3.11.3 ver. 8.1.1	Rossum and Drake, 2009[44]	https://www.python.org
Matplotlib ver. 3.7.0	Hunter, 2007[45]	RRID:SCR_008624
R ver. 4.2.2: 'pheatmap' package ver. 1.0.12	Kolde and Kolde, 2015[46]	N/A
R ver. 4.2.2: 'RiboWaltz' package	Lauriaetal., 2018[47]	N/A
Other		
Jess™ Automated Western Blot System	Protein Simple Bio-Techne	004-650
RiboCrypt tool	https://ribocrypt.org	N/A
G-Box imager	Syngene	N/A
Mendeley data	http://www.mendeley.com/	RRID:SCR_002750; https://doi.org/10.17632/zmdvhwggg8.1
Quantity One Software	Bio-Rad	RRID:SCR_014280