



HAL
open science

Translation inhibitory elements from Hoxa3 and Hoxa11 mRNAs use uORFs for translation inhibition

Fatima Alghoul, Laure Schaeffer, Gilbert Eriani, Franck Martin

► To cite this version:

Fatima Alghoul, Laure Schaeffer, Gilbert Eriani, Franck Martin. Translation inhibitory elements from Hoxa3 and Hoxa11 mRNAs use uORFs for translation inhibition. *eLife*, 2021, 10, 10.7554/eLife.66369 . hal-03257391

HAL Id: hal-03257391

<https://hal.science/hal-03257391>

Submitted on 10 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Translation Inhibitory Elements from *Hoxa3* and *Hoxa11* mRNAs use uORFs for translation inhibition

Fatima ALGHOUL, Laure SCHAEFFER, Gilbert ERIANI, Franck MARTIN*

Institut de Biologie Moléculaire et Cellulaire, “Architecture et Réactivité de l’ARN” CNRS UPR9002, Université de Strasbourg, 2 Allée Konrad Roentgen Strasbourg (France)

*Corresponding authors

Email for correspondence: f.martin@ibmc-cnrs.unistra.fr

Running title: *hoxa3* and *hoxa11* TIE translation inhibition

Keywords: Hox mRNA, Translation Inhibitory Elements, IRES, uORF

Abstract

During embryogenesis, Hox mRNA translation is tightly regulated by a sophisticated molecular mechanism that combines two RNA regulons located in their 5'UTR. First, an Internal Ribosome Entry Site (IRES) enables cap-independent translation. The second regulon is a Translation Inhibitory Element or TIE, which ensures concomitant cap-dependent translation inhibition. In this study, we deciphered the molecular mechanisms of mouse *Hoxa3* and *Hoxa11* TIE elements. Both TIEs possess an upstream Open Reading Frame (uORF) that is critical to inhibit cap-dependent translation. However, the molecular mechanisms used are different. In *Hoxa3* TIE, we identify a uORF which inhibits cap-dependent translation and we show the requirement of the non-canonical initiation factor eIF2D for this process. The mode of action of *Hoxa11* TIE is different, it also contains a uORF but it is a minimal uORF formed by an uAUG followed immediately by a stop codon, namely a 'start-stop'. The 'start-stop' sequence is species-specific and in mice, is located upstream of a highly stable stem loop structure which stalls the 80S ribosome and thereby inhibits cap-dependent translation of *Hoxa11* main ORF.

Introduction

Gene expression constitutes an indispensable cellular process for which the genetic information encodes a functional product, mainly proteins. This process named translation initiates by a cap-dependent mechanism for most cellular mRNAs. It involves a large number of auxiliary proteins termed eukaryotic Initiation Factors (eIFs) which are required for the recruitment of the ribosomes on the mRNA (Hinnebusch, 2014; Merrick and Pavitt, 2018; Pelletier and Sonenberg, 2019; Shirokikh and Preiss, 2018). To ensure fine-tuning of translation, this step is highly regulated. However, several mRNA subclasses are translated by non-canonical mechanisms. For instance, this is the case for homeobox (Hox) mRNAs. Hox genes encode a family of proteins that constitutes transcription factors. Their main function is to orchestrate specific sequential transcription processes during embryonic development. A wealth of experimental data over the last three decades has led to the identification of many *cis*-regulatory elements that control Hox gene transcriptional patterns, thus giving deeper insights into the expression of Hox mRNAs (Alexander et al., 2009). In fact, Hox homologues in *Drosophila*, precisely Antp and Ubx genes in the Hom-C cluster, have been suggested to be under translational control during embryonic development (Oh et al., 1992). A subgroup of mRNAs produced from the Antp and Ubx loci contain functional Internal Ribosome Entry Sites (IRES) that allow their translation using a cap-independent mechanism during development (Ye et al., 1997). More recently, the presence of other IRES elements in the 5'UTR of subsets of mice HoxA mRNAs (*Hoxa3*, *Hoxa4*, *Hoxa5*, *Hoxa9* and *Hoxa11*) have been demonstrated (Leppék et al., 2020; Xue et al., 2015). Some of these IRESes require the presence of the ribosomal protein Rpl38 in the ribosome to efficiently initiate translation, thereby explaining the tissue patterning defective phenotype observed with Rpl38 knockout mouse (Kondrashov et al., 2011). The IRES activity in these Hox mRNAs was found to be critical for their appropriate expression. Upon the discovery of the IRES elements in Hox mRNAs, other RNA regulons termed Translational Inhibitory Elements (TIE) were also found in the 5'UTR of Hox mRNAs (Xue et al., 2015). TIEs are located upstream of the previously described IRES. According to the study by Xue et al., these elements efficiently inhibit canonical cap-dependent translation in subsets of HoxA mRNAs (*Hoxa3*, *Hoxa4*, *Hoxa9* and *Hoxa11*) by an unknown mechanism (Xue et al., 2015). The action of TIE that ensures efficient blockage of cap-dependent translation, promotes IRES-mediated cap-independent translation. TIE and IRES act in synergy to ensure tightly regulated translation during organismal development. Indeed, it has been shown that Hox TIE elements ensure that

Hox mRNAs are translated solely by their IRES element. Thereby, TIE elements represent the first example of specific RNA elements dedicated to inhibiting specifically cap-dependent translation in Hox mRNAs. However, the mechanism of action of these elements as well as their structural characterization are still unknown. In this study, we investigate the functional mode of action of TIE elements from *Hoxa3* and *Hoxa11* mRNAs that were previously identified and characterized by *Xue et al.* (Xue et al., 2015). Our study was conducted on *Hoxa3* and *Hoxa11* for several reasons. First, among Hox mRNAs, the translation mechanism used for each is distinct (Xue et al., 2015), *Hoxa11* mRNA translation being RpL38-dependent while *Hoxa3* is RpL38-independent. Second, these two TIE elements have very efficient translation inhibition rate and third, their relatively short lengths were more suitable for our experiments. First, we determined their secondary structure by chemical probing assays and then, using cell-free translation extracts and *in vivo* assays, we deciphered their mode of action. Interestingly, the translation inhibitory mechanism that is mediated by *Hoxa3* TIE element is radically distinct from the one used by *Hoxa11*. *Hoxa3* TIE contains a uORF that is translated into a 9 KDa protein through the *Hoxa3* IRES and that requires the presence of the non-canonical translation initiation factor eIF2D. Indeed, eIF2D has been shown to be involved in diverse functions from translation initiation on specific mRNAs to reinitiation and recycling (Dmitriev et al., 2010; Skabkin et al., 2010; Weisser et al., 2017). On the contrary, *Hoxa11* TIE contains a long stable hairpin structure which is preceded by a start-stop codon combination. In mice, these three elements enable a highly efficient inhibition of cap-dependent translation by *Hoxa11* TIE that is achieved through a start-stop stalling mechanism of an 80S ribosome.

Results

TIE-mediated inhibition recapitulated in RRL

To dissect the mechanism of TIE-mediated inhibition, we first set-up an efficient cell-free assay. We used RRL and performed *in vitro* experiments using several reporter constructs. To analyse their translation efficiency, capped mRNAs were translated in RRL in the presence of [³⁵S]-methionine to verify the size of the expected synthesized protein and luciferase activity was measured to determine their translation efficiency. In order to standardize the measured translation activities, we inserted the TIE elements identified and characterized by Xue *et al.* (Xue *et al.*, 2015), in murine *Hoxa3* and mRNAs upstream of the well-characterized human β -globin (*Hbb-b1*) 5'UTR (**Figure 1A**). The *Hbb-b1* allows translation strictly by a cap-dependent mechanism (Fletcher *et al.*, 1990). Therefore, in order to evaluate the inhibition due to TIEs, we used a reporter mRNA containing the sole *Hbb-b1* 5'UTR placed upstream of Renilla luciferase coding sequence as a control reporter gene. First, we demonstrated that translation with these extracts is cap-dependent (**Figure 1-figure supplement 1A**). Second, we performed translation experiments with increasing amounts of the same mRNA and determined that with concentrations higher than 100 nM, the translation yield reached a plateau (**Figure 1-figure supplement 1A**). In order to avoid any titration effect due to excess of mRNA, we performed all translation assays at an mRNA concentration of 50 nM, which enables sub-saturating conditions. *Hoxa3* and *Hoxa11* TIE efficiently promote a translation inhibition by 79% and 91% respectively (**Figure 1A**). Therefore, our cell-free translation assay using RRL efficiently recapitulates the previously described TIE-mediated translation inhibition (Xue *et al.*, 2015). Likewise, TIE-mediated inhibition was efficiently recapitulated in other cell-free translation systems, which are partially or fully cap-dependent from other organisms like Wheat Germ Extracts, drosophila S2 cell extracts and human HeLa cell extracts (**Figure 1-figure supplement 1B**). Moreover, the same constructs were introduced in a plasmid that allowed us to monitor translation inhibition *in vivo*. We tested two cell lines that have been shown to express Hox genes, namely human HEK293T and murine mesenchymal stem cell line C3H10T1/2 (Phinney *et al.*, 2005). In these cells, translation inhibition by *Hoxa3* and *Hoxa11* TIE is 98% and 88% respectively (**Figure 1-figure supplement 1C**). Next, to define the minimal active domains of *Hoxa3* and *Hoxa11*, sequential 5' deletions were performed. In each experiment, translation for each construct was compared to the control (w/o TIE) (**Figure 1B**). First, we tested large deletions in order to roughly determine the functional domain (**Figure 1-figure supplement 1D**), then more

precise deletions to map exactly the 5' end of functional domains for *Hoxa3* and *Hoxa11* (**Figure 1B**). Deletion of 68 nts in *Hoxa3* TIE does not affect inhibition (81%). A further deletion of 113 nts completely abolishes the translation inhibition of *Hoxa3* TIE (10%). Therefore, the minimal *Hoxa3* element encompasses nucleotides 68 to 170. Concerning *Hoxa11* TIE, the situation is not as clear. In fact, the inhibition is completely lost when 161 nts are deleted. In this case, the deletions indicate that elements essential for translation inhibition are most likely located between residues 139 to 161. In conclusion, these experiments allowed the localisation of essential RNA elements required for *Hoxa3* and *Hoxa11* RNA regulons that are required to retain their full inhibitory function.

TIE elements have distinct secondary structures

To gain further insights into the structural and functional properties of TIE elements, we performed chemical probing using DMS and CMCT reagents for both *Hoxa3* TIE and *Hoxa11* TIE (**Figure 2**). Since modifications were performed in triplicates, the average of reactivity was calculated for each nucleotide at a specific position (**Figure 2-figure supplement 1**). With this reactivity, we built a secondary structural model for both *Hoxa3* and *Hoxa11* TIE. *Hoxa3* TIE contains a 5' proximal stem-loop and another bigger structure comprising two-way junctions (**Figure 2A**). For TIE *a11*, the higher GC content (64%) than *Hoxa3* TIE (45%) suggests more stable structure. Our probing experiments confirmed this statement. It comprises four stem-loops and a three-way junction structure (**Figure 2B**). To further validate our models, the secondary structures of both TIEs were also probed in the frame of the full-length Hox 5'UTR which contains the IRES element downstream of the TIE. The reactivities for each nucleotide obtained with isolated TIEs and with TIEs embedded in full-length 5'UTR were highly similar suggesting that the TIEs do fold independently from the IRES (**Figure 2-figure supplement 1**). Using these two models, we wanted to further characterize the structure-function relationships of both TIE elements.

***Hoxa3* TIE contains an upstream ORF that inhibits translation**

The 5' truncations experiments allowed us to pinpoint critical elements required for translation inhibition. Interestingly, the minimal *Hoxa3* contains putative uORFs starting from two putative uAUGs at positions 111-113 and 123-125 respectively (**Figure 3A**). In order to test their implication in the translation inhibition, both uAUGs were mutated independently into UAC thereby eliminating any possibility of AUG-like codon recognition. Interestingly, the mutation of AUG111 completely abolishes the inhibition by *Hoxa3* TIE thereby confirming its implication. On the contrary, this is not the case for AUG123 (**Figure 3A**). This is in good agreement with the fact that AUG111 has an optimal Kozak sequence (A at -3

and G at +4), thereby the ribosome initially recognizes it during scanning along the 5'UTR (Kozak, 1986). To further confirm the assembly of the ribosome on uAUG111, we mutated it into AUG-like codons. We tested previously described AUG-like codons such as AUU, CUG, GUG and ACG. None of the tested AUG-likes are used for translation inhibition indicating that *Hoxa3* TIE requires a genuine AUG start codon (**Figure 3-figure supplement 1A, left panel**). Since the AUG111 is involved in a stem, it is possible that the absence of inhibition with the AUG-likes is actually due to the disruption of the stem. To rule out this possibility, we mutated the AUG111 into CUG and GUG and inserted simultaneously the compensatory mutations in AUG123 into AGG and ACG respectively that enable the formation of the stem with the mutated CUG and GUG codons (**Figure 3-figure supplement 1A, right panel**). In both mutants, the inhibition was fully destroyed indicating that efficient inhibition requires a genuine AUG start codon independently of the secondary structure (**Figure 3-figure supplement 1A**). Next, we confirmed the use of uAUG111 by toe printing assay. As expected, a canonical +16 reverse transcription arrest from uAUG111 was clearly detected with GMP-PNP, a non-hydrolysable analogue of GTP that prevents subunit joining (Eliseev et al., 2018), and less efficiently with cycloheximide, a translocation inhibitor that binds the E-site of the ribosome (Garreau de Loubresse et al., 2014; Schneider-Poetsch et al., 2010). Accordingly, when the uAUG is mutated to UAC, the +16 toe-print disappears (**Figure 3B**). Altogether, this data confirm that a pre-initiation complex efficiently assembles on the uAUG111. We then wondered whether the ribosome assembled on this AUG codon proceeds to translation elongation resulting in a polypeptide synthesis encoded by the uORF. By performing sucrose gradient analysis, we detected polysome formation suggesting that efficient translation is occurring from uAUG111 (**Figure 3-figure supplement 1B**). Accordingly, mutation of uAUG111 drastically reduces the amount of polysomes and the formation of 48S complexes in the presence of GMP-PNP thereby corroborating that the translation is starting on uAUG111 (**Figure 3-figure supplement 1B**). In *Hoxa3* 5'UTR, the uORF starting from uAUG111 is extending through the full IRES. To check further whether the uORF is indeed translated through the full length *Hoxa3* 5'UTR, we first deleted a single nucleotide (G333) to change the frame and thereby producing a fusion protein formed by the peptide produced from the uORF and Renilla Luciferase. Indeed, with this single frameshifting point mutation, we could detect a longer protein demonstrating that the pre-initiation complex assembled on the uAUG111 indeed proceeds to translation elongation and is efficiently translating through the full length 5'UTR of *Hoxa3* mRNA (**Figure 4A**). We also verified the translation from uAUG111 with our reporter constructs containing only the

Hoxa3 TIE element. Likewise, the insertion of a single frameshifting nucleotide (A220) allows the detection of a longer fusion protein (**Figure 4-figure supplement 1A**). Remarkably, with a double mutant that combines the mutation of the uAUG111 to ACG with the (A220) insertion, a mixture of two proteins is detected, the fusion protein and the more abundant Renilla Luciferase protein. Indeed, this confirms that the ribosome requires the uAUG111 but still able to recognize, although less efficiently, the AUG-like ACG as a start codon (**Figure 4-figure supplement 1A**). Similarly, when we use the native full-length *Hoxa3* mRNA, we detect the *Hoxa3* uORF protein of 9 KDa size (**Figure 4-figure supplement 1B**). Altogether, our cell-free translation assays demonstrate that *Hoxa3* TIE translation is achieved by translation of a uORF that starts at AUG111 codon that extends through the whole *Hoxa3* 5'UTR. To confirm these results *in vivo*, we generated reporter constructs in the plasmid pmirGlo that contains TIE elements upstream of Renilla luciferase (hRluc). Values were normalized to that of control enhanced Firefly luciferase (luc2) to calculate the luciferase activity for each report (**Figure 4B**). As expected, wild type *Hoxa3* TIE blocks translation very efficiently in both HEK293T and C3H10T1/2 cells. When the uAUG111 is mutated to UAC, the inhibition is significantly affected in both HEK293T and C3H10T1/2 cells at respectively 57% and 45% compared to wild type *Hoxa3* TIE (**Figure 4B**). Although the inhibition is not fully abolished, these experiments confirmed that the codon AUG111 is critical for efficient translation inhibition in *Hoxa3 in vivo*.

***Hoxa11* TIE-mediated inhibition is mediated by a stalled 80S ribosome**

We next asked whether *Hoxa11* TIE has a similar inhibitory mechanism. Deletion experiments suggested that critical elements were located in the region 139 to 216. We also found that *Hoxa11* TIE contains two putative upstream AUGs at positions 84-86 and 159-161. Mutations of both AUG84 and AUG159 had no impact on translation inhibition (**Figure 5A**). According to our 2D model, a long GC-rich stable stem loop (SL) structure ($\Delta G = -25.00$ kcal/mol) spans nucleotides 104-154 (**Figure 2B**). This long hairpin comprises sixteen G-C base pairs that can putatively interfere with the progression of a pre-initiation scanning complex. To test the inhibitory efficiency of this stem-loop, we transplanted it in a strictly cap-dependent reporter mRNA containing the 5'UTR of *Hbb-b1* upstream of the luciferase coding sequence. The *Hoxa11* SL was inserted in the middle of the 5'UTR. In this construct the first 25 nucleotides from the 5' proximity are unfolded thereby ensuring proper access to the 5' cap. Interestingly, the translation of this mRNA was significantly abolished showing that this stem loop on its own is sufficient to inhibit cap-dependent translation when transplanted in another mRNA (**Figure 5-figure supplement 1**). Strikingly, in *Hoxa11* TIE,

uAUG84 codon is located 19 nts upstream of the inhibiting SL. This distance is compatible with the assembly of a pre-initiation complex on uAUG84 without clashing with the SL. Moreover, the distance between the uAUG and the SL is optimal for favouring AUG recognition by scanning arrest of the pre-initiation complex forced by the SL. Importantly, the uAUG84 is immediately followed by a stop codon UAG87. The sequence context of the uAUG84 (A at -3 and U at +4) is suboptimal compared to the consensus Kozak sequence. This unique combination of start-stop codon upstream of stem loop structure raised the question of whether the ribosome is forced to recognize this uAUG despite a suboptimal Kozak context. To address this hypothesis, we mutated the stop codon UAG87 to UGG thereby creating a uORF. When the stop is mutated, a small peptide is produced from the translation of uAUG84 through 5'UTR, which we called *Hoxa11* uORF (**Figure 5B**). This experiment demonstrates that uAUG84 is indeed efficiently used as a start codon despite its sub-optimal sequence context. Unfortunately, the presence of the highly stable SL is not compatible with the toe-printing assay. Indeed, premature RT arrests occur due to the presence of the highly stable SL, rendering the toe-printing assay on TIE *Hoxa11* impossible. Therefore, to further confirm that the ribosome is efficiently assembled on uAUG84, we performed sucrose gradient analysis with radiolabelled mRNAs. With wild type *Hoxa11* TIE, an 80S complex efficiently accumulates in the presence of cycloheximide as expected. However, an 80S complex is also detected without inhibitor indicating that the 80S complex is in fact a stalled ribosome (**Figure 5C, Figure 5-figure supplement 2**). The mutation of uAUG84 to UAC drastically reduces the amount of 80S complex; in contrast, mutation of uAUG159 to UAC does not affect 80S accumulation. This further confirms that the stalled 80S complex is indeed assembled on the uAUG84. Altogether, this data show that a stalled ribosome is indeed assembled on uAUG84 and the stalling is caused by the synergistic effect of a stop codon next to the AUG and a stable SL downstream of the start-stop of *Hoxa11*.

Mass spectrometry analysis pre-initiation complexes programmed by *Hoxa3* and *Hoxa11* TIE

To further characterize the two different modes of action employed by *Hoxa3* and *Hoxa11*, we identified the factors specifically acting in such mechanisms. For that, we purified pre-initiation complexes programmed by *Hoxa3* and *Hoxa11* TIE suitable for Mass Spectrometry analysis. Briefly, ribosomes were assembled on chimeric biotinylated mRNA-DNA molecules and immobilized on streptavidin-coated beads after incubation with RRL in the presence of cycloheximide or GMP-PNP. Complexes were then eluted by DNase treatment as previously described (Chicher et al., 2015; Prongidi-Fix et al., 2013) and analysed by Mass

Spectrometry. Three different mRNA constructs were used, *Hoxa3* TIE upstream of 5'UTR *Hbb-b1*, *Hoxa11* TIE upstream of 5'UTR *Hbb-b1*, and 5'UTR *Hbb-b1* as a negative control. Comparison between the three mRNAs blocked with cycloheximide enabled the identification of specific factors interacting with each RNA (**Figure 6**). Interestingly, for *Hoxa3* TIE, we identified a set of translation-related proteins. Among the strongest hits, we found eIF2D, a non-canonical GTP-independent initiation factor which has been shown to be involved in the initiation step on specific mRNAs (Dmitriev et al., 2010; Vaidya et al., 2017; Vasudevan et al., 2020), reinitiation on main ORF (Ahmed et al., 2018; Weisser et al., 2017) and recycling after translation termination (Skabkin et al., 2013, 2010; Young et al., 2018). Another interesting hit is methionine aminopeptidase MetAP1 which removes N-terminal methionine from nascent proteins in a co-translational manner (Varland et al., 2015).

Other factors that are linked to translation have been selected such as the scanning factor eIF1A and its isoform eIF1A-X, eIF3a, Arginyl- and Leucyl-tRNA synthetases QARS and LARS, DEAD-box helicases DHX36 and DDX39B, elongation factor HBS1L and RpL38 ribosomal protein. For *Hoxa11* TIE, we identified different translation-related proteins among which are ASAP1, a GTPase activator protein, MetAP1, RpL38, Valyl-tRNA synthetase and eIF3j, another subunit of initiation factor eIF3 usually dissociating at early stages of initiation to allow mRNA entry (Aylett et al., 2015; Fraser et al., 2007; Young and Guydosh, 2019) thereby, unlikely to be present in initiation complexes. When comparing *Hoxa3* TIE with *Hoxa11* TIE, we could detect some initiation factors and translation-related proteins specific for *Hoxa3* like eIF2D, eIF1A, eIF1A-X, eIF5B, LARS, DHX21 and HBS1L (**Figure 6, Supplemental file 1**). Interestingly, HBS1L is enriched in both *Hoxa3* and *Hoxa11* cycloheximide blocked complexes. HBS1L is a member of translational GTPase family which transports Pelota to stalled ribosomes with an empty A-site or an mRNA-occupied A-site in a codon independent manner (Lee et al., 2007; Shao et al., 2016; Shoemaker et al., 2010). Another interesting hit is eIF5B, this factor catalyses the joining of the large ribosomal subunit independently of its GTPase activity (Wang et al., 2019). The GTPase activity is only required for eIF5B release from the assembled 80S and thereby allow to proceed to elongation. EIF5B assists the correct positioning of Met-tRNA^{Met} in the P-site (Wang et al., 2020). Interestingly, RpL38 protein is enriched in both TIE-programmed complexes. In fact, all the ribosomal proteins are present in the purified complexes (**Figure 6, Supplemental file 1**) so this enrichment might be due to free RpL38 outside of the ribosome particle. This is particularly interesting because RpL38 is involved for Hox IRES elements (Xue et al., 2015).

Altogether, these results show a variation in the translation factors involved which hints two distinct TIE-mediated inhibitions.

Similarly, we purified programmed pre-initiation complexes blocked by GMP-PNP (**Figure 6-figure supplement 1, Supplemental file 2**). By comparing *Hoxa3* to either *Hbb-b1* or *Hoxa11* mRNAs, we also found an enrichment of eIF2D as *Hoxa3*-specific factor. Interestingly, we also found an enrichment of PKR protein (also called eIF2AK2) for both *Hoxa3* and *Hoxa11* (**Figure 6-figure supplement 1**). PKR is known to bind double-stranded RNA during viral infection which mediates its auto-activation and induces the phosphorylation of eIF2 α subunit (Adomavicius et al., 2019). This leads to the inhibition of mRNA translation. The enrichment of both eIF2D and PKR for *Hoxa3* raised the question of whether eIF2D is indeed mediating the translation of uORF by an alternative mechanism. To address this hypothesis and the involvement of eIF2D, we conducted a set of *in vivo* experiments.

Translation inhibition by TIE requires eIF2D

After specifically identifying eIF2D by Mass Spectrometry analysis for *Hoxa3* TIE-ribosomal complex, we were interested in getting more insights into how this factor might be involved. Previous studies have shown that eIF2D is a non-canonical translation initiation factor which delivers tRNA to the P-site of the ribosome in a GTP-independent manner (Dmitriev et al., 2010). It has been shown to be involved in the initiation step on specific mRNAs (Dmitriev et al., 2010; Vaidya et al., 2017; Vasudevan et al., 2020), reinitiation on main ORF (Ahmed et al., 2018; Weisser et al., 2017). Additionally, other studies showed that eIF2D is required for recycling after translation termination (Skabkin et al., 2013, 2010; Young et al., 2018). To confirm the involvement of eIF2D in *Hoxa3* TIE-mediated inhibition mechanism, we performed a co-transfection assay in HEK293T cell line using siRNA against either eIF2D or non-target pool of siRNAs (siRNA-NT) together with a reporter plasmid harbouring *Hoxa3* TIE. The plasmid used harbours two reporter genes, Renilla luciferase to test the impact of TIEs and Firefly luciferase for normalization purposes (**Figure 7A**). As a control, we used reporter plasmids with *Hoxa11* TIE and the same plasmid without TIE. After 48 hours incubation, cells were lysed for western blot analysis and luciferase assay (**Figure 7A**). For normalisation, Renilla luciferase values were compared to that of Firefly luciferase (RL/FL) in each construct. To determine the effect of eIF2D silencing, the RL/FL ratio was compared to that with the siRNA-NT, which serves as a control for siRNA effect on the cell line. Thereby, the comparison between the two values would determine the effect of eIF2D silencing with each TIE. Interestingly, silencing eIF2D in *Hoxa3*-transfected cells at an

efficiency of 76% drastically abolishes *Hoxa3*-mediated inhibition and RL/FL ratio (0.06) increases 6 times compared to that in siRNA-NT cells (0.01) (**Figure 7A**). On the contrary, the silencing of eIF2D at efficiency of 81% has no significant effect on *Hoxa11*-mediated inhibition (RL/FL ratio of 0.02 with both siRNAs). Similarly, silencing eIF2D at 83% efficiency has no effect on general translation (RL/FL ratio remains around 0.06 with both siRNAs). Altogether, this confirms the specific requirement of eIF2D for efficient *Hoxa3* inhibition but not for *Hoxa11*. This conclusion is in good agreement with our previous Mass Spectrometry analysis of *Hoxa3*-ribosomal complexes blocked by translation inhibitors, which show the specific presence of eIF2D in the initiation complexes programmed with *Hoxa3* TIE (**Figure 6, Figure 6-figure supplement 1**). The fact that we found eIF2D in pre-initiation complexes suggests that eIF2D is involved in the initiation of *Hoxa3* uORF. Since *Hoxa3* uORF is long and extends downstream the AUG start codon of the main CDS in native *Hoxa3* mRNA, which means that the *Hoxa3* uORF partially overlaps the main CDS, a mechanism using a reinitiation event after *Hoxa3* uORF translation is impossible. Therefore, we rather favour a model in which eIF2D is required for *Hoxa3* initiation. Previous studies have shown that an A-motif upstream of an uAUG has been shown to be important for proper eIF2D recruitment (Dmitriev et al., 2010). Interestingly, a closer look at the sequences in the *Hoxa3* TIE uORF revealed A-rich motifs upstream and downstream of uAUG111. We tested the implication of both A-motifs, AAAA107 upstream of the AUG and AAAAA147 downstream of the AUG, in *Hoxa3* TIE-mediated inhibition (**Figure 7B**). In order to avoid any side effect due to mutation in the sequence context around the uAUG, we kept an optimal Kozak sequence and mutated the As at position 107-110 to GGCC thereby keeping a purine residue at -3 position. The second A-motif downstream of AUG was similarly mutated to GGCCC. Interestingly, the mutation of upstream A-motif had a 2-fold reduction effect on translation inhibition of *Hoxa3* TIE. In contrast, mutation of the downstream A-motif does not affect *Hoxa3*TIE inhibition (**Figure 7B**). Additionally, we confirmed the implication of upstream A-motif *in vivo* with HEK293T cell line using plasmids with wild type *Hoxa3* and the mutant of AAAA107 into GGCC (**Figure 7B**). The AAAA107 mutant reduces the inhibitory efficiency of *Hoxa3* TIE by 2.5 times compared to wild type *Hoxa3*. Therefore, we show that the A-motif upstream of the AUG is critical for translation inhibition and most probably because it is required for eIF2D recruitment.

Discussion

Our study has shown that two HoxA mRNAs, *Hoxa3* and *Hoxa11*, are regulated by different mechanisms to ensure the inhibition of cap-dependent translation and allowed us to propose two distinct models for their mode of action (**Figure 8**). First, we have shown that TIE elements can function *in vitro* using cell-free translation extracts. We then confirmed the results obtained with these extracts using *in vivo* assays in several cell lines. Our findings suggest that *Hoxa3* inhibits translation by a uORF which is translated through the whole 5'UTR of *Hoxa3* mRNA producing a small protein of size 9 KDa. Interestingly, the alignment of *Hoxa3* TIE element shows a conservation of the uAUG111 (highlighted in red box) among different species. In contrast to the localisation of the uAUG that is highly conserved, the coding sequence of the uORF is not conserved among species (**Figure 8-figure supplement 1A**). Indeed, uORFs have been recognized as regulators of translation for number of cellular mRNAs (Barbosa et al., 2013). For instance, four uORFs in the 5' leader of yeast *GCN4* mRNA restrict the flow of scanning ribosomes from the cap site to the *GCN4* initiation codon (Dever et al., 2016; Hinnebusch, 1993). In plants, the uORF in *AdoMetDC* mRNA generates a nascent hexapeptide that interacts with its translating ribosome to suppress translation of *AdoMetDC* RNA in a cell-specific manner (Uchiyama-Kadokura et al., 2014). Interestingly, our data indicates that eIF2D is required in *Hoxa3* TIE mode of action (**Figure 6 & 7A**). A Recent study has shown that drosophila ATF4 mRNA translation is induced by eIF2D and its homologue DENR during integrated stress response (Vasudevan et al., 2020). In this case, eIF2D requires its RNA binding motif to mediate translation of ATF4 mRNA through its 5' leader sequence consisting of multiple uORFs (Vasudevan et al., 2020). Moreover, it has been shown that eIF2 α -phosphorylation during stages of embryonic development promotes translation from uORFs (Friend et al., 2015). Therefore, canonical cap-dependent translation initiation with eIF2 is not possible during embryonic development. The cap-dependent translation initiation of uORF from *Hoxa3* TIE might use eIF2D as an alternative to replace inactive phosphorylated eIF2 to promote uORF translation. So far, the only *cis*-acting sequence that has been clearly defined on an mRNA for specific eIF2D recruitment is an A-rich motif upstream of the start codon (Dmitriev et al., 2010). We showed that *Hoxa3* TIE contains such an A-motif upstream of the uORF that is critical for TIE function (**Figure 7**). Other reports showed that eIF2D would form initiation complexes on leaderless and A-rich 5'UTRs (Akulich et al., 2016; Dmitriev et al., 2010). In *Hoxa3* TIE, mass spectrometry analysis enabled us to demonstrate that eIF2D is present only in pre-

initiation complexes programmed with *Hoxa3* TIE. Importantly, two more factors, eIF5B and HBS1L have been specifically found in *Hoxa3* complexes, which would be interesting to investigate their involvement in eIF2D-dependent mechanism (**Figure 6**). Regarding *Hoxa11* TIE, the ribosome recognizes a combination of two *cis*-acting elements in the 5'UTR (**Figure 8**). (i) A start-stop codon combination located at positions 84-89 and (ii) a long highly stable GC-rich helical structure (SL) located at +20 downstream of the uAUG (by convention, the A from the AUG being +1). These two elements act in synergy to promote the stalling of an 80S complex upstream of the SL. This is reminiscent of a similar mechanism that has been described in the *Arabidopsis thaliana* NIP5.1 5'UTR mRNA that contains an AUG-stop that regulates translation of the main ORF through a ribosome stalling mechanism and mRNA degradation (Tanaka et al., 2016). This mechanism requires only the start-stop codons. In the case of *Hoxa11* TIE, the stable SL structure that is present downstream of the AUG-stop plays a major role. A ribosome that is stalled with a stop codon in the A site, in other words an empty A site without any tRNA, is usually the signal for recruitment of the release factors to dissociate the ribosomal subunits from the mRNA. With our cell-free translation assay and sucrose gradient analysis, we showed that the stalled 80S programmed with *Hoxa11* TIE is very stable and does not dissociate (**Figure 5C**). A possible explanation for the stability of this complex could be that the SL blocks the access of the release factors to the empty A site thereby preventing ribosome dissociation (Brown et al., 2015). Interestingly, the alignment of *Hoxa11* TIE among different species shows that the start-stop combination (highlighted in red) lacks conservation and remains species-specific (**Figure 8-figure supplement 1B**). Some species possess a substitution of the start codon AUG to AUG-like codons such GUG or a mutation of stop codon that leads to a longer uORF. In contrast, the SL (104-154) remains highly conserved suggesting its high functional significance (**Figure 8-figure supplement 1B**). Accordingly, we have shown that the sole SL is strong enough to impede scanning by the pre-initiation complex. Indeed, secondary structures in the 5'UTR have been shown to inhibit translation like the case of a conserved stem loop structure in the 5'UTR of TGF- β 1 mRNA (Jenkins et al., 2010). Therefore, in various species, *Hoxa11* TIE might use different combinations of *cis*-acting elements in order to block cap-dependent translation, the common *cis*-acting element between all species being the SL that is conserved.

As previously described, the *Hoxa11* IRES requires the presence of the ribosomal protein Rpl38 while *Hoxa3* IRES is Rpl38-independent (Kondrashov et al., 2011). Our probing experiments revealed that the folding of both *Hoxa3* and *Hoxa11* TIE elements is independent of the presence of IRES suggesting that their mode of action does not depend on the IRES.

TIE may have evolved in such a way to favour the translation from the downstream IRES hence justifying why there is variation in terms of sequence and mechanism but same inhibitory effect. This unique combination of an inhibitor of canonical translation mechanism and the activator of specialized translation sets an interesting point on how the 5'UTR elements confer ribosome specificity to translation (Xue et al., 2015). Importantly, the acquisition of these TIE elements in subsets of Hox mRNAs enables an additional layer of regulatory control between the canonical translation and the IRES-dependent one. One intriguing study would be to determine how other TIE elements a4 and a9 inhibit translation and whether there are common functional features amongst all Hox mRNAs. Beyond Hox mRNAs, our data on eIF2D suggests a specific role in translation initiation and it would be interesting to decipher its precise role at the molecular level. More precisely, determining how uORF length, codon composition and consensus sequences may influence the role of eIF2D in the initiation process on uORFs. Future studies will be required to fully understand the role of eIF2D in translation initiation of specific mRNAs.

Materials and Methods

Plasmids

For *in vitro* studies, murine *Hoxa3* TIE (170 nts) and *Hoxa11* TIE (216 nts) (sequences were kindly provided by Dr. Maria Barna, that were amplified from mouse E10.5–12.5 cDNA (Xue et al., 2015)), were placed upstream of 5'UTR of *Hbb-b1* (Accession number: KU350152) (50 nts) and *Renilla reniformis* Luciferase coding sequence (Accession number: M63501) (936 nts). These constructs were cloned in pUC19 vector in the *HindIII* site then used as a template for further PCR amplifications and site-directed mutagenesis.

For *in vivo* studies, we introduced an *EcoRI* restriction site upstream of hRLuc-neo fusion sequence in pmirGLO vector (Promega®) using Quick Change site-directed mutagenesis kit II XL (Thermo Fischer Scientific®). Subsequent cloning experiments of *Hoxa3* TIE, *Hoxa11* TIE and their mutants were performed in pmirGLO vector at *EcoRI* site using NEBuilder® *HiFi* DNA Assembly kit. All clones were checked by sequencing.

Cell lines

Two cell lines were used for our *in vivo* assays: Human embryonic kidney cell line HEK293FT (ATCC®) purchased from Invitrogen (ref R700-07) and murine mesenchymal stem cell line C3H10T1/2 (clone 8, ATCC® CCL26). HEK293FT cells were cultured in Dulbecco's modified Eagle medium (DMEM) with 2 mM of L-Glutamine and 10% Fetal Bovine Serum (FBS) supplemented with 100 units/ml of Penicillin/Streptomycin. Subcultures were performed after Trypsin-EDTA treatment for dissociation at sub-confluent conditions (70%-80%) 1:4 to 1:10 seeding at $2-4 \cdot 10^4$ cells/cm² according to manufacturer's instructions. C3H10T1/2 cells were cultured in basal DMEM medium supplemented with 2 mM Glutamine, 1.5 g/L sodium bicarbonate and 10% FBS supplemented with 40 µg/ml Gentamicine. Subcultures were performed after Trypsin-EDTA treatment for dissociation at sub-confluent conditions (60%-70%). Seeding dilutions were performed at 2000 cells/cm² one time per week.

RNA transcription

Transcription templates were generated by PCR amplification from the plasmids pUC19-TIE. The amplified templates were used for *in vitro* transcription with recombinant T7 RNA polymerase in the presence of either m⁷G_{ppp}G cap analog or non-functional cap analog A_{ppp}G (New England Biolabs®). To check RNA integrity, an aliquot was mixed with Formamide Dye and loaded on 4% denaturing polyacrylamide gel. The RNA was visualized under UV light after ethidium bromide staining. To eliminate unincorporated nucleotides, the RNA

sample was loaded on a gel filtration Sephadex G25-column (Pharmacia Fine Chemicals), proteins were then eliminated by phenol extraction and the RNA transcripts are precipitated with 0.25 M NaCl in ethanol. After centrifugation, RNA pellets were dried and resuspended in autoclaved milli-Q water. The concentration of purified RNA samples was determined by absorbance measurement at 260 nm.

***In vitro* translation assays in cell-free translation extracts**

In vitro translation was carried out using increasing concentrations of mRNA transcripts with self-made untreated Rabbit Reticulocyte Lysate (RRL), amino acid mixture containing all the amino-acids except methionine (1 mM of each), RNasin (Promega®), 75mM KCl, 0.5 mM MgCl₂, 3.8 mCi [³⁵S] methionine and autoclaved milli-Q water. Reaction mixture was incubated at 30°C for 1 hr. Aliquots of translation mixture were analysed by SDS PAGE (10%) (Laemmli, 1970) and translation products were visualized by phosphor-imaging. *In vitro* translation assays with wheat germ extract (Promega®) were performed according to manufacturer's instructions. *In vitro* translation assays with HeLa cell extract and drosophila S2 cell extract were performed as previously described (Thoma et al., 2004; Wakiyama et al., 2006).

Chemical probing

Probing experiments were performed as previously described (Alghoul et al., 2021).

- Probing with DMS

Modification by Dimethyl Sulfate (DMS) was performed on 2 pmoles of each RNA (*Hoxa3* TIE and *Hoxa11* TIE). The RNA is first incubated for 15 min in dimethylsulfate (DMS) buffer (50 mM Na Cacodylate (pH 7.5), 5 mM MgCl₂ and 100 mM KCl) and 1 µg of yeast total tRNA (Sigma- Aldrich®) and then modified with 1.25% DMS reagent (diluted with ethanol 100%) with 10 min incubation at 20°C and stopped on ice. Modified transcripts were precipitated with 0.25 M NaCl, 0.1 mg/ml glycogen in ethanol. RNA pellets were dried and resuspended in autoclaved milli-Q water. Modified nucleotides were detected by primer extension arrests that were quantified. The intensity of the RT stops is proportional to the reactivity for each nucleotide.

- Probing with CMCT

Similarly, modification by 1-cyclohexyl-3-(2-morpholinoethyl) carbodiimide metho-p-toluene sulfonate (CMCT) was performed on 2 pmoles of each RNA (*Hoxa3* TIE and *Hoxa11* TIE). Each RNA is incubated for 20 min in CMCT buffer (50 mM Na borate (pH 8.5); 5 mM MgCl₂; 100 mM KCl) and 1 µg of yeast total tRNA. Then modifications were performed with 10.5 g/L CMCT reagent with 20 min incubation at 20°C and stopped on ice. Modified

transcripts were precipitated with 0.25 M NaCl, 0.1 mg/ml glycogen in ethanol. RNA Pellets were dried and resuspended in autoclaved milli-Q water. Modified nucleotides were detected by primer extension arrests that were quantified. The intensity of the RT arrests is proportional to the reactivity for each nucleotide.

Primer extension

Reverse transcription was carried out in a 20 µl-reaction with 2 pmoles of RNA and 0.9 pmoles of 5' fluorescently labelled primers. We used fluorescent Vic and Ned primers (Thermo Fischer Scientific) of same sequence for all reverse transcription reactions which are complementary to the *Hbb-b1* 5'UTR from nucleotides 6-37 :

5'-GGTTGCTAGTGAACACAGTTGTGTCAGAAGC-3'.

First, the RNA was unfolded by a denaturation step at 95°C for 2 min. Then, fluorescent primers were annealed for 2 min at 65°C followed by incubation on ice for 2 min. Samples were incubated in a buffer containing 83 mM KCl, 56 mM Tris-HCl (pH 8.3), 0.56 mM each of the four deoxynucleotides (dNTP), 5.6 mM DTT and 3 mM MgCl₂. Reverse transcription was performed with 1 unit of Avian Myoblastosis Virus (AMV) reverse transcriptase (Promega®) at 42°C for 2 min, 50°C for 30 min and finally 65°C for 5 min. In parallel, sequencing reactions were performed in similar conditions, but supplemented with 0.5 mM dideoxythymidine or or dideoxycytidine triphosphate (ddTTP or ddCTP) (protocol adapted from (Gross et al., 2017)). The synthesized cDNA were phenol–chloroform extracted and precipitated. After centrifugation, the cDNA pellets were washed, dried and resuspended in 10 µl deionized Hi-Di formamide (freshly prepared highly deionized formamide). Samples were loaded on a 96-well plate for sequencing on an Applied Biosystems 3130xl genetic analyzer. The resulting electropherograms were analyzed using QuSHAPE software (Karabiber et al., 2013), which aligns signal within and across capillaries, as well as to the dideoxy references of nucleotide at specific position and corrects for signal decay. Normalized reactivities range from 0 to 2, with 1.0–2.0 being the range of highly reactive positions. A preliminary secondary structure model was first initiated by mfold (Mathews et al., 2016) and then edited according to reactivity values.

Sucrose gradient analysis

To analyse the assembly of ribosomal preinitiation complexes on the RNA of interest, the complexes were loaded on 7-47% sucrose gradients containing 5 mM MgCl₂, 25 mM Tris-HCl (pH 7.5), 1 mM DTT and 50 mM KCl. We used the specific translation inhibitors GMP-PNP (4 mM), cycloheximide (1 mg/mL), geneticin (0.7 mM), hygromycin (0.5 mg/mL) and edeine (10 mM), they were added to the RRL with a mix containing the 20 amino acids at 1.5

mM each, RNasin (Promega®), 35 mM KCl and 0.24 mM or 2.4 mM MgCl₂, prior to incubation with the 5' capped radioactive mRNA of interest. The assembled pre-initiation complexes were formed by incubation in RRL at 30°C for 5 min. Then, 8 mM MgAc₂ was added and one volume of 7% sucrose. Samples were then layered on the surface of 11-mL 7-47% sucrose gradient and centrifuged for 2h30 in an SW41 rotor at 37,000 rpm at 4°C. After centrifugation, the whole gradient was fractionated, and the mRNA was localized by measuring radioactivity in each collected fraction by Cerenkov counting in a scintillation counter.

Mass spectrometry and data processing

Protein extracts were digested with sequencing-grade trypsin (Promega®) as previously described (Chicher et al., 2015; Prongidi-Fix et al., 2013). Peptide digests were analysed by nano LC-MS/MS and MS data were searched by the Mascot algorithm against the UniProtKB database from *Oryctolagus cuniculus* (rabbit). Identifications were validated with a protein False Discovery Rate (FDR) of less than 1% using a decoy database strategy. The total number of MS/MS fragmentation spectra was used to quantify each protein from three independent biological replicates. This spectral count was submitted to a negative-binomial test using an edge R GLM regression through the R-package. For each identified protein, an adjusted P-value corrected by Benjamini–Hochberg was calculated, as well as a protein fold-change (FC is the ratio of the average of spectral counts from a specific complex divided by the average of spectral counts from a reference protein complex). The results were presented in a volcano plot using protein log₂ FC and their corresponding adjusted log₁₀ P-values. The proteins that were up-regulated in each condition are shown in red (*Hoxa3* TIE versus *Hbb-b1* mRNA, *Hoxa11* TIE versus *Hbb-b1* mRNA and *Hoxa3* TIE versus *Hoxa11* TIE).

***In vivo* luciferase assay**

For *in vivo* luciferase assay, HEK293T cells and C3H10T1/2 cells were transfected in 6-well plates with various constructs of pmirGlo vector (Promega®). Transfection was performed using Turbofect transfection reagent (Thermo Fischer scientific®) according to manufacturer's instructions. Cells were collected 24 hrs post-transfection. Luciferase assay was performed using Dual-Glo luciferase kit (Promega®) according to manufacturer's instructions. Firefly Luciferase activities were measured to monitor transfection efficiency in order to normalize Renilla luciferase activities for each construct.

Co-transfection assay of siRNAs and reporter plasmid in HEK293T cells.

HEK293FT cells were used to test the effect on inhibition of eIF2D knock-down and of a non-target siRNA pool, as a negative control, on Renilla luciferase expression (*hRluc-neo*) in

pmirGLO vectors. We used ON-TARGETplus human siRNAs against eIF2D (catalog number L-003680-01-00) and non-target pool of siRNAs (catalog number D-001810-10-05) purchased from Horizon Discovery. HEK293T cells were cultured according to manufacturer's instructions (ATCC®) for 24 hr and used for co-transfection by siRNAs and reporter plasmid. We used 2×10^5 cells in 1 ml of culture medium without antibiotics. Upon reaching 70% confluence, cells were transfected by 5 pmoles of siRNAs in different wells with 500 ng of reporter plasmid. Transfections were performed with Lipofectamine 2000 (Invitrogen®) according to manufacturer's instructions. After 48h, cells were washed twice by phosphate buffered saline (PBS1X) and incubated with Passive Lysis Buffer (PLB) 1X for 15 min. Luciferase assay was performed according to manufacturer's instructions. Protein concentration was measured by Bradford's assay. The impact of silencing on TIE-mediated translation inhibition was measured by luciferase assay according to manufacturer's instructions as previously mentioned.

Western Blot against eIF2D and GAPDH

The silencing efficiency was quantified by western blots using rabbit polyclonal anti-eIF2D antibody (12840-1-AP) from Proteintech and mouse polyclonal anti-GAPDH antibody (sc-1377179) from Santa Cruz Biotechnology. For that, 20 µg of protein extracts were loaded on 10% polyacrylamide SDS PAGE. After migration, proteins were transferred to an Immobilon-P membrane (Millipore®) at 10 Volts for 1 h in a semi-dry apparatus (Trans-Blot® SD) on a PVDF membrane (PolyVinylidene Fluoride) that had been previously activated with 100% methanol for few seconds and a transfer buffer pH 8 (25 mM Tris; 200 mM glycine; 20% ethanol). After transfer, the membrane was saturated for 2h by blocking buffer (5% milk, 0.05% Tween-20, PBS 1X). Primary antibodies were added at dilutions recommended by the manufacturers in blocking buffer, the membranes were incubated overnight at 4°C. Then, the membranes were washed three times by PBST (PBS 1X; 0.05% Tween-20) to remove the excess of primary antibodies. Then, membranes were incubated with secondary HRP-conjugated antibody for 1h at room temperature followed by three washing steps. The signal produced by reaction between HRP and ECL (Kit ECL Plus Western Blotting Detection System, GE Healthcare®) was detected by chemiluminescence using imaging Chemidoc (Biorad®).

Ribosome toe printing assay

Toe printing assay was adapted from previously established protocols (Martin et al., 2016, 2011). Briefly, RRL were incubated for 5 min at 30°C then 10 min on ice with buffer containing 1U/µl of recombinant RNasin (Promega®), 75 mM KCl 0.5 mM MgCl₂, and 1.3

mM of puromycin prior to initiation complex assembly. Then, the pre-initiation complexes were formed by incubation with 500 nM of the RNA of interest in the presence of specific inhibitors such as cycloheximide (1 mg/ml) or GMP-PNP (4 mM) for 5 min at 30°C and then 20 min on ice. Then, the pre-initiation complexes were complemented with one volume of ice-cold buffer A containing 20 mM Tris-HCl (pH 7.5), 100 mM KAc, 2.5 mM MgAc₂, 2 mM DTT, 1 mM ATP and 0.25 mM spermidine and placed on ice. In order to separate ribosomal complexes from the non-ribosomal fraction, samples were centrifuged at 88,000 rpm in S100AT3 rotor (Sorvall-Hitachi) at 4°C for 1h. After centrifugation, the pellets containing the pre-initiation complexes were resuspended in 30 µl of ice-cold buffer A and incubated with 5' radioactively labelled DNA oligonucleotide complementary to nts 22-51 of Renilla luciferase coding sequence for 3 min at 30°C. Then, 1 µl of a 320 mM Mg(Ac)₂, 4 µl of a dNTP mixture (containing 5 mM of dATP, dGTP, dTTP and dCTP), 10 units of recombinant RNAsin (Promega®) and 1 unit/µl Avian Myoblastosis virus reverse transcriptase (Promega®) were added and incubated for 1h at 30°C. The synthesized cDNAs were analysed on 8% PAGE next to sequencing ladders.

Acknowledgments

This work was funded by 'Agence Nationale pour la Recherche' (Ribofluidix, ANR-17-CE12-0025-01, by University of Strasbourg and by the 'Centre National de la Recherche Scientifique'. We would like to thank Dr. Maria Barna for sharing sequences of *Hoxa3* and *Hox11* transcripts. We are grateful to Dr. Christine Allmang, Hassan Hayek, Antonin Tidu, Aurélie Janvier, Aurélie Durand for technical assistance, Philippe Hamman, Johana Chicher, Lauriane Kuhn for Mass spectrometry analysis, Dr. Mireille Baltzinger and Dr. Pascale Romby for support and useful discussions on the project. We would also like to thank Dr. Sebastien Pfeffer for the pmirGLO plasmid.

References

- Adomavicius T, Guaita M, Zhou Y, Jennings MD, Latif Z, Roseman AM, Pavitt GD. 2019. The structural basis of translational control by eIF2 phosphorylation. *Nat Commun* **10**:2136. doi:10.1038/s41467-019-10167-3
- Ahmed YL, Schleich S, Bohlen J, Mandel N, Simon B, Sinning I, Teleman AA. 2018. DENR–MCTS1 heterodimerization and tRNA recruitment are required for translation reinitiation. *PLoS Biol* **16**:e2005160. doi:10.1371/journal.pbio.2005160
- Akulich KA, Andreev DE, Terenin IM, Smirnova V V., Anisimova AS, Makeeva DS, Arkhipova VI, Stolboushkina EA, Garber MB, Prokofjeva MM, Spirin P V., Prassolov VS, Shatsky IN, Dmitriev SE. 2016. Four translation initiation pathways employed by the leaderless mRNA in eukaryotes. *Sci Rep* **6**:37905. doi:10.1038/srep37905
- Alexander T, Nolte C, Krumlauf R. 2009. Hox Genes and Segmentation of the Hindbrain and Axial Skeleton. *Annu Rev Cell Dev Biol* **25**:431–456. doi:10.1146/annurev.cellbio.042308.113423
- Alghoul F, Eriani G, Martin F. 2021. RNA Secondary Structure Study by Chemical Probing Methods Using DMS and CMCT. *Methods Mol Biol* **2300**:241–250. doi:10.1007/978-1-0716-1386-3_18
- Aylett CHS, Boehringer D, Erzberger JP, Schaefer T, Ban N. 2015. Structure of a yeast 40S-eIF1-eIF1A-eIF3-eIF3j initiation complex. *Nat Struct Mol Biol* **22**:269–71. doi:10.1038/nsmb.2963
- Barbosa C, Peixeiro I, Romão L. 2013. Gene Expression Regulation by Upstream Open Reading Frames and Human Disease. *PLoS Genet*. doi:10.1371/journal.pgen.1003529
- Brown A, Shao S, Murray J, Hegde RS, Ramakrishnan V. 2015. Structural basis for stop codon recognition in eukaryotes. *Nature* **524**:493–496. doi:10.1038/nature14896
- Chicher J, Simonetti A, Kuhn L, Schaeffer L, Hammann P, Eriani G, Martin F. 2015. Purification of mRNA-programmed translation initiation complexes suitable for mass spectrometry analysis. *Proteomics* **15**:2417–2425. doi:10.1002/pmic.201400628
- Dever TE, Kinzy TG, Pavitt GD. 2016. Mechanism and regulation of protein synthesis in *Saccharomyces cerevisiae*. *Genetics*. doi:10.1534/genetics.115.186221
- Dmitriev SE, Terenin IM, Andreev DE, Ivanov PA, Dunaevsky JE, Merrick WC, Shatsky IN. 2010. GTP-independent tRNA delivery to the ribosomal P-site by a novel eukaryotic translation factor. *J Biol Chem* **285**:26779–26787. doi:10.1074/jbc.M110.119693
- Eliseev B, Yeramala L, Leitner A, Karuppasamy M, Raimondeau E, Huard K, Alkalaeva E,

- Aebersold R, Schaffitzel C. 2018. Structure of a human cap-dependent 48S translation pre-initiation complex. *Nucleic Acids Res* **46**:2678–2689. doi:10.1093/nar/gky054
- Fletcher L, Corbin SD, Browning KS, Ravel JM. 1990. The absence of a m7G cap on beta-globin mRNA and alfalfa mosaic virus RNA 4 increases the amounts of initiation factor 4F required for translation. *J Biol Chem* **265**:19582–7.
- Fraser CS, Berry KE, Hershey JWB, Doudna JA. 2007. eIF3j Is Located in the Decoding Center of the Human 40S Ribosomal Subunit. *Mol Cell* **26**:811–819. doi:10.1016/j.molcel.2007.05.019
- Friend K, Brooks HA, Propson NE, Thomson JA, Kimble J. 2015. Embryonic stem cell growth factors regulate eIF2 α phosphorylation. *PLoS One* **10**:e0139076. doi:10.1371/journal.pone.0139076
- Garreau de Loubresse N, Prokhorova I, Holtkamp W, Rodnina M V., Yusupova G, Yusupov M. 2014. Structural basis for the inhibition of the eukaryotic ribosome. *Nature* **513**:517–522. doi:10.1038/nature13737
- Gross L, Vicens Q, Einhorn E, Noireterre A, Schaeffer L, Kuhn L, Imler J-L, Eriani G, Meignin C, Martin F. 2017. The IRES5'UTR of the dicistrovirus Cricket Paralysis Virus is a type III IRES containing an essential pseudoknot structure. *Nucleic Acids Res* **45**:8993–9004.
- Hinnebusch AG. 2014. The scanning mechanism of eukaryotic translation initiation. *Annu Rev Biochem* **83**:779–812.
- Hinnebusch AG. 1993. Gene-specific translational control of the yeast GCN4 gene by phosphorylation of eukaryotic initiation factor 2. *Mol Microbiol* **10**:215–23.
- Jenkins RH, Bennagi R, Martin J, Phillips AO, Redman JE, Fraser DJ. 2010. A conserved stem loop motif in the 5'untranslated region regulates transforming growth factor- β 1 translation. *PLoS One* **5**:e12283. doi:10.1371/journal.pone.0012283
- Karabiber F, McGinnis JL, Favorov O V, Weeks KM. 2013. QuShape: rapid, accurate, and best-practices quantification of nucleic acid probing information, resolved by capillary electrophoresis. *RNA* **19**:63–73. doi:10.1261/rna.036327.112
- Kearse MG, Goldman DH, Choi J, Nwaezeapu C, Liang D, Green KM, Goldstrohm AC, Todd PK, Green R, Wilusz JE. 2019. Ribosome queuing enables non-AUG translation to be resistant to multiple protein synthesis inhibitors. *Genes Dev* **33**:871–885. doi:10.1101/gad.324715.119
- Kondrashov N, Pusic A, Stumpf CR, Shimizu K, Hsieh AC, Xue S, Ishijima J, Shiroishi T, Barna M. 2011. Ribosome-mediated specificity in Hox mRNA translation and vertebrate

- tissue patterning. *Cell* **145**:383–97. doi:10.1016/j.cell.2011.03.028
- Kozak M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**:283–292. doi:10.1016/0092-8674(86)90762-2
- Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**:680–685.
- Lee HH, Kim YS, Kim KH, Heo I, Kim SK, Kim O, Kim HK, Yoon JY, Kim HS, Kim DJ, Lee SJ, Yoon HJ, Kim SJ, Lee BG, Song HK, Kim VN, Park CM, Suh SW. 2007. Structural and Functional Insights into Dom34, a Key Component of No-Go mRNA Decay. *Mol Cell* **27**:938–950. doi:10.1016/j.molcel.2007.07.019
- Leppek K, Fujii K, Quade N, Susanto TT, Boehringer D, Lenarčič T, Xue S, Genuth NR, Ban N, Barna M. 2020. Gene- and Species-Specific Hox mRNA Translation by Ribosome Expansion Segments. *Mol Cell* **S1097-2765**:30730–9.
- Martin F, Barends S, Jaeger S, Schaeffer L, Prongidi-Fix L, Eriani G. 2011. Cap-assisted internal initiation of translation of histone H4. *Mol Cell* **41**:197–209. doi:10.1016/j.molcel.2010.12.019
- Martin F, Ménétret JF, Simonetti A, Myasnikov AG, Vicens Q, Prongidi-Fix L, Natchiar SK, Klaholz BP, Eriani G. 2016. Ribosomal 18S rRNA base pairs with mRNA during eukaryotic translation initiation. *Nat Commun* **7**. doi:10.1038/ncomms12622
- Mathews DH, Turner DH, Watson RM. 2016. RNA secondary structure prediction. *Curr Protoc Nucleic Acid Chem* **2016**:11.2.1-11.2.19. doi:10.1002/cpnc.19
- Merrick WC, Pavitt GD. 2018. Protein synthesis initiation in eukaryotic cells. *Cold Spring Harb Perspect Biol* **10**:a033092. doi:10.1101/cshperspect.a033092
- Oh SK, Scott MP, Sarnow P. 1992. Homeotic gene Antennapedia mRNA contains 5'-noncoding sequences that confer translational initiation by internal ribosome binding. *Genes Dev* **6**:1643–53.
- Pelletier J, Sonenberg N. 2019. The Organizing Principles of Eukaryotic Ribosome Recruitment. *Annu Rev Biochem* **88**:307–335. doi:10.1146/annurev-biochem-013118-111042
- Phinney DG, Gray AJ, Hill K, Pandey A. 2005. Murine mesenchymal and embryonic stem cells express a similar Hox gene profile. *Biochem Biophys Res Commun* **338**:1759–1765. doi:10.1016/j.bbrc.2005.10.140
- Prongidi-Fix L, Schaeffer L, Simonetti A, Barends S, Ménétret JF, Klaholz BP, Eriani G, Martin F. 2013. Rapid purification of ribosomal particles assembled on histone H4

- mRNA: A new method based on mRNA-DNA chimaeras. *Biochem J* **449**:719–728.
doi:10.1042/BJ20121211
- Schneider-Poetsch T, Ju J, Dang Y, Bhat S, Green R, Shen B, Merrick WC, Eyler DE, Liu JO. 2010. Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. *Nat Chem Biol* **6**:209–217. doi:10.1038/nchembio.304
- Shao S, Murray J, Brown A, Taunton J, Ramakrishnan V, Hegde RS. 2016. Decoding Mammalian Ribosome-mRNA States by Translational GTPase Complexes. *Cell* **167**:1229–1240.e15. doi:10.1016/j.cell.2016.10.046
- Shirokikh NE, Preiss T. 2018. Translation initiation by cap-dependent ribosome recruitment: Recent insights and open questions. *Wiley Interdiscip Rev RNA* **9**:e1473.
doi:10.1002/wrna.1473
- Shoemaker CJ, Eyler DE, Green R. 2010. Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop-off to initiate no-go decay. *Science (80-)* **330**:369–372.
doi:10.1126/science.1192430
- Skabkin MA, Skabkina O V., Dhote V, Komar AA, Hellen CUT, Pestova T V. 2010. Activities of Ligatin and MCT-1/DENR in eukaryotic translation initiation and ribosomal recycling. *Genes Dev* **24**:1787–801. doi:10.1101/gad.1957510
- Skabkin MA, Skabkina O V., Hellen CUT, Pestova T V. 2013. Reinitiation and other unconventional posttermination events during eukaryotic translation. *Mol Cell* **51**:249–64. doi:10.1016/j.molcel.2013.05.026
- Tanaka M, Sotta N, Yamazumi Y, Yamashita Y, Miwa K, Murota K, Chiba Y, Hirai MY, Akiyama T, Onouchi H, Naito S, Fujiwara T. 2016. The Minimum Open Reading Frame, AUG-Stop, Induces Boron-Dependent Ribosome Stalling and mRNA Degradation. *Plant Cell* **28**:2830–2849. doi:10.1105/tpc.16.00481
- Thoma C, Bergamini G, Galy B, Hundsdoerfer P, Hentze MW. 2004. Enhancement of IRES-mediated translation of the c-myc and BiP mRNAs by the poly(A) tail is independent of intact eIF4G and PABP. *Mol Cell* **15**:925–935.
- Uchiyama-Kadokura N, Murakami K, Takemoto M, Koyanagi N, Murota K, Naito S, Onouchi H. 2014. Polyamine-responsive ribosomal arrest at the stop codon of an upstream open reading frame of the AdoMetDC1 gene triggers nonsense-mediated mRNA decay in *Arabidopsis thaliana*. *Plant Cell Physiol* **55**:1556–1567.
doi:10.1093/pcp/pcu086
- Vaidya AT, Lomakin IB, Joseph NN, Dmitriev SE, Steitz TA. 2017. Crystal Structure of the C-terminal Domain of Human eIF2D and Its Implications on Eukaryotic Translation

- Initiation. *J Mol Biol* **429**:2765–2771. doi:10.1016/j.jmb.2017.07.015
- Varland S, Osberg C, Arnesen T. 2015. N-terminal modifications of cellular proteins: The enzymes involved, their substrate specificities and biological effects. *Proteomics* **15**:2385–2401. doi:10.1002/pmic.201400619
- Vasudevan D, Neuman SD, Yang A, Lough L, Brown B, Bashirullah A, Cardozo T, Ryoo HD. 2020. Translational induction of ATF4 during integrated stress response requires noncanonical initiation factors eIF2D and DENR. *Nat Commun* **11**:4677. doi:10.1038/s41467-020-18453-1
- Vicens Q, Westhof E. 2003. Crystal structure of geneticin bound to a bacterial 16 S ribosomal RNA A site oligonucleotide. *J Mol Biol* **326**:1175–1188. doi:10.1016/S0022-2836(02)01435-3
- Wakiyama M, Kaitsu Y, Yokoyama S. 2006. Cell-free translation system from Drosophila S2 cells that recapitulates RNAi. *Biochem Biophys Res Commun* **343**:1067–1071. doi:10.1016/j.bbrc.2006.03.076
- Wang J, Johnson AG, Lapointe CP, Choi J, Prabhakar A, Chen DH, Petrov AN, Puglisi JD. 2019. eIF5B gates the transition from translation initiation to elongation. *Nature* **573**:605–608. doi:10.1038/s41586-019-1561-0
- Wang Jinfan, Wang Jing, Shin BS, Kim JR, Dever TE, Puglisi JD, Fernández IS. 2020. Structural basis for the transition from translation initiation to elongation by an 80S-eIF5B complex. *Nat Commun* **11**. doi:10.1038/s41467-020-18829-3
- Weisser M, Schäfer T, Leibundgut M, Böhringer D, Aylett CHS, Ban N. 2017. Structural and Functional Insights into Human Re-initiation Complexes. *Mol Cell* **67**:447-456.e7. doi:10.1016/j.molcel.2017.06.032
- Xue S, Tian S, Fujii K, Kladwang W, Das R, Barna M. 2015. RNA regulons in Hox 5' UTRs confer ribosome specificity to gene regulation. *Nature* **517**:33–38. doi:10.1038/nature14010
- Ye X, Fong P, Iizuka N, Choate D, Cavener DR. 1997. Ultrabithorax and Antennapedia 5' untranslated regions promote developmentally regulated internal translation initiation. *Mol Cell Biol* **17**:1714–21.
- Young DJ, Guydosh NR. 2019. Hcr1/eIF3j Is a 60S Ribosomal Subunit Recycling Accessory Factor In Vivo. *Cell Rep* **28**:39-50.e4. doi:10.1016/j.celrep.2019.05.111
- Young DJ, Makeeva DS, Zhang F, Anisimova AS, Stolboushkina EA, Ghobakhlou F, Shatsky IN, Dmitriev SE, Hinnebusch AG, Guydosh NR. 2018. Tma64/eIF2D, Tma20/MCT-1, and Tma22/DENR Recycle Post-termination 40S Subunits In Vivo. *Mol*

Figure Legends

- **Figure 1: TIE-mediated inhibition is recapitulated in RRL and does not require full length TIEs.**

A. Three capped mRNAs were used to test TIE-mediated inhibition *in vitro*. *Hoxa3* TIE and *Hoxa11* TIE were placed upstream of the 5'UTR of *Hbb-b1* and the Renilla luciferase coding sequence. Translation assays were performed *in vitro* using RRL at an mRNA concentration of 50 nM, which enables sub-saturating conditions. The relative expression of Luciferase protein reflects the efficiency of translation inhibition by *Hoxa3* TIE and *Hoxa11* TIE. Values were normalized to that of the control (w/o TIE) which corresponds to normal expression without inhibition and was set to 100%. **p<0.01 (t-test as compared to w/o TIE) n=3. Experiments were performed in triplicates. **B.** Sequential deletions in the 5' extremity of *Hoxa3* TIE and *Hoxa11* TIE constructs were performed to assay their effect on translation. Values of translation expression were normalized to that of the control (w/o TIE). Experiments were performed in triplicates. The percentages of inhibition for each TIE are indicated in the histogram. **p<0.01 (t-test as compared to w/o TIE) n=3. Experiments were performed in triplicates.

- **Figure 2: the secondary structural models of *Hoxa3* TIE and *Hoxa11* TIE reveal distinct structures.**

The structures of (A) *Hoxa3* TIE (170 nucleotides) and (B) *Hoxa11* TIE (216 nucleotides) were obtained by chemical probing using base-specific reagents, DMS and CMCT. After modifications, reverse transcription was performed using fluorescently labelled primers to determine the position of modified nucleotides. Experiments were performed in triplicates. Reactivities are shown as average reactivity from three independent experiments. A representation of reactivities is assigned as colour code depending on a range of values as shown in the figure legend on the right. Reactivity values for each nucleotide with corresponding standard deviations are shown in figure supplements (S2A, S2B, S2C, and S2D).

- **Figure 3: upstream AUG111 in *Hoxa3* TIE is essential for inhibition.**

(A) Substitution mutations in uAUG111 and uAUG123 to UAC in *Hoxa3* TIE were performed. Constructs with the corresponding mutations were translated in RRL and

luciferase assay was performed to evaluate the effect of mutation on translation efficiency as previously described. **p<0.01 (t-test as compared to construct w/o TIE). n=3. Experiments were performed in triplicates **(B)** Toe printing analysis of ribosomal assembly on two mRNAs, *Hoxa3* TIE Wt and the mutant of upstream (AUG/UAC)111. Initiation complexes were assembled in RRL extracts in the absence or presence of translation inhibitors: cycloheximide and GMP-PNP. Reaction samples were separated on 8% denaturing PAGE together with the appropriate sequencing ladder. Toe-print positions were counted starting on the A+1 of the AUG codon at +16 position. A, U and G nucleotides of the start codon are marked by black dots. Full-length cDNAs are indicated by an arrow at the top of the gel.

- **Figure 4: the uAUG111 in *Hoxa3* TIE is translated through 5'UTR of *Hoxa3*.**

(A) Three transcripts were used for this experiment: full length 5'UTR of *Hoxa3*, a deletion mutant at nucleotide G₃₃₃ in *Hoxa3* IRES and a control transcript without TIE. To test the translation of uORF in *Hoxa3* TIE starting from uAUG111, a deletion of G in *Hoxa3* IRES at position 333 was performed to create a longer uORF that is in the same frame as the ORF of Renilla luciferase to create an N-terminally extended luciferase. FL₃ mRNA was used as a control to correspond to the full length *Hoxa3* 5'UTR (TIE+IRES). Transcripts were translated *in vitro* in RRL and products were loaded on 10% SDS-PAGE in the presence of ³⁵S-Methionine. **(B)** *In vivo* luciferase assays in two embryonic cell lines; HEK293FT (left) and C3H10T1/2 (right). Reporter constructs in pmirGlo containing *Hoxa3* TIE, u(AUG/UAC)111, and without TIE, were transfected in the two indicated cell lines. Renilla luciferase expression was normalized to the control (w/o TIE), which was set to 100%. **p<0.01 (t-test as compared to empty plasmid (w/o TIE)). Experiments were performed in triplicates n=3

- **Figure 5: a start-stop uORF in *Hoxa11* TIE stalls an 80S upstream of a highly stable structure.**

(A) Mutational analysis of uAUGs in *Hoxa11* TIE. Three transcripts with AUG/UAC mutations were used: M1: (AUG/UAC)84+ (AUG/UAC)159, M2: (AUG/UAC)84 and M3: AUG/UAC159. Transcripts were translated in RRL at 50 nM concentrations and the luciferase expression was normalized to the control (w/o TIE) as previously described. **p<0.01 (t-test as compared to construct w/o TIE). n=3. Experiments were performed in triplicates. **(B)** A single substitution (A/G)88 mutation destroys the stop codon UAG87 to UGG in *Hoxa11* TIE, *Hoxa11* TIE wt and control (w/o TIE) were also translated as references

in RRL. Translation products were loaded on 15% SDS-PAGE. **(C)** Ribosomal pre-initiation complexes were assembled and analysed on 7-47% sucrose gradient with [α - 32 P]GTP-radiolabeled *Hoxa11* TIE as well as the two mutants of uAUG/UAC at the previously indicated positions in the absence or presence of cycloheximide. Heavy fractions correspond to polysomes and lighter fractions correspond to free RNPs. The coloured sections corresponding to RNP, 48S, 80S and polysomes have been assigned according to the OD_{254nm} profiles (see also Figure supplement S6).

- **Figure 6: distinct profiles for factors involved in TIE-mediated inhibition blocked with cycloheximide.**

Mass spectrometry analysis of cycloheximide-blocked translation initiation complexes and on three transcripts: *Hoxa3* TIE, *Hoxa11* TIE placed upstream of 5'UTR of *Hbb-b1* and the 5'UTR of *Hbb-b1* (control). Graphical representation of proteomics data: protein \log_2 spectral count fold changes (on the x-axis) and the corresponding adjusted \log_{10} P-values (on the y-axis) are plotted in a pair-wise volcano plot. The significance thresholds are represented by a horizontal dashed line (P-value = 1.25, negative-binomial test with Benjamini–Hochberg adjustment) and two vertical dashed lines (-1.0-fold on the left and +1.0-fold on the right). Data points in the upper left and upper right quadrants indicate significant negative and positive changes in protein abundance. Protein names are labelled next to the off-centred spots and they are depicted according to the following color code: red spots are significant hits and black are non-significant with <10 spectra. Data points are plotted based on the average spectral counts from triplicate analysis. Three profiles were produced by comparing the proteomics of two transcripts.

- **Figure 7: Co-transfection assay of TIE plasmids and siRNA against eIF2D confirms its implication in *Hoxa3*-mediated inhibition. Mutational analysis of A-motif sequences in *Hoxa3* TIE shows a requirement for an upstream A-motif for efficient inhibition.**

(A) Co-transfection assay was performed using pmirGLO plasmids with inserts of *Hoxa3* TIE, *Hoxa11* TIE and w/o TIE (pmirGLO) with siRNA against eIF2D (siRNA-eIF2D), and non-target pool of siRNA (siRNA-NT) in HEK293 cell line. The plasmids harbour two reporter genes, Renilla luciferase for monitoring the impact of TIEs and Firefly luciferase for normalization. Plates were incubated for 48 hours followed by cell lysis. Cell lysate was analysed by western blot and luciferase assay. The histograms present the Renilla to Firefly

Luciferase activity ratio. **p<0.01 (t-test as compared to construct w/o TIE). n=3. Experiments were performed in triplicates. Efficiency of silencing of each protein was quantified by western blot analysis for *Hoxa3* TIE and *Hoxa11* TIE samples **(B)** Two sets of mutations were performed on distinct A-motif sequences in *Hoxa3* TIE. The first mutation is AAAA/GGCC107 and the second mutation AAAAA/GGCC147. The transcripts were *in vitro* translated in RRL with *Hoxa3* TIE Wt and control (w/o TIE). Results were confirmed *in vivo* in HEK293FT cell line. Luciferase activities were normalized to control (w/o TIE). **p<0.01 (t-test as compared to construct w/o TIE). n=3. Experiments were performed in triplicates.

- **Figure 8: two distinct models for translational inhibition by *Hoxa3* TIE and *Hoxa11* TIE.**

A model for *Hoxa3* TIE suggests a ribosomal assembly on the uAUG111 with a requirement of eIF2D initiation factor. The model for *Hoxa11* TIE suggests a stalled 80S ribosome on AUG-stop codons combination upstream of a highly stable structure.

Figure supplement and Tables Legends

- **Supplemental File 1: Mass spectrometry analysis of pre-initiation complexes programmed with *Hoxa3* TIE, *Hoxa11* TIE and 5'UTR *Hbb-b1* in the presence of cycloheximide**
- **Supplemental File 2: Mass spectrometry analysis of pre-initiation complexes programmed with *Hoxa3* TIE, *Hoxa11* TIE and 5'UTR *Hbb-b1* in the presence of GMP-PNP**
- **Figure 1-figure supplement 1: TIE-mediated inhibition recapitulated in different *in vitro* and *in vivo* systems.**

(A) Translation efficiency was measured with a 5'UTR *Hbb-b1* Renilla luciferase reporter without TIE. A-capped and m7G-capped mRNAs were first used to assess cap-dependency of the RRL extracts. Then, the translation efficiency was measured with increasing concentrations of m7G-capped 5'UTR *Hbb-b1* Renilla. The intensity of light emitted by Renilla Luciferase protein as a function of control mRNA (w/o TIE) concentration (nM). Emission of light was measured under sub-saturating conditions (50-100 nM mRNA concentration). **(To the left)** The intensity of light emission by RLuc protein relative to different concentrations (50 nM and 100 nM respectively) of tested mRNA samples *Hoxa3* TIE, *Hoxa11* TIE and control (w/o TIE). **(To the right)** Analysis of translation products after *in vitro* translation in RRL on 10% SDS PAGE to monitor RLuc expression. Visualization of protein bands was achieved by incorporation of radiolabelled ³⁵S-methionine, which are detected by autoradiography. **(B)** *In vitro* translation of *Hoxa3* TIE and *Hoxa11* TIE transcripts in the presence of m⁷G_{ppp}G cap or a non-functional analog A_{ppp}G. Three *in vitro* systems were used: Wheat Germ Extract (WGE), drosophila embryonic cell extract (S2) and HeLa cell extract. All mRNAs were translated *in vitro* at 50 nM concentrations and RLuc expression was normalized to control (w/o TIE) in each condition. **(C)** Transfection of reporter plasmids with *Hoxa3* TIE or *Hoxa11* TIE in two embryonic cell lines, kidney HEK293FT and mesenchymal C3H10T1/2 cell lines. Renilla luciferase expression was normalized to the control (w/o TIE). **p<0.01 (t-test as compared to construct w/o TIE). n=3. Experiments were performed in triplicates. **(D)** SDS PAGE of Renilla expression with 5' deletions in *Hoxa3* TIE and *Hoxa11* TIE.

- **Figure 2-figure supplement 1: average of reactivities of DMS and CMCT for *Hoxa3* TIE and *Hoxa11* TIE.**

Probing experiments by DMS and CMCT were performed in triplicates. Figures S2A and S2B represent averages of reactivities of DMS and CMCT for *Hoxa3* TIE respectively. Figures S2C and S2D represent averages of reactivities of DMS and CMCT for *Hoxa11* TIE respectively. A scale of 40 nucleotides range was included to determine position of every nucleotide. Standard deviations are shown.

- **Figure 3-figure supplement 1: upstream AUG111 in *Hoxa3* TIE is essential for inhibition.**

(A) (To the left) AUG-like mutations (AUU, GUG, CUG and ACG) were tested *in vitro* in RRL compared to wt *Hoxa3*. **(To the right)** Representation of uAUG111 context in the secondary structure of *Hoxa3* TIE is shown. A111 and U112 base pair with the A123 and U124 of the second AUG123. Mutants of the uAUG into CUG and GUG with the corresponding compensatory mutagenesis were tested. RLuc expression with different transcripts was normalized to control (w/o TIE). ** $p < 0.01$ (t-test as compared to construct w/o TIE). $n=3$. Experiments were performed in triplicates. **(B)** Sucrose gradient analysis of *Hoxa3* TIE and mutant of upstream (AUG/UAC)111. Ribosomal pre-initiation complexes were assembled and analysed on 7-47% sucrose gradient with radioactive m^7G -capped *Hoxa3* TIE and the mutants of upstream (AUG/UAC)111 in the absence (upper panel) or in the presence of GMP-PNP (lower panel). Heavy fractions correspond to polysomes while the lightest correspond to free RNPs.

- **Figure 4-figure supplement 1: translation of *Hoxa3* TIE uORF in different constructs.**

(A) A single nucleotide deletion of A220 in *Hoxa3* TIE is performed to frameshift the uORF frame to the same as the main RLuc ORF. Another mutation was performed combining this deletion with uAUG mutation of (U/C)112. Transcripts were *in vitro* translated in RRL and products were loaded on 10% SDS PAGE. **(B)** Translation of *Hoxa3* TIE uORF in *Hoxa3* mRNA. *Hoxa3* mRNA (5'UTR+CDS) was translated *in vitro* to check uORF translation in this context. Products were analysed on 15% SDS-PAGE. The experiment was performed in duplicate samples.

- **Figure 5-figure supplement 1: transplanting *Hoxa11* TIE stem loop structure in *Hbb-b1* 5' UTR efficiently inhibits translation of RLuc mRNA.**

The stem loop structure of *Hoxa11* TIE (104-154) was transplanted in the 5'UTR of *Hbb-b1* with 25 nucleotides spanning from each extremity. The three shown transcripts were *in vitro* translated in RRL at two concentrations, 25 nM and 50 nM. Results are represented in relative light unit of RLuc expression.

- **Figure 5-figure supplement 2: polysome fractionation on 7-47% sucrose gradient of pre-initiation complexes**

Polysome fractionation of pre-initiation complexes programmed with m⁷G-capped radioactive mRNAs containing Wt *Hoxa3* TIE and the mutant (AUG/UAC)₈₄ in RRL previously treated with cycloheximide. The Y axis represent the percentage of radioactive RNA present in the fractions. The coloured sections corresponding to RNP, 48S, 80S and polysomes are assigned according to the OD_{254 nm} profile (dashed line).

- **Figure 6-figure supplement 1: distinct profiles for factors involved in TIE-mediated inhibition blocked with GMP-PNP.**

Mass spectrometry analysis of GMP-PNP-blocked translation initiation complexes and on three transcripts: *Hoxa3* TIE, *Hoxa11* TIE placed upstream of 5'UTR of *Hbb-b1* and the 5'UTR of *Hbb-b1* (control). Graphical representation of proteomics data: protein log₂ spectral count fold changes (on the x-axis) and the corresponding adjusted log₁₀ P-values (on the y-axis) are plotted in a pair-wise volcano plot. The significance thresholds are represented by a horizontal dashed line (P-value = 1.25, negative-binomial test with Benjamini–Hochberg adjustment) and two vertical dashed lines (−1.0-fold on the left and +1.0-fold on the right). Data points in the upper left and upper right quadrants indicate significant negative and positive changes in protein abundance. Protein names are labelled next to the off-centred spots and they are depicted according to the following color code: red spots are significant hits and black are non-significant with <10 spectra. Data points are plotted based on the average spectral counts from triplicate analysis. Three profiles were produced by comparing the proteomics of two transcripts.

- **Figure 8-figure supplement 1: Alignment of *Hoxa3* TIE and *Hoxa11* TIE among different species shows variation in the conservation of the inhibitory elements.**

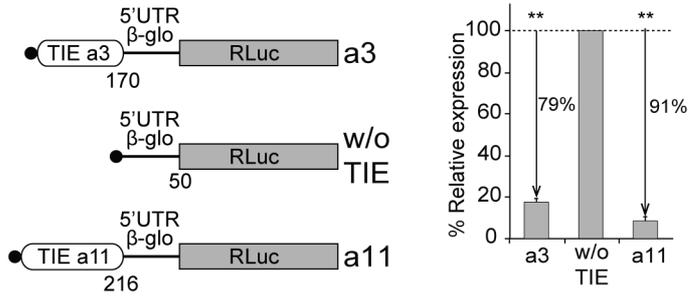
(A) Nucleotides 1-170 from the mouse *Hoxa3* TIE sequence were aligned with different

species. The position of uAUG in *Mus musculus* is highlighted in a red box. Accession numbers: *F. catus*: XR_002740526.1, *M. musculus*: NM_010452.3, *M. flaviventris*: XM_0279, *H. sapiens*: XM_011515343.3, *M. mulatta*: XM_028845931.1, *P. Anubis*: XM_017956197.2, *E. caballus*: XM_023639416.1, *U. horribilis*: XM_026508219.1, *C. dingo*: XM_025464436.1 **(B)** Nucleotides 1-216 from the mouse *Hoxa11* TIE sequence were aligned with different species. The position of uAUG-UAG in *Mus Musculus* and the insertion of AUG-UAG in *Homo sapiens* are highlighted in red boxes. Accession numbers: *H. sapiens*: AF071164.1, *M. Mulata*: XM_015133828.2, *C. dingo*: XR_003143092.1, *C. griseus*: XM_027436289.1, *R. norvegicus*: XM_008762951.2, *E. caballus*: XM_023639423.1, *M auratus*: XM_00508585871.3, *M. musculus*: NM_010450.

-

Figure 1

A



B

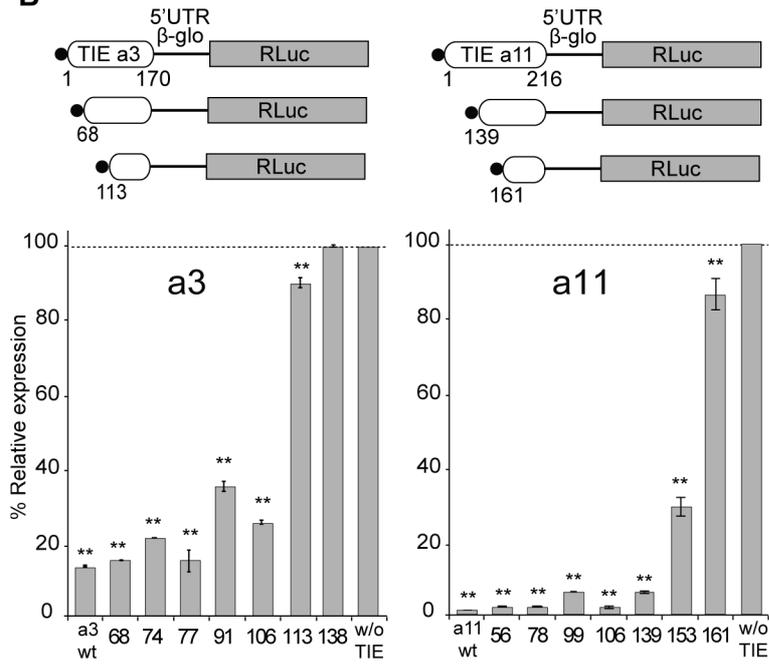
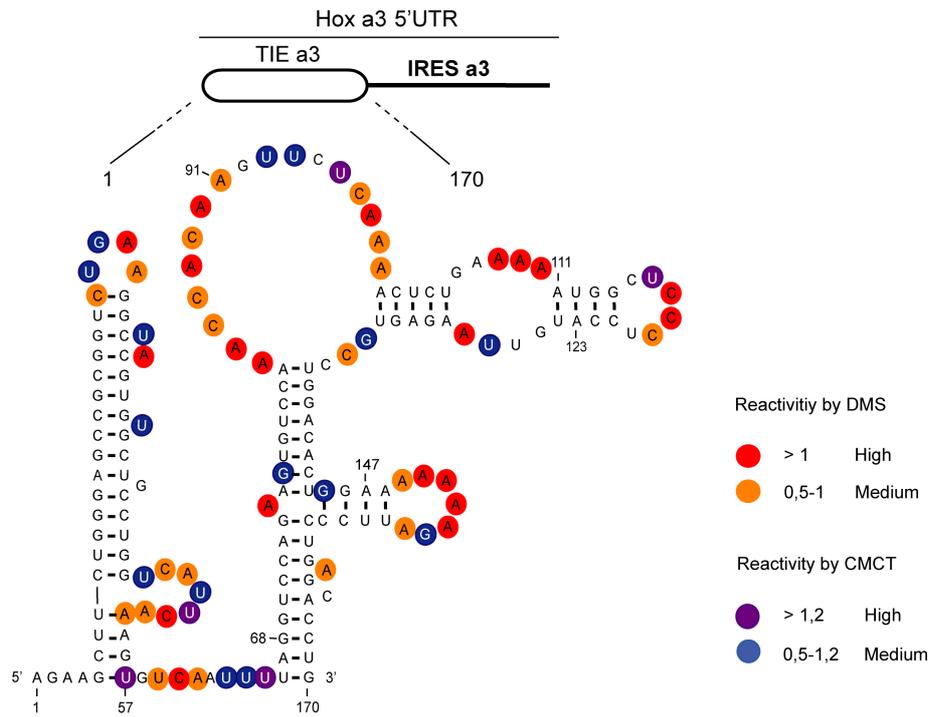


Figure 2

A



B

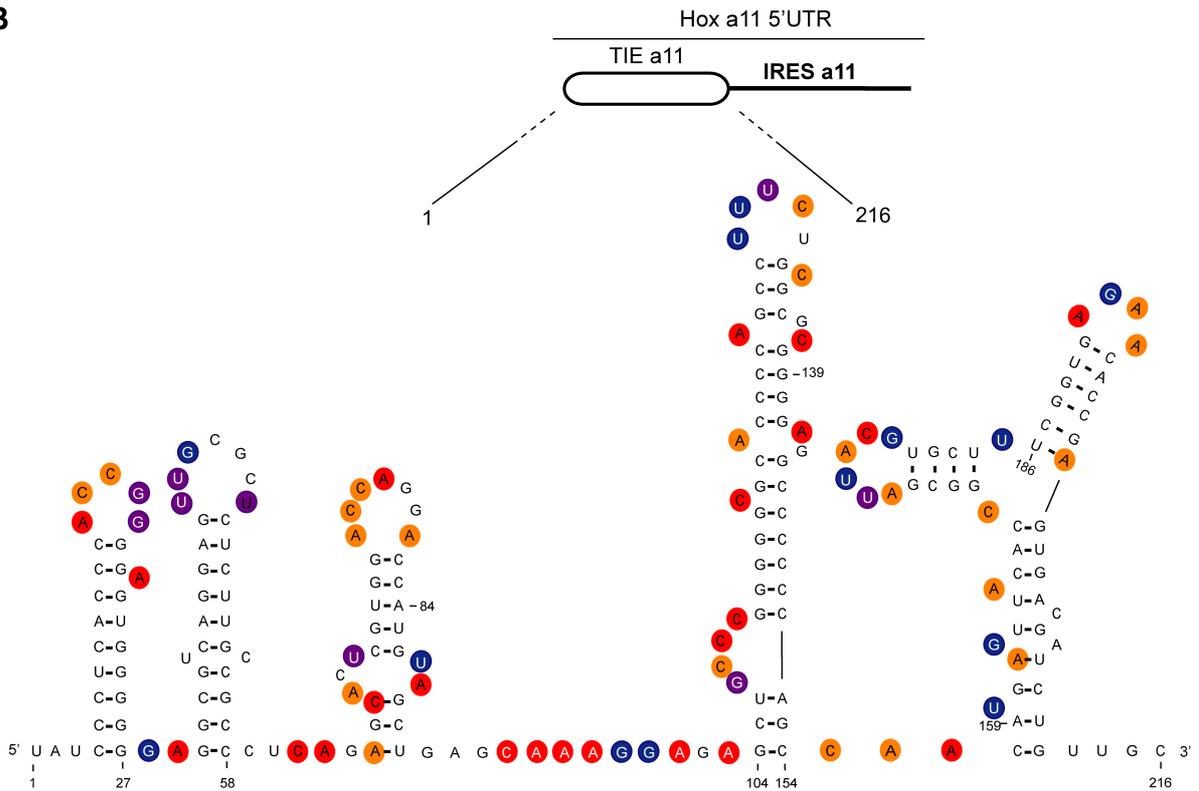
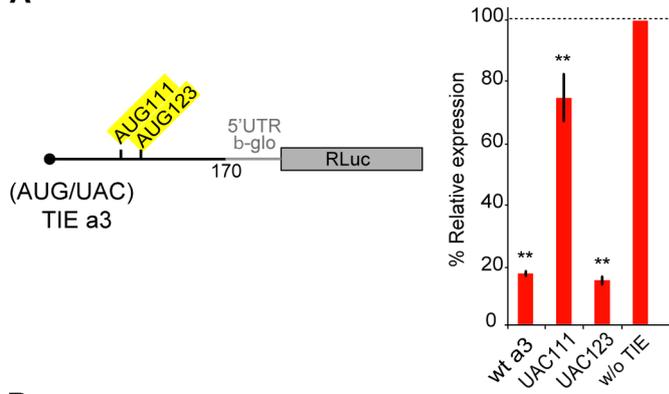


Figure 3

A



B

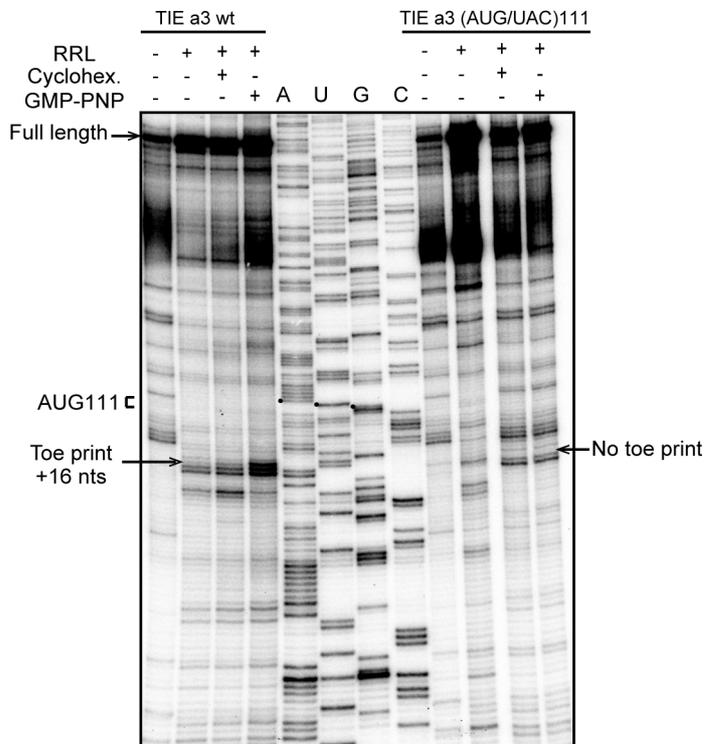
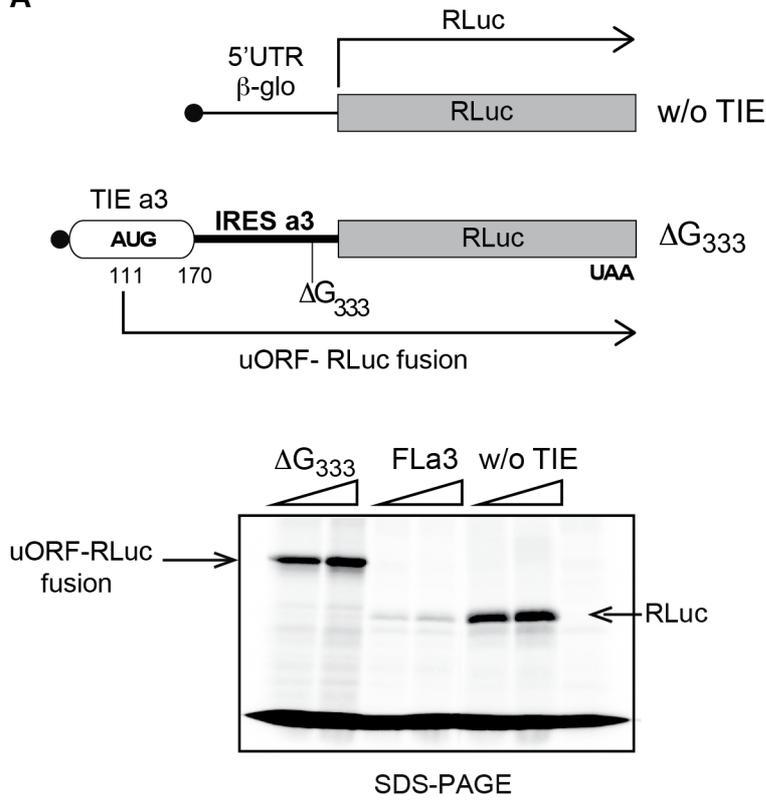


Figure 4

A



B

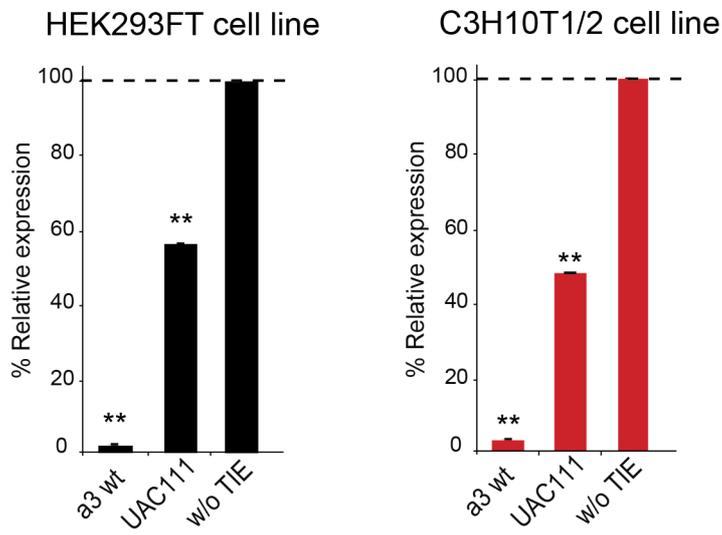


Figure 5

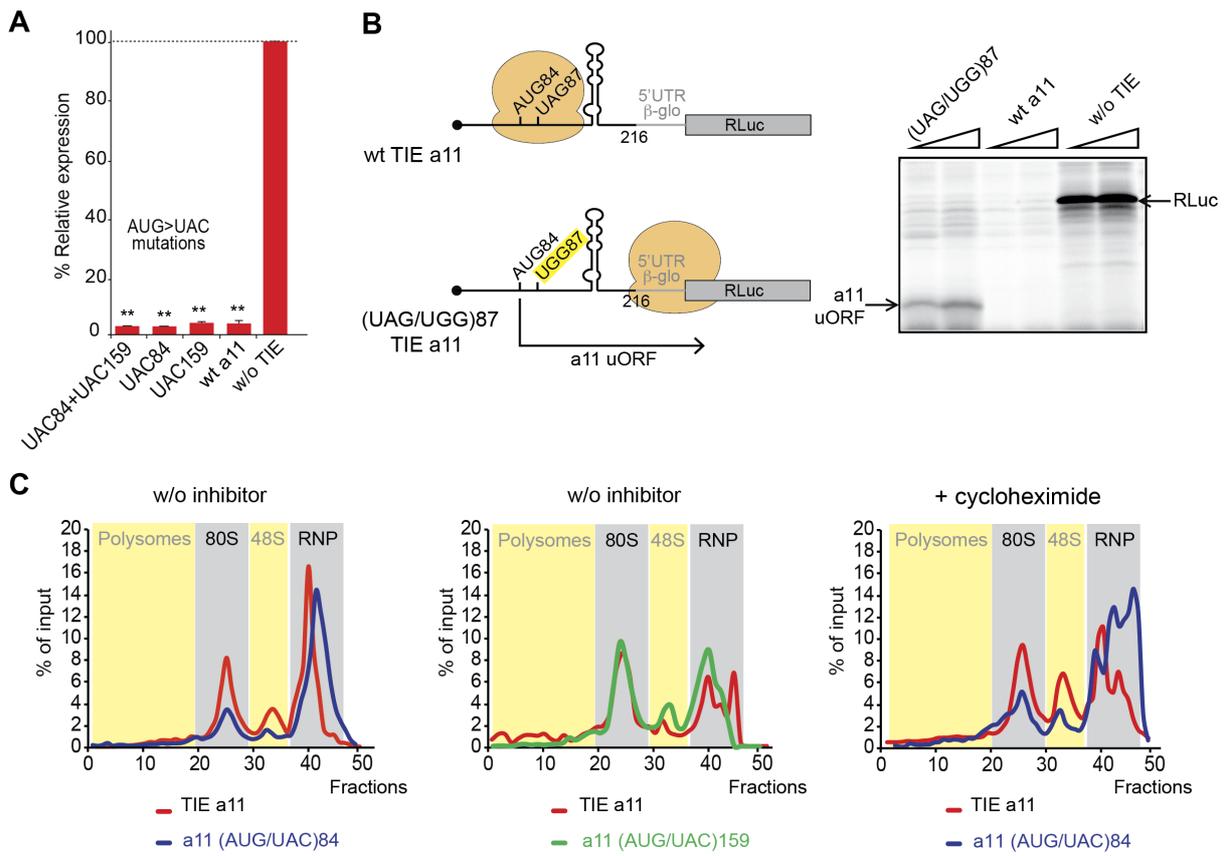


Figure 6

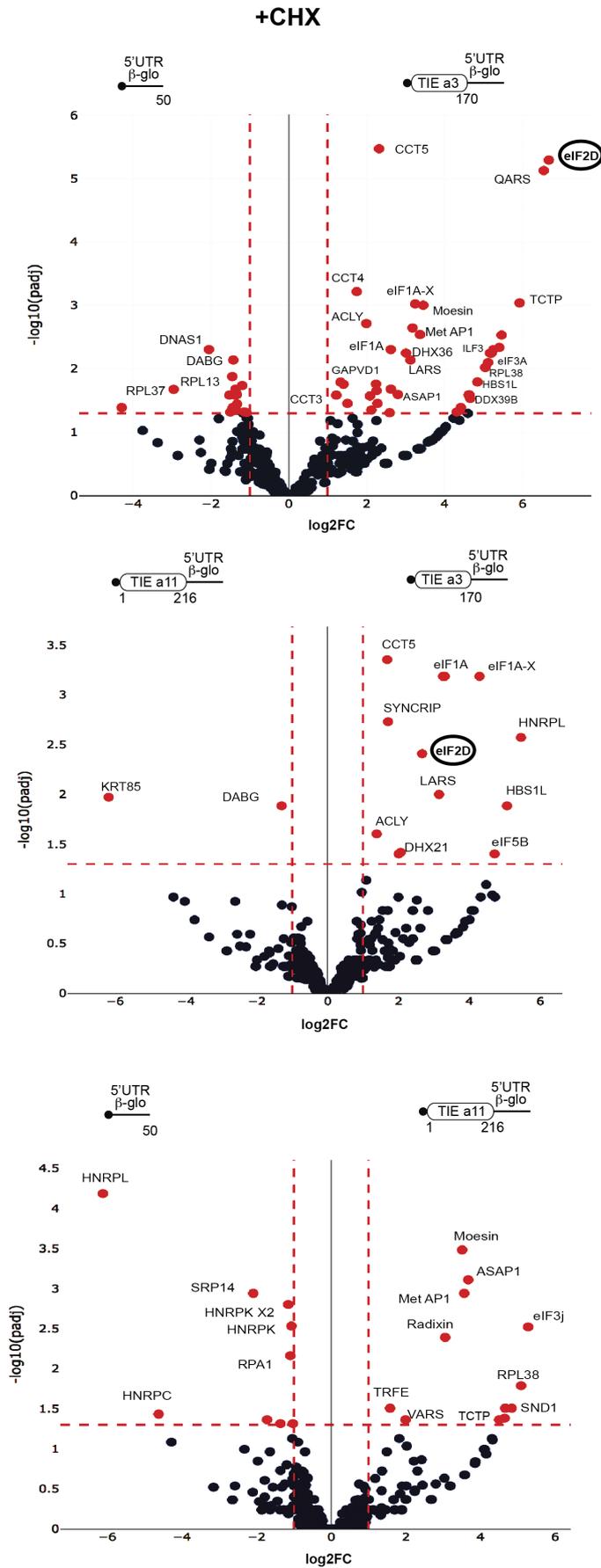
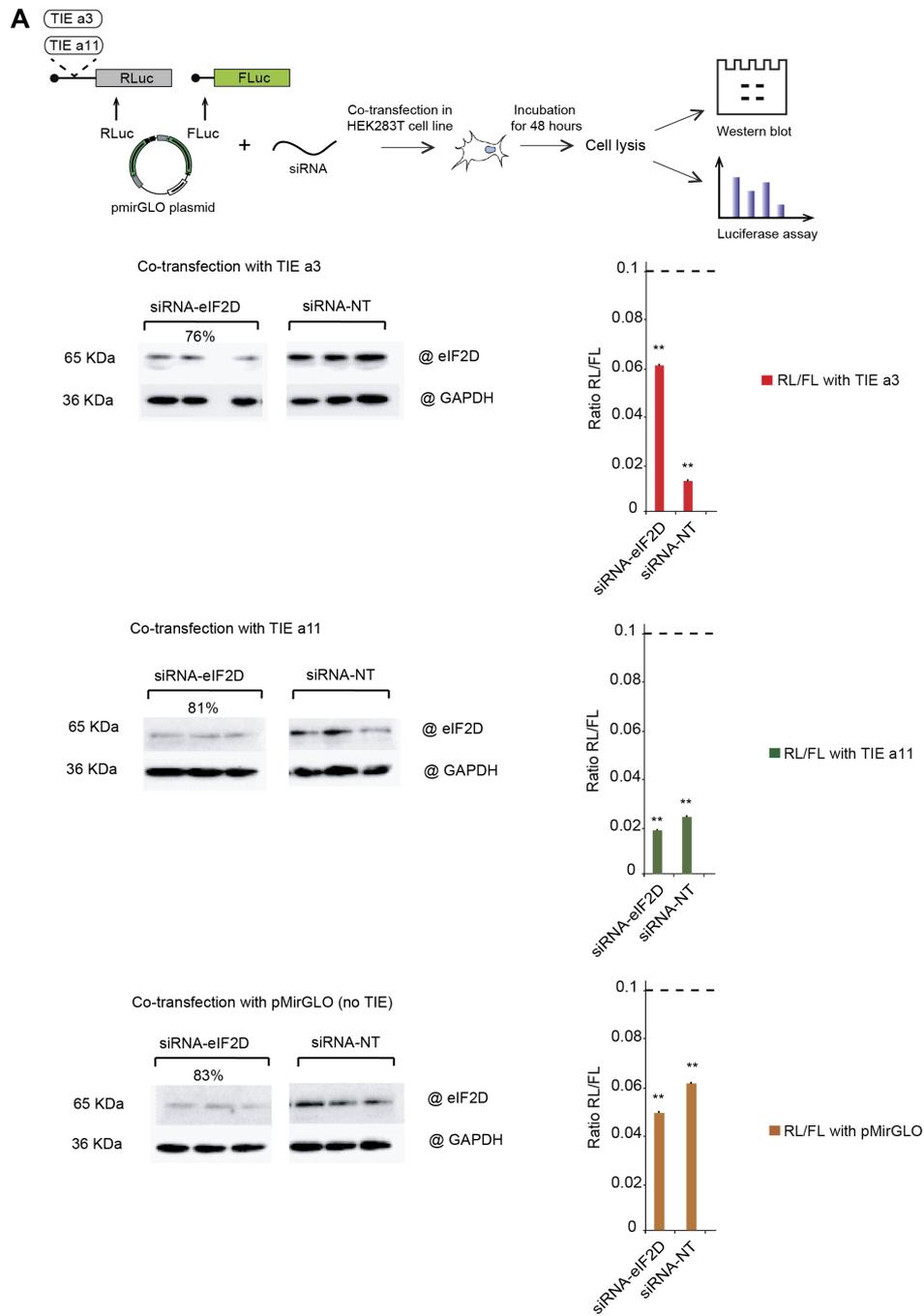


Figure 7



B

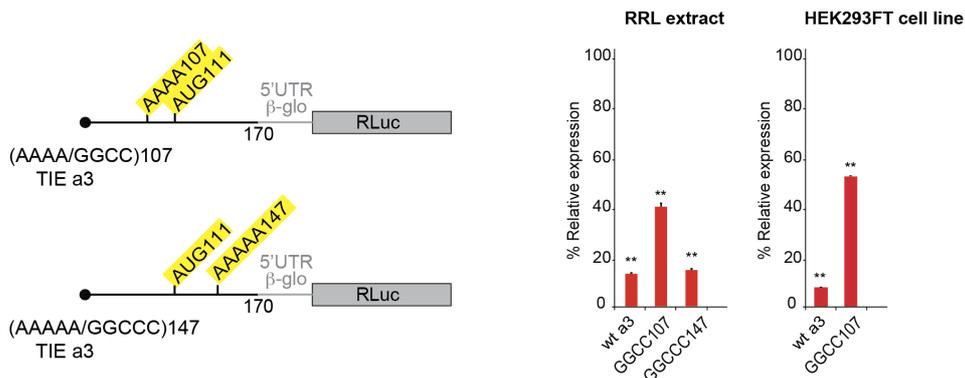


Figure 8



Figure 1-figure supplement 1

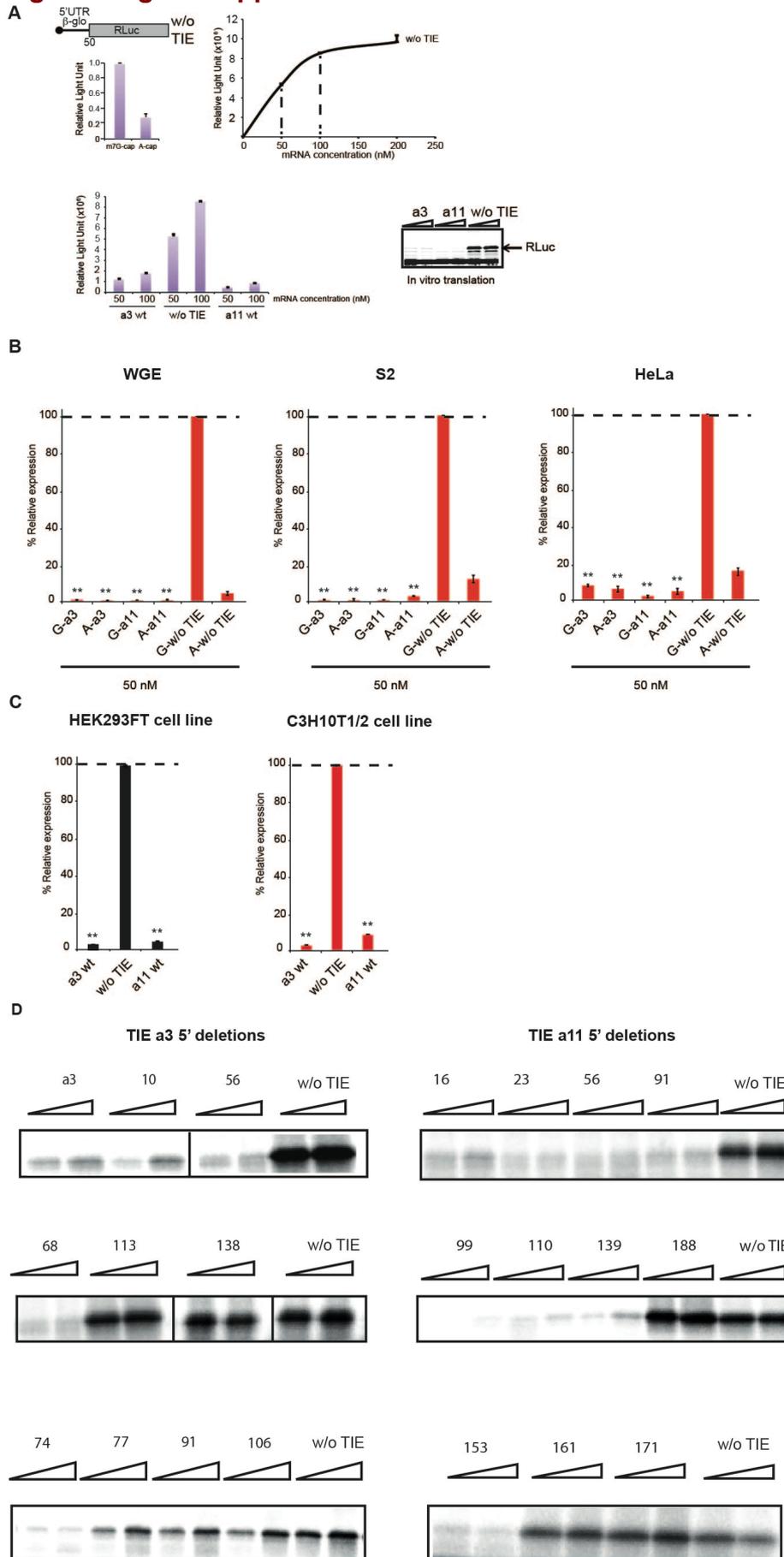


Figure 2-figure supplement 1

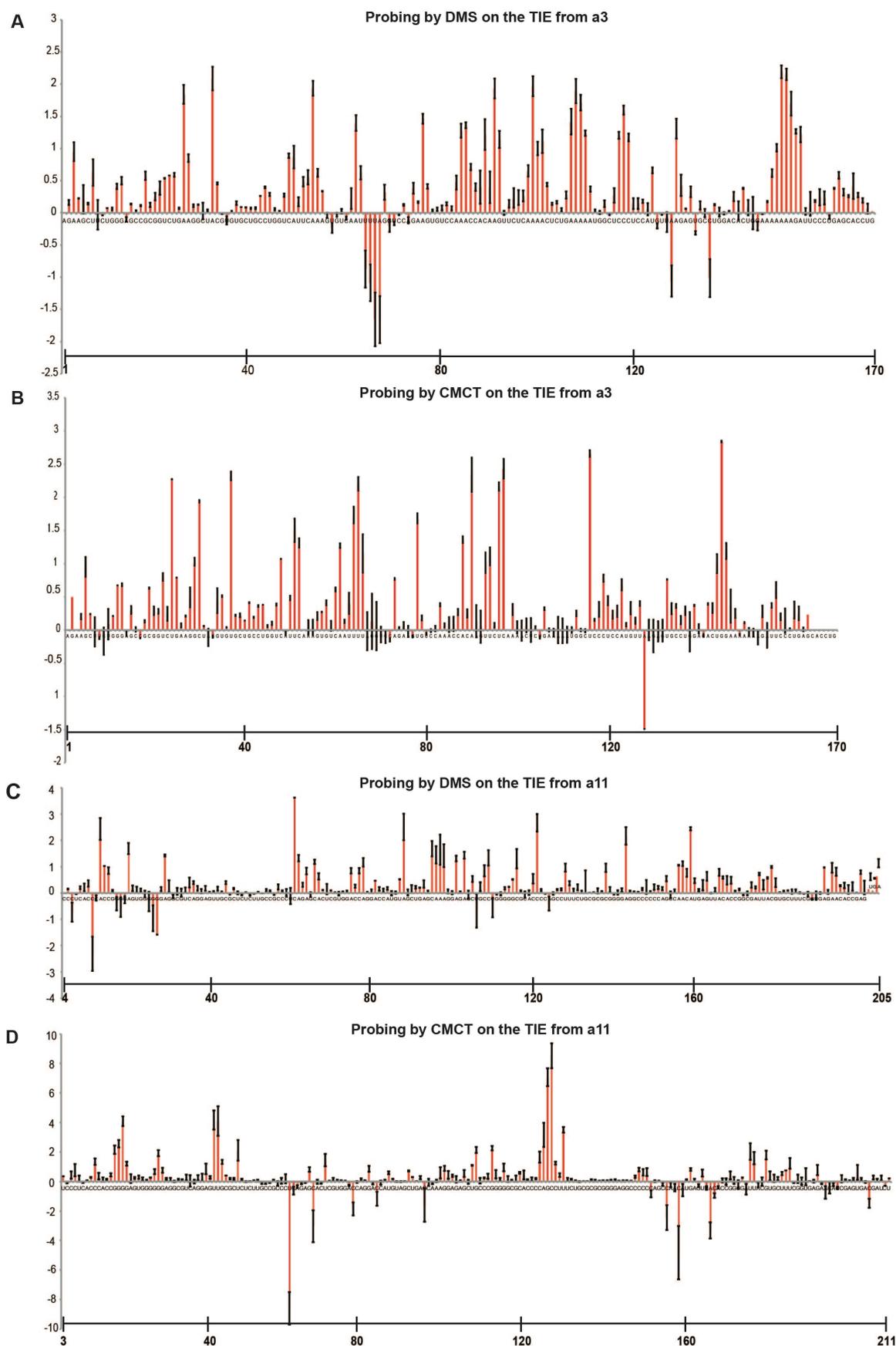
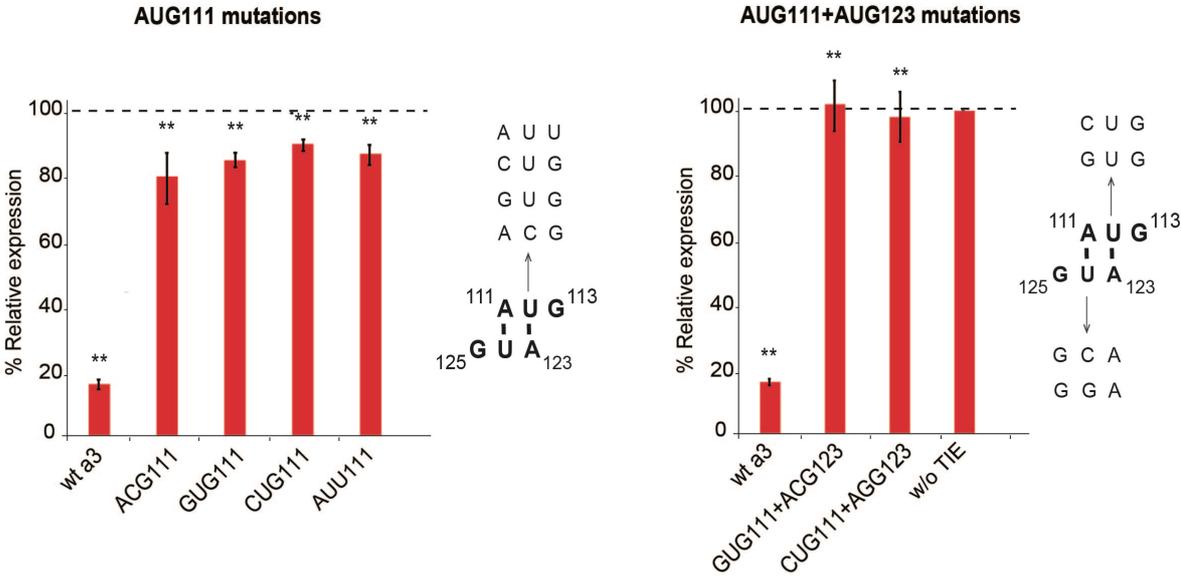


Figure 3-figure supplement 1

A



B

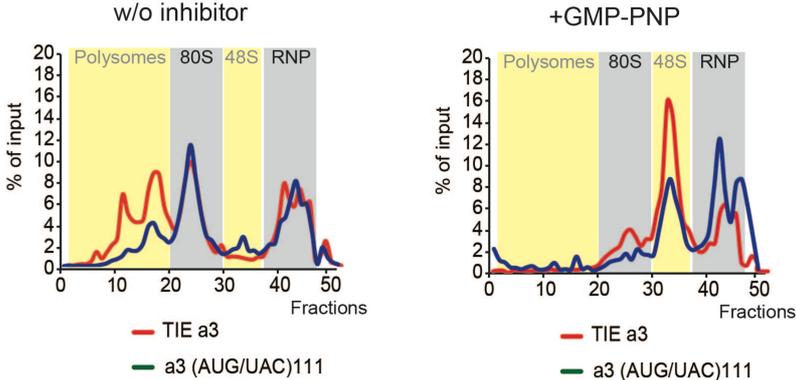
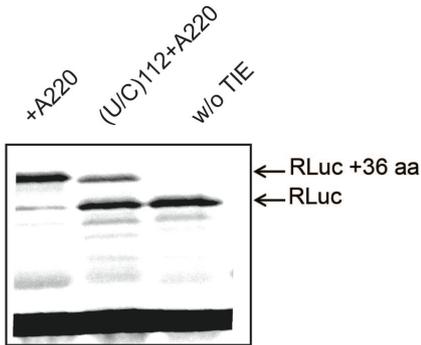
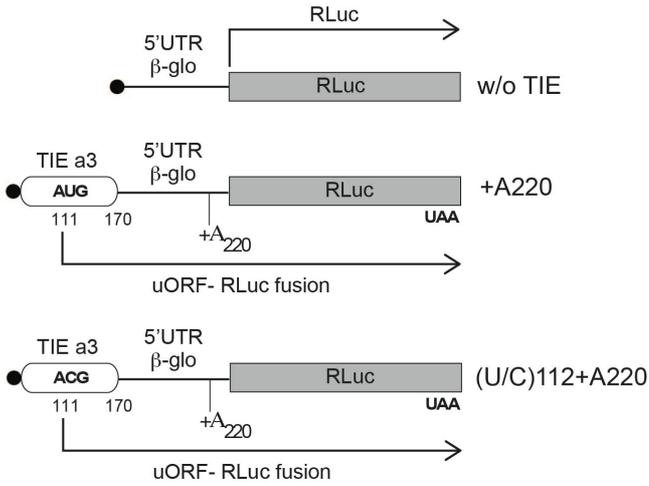


Figure 4-figure supplement 1

A



B

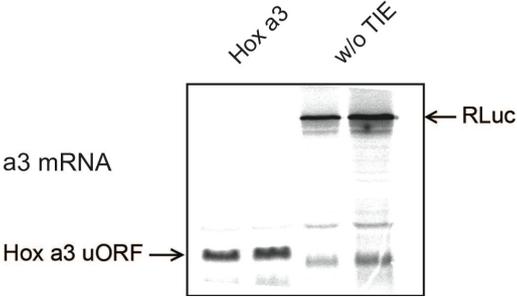
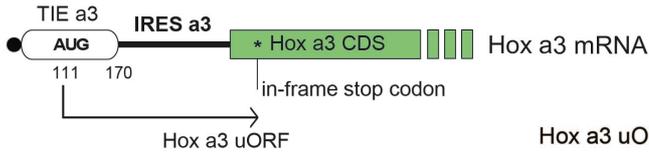


Figure 5-figure supplement 1

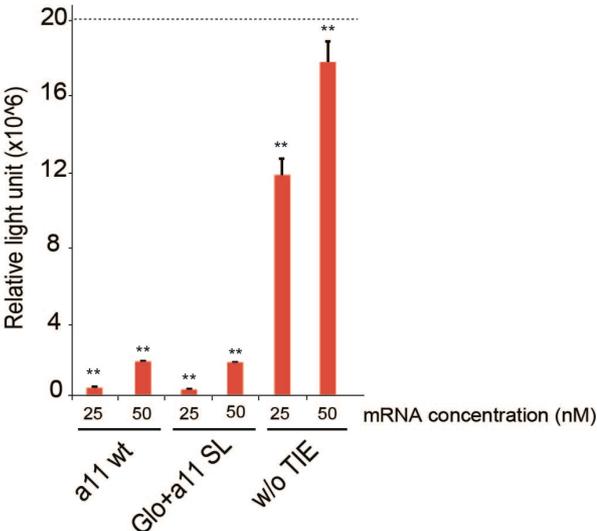
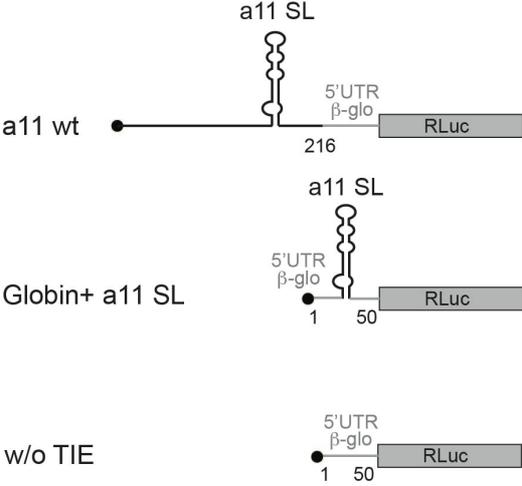


Figure 5-figure supplement 2

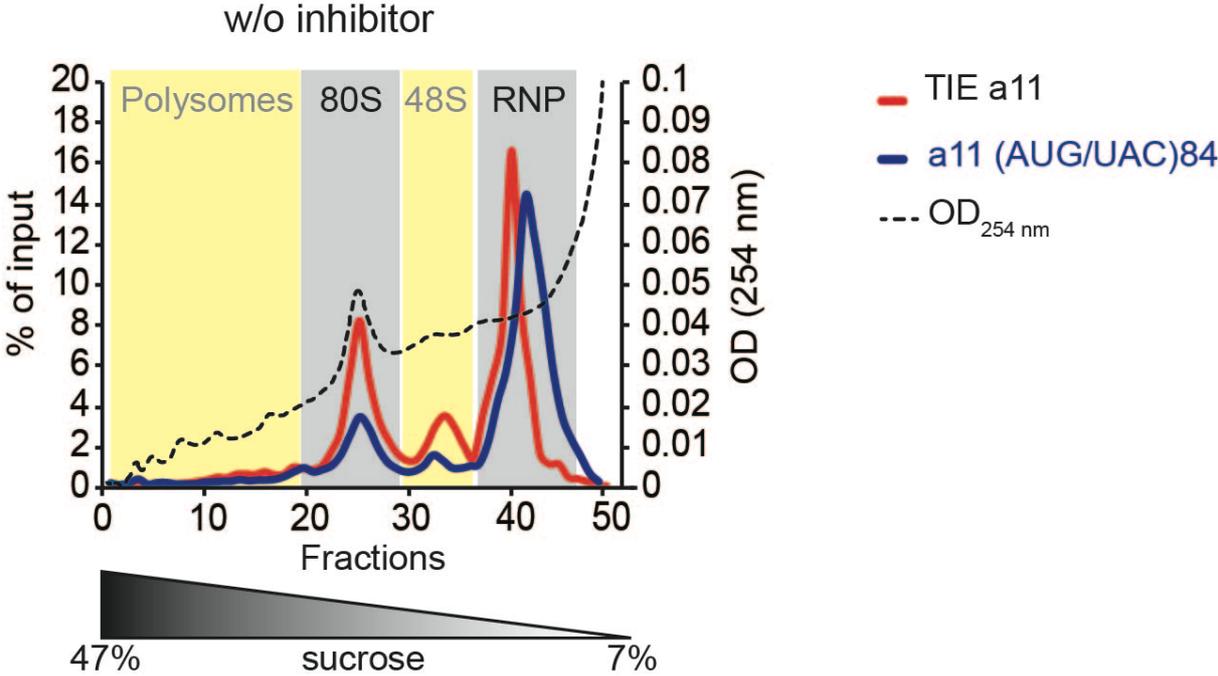


Figure 6-figure supplement 1
+GMP-PNP

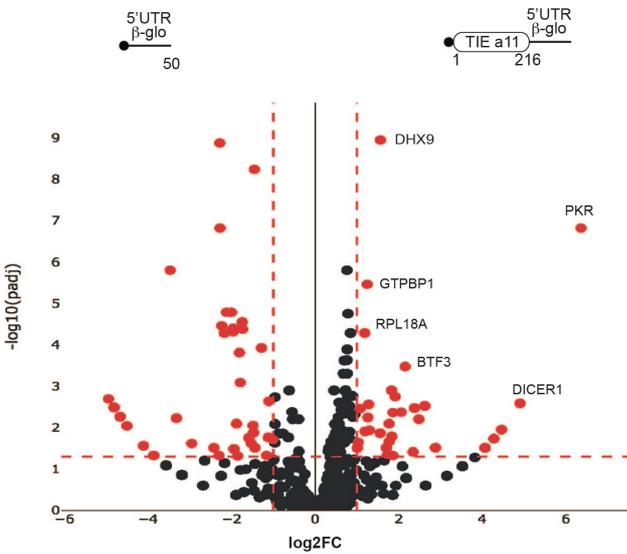
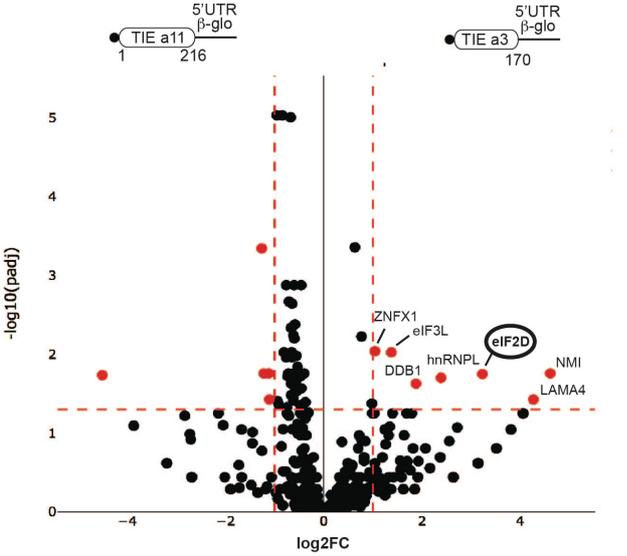
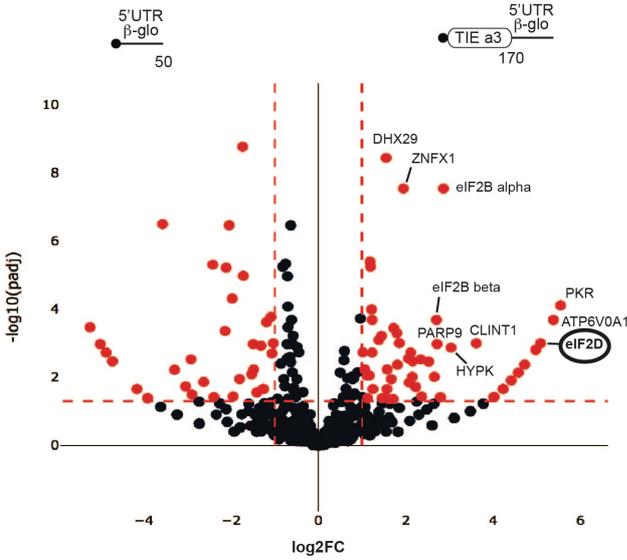


Figure 8-figure supplement 1

A

Consensus	AAA c t CTCTG	AAAAATGGCT	CCCTCCGAG -	- - - - -
Conservation				
F. catus	AAAC TCTCTG	AAAAATGGCT	CCCTCCGAG -	- - - - -
M. musculus	AAA - - CTCTG	AAAAATGGCT	CCCTCCATG -	- - - - -
M. flaviventris	AAA - - CTCTG	AAAAATGGCT	CCCTCCGAG -	- - - - -
H. sapiens	AAAC TCTCTG	AAAAATGGCT	CCCTCCGAGT	TAA GCAATTC
M. mulatta	AAA - - CTCTG	AAAAATGGCT	CCCTCCGAGT	TAA GCAATTC
P. anubis	AAA - - CTCTG	AAAAATGGCT	CCCTCCGAG -	- - - - -
E. caballus	AAAC TCTCTG	AAAAATGGCT	CCCTCCGAG -	- - - - -
U. horribilis	AAAC TCTCTG	AAAAATGGCT	CCCTCCGAC -	- - - - -
C. dingo	AAAC TCTCTG	AAAAATGGCT	CCCTCCGAG -	- - - - -

B

Consensus	51 - - - - - gTT	61 g c g c T c T c T c	71 a C C g c C c T c -	81 a g c g c a c T c g	91 t G g a C c a g g A
Conservation					
H. sapiens	GGTCTCTGTT	AATTTTTTTT	TCCTTCATT -	- - - - - CTCC	TGTTCTGAAA
C. griseus	- - - - - GTT	GCGCTCTCTC	ACCGCCCTC -	AGCGCATTCC	TGGACCAGGA
M. auratus	- - - - - CT	GCGCTCTCTC	GCGCCCTCA	AGCGCATTCC	TGGACCAGGA
M. musculus	- - - - - GTT	GCGCTCTCTT	GCGCCCTC	AGAGCACTCC	TGGACCAGGA
R. norvegicus	- - - - - GTT	GCGCTCTCTT	GCGCCCTC	AGAGCGCTCC	TGGACCAGGA
U. maritimus	- - - - - GTT	GAGCTCTCTC	ACCGCCCTC -	GGCGCACTTG	AGGGCCAGGA
C. dingo	- - - - - GTT	GAGCTCTCTC	ACCGCCCTC -	GGCGCACTTG	AGGGCCAGGA
E. caballus	- - - - - - - -	- - - - - - - -	- - - - - - - -	- - - - - - - -	- - - - - - - -
Consensus	101 C C a - - - - -	111 - - - - -	121 - - - t G c a G c T	131 G A G C A - - a a g	141 G a g A g c t G C t
Conservation					
H. sapiens	CCA G A T T T T G	A C T T G A C G A T	C A G T G A G G T T	G A G C A T G C G G	G A C A G T T G C -
C. griseus	CCA - - - - -	- - - - -	- - - G G T A G C T	G A G C A - - A A G	G A G A G C T G C T
M. auratus	CCA - - - - -	- - - - -	- - - G G T A G C T	G A G C A - - A A G	G A G A G C T G C T
M. musculus	CCA - - - - -	- - - - -	- - - T G T A G C T	G A G C A - - A A G	G A G A G C T G C C
R. norvegicus	CCA - - - - -	- - - - -	- - - T G C A G C T	G A G C A - - A A G	G A A A G C T G C T
U. maritimus	CCG - - - - -	- - - - -	- - - T G C A G C T	G A G C A - - C A C	G C G A G T G G C C
C. dingo	CCG - - - - -	- - - - -	- - - T G C A G C T	G A G C A - - C A G	G C G A G C G G C A
E. caballus	- - - - - - - -	- - - - - - - -	- - - - - - - -	- - - - - - - -	- - - - - - - -
Consensus	151 t - g G G G C G C A	161 C C c C A G C C T t	171 t C t G - - - C G	181 C G C g g G G A G G	191 C C C C C C - - - -
Conservation					
H. sapiens	- - - A G G C G C T	T C T C T T T G T T	A A T G T A G A C G	C T G A A G A A G A	A C T C C C G T T C
C. griseus	T G G G G G C G C A	C C G C A G C C T T	T C T G - - - C G	C G C G G G A G G	C C C C C C - - - -
M. auratus	T G G G G G C G C A	C C G C A G C C T T	T C T G - - - C G	C G C G G G A G G	C C C C C C - - - -
M. musculus	C G G G G G C G C A	C C C C A G C C T T	T C T G - - - C G	C G C G G G A G G	C C C C C C - - - -
R. norvegicus	T G G G G G C G C A	C C C C A G C C T T	T C T G - - - C G	A G C G G G A G G	C C C C C C - - - -
U. maritimus	T - G G G G C G C A	C C T C A G C C T C	C C C G - - - C G	C G C C T G G A G G	C C C C C C - - - -
C. dingo	T - G G G G C G C A	C C C C A G C C T C	G C C G - - - C G	C G C C G G A G G	C C C C C C - - - -
E. caballus	- - - G G G C G C A	C C C C A G C C T C	G C C G - - - C G	C G C C G G A G G	C C C C C C - - - -