Published in final edited form as:

Nat Struct Mol Biol. 2018 August ; 25(8): 677–686. doi:10.1038/s41594-018-0091-z.

Analyses of mRNA structure dynamics identify embryonic gene regulatory programs

Jean-Denis Beaudoin[1,8,*], **Eva Maria Novoa**[2,3,4,5,8], **Charles E. Vejnar**[1], **Valeria Yartseva**[1], **Carter M. Takacs**[1,6], **Manolis Kellis**[2,3], **Antonio J. Giraldez**[1,7,*]

[1]Department of Genetics, Yale University School of Medicine, New Haven, CT, USA.

[2]Computer Science and Electrical Engineering Department, Massachusetts Institute of Technology, Cambridge, MA, USA.

[3]The Broad Institute of MIT and Harvard, Cambridge, MA, USA.

[4]Department of Neuroscience, Garvan Institute of Medical Research, Darlinghurst, New South Wales, Australia.

[5]School of Medicine, University of New South Wales, Sydney, New South Wales, Australia.

[6]College of Arts and Sciences, University of New Haven, West Haven, CT, USA.

[7]Yale Stem Cell Center, Yale University School of Medicine, New Haven, CT, USA.

[8]These authors contributed equally: Jean-Denis Beaudoin, Eva Maria Novoa.

Abstract

RNA folding plays a crucial role in RNA function. However, knowledge of the global structure of the transcriptome is limited to cellular systems at steady state, thus hindering the understanding of RNA structure dynamics during biological transitions and how it influences gene function. Here, we characterized mRNA structure dynamics during zebrafish development. We observed that on a global level, translation guides structure rather than structure guiding translation. We detected a decrease in structure in translated regions and identified the ribosome as a major remodeler of RNA structure in vivo. In contrast, we found that 3' untranslated regions (UTRs) form highly folded structures in vivo, which can affect gene expression by modulating microRNA activity. Furthermore, dynamic 3'-UTR structures contain RNA-decay elements, such as the regulatory elements in *nanog* and *ccna1*, two genes encoding key maternal factors orchestrating the maternal-

[*]**Correspondence and requests for** materials should be addressed to J.-D.B. or A.J.G. jean-denis.beaudoin@yale.edu; antonio.giraldez@yale.edu.

Author contributions

J.-D.B. and A.J.G. conceived the project. J.-D.B. performed the experiments. J.-D.B., E.M.N. and C.E.V. performed data processing. V.Y. identified the regulatory element in the *nanog* 3′ UTR and built the reporter constructs. C.M.T. performed the KHSRP iCLIP experiment. J.-D.B. and E.M.N. performed data analysis and, together with A.J.G., interpreted the results. A.J.G. supervised the project, with the contribution of M.K. J.-D.B., E.M.N. and A.J.G. wrote the manuscript with input from the other authors.

Competing interests

The authors declare no competing interests.

Supplementary information is available for this paper at https://doi.org/10.1038/s41594-018-0091-z.

to-zygotic transition. These results reveal a central role of RNA structure dynamics in gene regulatory programs.

RNA carries out a broad range of functions[1], and RNA structure has emerged as a fundamental regulator)" mechanism modulating various post-transcriptional events, including splicing[2], subcellular localization[3], translation[4,5] and decay[1]: RNA-probing reagents combined with high-throughput sequencing[7] allow for interrogation of RNA structure in a transcriptome-wide manner in vitro and in vivo[8,9]. These approaches have revealed common features of RNA structures and their regulation at steady state[10–14] that suggest that RNAs tend to be unfolded in vivo[13,15]. However the cellular factors remodeling RNA folding in vivo remain unknown.

Although the ribosome possesses a constitutive mRNA helicase activity[16], stable RNA structures found within coding regions can decrease the rate of translation in vitro[17,18]. Similarly, the intrinsic structure adopted by coding regions affects the rate of translation in bacteria, in which sequences forming stable structures show decreased translation in vivo[19]. In addition, stable structures located in the 5' UTR or around the AUG initiation codon repress translation and modulate protein output in bacteria and eukaryotes[20–23]. In plants and yeast, highly structured mRNA in vitro correlates with higher ribosome density and protein output[9,24]; however, other studies have found that RNA structures in yeast are not correlated with translation efficiency[13].

Previous studies have focused on the global analysis of RNA structures in cells at steady state. However, in an organism, critical developmental and metabolic decisions are often made when cells transition between cellular states. Consequently, to understand the relationships among RNA structure, regulation and function in vivo, it is paramount to determine how RNA structures change in dynamic cellular environments. During the maternal-to-zygotic transition (MZT) in animals, the embryo is initially transcriptionally silent, and gene expression is primarily orchestrated by post-transcriptional regulation of maternal mRNA translation and decay[25,26]. These regulatory pathways are central to embryogenesis, controlling mRNA translation, the cell cycle and the activation of the zygotic genome[25,27]. Among the set of maternally deposited mRNAs, the transcription-factor-encoding genes *nanog, oct4* (official symbol *pou5f3*) and *sox19b* are highly translated. Their combined function is required to activate a large fraction of the zygotic program[28,29]. One of the first zygotically transcribed RNAs is the micsroRNA (miRNA) miR-430 (refs.[28,30]), which causes translation repression, deadenylation and clearance of hundreds of maternal mRNAs during the MZT[31–33]. Thus, the MZT provides an ideal system to understand how the entire cellular population of RNA structures is remodeled across changing cellular states, and how these changes may affect gene regulation. Here, we analyzed mRNA structural dynamics during zebrafish embryogenesis to investigate the relationship between RNA structure and gene regulation in vivo.

Results

Dynamic structure reflects translation changes.

We analyzed the structure of the zebrafish transcriptome during the MZT (Fig. 1a) by using dimethyl sulfate sequencing (DMS-seq[13]), which we validated as an accurate readout of RNA structure in zebrafish embryos (Supplementary Fig. 1 and 'DMS-seq controls' section in Methods). Our analysis revealed global changes in mRNA structure during the MZT (Supplementary Fig. 1i–k). Changes in translation, analyzed through ribosome footprinting[34], were correlated with global mRNA accessibility changes in coding sequences (CDS) and 5' UTRs, but not in 3' UTRs (Fig. 1b and Supplementary Fig. 1l). We analyzed differentially structured regions over development as 100-nt sliding windows (Fig. 1c) and observed that maternal mRNAs with decreasing rates of translation between 2 and 6 hours postfertilization (hpf) increased in structure (orange), and those whose translation increased became less structured (turquoise) (Fig. 1d). These results indicate that translation and mRNA structure are anticorrelated in vivo, in contrast to previously reported findings[9,13,24,35]. Moreover, we compared the accessibility of the different mRNA regions (5′ UTR, CDS and 3′ UTR) and found that highly translated mRNAs displayed relatively greater accessibility in their CDS ($P = 2.3 \times 10^{-118}$) and, to a lesser extent, in their 5′ UTRs ($P = 1.9 \times 10^{-11}$; Fig. 1e and Supplementary Fig. 2a).

Preexisting differences in RNA structure might determine mRNA translation, wherein less structured mRNAs would be more accessible and consequently would be translated more effectively (Fig. 2a). Alternatively, high translation rates might lead to lower structure in vivo, owing to constant mRNA unfolding by the ribosome (Fig. 2a). To distinguish between these scenarios, we compared the accessibility of transcripts analyzed in vitro, binned by their translation efficiency in vivo (Methods) (Fig. 2b and Supplementary Fig. 2a). We found that weakly and highly translated mRNAs showed similar CDS accessibility in vitro, and the results were distinct from those in vivo (Fig. 1e, Fig. 2b,c and Supplementary Fig. 2a–c), thus suggesting that the differences in mRNA accessibility between highly and weakly translated mRNAs are not intrinsic to the nucleotide sequence. Although very stable structures in the 5′ UTR clearly disrupt translation[20,22,23], we observed no correlation between the translation and structure of 5′ UTRs or the AUG initiation codons in vitro (Fig. 2d,e and Supplementary Fig. 2d–g). Thus, on a global level, RNA structure in these regions is not a major determinant of translation across the transcriptome in vertebrate embryos.

Ribosomes shape mRNA structures.

The helicase activity of the ribosome unfolds RNAs in vitro during translocation[16–18]. However, in yeast, there is no correlation between RNA structure and translation efficiency[13], thus suggesting that ribosome activity does not account for the differences between mRNA structures in vivo and in vitro. To determine the role of ribosomes in structural remodeling in vivo, we examined RNA structural changes in embryos incubated with inhibitors of translation initiation (pateamine A; PatA)[36,37] or elongation (cycloheximide; CHX)[38]. Ribosome profiling of 64-cell embryos (2 hpf) treated with PatA revealed prominently lower (~ 34-fold) translation efficiency than that in untreated embryos (Fig. 3a,b). mRNAs most sensitive to PatA treatment had longer 5′ UTRs ($P = 2.1 \times 10^{-42}$),

in agreement with a loss of the PatA target eIF4A during scanning by the preinitiation complex[36,37] (Fig. 3a,c).

In contrast, mitochondrial transcripts showed no change in translation (Fig. 3d), in agreement with the observation that their translation is independent of eIF4A[39]. In PatA-treated embryos, we found an overall decrease in accessibility of the CDS regions of highly translated mRNAs ($P = 5.3 \times 10^{-35}$; Fig. 3e and Supplementary Fig. 3a). Similarly, regions with higher ribosome-footprint densities formed more stable RNA conformations after PatA treatment, similarly to those observed in vitro (Fig. 3f and Supplementary Fig. 3b, c). In contrast to PatA treatment, CHX treatment did not decrease the average CDS accessibility (Supplementary Fig. 3d), thus suggesting that even when the ribosomes are stalled[38], the residency of the ribosome on the mRNA during translation plays a major role in remodeling mRNA structure. Therefore, translation initiation and ribosomal entry are needed to locally favor the formation of less stable alternative RNA structures.

To evaluate the effect of the ribosome remodeling activity on mRNA structures, we compared the RNA conformations adopted by full-length transcripts in each probing condition by using SeqFold[40]. We found that the same transcript favored different conformations depending on its translation status (Fig. 3g and Supplementary Fig. 3e). Principal component analysis (PCA) on the structural data from 1,143 transcripts (Methods) revealed that in the absence of ribosomes (in vitro or in the presence of PatA), individual mRNAs exhibited similar CDS and 3′-UTR structures distinct from those formed in the presence of ribosomes (untreated and CHX treated; Fig. 3h,i). Finally, translation of upstream open reading frames (uORFs) located in 5′ UTRs led to a restructuring of 5′-UTR structures in vivo (Supplementary Fig. 4). Thus, the ribosome increases mRNA accessibility and promotes alternative RNA conformations throughout the mRNA.

Translation-dependent mRNA structure.

We then evaluated translation control mechanisms during the MZT for their effect on mRNA structure regulation (Fig. 4a). Polyadenylation increases translation efficiency during early embryogenesis[33,41] (Fig. 4a). We found that mRNAs with longer poly(A) tails showed higher accessibility in the CDS at 2 and 4 hpf (ρ of 0.42 and 0.36, respectively) but not in the 3′ UTR (Supplementary Fig. 5a). This effect was decreased after PatA treatment and was lost in vitro (ρ of 0.29 and 0.01, respectively) (Supplementary Fig. 5b,c), in agreement with a translation effect specific to the CDS. Moreover, mRNAs with extended poly(A) tails during the MZT displayed a decrease in structure, whereas deadenylated transcripts showed an increase in structure (Fig. 4b,c). These results suggest that changes in polyadenylation during development influence mRNA structure, an effect that depends on the translation of the mRNA.

To understand the effect of miRNA-mediated repression on mRNA structural changes, we analyzed miR-430, which regulates hundreds of maternal mRNAs during the MZT (Figs. 1a and 4a). The miR-430 targets are initially translationally repressed at 4 hpf without substantial signs of decay[42]. If ribosomes play a crucial role in shaping mRNA structure in vivo, mRNAs targeted by miR-430 should reach an intermediate state at 4 hpf, which is characterized by an increase in structure due to depletion of ribosomes. To test this

hypothesis, we monitored structural changes in 483 miR-430 targets and a control set of 1,495 mRNAs not targeted by miR-430. We observed a global increase in structure (orange) for miR-430 targets compared with the control set, which was characterized by a decrease in structure (turquoise) between 2 and 4 hpf (Fig. 4d,e and Supplementary Fig. 5d). These results support a central role of the ribosome in orchestrating mRNA structure in the cell and suggest that biological processes regulating translation can have a broad effect on directing mRNA folding.

3′ UTRs have a distinct folding landscape in the cell.

We found that zebrafish transcripts were globally more accessible in vivo than in vitro (Supplementary Fig. 6a). To investigate whether this global unfolding was uniformly distributed along the transcript, we performed a metagene analysis of the structure, for both in vivo and in vitro conditions, by using a sliding-window approach (Fig. 1b). Our results revealed that regions with decreased structure in vivo were mainly located in the 5′-UTR and CDS of highly translated transcripts (Fig. 5a,b and Supplementary Fig. 6b), in agreement with the remodeling role of the ribosome. In contrast, 3′ UTRs showed increased structure in vivo. This increase in RNA structure in vivo did not seem to primarily result from RNA-binding-protein footprints, because no decrease in accessibility was observed over the binding sites of the AU-rich element-binding protein KHSRP and surrounding regions recognized by the Ago2-miR-430 complex (Supplementary Fig. 7 and ‘RNA-RNA versus RNA-protein interaction analysis’ section in Methods). KHSRP-bound regions maintained a higher accessibility in vivo than did their flanking regions, thus suggesting that KHSRP traps the RNA in a single-stranded region and that our approach maps the RNA structure of the RNA-protein complex (Supplementary Fig. 7a–d). Although this possibility is not supported by our controls, we cannot exclude that other proteins might cause a local increase in RNA structure by directly interacting with the DMS-methylation site. Finally, the increased structure of 3′ UTRs in vivo was anticorrelated with translation efficiency and spread into the CDS of poorly translated mRNAs, thus suggesting that the increased structure was specific to ribosome-free regions (Fig. 5b and Supplementary Fig. 6b).

To evaluate the importance of the 3′-UTR RNA folding landscape in the cell, we analyzed the effects of RNA structures on miR-430 activity. Because the accessibility of miRNA-binding sites has been shown to affect their activity[9,43,44], we analyzed how miRNA targeting efficiency compared to the stability (ΔG) of the predicted structure in silico, in vitro and in vivo (Fig. 6a,b). We found that the stability of the RNA structure (ΔG) according to in vivo probing was the best predictor of miR-430 activity ($\rho = -0.67$), outperforming in silico ($\rho = -0.34$) and in vitro ($\rho = -0.60$) structure predictions. In agreement with these results, endogenous miRNA targets with higher free energy in vivo were more strongly regulated by miR-430 than those with lower free energy (Fig. 6c,d,g,h and Supplementary Fig. 8). To confirm the regulatory roles of those in vivo-specific RNA conformations, we compared the activity of endogenous targets with different free energy and analyzed the effects of altering their structure on their regulation. To this end, we quantified the repression of a GFP reporter mRNA in wild-type and mutant embryos in which the miR-430 locus was deleted (MZ*miR430*). As a control, we co-injected a DsRed reporter mRNA not targeted by miR-430. Modifying the stability of the in vivo structure in the target site altered the

regulatory strength of each site. For example, destabilizing the in vivo structure of the *rab33ba* target site, which was poorly regulated by miR-430 in vivo (Fig. 6d), increased miR-430-mediated regulation (Fig. 6e,f). In contrast, stabilizing the in vivo structure of *faml71al*, which was strongly regulated by miR-430 (Fig. 6h), abolished miR-430 activity (Fig. 6i,j). Altogether, these results demonstrate that there are increases in 3′-UTR structures in vivo, which can regulate gene expression by modulating miRNA activity.

Dynamic 3′ UTRs are enriched in functional elements.

During the MZT, maternally deposited factors (Nanog, Oct4 and SoxBl) mediate activation of the zygotic genome[28], after which different programs mediate post-transcriptional regulation of maternal mRNAs (Fig. 1a). This process is required for the embryo to transition from the maternal to the zygotic state, and it is characterized by changes in the transcriptional and epigenetic landscapes, cellular motility and the cell cycle, with an extension of G_1 and G_2(ref.[27]).

Analysis of the structurally dynamic regions in the 3′ UTR between 4 and 6 hpf revealed an enrichment in conserved sequences (Fig. 7a and Supplementary Fig. 9a), thus suggesting a potential role in the post-transcriptional regulation of the maternal mRNAs. To test this hypothesis, we performed a parallel reporter assay (RNA-element selection assay (RESA))[45] to measure the regulatory role of 71 different 3′-UTR regions derived from maternal mRNAs subject to decay during the MZT (Fig. 7b). Out of these 71 regions, 53 contained dynamic RNA structures, whereas 18 showed no structural changes during MZT. These sequences, inserted into the 3′ UTR of a reporter gene, were in vitro transcribed and injected in one-cell-stage embryos. Then, the abundance of each reporter was measured by high-throughput sequencing at 2 and 6 hpf (Fig. 7b). To identify zygotic-dependent regulation, we compared the levels of each fragment in the presence or absence of zygotic transcription (i.e., with or without α-amanitin). Quantification of mRNA reporters identified decay elements in the 3′ UTRs of the cyclin A1 (*ccna1*) and *nanog* mRNAs and an enrichment in destabilizing elements in regions that changed in structure during development compared with regions that did not change ($P = 0.025$; Fig. 7c,d and Supplementary Fig. 9b–e). These results suggest that regions with dynamic structures are enriched in functional regulatory elements.

One such dynamic regulatory region is located in the *ccna1* mRNA. Cyclin A1 controls the cell cycle and is actively degraded during the MZT. Time-course analyses of mRNA expression during the MZT through RNA-seq revealed that *ccna1* mRNA decayed in wild-type embryos but not in embryos in which zygotic transcription was blocked with a-amanitin (Fig. 7e and Supplementary Fig. 9f). The structurally dynamic region of *ccna1* tested in RESA was also found to possess a decay activity relying on the activation of the zygotic program (Fig. 7f). This result is consistent with the regulation of *ccna1* in vivo and the change in cell-cycle length after zygotic genome activation. Another dynamic region with regulatory activity was found in the *nanog* mRNA, which encodes a transcription factor that is required for genome activation and miR-430 expression, and is degraded after zygotic genome activation (Fig. 7e and Supplementary Fig. 9f). We predicted the secondary structure of the *nanog* 3′ UTR at 2 and 6 hpf by using in vivo DMS-seq accessibility values

as constraints (Fig. 7g). We identified two domains with constant structure (I and III) that flank a region with differential structure (II). Domain II overlaps with the dynamic region that was found to have regulatory activity in vivo (Fig. 7d,h). To test the regulatory activity of this region in the context of the 3′ UTR, we compared the regulation provided by the wild type *nanog*3′ UTR and a deletion-mutant reporter that disrupted the stem region (II) with a 20-nt deletion (Fig. 7g,h). The full-length *nanog* 3′ UTR had a decay activity similar to that of the differentially structured region tested in RESA (Fig. 7h). However, a deletion of 20 nt in domain II stabilized the reporter mRNA, thus suggesting that this region is required to confer regulation in vivo (Fig. 7g,h). Altogether, these results demonstrate that following mRNA structure dynamics through biological transitions can reveal functional elements in 3′ UTRs.

Discussion

Here, we examined the mRNA structure dynamics during the MZT, a universal regulatory transition in animal embryogenesis. Our results revealed that 3′ UTRs and untranslated mRNAs are more structured in the embryo than in vitro. Dynamic 3′-UTR structures in vivo correspond to novel regulatory regions controlling maternal mRNA decay. Finally, we demonstrated that the ribosome is one of the principal factors unwinding mRNA, and although structure did not emerge as a global regulator of translation, translation was found to be a major driver shaping the mRNA structure landscape during early embryogenesis.

Previous efforts to characterize the mRNA folding landscape have reported various structural features of UTRs, CDS and regions surrounding translation start and stop codons, both in vitro[9,11,35,45,47] and in vivo[10,13,14]. Indeed, UTRs are more structured than coding regions in *Drosophila melanogaster, Caenorhabditis elegans*[46] and humans[47] in vitro, whereas the inverse is observed in *Saccharomyces cerevisiae*[11] and *Arabidopsis thaliana*[10,35]. Although very informative, the analysis of steady-state conditions provides limited information on how the mRNA folding landscape changes in the cellular environment or during biological transitions. Thus, to study mRNA structure dynamics, we focused our analysis on the differences in RNA structure across conditions and developmental stages. The comparison of in vivo and in vitro RNA structures allowed us to capture the effect of the cellular environment on the folding of mRNAs. We found that untranslated mRNAs and 3′ UTRs were more structured in vivo than in vitro (Fig. 5a,b and Supplementary Fig. 6b), a finding contrary to work showing that mRNAs are mostly unfolded in the cell[13,15]. Previous analyses comparing mRNA structures in vitro and in vivo had been carried out in yeast, in which the average 3′-UTR length is much shorter than that in vertebrate mRNAs[13]. In the absence of constitutive unwinding by the ribosome, both molecular crowding and RNA-binding proteins are likely to favor higher-order RNA structures[48,49]. Remodeling of 3′-UTR structures in vivo can also affect regulatory structural elements modulating 3′-end processing and mRNA stability[50]. Our data support a view in which cellular factors explicitly affect 3′-UTR structures, modulating the regulatory activity of miRNAs and gene expression during vertebrate development.

Our analysis of dynamic mRNA structures revealed regulatory regions during the MZT in *ccnal* and *nanog* mRNAs, which encode two maternal factors required to regulate the cell

cycle and zygotic transcription, and are rapidly degraded during the MZT. We found that the characterization of structurally dynamic regions during cellular or developmental transitions allowed us to identify potential regulatory elements in vivo (Fig. 7b–d and Supplementary Fig. 9b,c). Interestingly, the analysis of the RNA folding landscape in plant nuclear extracts has unveiled differential RNA structures in 3′-UTR regions across cell types[51]. These results are in agreement with our finding that dynamic 3′-UTR structures during developmental transitions are enriched in functional elements, and this enrichment may potentially be conserved in plants.

Global RNA unwinding in yeast is an active process[13]. The cellular factors responsible for this structural remodeling and the specific regions that are remodeled remain unknown. Although the ribosome has been shown to have helicase activity in vitro[16], previous experiments in exponentially growing yeast have not revealed a relationship between translation and RNA unfolding[13]. Here, we followed the dynamics of mRNA and demonstrated that the ribosome is a major engine for mRNA structure remodeling, promoting alternative RNA conformations for thousands of endogenous transcripts in vertebrate cells. This effect has also been observed in 5′ UTRs containing translated uORFs (Supplementary Fig. 4), thus indicating that the ribosome helicase activity has the potential to regulate elements that rely on specific RNA conformations[52,53]. Several mechanisms regulating mRNA translation, such as changes in poly(A)-tail length and miRNA-mediated repression, also result in the remodeling of mRNA structure during embryogenesis.

RNA structure has generally been thought to have a major effect on mRNA translation rates[19], on the basis of early experiments in bacteria and eukaryotes, in which strong hairpin structures in mRNAs have been found to decrease translation[20–23]. Here, we demonstrated precisely the opposite effect; i.e., translation modulates mRNA structure, and mRNA structure does not have a global effect on translation during embryogenesis (Fig. 2 and Supplementary Fig. 2). Although individual examples in eukaryotes show that stable 5′-UTR structures can block translation[20,22], at the level of the transcriptome, complex eukaryotic cells have efficient helicase activity that greatly diminishes the global effects of 5′-UTR structure on translation. We propose that eukaryotic systems have evolved powerful helicases that, together with sophisticated mechanisms regulating translation, such as uORFs, miRNAs, RNA-binding proteins and longer 3′ UTRs, weaken the transcrip-tome-wide effects that 5′-UTR and CDS structures might have on translation. Our results prompt a conceptual shift in understanding of the relationship between mRNA structure and translation, and suggest that on a global level, translation guides structure rather than structure guiding translation.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/S41594-018-0091-z.

Methods

Zebrafish maintenance.

Wild-type zebrafish *(Danio rerio)* embryos were obtained through natural mating of the TU-AB strain of mixed ages (5–18 months). Mating pairs were randomly chosen from a pool of 60 males and 60 females allocated for each day of the month. Embryos and adult fish were maintained at 28 °C. Fish lines were maintained according to the International Association for Assessment and Accreditation of Laboratory Animal Care research guidelines, and protocols were approved by the Yale University Institutional Animal Care and Use Committee (IACUC).

In vivo and in vitro DMS modification.

For in vivo DMS modification, 150 embryos at the specified stage were transferred to 5-mL tubes containing 400 μL of system water from the fish facility. 100% DMS (Sigma-Aldrich) was diluted in 100% ethanol to obtain a 20% DMS stock solution. The DMS stock solution was used to generate a master mix containing 6% DMS and 600 mM Tris-HCl, pH 7.4 (AmericanBio), in system water from the fish facility. The master-mix solution was immediately mixed vigorously, and 200 μL was added to each tube to reach a final concentration of 2% DMS and 200 mM Tris-HCl, pH 7.4. Embryos were incubated at room temperature for l0 min with occasional gentle mixing. The DMS solution was then quickly removed from the tubes, and the embryos were flash frozen in liquid nitrogen. Frozen embryos were thawed and actively lysed with 800 μL of TRIzol (Life Technologies) supplemented with 0.7M β-mercaptoethanol (Sigma-Aldrich) to quench any remaining trace of DMS. After a 2-min incubation, TRIzol was added to reach a final volume of 4 mL, and total RNA was extracted according to the manufacturer's protocol. Poly(A)$^+$mRNAs were isolated with oligo d(T)$_{25}$ magnetic beads (New England BioLabs) according to the manufacturer′s protocol and eluted in 20 μL of water.

For in vitro DMS modification, embryos were collected, and total RNA was extracted as mentioned above, with DMS and β-mercaptoethanol omitted. For each replicate, 20 μg of total RNA was resuspended in RNA folding buffer (57 mM Tris-HCl, pH 7.4, and 114mM KC1) in a final volume of 44 μL. RNA samples were incubated at 65 °C for 2min and slowly brought back to 28 °C over ~ 45 min in a heating block Then, 5 μL of 100 mM $MgCl_2$was added to each tube, and samples were incubated at 28 °C for 3 min. RNA was modified by the addition of 1 μL of 25% DMS, previously diluted in ethanol (1:3 dilution), to reach a final DMS concentration of 0.5%. The reaction was incubated at 28 °C for 10 min in a PCR machine and quenched by the addition of 59 pi of stop solution (3 M β-mercaptoethanol, 508 mM sodium acetate and 15 μg glycoblue). After a 5-min incubation at room temperature, modified RNA samples were ethanol precipitated by the addition of 220 μL ethanol. Poly(A)$^+$mRNAs were isolated as described above.

All DMS-seq experiments (in vivo, in vitro, untreated, PatA treated and CHX treated at 2 hpf, as well as 4 and 6 hpf) were performed on biological duplicates from embryos originating from different crosses and collected on different days.

Translation-inhibitor treatments.

PatA decreases the levels of functional eIF4F initiation complex, which is essential for cap-dependent translation[54], by trapping eIF4A on mRNAs and ectopically enhancing its RNA helicase activity[36,37] (Fig. 3a). CHX binds the E site of the 60S ribosome subunit and inhibits ribosomal translocation during translation elongation[38] For DMS-seq and ribosome profiling experiments, 16-cell-stage embryos were bathed in either 10 μM pateamine A (PatA, purchased from D. Romo at Baylor University) or 50 μg/mL CHX (Sigma Aldrich). Embryos were collected after untreated embryos from the same clutches had reached 64-cell stage (2 hpf) (~ 45 min after the addition of inhibitors). For DMS-seq samples, 150 embryos were collected in a 5-mL tube for each condition (untreated, PatA and CHX) and treated with DMS. Translation-inhibitor concentrations were maintained during the DMS-modification step. For ribosome profiling samples, 55 embryos per condition and per replicate were collected and flash frozen in liquid nitrogen.

DMS-seq library preparation.

DMS-seq experiments were performed as previously described by Rouskin et al.[13], with minor modifications. Briefly, DMS-treated poly(A)$^+$ RNA samples were denatured at 95 °C for 2min and fragmented at 95 °C for 1.5 min in 1× RNA fragmentation buffer (Zn^{2+} based, from Ambion). The reaction was stopped by the addition of 0.1 volume of a 10×Stop solution (Ambion). One volume of formamide loading dye (95% formamide, 10 mM EDTA, 0.025% bromophenol blue and 0.025% xylene cyanol) was added to fragmented RNAs, and samples were separated through 10% TBU (Tris borate, 8M urea) PAGE. Fragments were visualized with blue light (Clare Chemical Research), and RNAs of 60–70 nt were excised. Gel pieces were shredded and then extracted in 300 μL of nuclease-free water (AmericanBio) at 70°C for 10 min with vigorous shaking. Eluted RNA was purified with Spin-X tube filters (Sigma-Aldrich) and ethanol precipitated by addition of 33 μL 3M sodium acetate, 1.5 μL glycoblue and 700 μL ethanol. RNA samples were dissolved in 6 μL of nuclease-free water, and 3′ phosphates were removed by incubation at 37 °C for 1 h with 1 μL 10× PNK buffer (NEB), 1 μL SUPERase Inhibitor (Ambion) and 2 μL of T4 PNK enzyme (NEB). Samples were then directly ligated to 1 μg of our in-house barcoded 3′ adapters (adaptor sequences in Supplementary Table 1) by addition of 1 μL 0.1M DTT, 6 μL 50% PEG, 1 μL 10× ligase2 buffer, 2 μL T4 RNA ligase2 and truncated K227Q (NEB), and incubated at 25 °C for 1.5 h. Each replicate was ligated to a different barcode set, and replicates could therefore be pooled after ligation. 20 μL of formamide dye was added to each tube, and ligated products were run in 10% TBU PAGE for ~45 min, visualized with blue light and separated from unligated adapters by gel extraction as described above. Reverse transcription was performed in 20 μL at 52 °C for 45 m in with Superscript III (Invitrogen) and reverse-transcription primer (primer sequence in Supplementary Table 1), and was followed by RNase H treatment at 37 °C for 15 min. cDNAs were separated in 10% TBU PAGE for 1.5h, and truncated reverse-transcription products of 25–45 nt above the size of the reverse-transcription primer were extracted by gel purification. Samples were then circularized with CircLigase (Epicentre), ethanol precipitated and PCR amplified with Illumina sequencing adapters, keeping the number of cycles to the minimum needed for the detection of amplified products (9–11 cycles).

Tetrahymena ribozyme spike-in control.

Tetrahymena ribozyme sequence was ordered from IDT DNA. RNA was generated with a AmpliScribe-T7-Flash transcription kit (Epicentre) and DMS probed in vitro as described above. DMS-modified *Tetrahymena* ribozyme was then spiked into an untreated poly(A)$^+$ RNA sample extracted from 2-hpf embryos, random fragmented and subjected to the DMS-seq library-preparation protocol described above.

DMS-seq controls.

To determine how mRNA structure changes during cellular transitions in vivo, we analyzed the structure of the zebrafish transcriptome during the MZT with DMS-seq (at 2,4 and 6 hpf). DMS methylales accessible N1 of adenine and N3 of cytidine bases in single-stranded regions, at the end of a stem, and at base pairs flanking G-U wobble interactions[55], thus causing reverse transcriptase drop-off during reverse transcription. The truncated RNAs are then captured, analyzed by high-throughput sequencing and normalized per transcript, thereby providing per-nucleotide accessibility values between 0 and 1 (‘DMS-seq profiles and accessibility’Methods section). Three lines of evidence indicated that we were able to effectively probe RNA structure during embryogenesis. First, we observed a robust and reproducible enrichment of reads mapping to A and C bases (74–83%), as compared with reads from DAiS-untreated samples (43%) (Supplementary Fig 1c), as well as reproducible counts between biological replicates (r of 0.989) (Supplementary Fig 1d,e). Second, the DMS-seq accessibility profile of the *Tetrahymena* ribozyme, used as an exogenous control, was in close agreement ($P = 6.3 \times 10–18$) with the previously reported structure[56] (Supplementary Fig 1f). Third, analysis of conserved secondary structures, such as those found in *dgcr8* (ref.[57]), *selenotla* and *selenot2* mRNAs showed that paired A and C bases exhibited lower accessibility than single-stranded regions ($P = 1.8 \times 10–05$) (Supplementary Fig 1g,h). These results indicate that our DMS-seq analysis provides a robust transcriptome-wide map of RNA structure dynamics in a vertebrate embryo.

Ribosome profiling experiments.

Ribosome profiling was performed with an ARTseq Ribosome Profiling Kit Mammalian (Epicentre), as previously described[58] with minor changes. For each condition, two biological replicates were generated from embryos originating from different crosses and collected on different days. For each replicate, 55 embryos were lysed in 800 μL of lysis buffer (1 × polysome buffer, 1% Triton X-100,1 mM DTT, 25 U/mL DNase land 100 μg/mL CHX) according to the manufacturer’s protocol (Epicentre). Lysates were centrifuged for 10 min at 20,000 *g* at 4 °C. 4.5 μL of ARTseq nuclease was added to 600 μL of lysate supernatant and incubated at25 °C for 45 min with gentle mixing Nuclease digestion was stopped by the addition of 22.5 μL of SUPERase-In RNase Inhibitor (Life Technologies) and chilled on ice for 5 min. Ribosomes were purified through Sephacryl S400 spin column chromatography (GE Healthcare) according to the manufacturer’s protocol. Before RNA purification, 7.5 μL of a mix of four different 28-nt spike-in RNA oligonucleotides (at the following concentrations: spike-in 1, 3.5×10^{-10}; spike-in 2, $6.9 \times l0^{-11}$ spike-in 3, 1.4×10^{-11}; spike-in 4, 2.8×10^{-12}; sequences in Supplementary Table 1) was added to the purified ribosome samples. Purified ribosome-protected fragments were separated in a 15%

TBU gel, and fragments of 28–30 nt were extracted. Adapters were ligated at the 3′ end of the fragments, reverse transcribed, gel purified and circularized as described above for DMS-seq fragments.

For RNA-seq samples (input), 75 ng of *S. cerevisiae* RNA was added as a spike-in to 175 μL of the remaining lysate supernatant, and total RNA was then extracted with TRIzol. Total RNA samples were sent to the Yale Center for Genome Analysis, and strand-specific TruSeq Illumina RNA sequencing libraries were constructed. Before sequencing samples were treated with Epicentre Ribo-Zero Gold, to deplete ribosomal RNAs, according to the manufacturer's protocol.

RNA-seq time-course experiments.

To quantify changes in mRNA abundance and to measure the effects of zygotic factors and miR-430 activity, we performed an RNA-seq time-course experiment in wild-type conditions, in the presence of α-amanitin (to inhibit zygotic transcription activation) and in the presence of a miR-430 inhibitor. To inhibit transcription activation of the zygotic genome, we injected embryos with 2 ng of α-amanitin (Sigma Aldrich), an inhibitor of RNA polymerase II. To inhibit miR-430 activity, we injected embryos with 1 nL of 10 μM of tiny locked nucleic acid (tinyLNA), complementary to the seed region of miR-430 (5′-TAGCACTT-3′ (Exiqon)). Approximately 25 embryos were collected per stage and per condition. Embryos were lysed in TRIzol and spiked with 75 ng of *S. cerevisiae* RNA for normalization purposes. Total RNA was extracted and sent to the Yale Center for Genome Analysis, and strand-specific TruSeq Ilium ina RNA-sequencing libraries were constructed for each sample. Before sequencing, samples were treated with either Epicentre Ribo-Zero Gold, to deplete ribosomal RNAs, or oligo(dT) beads, to enrich in poly(A)$^+$ RNA, according to the manufacturer's protocol.

Read sequencing and mapping.

DMS-seq, ribosome profiling and RNA-seq samples were sequenced on Illumina HiSeq 2000/2500 machines producing singleend 76-nt reads. This procedure yielded more than 1.4 billion reads, 500 million reads, 500 million reads and 200 million reads for DMS-seq, KHSRP iCLIR RNA-seq and ribosome profiling experiments, respectively. Sequencing samples are summarized in Supplementary Table 2.

With the library preparation protocol for DMS-seq and ribosome profiling samples, raw reads contained the following features: NNNN-insert-NN-barcode(4-mer)-adaptor, where the 6N (NNNN+NN) sequence composes the unique molecular identifier (UMI), 'barcode' is the sample 4-mer in-house barcode, and 'adaptor' is the 3′ Illumina adaptor. The UMI was used to discard PCR duplicates and count single ligation events. The barcode was used to mark individual replicates after the 3′-adaptor ligation step. Base calling was performed in CASAVA-1.8.2. The Illumina TruSeq index adaptor sequence was then trimmed by aligning its sequence, requiring a 100% match of the first 5 bp and a minimum global alignment score of 60 (matches, 5; mismatches, −4; gap opening −7; gap extension, −7; cost-free end gaps). Trimmed reads were demultiplexed according to the sample's in-house barcode, and the UMI was clipped from the 5′ and 3′ ends and kept within the read name

for marking PCR duplicates. Reads were then depleted of rRNA, tRNA, snRNA, snoRNA and miscRNA, with Ensembl78 annotations[59] as well as RepeatMasker annotations, through strand-specific alignment with Bowtie2 v2.2.4 (ref.[60]). The remaining reads were aligned to the zebrafish Zv9 genome assembly with STAR version 2.4.2a[61] with the following nondefault parameters: --alignEndsType EndToEnd --outFilterMultimapNmax 100 --seedSearchStartLmax 15 --sfbdScore 10 --outSAMattributes All. Genomic sequence indices for STAR were built including exon-junction coordinates from Ensembl 78. Only reads of unique UMIs were kept at each genomic coordinate for DMS-seq and ribosome profiling experiments. For ribosome profiling samples, reads were also aligned to the 256 sequences composing the four partially degenerated 28-nt RNA spike-ins using Bowtie2 v2.2.4.

Raw reads from RNA-seq experiments were processed with the same pipeline, by omitting the adaptor-trimming, barcoding-demultiplexing and UMI-clipping steps. The filtered reads were aligned onto Zebrafish Zv9 and *S. cerevisiae* r64d1d1 genome assemblies in STAR, with the same parameters as described above. For both species, STAR genomic-sequence indices were built, including exon-junction coordinates from Ensembl 78.

DMS-seq profiles and accessibility.

Per-transcript profiles were computed with uniquely mapped reads overlapping at least 10 nt with the transcripts' annotation. Each read count was attributed to the nucleotide at position −1 of the read's 5′ end within the transcript coordinate, to correct for reverse transcription stopping 1 nt before the DMS-modified nucleotide. To determine the read distributions for each nucleotide (Supplementary Fig. 1c), only transcripts with a minimum of 100 counts were considered. In the same figure, 'transcriptome' represents the frequency of each nucleotide for the same subset of transcripts (Supplementary Fig. 1c).

Accessibility was calculated according to the 2–8% rule[62], i.e„ by normalizing the read counts proportionally to the most reactive A and C bases within the region after the removal of outliers. More specifically, the 2% most reactive A and C bases were discarded, and each position was divided by the average of the next 8% most reactive A and C bases. Accessibility values greater than 1 were set to 1, and accessibility values for G and T were set to 0.

Calculating translation efficiency.

Translation efficiency was calculated by division of the ribo-seq reads per kilobase per million reads of spike-in (RPKM) by the RNA-seq RPKM for each coding sequence, excluding the first and last three codons (effective CDS), as described in Bazzini et al.[58]. Using the effective CDS of each transcript allows for computation of translation efficiency from actively translating ribosomes.

For ribosome profiling samples, the effective CDS annotation was shifted 12 nt upstream to position each read at the ribosome P-site location. RPKMs were computed by summing the effective CDS counts, including reads matching up to five times in the genome (each mapping site counting 1/number of mapping sites) and normalizing to the effective CDS length and the total number of reads mapped to the spike-in. For RNA-seq, RPKM values were calculated with the same pipeline applied to obtain ribosome profiling RPKMs, with

the following differences: the effective CDS annotation remained unshifted, counts from reads overlapping the effective CDS by a minimum of 10 nt were added, and the total number of reads mapped to yeast RNAs was used as normalization spike-in. Finally, translation efficiency was computed by dividing ribosome profiling RPKM values by RNA-seq RPKM values for each transcript's effective CDS. For each sample, this analysis was performed by combining reads of both independent replicates.

Identification of conserved RNA structures.

Conserved RNA structures (Supplementary Fig. 1g,h) were identified using the RNAz software[63] on the UCSC *D. rerio* multiz eight-way whole genome alignment (http://hgdownload.cse.ucsc. edu/goldenPath/danRer7/multiz8way/), with default parameters.

mRNA accessibility analysis.

To investigate global mRNA structure, the average accessibility of the 5′ UTR, CDS and 3′ UTR was individually calculated and compared across conditions and/or between translation statuses. To this end, we first identified the major transcript isoform of each gene, by selecting the one with the highest DMS-seq read density and with UTRs longer than 75 nt (total of 11,404 mRNAs). If no isoform met the minimum UTR length criteria, no isoform was selected for that gene. Transcripts with average DMS-seq counts per A and C bases ≥ 5 (except when specified differently), and with counts covering at least 85% of A and C bases in each region (UTRs and CDS) were selected for further analysis. When the analysis was performed on full-length UTRs (Fig. 4c and Supplementary Figs. 1l, 2b,c, 3a, 5a and 6a), DMS-seq data were used to determine the final UTR lengths. Specifically, UTR length was extended one nucleotide at a time, starting at position 76, until the A and C base count coverage fell below 85%. The longest 5′ and 3′ UTRs with at least 85% coverage were then associated with this transcript Accessibility values were calculated within each region (5′ UTR, CDS and 3′ UTR) of each transcript, as described above, and averaged (for A and C bases only) to obtain a single value per region. For all global-accessibility mRNA-region analyses (cumulative plots), the statistical significance between highly (red) and weakly (blue) translated mRNAs within the same condition was calculated with a onesided Mann-Whitney *U* test, and the statistical significance between the same subset of mRNAs, but across two different conditions, was calculated with a two-sided Wilcoxon signed-rank test.

Definition and selection of differentially translated mRNAs.

To select weakly and highly translated mRNAs, transcripts were binned in quintiles on the basis of translation efficiency. The quintile with the lowest translation efficiency was labeled as 'low translation', and the one with the highest translation efficiency was labeled as 'high translation'.

Sliding-window analysis.

To decipher which regions of the transcript significantly changed in RNA structure across conditions and developmental stages, per-transcript DMS-seq counts were subdivided into overlapping 100-nt sliding windows, offset by 1 nt. Only windows with a minimum coverage of 250 counts were considered for further analysis. For each window, two metrics

were calculated to identify changes in RNA structure across conditions. First, we performed a Kolmogorov-Smirnov test (KS test), a nonparametric test sensitive to differences in both the location and shape of the empirical cumulative distribution functions of two samples. Second, to identify windows with increased or decreased RNA structures, we used a ratio of Gini indices between the two compared conditions, which has previously been used to compare RNA structures[13] because of its ability to capture inequalities in a given distribution (Fig. 1c). All analyses were performed separately for each replicate, and KS-test *P* values and ratios of Gini indices were combined with Fisher's method (R package 'metap') and averaged, respectively. Differential structured windows between two conditions were defined as those that were significant (combined *P* value < 0.05 in the KS test) and that had an average Gini fold change across replicates>1.1 or < 0.9 (Fig 1c). Meta plots of the distribution of differentially structured windows along the transcript were normalized to the region lengths and the total number of windows analyzed in each region (5′ UTR, CDS and 3′ UTR) (Figs. 1d, 4b,e and 5 and Supplementary Fig. 6b).

Of note, on average, 100-nt windows contained ~ 50 As and Cs, which we found to be the minimal window length to have sufficient data points for statistical analysis, as well as to ensure the presence of both paired and unpaired nucleotides in the window, while still identifying individual regions that maybe changing in structure.

Sequence conservation analysis.

To determine the sequence conservation, we used PhastCons regional conservation[64] scores derived from the Multiz alignment of eight vertebrate species, which included five fish species (ftp://hgdownload.cse.ucsc.edu/goldenPath/danRer7/phastCons8way/).

uORF analysis.

A single transcript isoform per gene was selected as described above except for the minimal UTR length, which was set to 12 nt. We then used the zebrafish uORF dataset from Johnstone et al.[66], which contains location and translation efficiency information, to identify weakly and highly translated uORFs at 2 hpf as described above for mRNAs. We calculated per-uORF DMS-seq accessibility profiles by applying the 2–8% rules described above, for uORFs longer than 150 nt For shorter uORFs, the DMS accessibility was computed with 150-nt regions starting at the uORF AUG initiation codon. Global uORF accessibility values were derived as described above. uORFs containing fewer than ten A and C bases were removed from this analysis.

Poly(A)-tail length and miR-430-mediated repression analyses.

Datasets of poly(A)-tail lengths were obtained from Subtelny et al.[33] who developed a technique, called poly(A)-tail-length profiling by sequencing (PAL-seq), to measure the poly(A)-tail length of transcripts from various organisms, including 2-, 4- and 6-hpf zebrafish embryos. Only transcripts with at least 50 poly(A) tags were considered in our poly(A)-tail-length analyses (Fig. 4b,c and Supplementary Fig. 5a–c). To analyze the effect of miR-430-mediated repression on mRNA structures (Fig. 4d,e and Supplementary Fig. 5d), we selected a set of 483 mRNAs targeted by miR-430 with the following criteria: (i) minimum twofold decrease in mRNA level between 2 and 6 hpf (6/2); (ii) minimum 1.5-fold

increase in mRNA level between 6-hpf tinyLNA and 6-hpf wild type (6 tinyLNA/6 wt); and (iii) at least one window with more than 250 reads in DMS-seq experiments at both 2 and 4 hpf. For comparison, a control set of 1,495 mRNAs was selected, corresponding to all the remaining transcripts with sufficient DMS-seq coverage and without miR-430 octamers or heptamers in their 3′ UTRs. For genes with multiple transcript isoforms, the one with UTRs> 50-nt in length and with the highest coverage in our DMS-seq experiment at 2 hpf was selected. If no transcript isoform met those criteria for a specific gene, no transcript from that gene was included in the analysis.

RNA secondary-structure analyses.

Predicted RNA secondary structures were computed with either the *Fold* executable of the RNA structure package[67] (version 5.6) or SeqFold software[40]. For the *Fold* software, default parameters were used except for the temperature parameter, which was set to 301.15 K (28 °C)[68]. The DMS-seq accessibility values of A and C bases were inferred as soft constraints with the --dms option. The Gibbs energy of folding (ΔG) of each predicted RNA structure was calculated with the *efn2* executable with default parameters except for the temperature parameter, which was set to 301.15 K (28 °C)[69]. Arc plots representing RNA secondary structure were drawn with R-chie software with default parameters[70]. The analysis of the RNA structure surrounding AUG initiation codons was performed with 100-nt windows centered on each AUG with a minimum DMS-seq depth of five reads on average per A and C bases (Fig. 2e and Supplementary Fig. 2d–g; $n = 2{,}360$). The analysis of the effect of ribosomes on local RNA structure was limited to 100-nt windows located in CDS regions and with a minimum DMS-seq depth of five reads on average per A and C bases (Fig. 3f and Supplementary Fig. 3b,c). Windows with ribosome footprints corresponded to those with an average number of ribo-seq reads between replicates greater than seven. Windows without ribosome footprints corresponded to those with an average of less than one ribo-seq read and found in mRNAs with translation efficiency< 1.65. Each subset contains atotal of 728 nonoverlapping windows covering 332 and 312 different genes for windows with and without ribosome footprints, respectively (Supplementary Fig. 3c).

For the SeqFold analysis, only transcripts with more than seven reads on average per A and C bases in all four tested conditions (untreated, in vitro, PatA treated or CHX treated) were considered, thus resulting in a set of 1,143 transcripts. The DMS-seq accessibility of A and C bases was inferred through the same protocol as that used for SHAPE accessibility, and energies of folding were directly retrieved from the SeqFold output. To determine the global similarity of predicted RNA structures across conditions, per-transcript Gini indices were computed from the SeqFold outputs with the reldist library in R. PCA was then used to reduce the dimensionality of the dataset A biplot of the PCA loadings (i.e., conditions) of the two first principal components was used to determine which conditions were most similar. The same transcriptome-wide PCA was done separately for CDS and 3′-UTR regions. The cumulative proportion of variance explained by the two first components was 92% (PC1) and 3.1% (PC2) in CDS regions, and 94.5% (PC1) and 2.3% (PC2) in 3′-UTR regions (Fig. 3h,i).

miR-430-site structural analysis.

To assess the effect of RNA structure on miRNA activity, we selected 18 different endogenous miR-430-binding sites with DMS-seq coverage of five or more reads per A and C bases and obtained their decay activity from Yartseva et al.[45]. The stability (minimum free energy, ΔG) of the structure formed by the 120-nt region centered on the miR-430 seed was calculated in different conditions (in silico, in vitro and in vivo at 2 hpf, before miR-430 expression), and the correlation between the RNA structure stability and decay strength was determined with Spearman correlations (Fig. 6b). The predicted secondary structure and stability of a 200-nt region centered on the miR-430 seed were computed for additional endogenous miR-430 sites found in maternal mRNA 3′ UTRs (Fig. 6c,g and Supplementary Fig. 8). For each maternal mRNA containing miR-430 sites, decay patterns were analyzed in the wild-type condition (with miR-430, WT) or when miR-430 activity was inhibited (-miR-430) with a tiny LNA complementary to the seed region of miR-430. An increase in maternal mRNA stability in the –miR-430 condition suggests miR-430-mediated repression of those maternal mRNAs in WT condition (Fig. 6d,h and Supplementary Fig. 8).

Next, GFP reporters containing endogenous miR-430-binding sites in their 3′ UTRs were used, and GFP expression was quantified in both wild-type and MZ*mir430*-mutant fish[71], in which the miR-430 locus on chromosome 4 was deleted[71] (Fig. 6e,f,i,j). In addition to the 200-nt wild-type sequence of each endogenous miR-430 site, mutated versions that either stabilized or destabilized the in vivo structure were also built. Briefly, the GFP coding sequence was inserted in pCS2^{+} plasmid between the BamHI and XhoI sites, and the miR-430 sites were inserted in the XbaI sites (oligonucleotide sequences in Supplementary Table 1). The resulting plasmids were linearized with NotI digestion, and the products were in vitro transcribed with an Sp6 mMessage mMachine kit (Thermo Fisher, AM1340). Zebrafish embryos were injected with 100 pg of mRNA GFP reporters and 100 pg of DsRed reporter (internal control) at the one-cell stage. Fluorescence was quantified at 24 hpf with ImageJ. miR-430 activity was calculated as the GFP/DsRed ratio for both genetic backgrounds (wild type and MZ*mir430*). Wild-type GFP/DsRed was then normalized to the GFP/DsRed ratio from MZ*mir430*, thus resulting in miR-430-mediated repression values for each reporter. The reported values came from a minimum of three independent replicates (Fig. 6f,j).

Finally, to determine the effect of binding of the Ago2–miR-430 complex on the RNA accessibility of the recognition site, we analyzed the per-nucleotide accessibility values of 100-nt windows centered on individual miR-430 seeds (octamers and heptamers) found in the 3′ UTRs of miR-430 targets (identified as described above), for each developmental stage and condition (Supplementary Fig. 7e–g). For comparison, per-nucleotide accessibility values of randomly chosen 100-nt windows within the same set of 3′ UTRs were used. Global seed and control accessibility values were calculated by averaging the accessibility of each A and C base found in the 8-nt sequence located in the center of each window (Supplementary Fig. 7f). Only windows with a minimum of five reads on average per A and C bases were considered.

RNA-element selection assay.

To quantify the regulatory activity of maternal mRNA 3′-UTR sequences during the MZT, we used RESA, as described in Yartseva et al.[45], with minor changes (Fig. 7b). Maternal 3′-UTR regions of 200 nt showing changes in RNA structure during the MZT (on the basis of KS tests across stages, $P < 0.05$) or no change (as controls, $P > 0.05$) were amplified from cDNA of 2-hpf embryos (primer sequences in Supplementary Table 3). In total, 74 dynamic and 25 control regions were amplified and inserted in the 3′ UTR of a GFP reporter by assembly PCR (primer sequences in Supplementary Table 1) with Phusion High Fidelity DNA polymerase (NEB, M0530) and purified on an agarose gel. The final library constructs consisted of an Sp6 promoter, a GFP coding sequence, a 3′ UTR with various inserts flanked by part of the Illumina 5′ and 3′ adaptor sequences, and an SV40 polyadenylation signal. The library was in vitro transcribed with an Sp6 mMessage mMachine kit (Thermo Fisher, AM1340), and the resulting RNA library was injected (10 pg) in one-cell-stage embryos, with or without the presence of α-amanitin, an inhibitor of RNA polymerase II that blocks the activation of the zygotic genome. Approximately 25 injected embryos were collected at 2 and 6 hpf (with and without α-amanitin), and total RNA was extracted with TRIzol reagent according to the manufacturer's protocol. Poly(A)$^+$ mRNA was selected with NEB oligo(dT)$_{25}$ magnetic beads (NEB, S1419S) according to the manufacturer's protocol. Poly(A)$^+$ mRNA was recovered with ethanol precipitation and dissolved in 11 μL of water. Concentrated poly(A)$^+$ mRNA was reverse-transcribed with a reporter-specific primer with superscript III. Illumina adapters were added, and the library was amplified with ~20 cycles of Phusion PCR (primer sequences in Supplementary Table 1). The RESA experiment was performed in five independent biological replicates.

RESA libraries were sequenced with Illumina HiSeq 2000/2500 machines producing single-end 76-nt reads. Reads were processed and aligned as mentioned above for DMS-seq reads, with the adaptor-trimming, barcoding-demultiplexing and UMI-clipping steps omitted. The number of reads corresponding to each insert was counted and normalized to the total number of reads in each sample. Changes in reporter abundance during the MZT were calculated as the ratio of counts (6/2 hpf or 6 wt/6 aAm) for each insert, for those reporters with sufficient coverage at 2 hpf (average count≥ 250). Then, the average of the replicates was calculated after removal of outliers (lower or higher of the mean ± s.d.) for each insert. When multiple inserts corresponding to dynamic structure or control were present in a single 3′ UTR, the insert with the greatest effect on gene expression (positive or negative) was kept for each category. This analysis allowed us to quantify the effects of 53 and 18 3′-UTR sequences with dynamic structures and with no structural change, respectively, in regulating mRNA abundance during the MZT (Fig. 7c,d,f,h and Supplementary Fig. 9b,c,e).

Nanog 3′-UTR analysis.

The RNA secondary structure of the full-length *nanog* 3′ UTR was predicted with the *Fold* software and DMS-seq accessibility values as described above. The predicted structures of the 544-nt *nanog* 3′ UTR at 2 and 6 hpf are shown in Fig. 7g. Wild-type and 20-nt-deletion (Fig. 7g,h) versions were cloned downstream of the GFP in the pCS2$^+$GFP plasmid in the XbaI sites (oligonucleotide sequences in Supplementary Table 1). Reporters were in vitro transcribed and co-injected with DsRed reporter in one-cell-stage embryos as described

above for miR-430 reporters. Injected embryos were collected at 2 and 6 hpf, and RNA was extracted with TRIzol reagent and reverse transcribed with random primers. GFP and DsRed reporters were quantified by qPCR (primer sequences in Supplementary Table 1), and decay activity was calculated by comparison of levels at 2 and 6 hpf (Fig. 7h, from four independent experiments).

KHSRP iCLIP experiment and analysis.

To identify the regions bound by KHSRP during early embryogenesis, we performed iCLIP experiments as described in Huppertz et al.[72] with minor changes. Embryos at sphere stage (4 hpf) were collected and irradiated with 254-nm UV light to induce cross-linking (embryos were snap-frozen and stored in batches to yield a total of 1,000 embryos per condition). Frozen embryos were then thawed and homogenized on ice in iCLIP lysis buffer[72]. An affinity-purified rabbit polyclonal antibody raised against zebrafish KHSRP (generated by YenZym Antibodies) was used to isolate RNA–protein complexes. Briefly, 200 µl protein G Dynabeads was added to 50 µg antibody in lysis buffer. Beads were incubated for 1 h, washed three times with lysis buffer and added to the lysates. As a control, a parallel experiment was performed without the addition of antibody. Subsequent steps in the iCLIP protocol were performed as described in Huppertz et al.[72], except for the 3′ adaptors containing in-house barcode, the reverse-transcription primer and PCR primers, which were identical to those used for the DMS-seq library preparation. Twenty cycles of amplification of cDNA were used for final library construction. Barcoded PCR-amplified libraries were size-selected in a 6% TBE gel and combined for Illumina sequencing. Both the KHSRP pulldown and the no-antibody control were performed in triplicate. Raw reads from iCLIP experiments were processed with the same pipeline as that for the DMS-seq experiments.

To define the binding preference of KHSRP, bound windows were defined for each hexamer with at least ten reads and extended 5 nt upstream and 10 nt downstream. For each bound window, the frequency of each hexamer was computed. These frequencies were normalized to the average hexamer frequencies observed in the no-antibody controls, which were calculated as mentioned above except for the read-depth requirement, which was set to two instead of ten reads to compensate for differences in sequencing depth. To build the logo representation, the top ten most frequent normalized hexamers were aligned with MAFFT with the following parameters: --reorder --lop −10 --lexp −10 –localpair –genafpair --maxiterate 1000. Multiple sequence alignment positions with more than 50% gaps were trimmed on both ends. Nucleotide frequencies were calculated with the final trimmed alignment and are represented as logos (Supplementary Fig. 7a).

iCLIP peaks were called by testing scanning 30-nt windows with Poisson’s law. Overlapping significant windows ($P < 0.01$) were then merged to obtain the peaks. To obtain robust Poisson parameter estimates, we analyzed only transcripts with a minimum of 20 positions with at least one read. The final KHSRP peaks corresponded to overlaps of peaks found in at least two out of three replicates. To analyze the RNA structure of KHSRP-bound and KHSRP-unbound regions, we defined bound regions as 100-nt windows centered on the middle point of each KHSRP peak. For each bound region, an unbound control within the same 3′ UTR and overlapping by less than 25% with the bound region was selected. The

accessibility of each region was computed by averaging the accessibility of A and C bases. Finally, only transcripts and regions with a minimum of five reads on average per A and C bases were considered (Supplementary Fig. 7c).

RNA-RNA versus RNA-protein interaction analysis.

Whereas a high DMS-seq signal implies a single-stranded conformation, a low DMS-seq signal signifies either RNA-RNA or RNA-protein interactions involving the Watson–Crick face of the base[73]. To test whether DMS probing mainly analyzed RNA-RNA pairing versus protein footprinting in vivo, we analyzed the structure over two types of regions: one bound by the RNA-binding protein KHSRP, identified by iCLIP, and another occupied by the Ago2–miR-430 complex (Supplementary Fig. 7). We found that, in vivo, KHSRP-bound 3′-UTR regions were more accessible than control regions from the same 3′ UTRs, and the results were comparable to those in vitro, thus suggesting that KHSRP binding helps to maintain a single-stranded RNA conformation in the cell (Supplementary Fig. 7a–d). In addition, decreased accessibility was observed at miR-430 seed regions when the Ago2–miR-430 complex was present, but not in the flanking sequences or in the absence of miR-430 expression, thus suggesting that seed-miRNA interactions were measured (Supplementary Fig. 7e–g). Analyses from two different types of trans factors (KHSRP and Ago2–miR-430 complex) suggest that our approach measures the RNA structure promoted by those factors rather than the protein footprints limiting DMS's access to RNA. However, we cannot exclude the possibility that part of the observed decrease in DMS accessibility values in vivo, especially in 3′ UTRs, is the result of RNA-binding proteins interacting directly with the DMS-methylation sites of A and C bases.

Statistics.

Mann–Whitney *U* tests, Wilcoxon signed-rank tests, and Spearman correlations and corresponding asymptotic *P* values were computed with the Python Scipy package. Student two-tailed *t* tests were calculated with the GraphPad Prism software.

Reporting Summary.

Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data availability.

Raw reads that support the findings of this study are publicly available in the Sequence Read Archive. SRP114782 contains demultiplexed and raw reads for DMS-seq, ribo-seq and raw reads for RNA-seq and RESA; SRP149556 contains raw reads for the developmental mRNA-seq time course; SRP149368 contains demultiplexed and raw reads for endogenous KHSRP iCLIP. For multiplexed data, a column 'barcode' in SRA annotation indicates which internal barcodes were used (described above; additional data at https://data.giraldezlab.org/). RNA-seq and ribo-seq tables are provided in Supplementary Datasets 1 and 2, respectively. CDS annotation, per-transcript DMS-seq counts and accessibility profiles for 2-, 4- and 6-hpf developmental stages are provided in Supplementary Datasets 3, 4 and 5, respectively. Source data for Figs. 3h–i, 6b,d,f,h,j and 7c,d,f,h and Supplementary

Figs. 1j,k and 9c,e are available online. DMS-seq processed data for the other conditions and stages (for example, 2 hpf in vitro and in the presence of the translation inhibitors) are available at https://data.giraldezlab.org/. Raw data of the sliding-window analyses are available at https://data.giraldezlab.org/. For convenience, DMS-seq profiles, ribo-seq tables and RNA-seq tables mapped on the latest zebrafish genome (GRCz11) and annotation (Ensembl 92, all genes) are also available at https://data.giraldezlab.org/for all developmental stages and conditions. Other data that support the findings of this study are available from the corresponding authors upon reasonable request.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank K. Bilguvar, S. Mane and C. Castaldi for sequencing support. We thank A. Bazzini and members of the laboratories of A.J.G. and M.K. for intellectual and technical support. This research was supported by the Fonds de Recherche du Québec–Santé (postdoctoral fellowship to J.-D.B.); the Human Frontier Science Program (LT000307/2013-l to E.M.N.); the Australian Research Council (DE170100506 to E.M.N.); and the National Institute of Health (grants R01 HD074078, GM103789, GM102251, GM101108 and GM081602), the Pew Scholars Program in the Biomedical Sciences, the March of Dimes (1-FY12-230), the Yale Scholars Program and Whitman fellowship funds provided by E. E. Just, Lucy B. Lemann, Evelyn and Melvin Spiegel, H. Keffer Hartline and Edward F. MacNichol, Jr. of the Marine Biological Laboratory, Woods Hole, Massachusetts, USA, to A.J.G.

References

1. Sharp PA The centrality of RNA. Cell 136, 577–580 (2009). [PubMed: 19239877]
2. Warf MB & Berglund JA Role of RNA structure in regulating pre-mRNA splicing. Trends Biochem.Sci 35,169–178 (2010). [PubMed: 19959365]
3. Martin KC & Ephrussi A mRNA localization: gene expression in the spatial dimension. Cell 136, 719–730 (2009). [PubMed: 19239891]
4. Ray PS et al. A stress-responsive RNA switch regulates VEGFA expression. Nature 457, 915–919 (2009). [PubMed: 19098893]
5. Weingarten-Gabbay S et al. Systematic discovery of cap-independent translation sequences in human and viral genomes. Science 351, aad4939 (2016). [PubMed: 26816383]
6. Kedde M et al. A Pumilio-induced RNA structure switch in p27-3′ UTR controls miR-221 and miR-222 accessibility. Nat. Cell Biol. 12, 1014–1020 (2010). [PubMed: 20818387]
7. Lucks JB et al. Multiplexed RNA structure characterization with selective 2′ -hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). Proc. Natl. Acad. Sd. USA 108, 11063–11068 (2011).
8. Kwok CK, Tang Y, Assmann SM & Bevilacqua PCThe RNA structurome: transcriptome-wide structure probing with next-generation sequencing. Trends Biochem. Sci 40, 221–232 (2015). [PubMed: 25797096]
9. Mortimer SA, Kidwell MA & Doudna JA Insights into RNA structure and function from genome-wide studies. Nat. Rev. Genet 15, 469–479 (2014). [PubMed: 24821474]
10. Ding Y et al. In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. Nature 505, 696–700 (2014). [PubMed: 24270811]
11. Kertesz M et al. Genome-wide measurement of RNA secondary structure in yeast. Nature 467, 103–107 (2010). [PubMed: 20811459]
12. Lu Z et al. RNA duplex map in living cells reveals higher-order transcriptome structure. Cell 165, 1267–1279 (2016). [PubMed: 27180905]

13. Rouskin S, Zubradt M,, Washietl, S., Kellis, M. & Weissman, J. S. Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. Nature 505, 701–705 (2014). [PubMed: 24336214]
14. Spitale RC et al. Structural imprints in vivo decode RNA regulatory mechanisms. Nature 519, 486–490 (2015). [PubMed: 25799993]
15. Guo JU & Bartel DP RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. Science 353, aaf5371 (2016). [PubMed: 27708011]
16. Takyar S, Hickerson RP & Noller H E mRNA helicase activity of the ribosome. Cell 120, 49–58 (2005). [PubMed: 15652481]
17. Wen J-D et al. Following translation by single ribosomes one codon at a time. Nature 452, 598–603 (2008). [PubMed: 18327250]
18. Qu X et al. The ribosome uses two active mechanisms to unwind messenger RNA during translation. Nature 475, 118–121 (2011). [PubMed: 21734708]
19. Burkhardt DH et al. Operon mRNAs are organized into ORF-centric structures that predict translation efficiency. elife 6, e22037 (2017). [PubMed: 28139975]
20. Kozak M Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs. Mol. Cell. Biol 9, 5134–5142 (1989). [PubMed: 2601712]
21. de Smit MH & van Duin J Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. Proc. Natl Acad. Sci. USA S7, 7668–7672 (1990).
22. Gray NK & Hentze MW Regulation of protein synthesis by mRNA structure. Mol. Biol. Rep 19, 195–200 (1994). [PubMed: 7969107]
23. Dvir S et al. Deciphering the rules by which 5′ -UTR sequences affect protein expression in yeast. Proc. Natl Acad. Sci. USA 110, E2792–E2801 (2013). [PubMed: 23832786]
24. Zur H & Tuller T Strong association between mRNA folding strength and protein abundance in S. cerevisiae. EMBO Rep 13, 272–277 (2012). [PubMed: 22249164]
25. Tadros W & Lipshitz HD The maternal-to-zygotic transition: a play in two acts. Development 136, 3033–3042 (2009). [PubMed: 19700615]
26. Yartseva V & Giraldez AJ The maternal-to-zygotic transition during vertebrate development: a model for reprogramming. Curr. Top. Dev. Biol 113, 191–232 (2015). [PubMed: 26358874]
27. Lee M T,, Bonneau, A. R. & Giraldez, A. J. Zygotic genome activation during the maternal-to-zygotic transition. Annu. Rev. Cell Dev. Biol 30, 581–613 (2014). [PubMed: 25150012]
28. Lee MT et al. Nanog, Pou5f1 and SoxB1 activate zygotic gene expression during the maternal-to-zygotic transition. Nature 503, 360–364 (2013). [PubMed: 24056933]
29. Leichsenring M, Maes J, Mössner R, Driever W & Onichtchouk D Pou5f1 transcription factor controls zygotic gene activation in vertebrates. Science 341, 1005–1009 (2013). [PubMed: 23950494]
30. Heyn P et al. The earliest transcribed zygotic genes are short, newly evolved, and different across species. Cell Rep 6, 285–292 (2014). [PubMed: 24440719]
31. Giraldez AJ et al. Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science 312, 75–79 (2006). [PubMed: 16484454]
32. Lund E, Liu M, Hartley RS, Sheets MD & Dahlberg JE Deadenylation of maternal mRNAs mediated by miR-427 in *Xenopus laevis* embryos. RNA 15, 2351–2363 (2009). [PubMed: 19854872]
33. Subtelny AO, Eichhorn SW, Chen GR, Sive H & Bartel DP Poly(A)-tail profiling reveals an embryonic switch in translational control. Nature 508, 66–71 (2014). [PubMed: 24476825]
34. Bazzini AA et al. Codon identity regulates mRNA stability and translation efficiency during the maternal-to-zygotic transition. EMBO J 35, 2087–2103 (2016). [PubMed: 27436874]
35. Li F et al. Regulatory impact of RNA secondary structure across the *Arabidopsis* transcriptome. Plant Cell 24, 4346–4359 (2012). [PubMed: 23150631]
36. Low W-K et al. Inhibition of eukaryotic translation initiation by the marine natural product pateamine A. Mol. Cel 20, 709–722 (2005).
37. Bordeleau M-E et al. RNA-mediated sequestration of the RNA helicase eIF4A by Pateamine A inhibits translation initiation. Chem. Biol 13, 1287–1295 (2006). [PubMed: 17185224]

38. Schneider-Poetsch T et al. Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. Nat. Chem. Biol 6, 209–217 (2010). [PubMed: 20118940]
39. Kuzmenko A et al. Mitochondrial translation initiation machinery: conservation and diversification. Biochimie 100,132–140 (2014). [PubMed: 23954798]
40. Ouyang Z, Snyder MP & Chang HYSeqFold: genome-scale reconstruction of RNA secondary structure integrating high-throughput sequencing data. Genome Res 23, 377–387 (2013). [PubMed: 23064747]
41. Lim J, Lee M, Son A, Chang H & Kim VN mTAIL-seq reveals dynamic poly(A) tail regulation in oocyte-to-embryo development. Genes Dev 30, 1671–1682(2016). [PubMed: 27445395]
42. Bazzini AA, Lee MT & Giraldez AJ Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish. Science 336, 233–237 (2012). [PubMed: 22422859]
43. Kertesz M, Iovino N, Unnerstall U,, Gaul, U. & Segal, E. The role of site accessibility in microRNA target recognition. Nat. Genet 39, 1278–1284 (2007). [PubMed: 17893677]
44. Long D et al. Potent effect of target structure on microRNA function. Nat. Struct. Mol. Biol 14, 287–294 (2007). [PubMed: 17401373]
45. Yartseva V, Takacs CM Vejnar CE, Lee MT & Giraldez AJ RESA identifies mRNA-regulatory sequences at high resolution. Nat. Methods 14, 201–207 (2017). [PubMed: 28024160]
46. Li F et al. Global analysis of RNA secondary structure in two metazoans. Cell Rep 1, 69–82 (2012). [PubMed: 22832108]
47. Wan Y et al. Landscape and variation of RNA secondary structure across the human transcriptome. Nature 505, 706–709 (2014). [PubMed: 24476892]
48. Nakano S, Miyoshi D & Sugimoto N Effects of molecular crowding on the structures, interactions, and functions of nucleic acids. Chem. Rev 114, 2733–2758 (2014). [PubMed: 24364729]
49. Tyrrell J, McGinnis JL, Weeks KM, & Pielak GJ. The cellular environment stabilizes adenine riboswitch RNA structure. Biochemistry 52 8777–8785 (2013). [PubMed: 24215455]
50. Wu X & Bartel DP Widespread influence of 3′-end structures on mammalian mrna processing and stability. Cell 169, 905–917.ell (2017). [PubMed: 28525757]
51. Foley SW et al. A global view of RNA-protein interactions identifies post-transcriptional regulators of root hair cell fate. Dev. Cell 41, 204–220.e5 (2017). [PubMed: 28441533]
52. Yaman I et al. The zipper model of translational control: a small upstream ORF is the switch that controls structural remodeling of an mRNA leader. Cell 113,519–531 (2003). [PubMed: 12757712]
53. Halstead JM et al. An RNA biosensor for imaging the first round of translation from single cells to living animals. Science 347, 1367–1671 (2015). [PubMed: 25792328]
54. Gingras AC, Raught B & Sonenberg N eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. Annu. Rev. Biochem 68, 913–963 (2003).
55. Mathews DH et al. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. Proc. Natl. Acad. Sci. USA 101, 7287–7292 (2004). [PubMed: 15123812]
56. Golden BL, Gooding AR Podell ER & Cech TR A preorganized active site in the crystal structure of the *Tetrahymena* ribozyme. Science 282, 259–264 (1998). [PubMed: 9841391]
57. Han J et al. Posttranscriptional crossregulation between Drosha and DGCR8. Cell 136, 75–84 (2009). [PubMed: 19135890]
58. Bazzini AA et al. Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation. EMBO J 33,981–993 (2014). [PubMed: 24705786]
59. Yates A et al. Ensembl 2016. Nucleic Acids Res 44, D710–D716 (2016). [PubMed: 26687719]
60. Langmead B & Salzberg SL Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359 (2012). [PubMed: 22388286]
61. Dobin A et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2013). [PubMed: 23104886]
62. Deigan KE, Li TW, Mathews DH & Weeks KM Accurate SHAPE-directed RNA structure determination. Proc. Natl. Acad. Sci. USA 106,97–102 (2009). [PubMed: 19109441]

63. Gruber AR, Findei S, Washietl S, Hofacker IL & Stadler PF. RNAz 2.0: improved noncoding RNA detection. Pac. Symp. Biocomput (2010).

64. Siepel A et al. Evolutionary conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res 15, 1034–1050 (2005). [PubMed: 16024819]

65. Blanchette M et al. Aligning multiple genomic sequences with the threaded blockset aligner. Genome Res 14, 708–715 (2004). [PubMed: 15060014]

66. Johnstone TG Bazzini AA & Giraldez AJ Upstream ORFs are prevalent translational repressors in vertebrates. EMBO J 35,706–723 (2016). [PubMed: 26896445]

67. Reuter JS & Mathews DH RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinforma 11, 129 (2010).

68. Mathews DH Using an RNA secondary structure partition function to determine confidence in base pairs predicted by free energy minimization. RNA 10, 1178–1190 (2004). [PubMed: 15272118]

69. Mathews DH, Sabina J, Zuker M & Turner DH Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J. Mol. Biol 288, 911–940 (1999). [PubMed: 10329189]

70. Lai D Proctor JR, Zhu JYA & Meyer IM R-CHIE: a web server and R package for visualizing RNA secondary structures. Nucleic Acids Res 40, e95 (2012). [PubMed: 22434875]

71. Liu Y et al. Inheritable and precise large genomic deletions of non-coding RNA genes in zebrafish using TALENs. PLoS One 8, e76387 (2013). [PubMed: 24130773]

72. Huppertz I et al. iCLIP: protein-RNA interactions at nucleotide resolution. Methods 65, 274–287 (2014). [PubMed: 24184352]

73. Tijerina P, Mohr S & Russell R DMS footprinting of structured RNAs and RNA-protein complexes. Nat. Protoc 2, 2608–2623 (2007). [PubMed: 17948004]

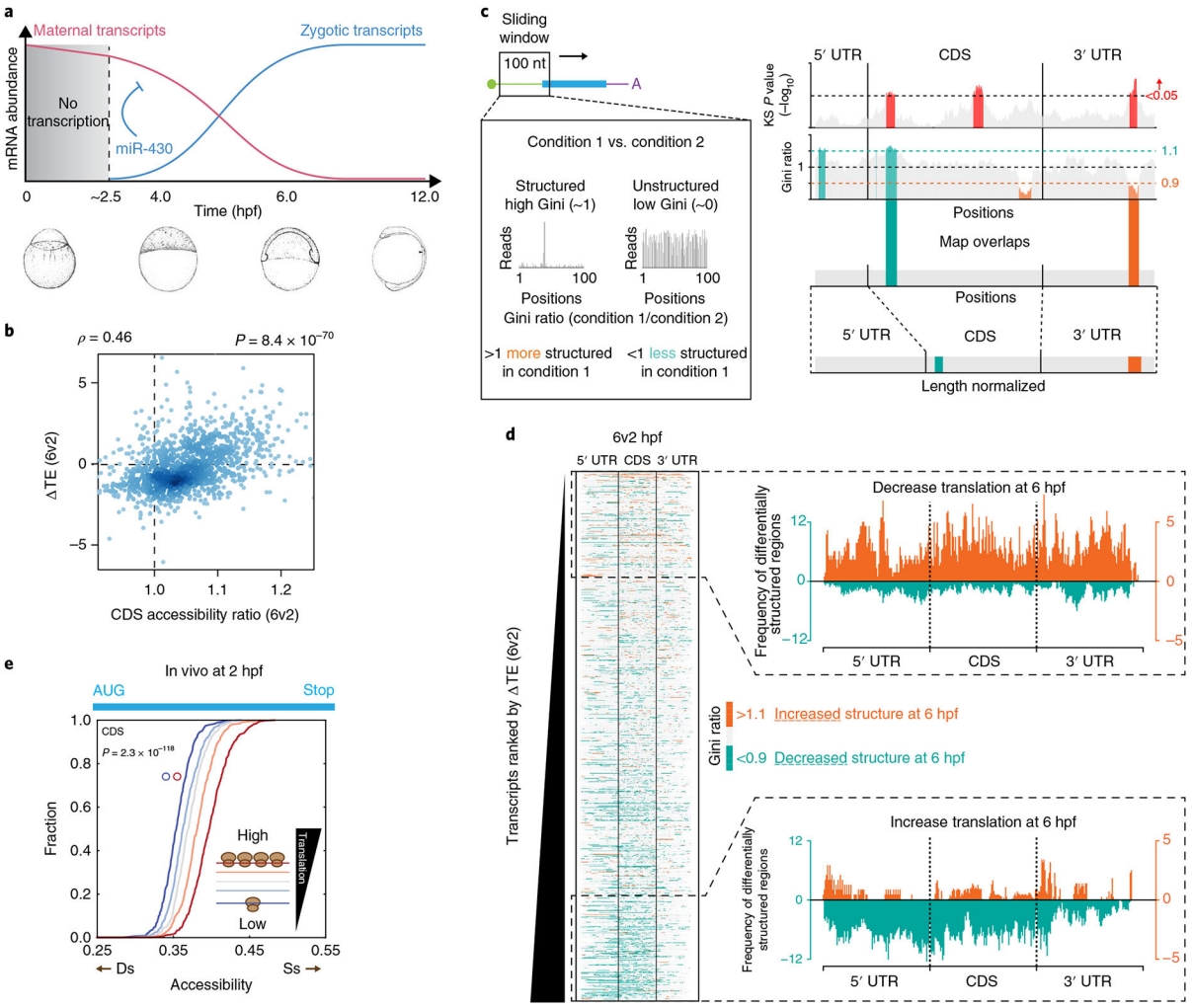

Fig. 1 |. Relationship between mRNA structure and translation during the MZT.

a, Schematic view of the transcriptomic remodeling that occurs during the MZT. After a transcriptionally silent period (gray), the maternal program (pink) transitions to a zygotic program (blue). **b**, Correlation between the per-transcript changes at 6 hpf versus 2 hpf (6v2) in translation efficiency (**Δ**TE, 6-2 hpf) and CDS accessibility (6/2 hpf) for 1,337 coding sequences with high coverage at both stages. The Spearman correlation coefficient (ρ) and the corresponding asymptotic P value are shown. **c**, Methodology used to identify local RNA structure changes between conditions. Differentially structured (DS) sliding windows are defined as those that are significantly changing ($P < 0.05$) on the basis of the Kolmogorov–Smirnov test, and show an increase (> 1.1) or decrease (< 0.9) in their Gini-index ratio. **d**, Left, structure changes of each 100-nt window (Gini ratio 6/2 hpf) along each transcript (total of 1,273). Transcripts are ranked according to their changes in translation efficiency during the MZT (6 - 2 hpf). Each transcript region (5′ UTR, CDS and 3′ UTR) is normalized to its length. Right, cumulative distributions of differentially structured windows for transcripts with increased (top 20%, top) and decreased (bottom 20%, bottom) translation. **e**, Cumulative distribution of global CDS accessibility in vivo, showing that highly translated mRNAs exhibit increased CDS accessibility in vivo at 2 hpf, a stage at

which the rate of translation spans ~10,000-fold between highly and weakly translated mRNAs. Ss, single-stranded; ds, double-stranded. Transcripts (n = 2,526) were binned into quintiles according to their translation efficiency. P value was computed with a one-sided Mann–Whitney U test.

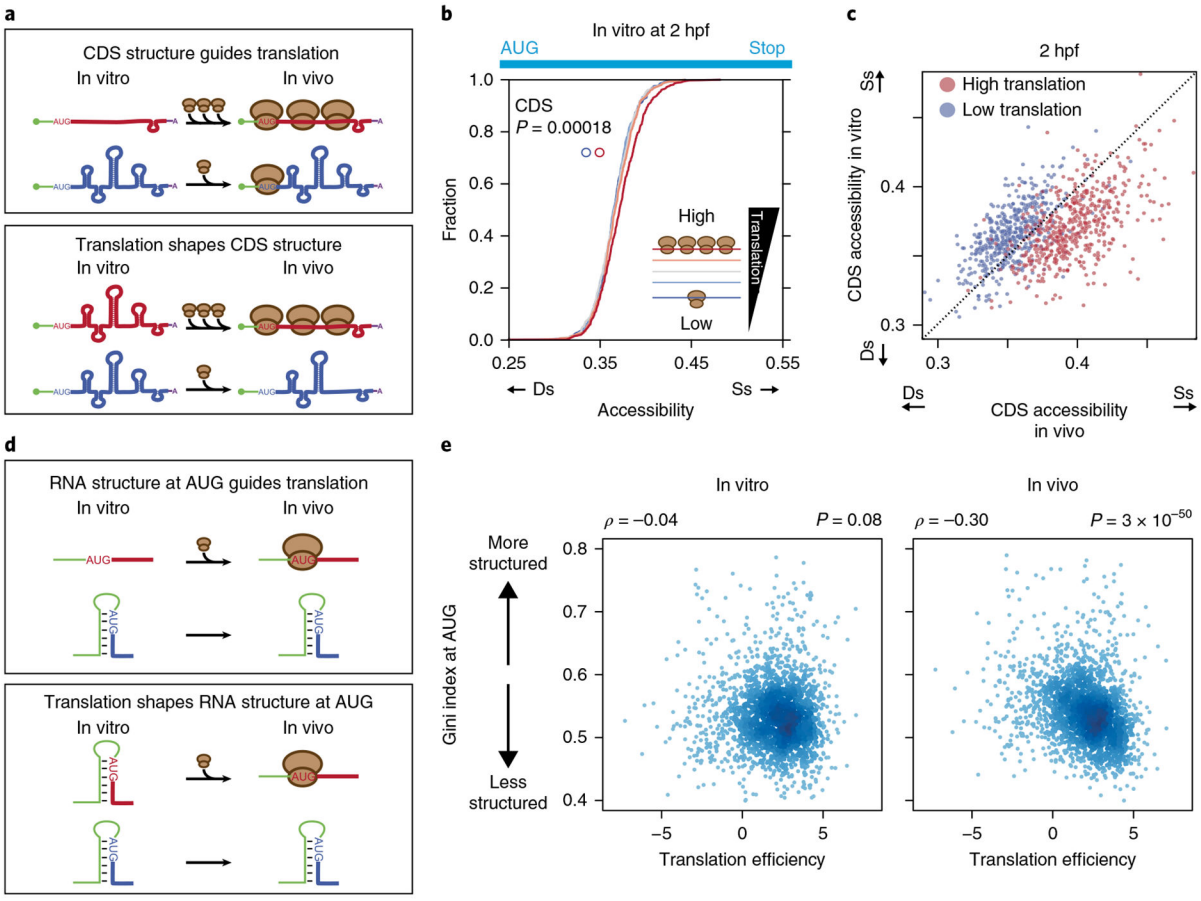

Fig. 2 |. Globally, mRNA intrinsic structure is not a main driver of translation.

a, Schematic representation of the two models that might explain the observed anticorrelation between translation efficiency and RNA structure in CDS regions in vivo. In the first model, preexisting CDS RNA structures guide translation (top; this is not supported by our analysis). In a second model, translating ribosomes are responsible for the unfolding of CDS RNA structures (bottom; consistent with our analysis). **b**, Cumulative distributions of global CDS RNA accessibility in vitro at 2 hpf, displaying no difference among mRNAs with different translation efficiency. Transcripts (n = 2,526) were binned into quintiles according to their translation efficiency. P value was computed with a one-sided Mann–Whitney U test. Although P was < 0.05, we considered this change nonsignificant, given the magnitude of the change and the large sample size; this conclusion is further supported by the scatter-plot analysis in Supplementary Fig. 2c. **c**, Scatter plot between in vivo and in vitro CDS accessibility at 2 hpf. In vivo, the accessibility was higher for highly translated mRNAs (red dots) and lower for weakly translated mRNAs (blue dots). **d**, Schematic view of the two models in which RNA structure at the AUG initiation codon is either the cause (top) or the result (bottom) of mRNA translation. **e**, Correlation between translation efficiency and the structure at AUG regions (n = 2,360), for both in vitro (left) and in vivo (right) conditions. Spearman correlation coefficients (ρ) and the corresponding asymptotic P values are shown.

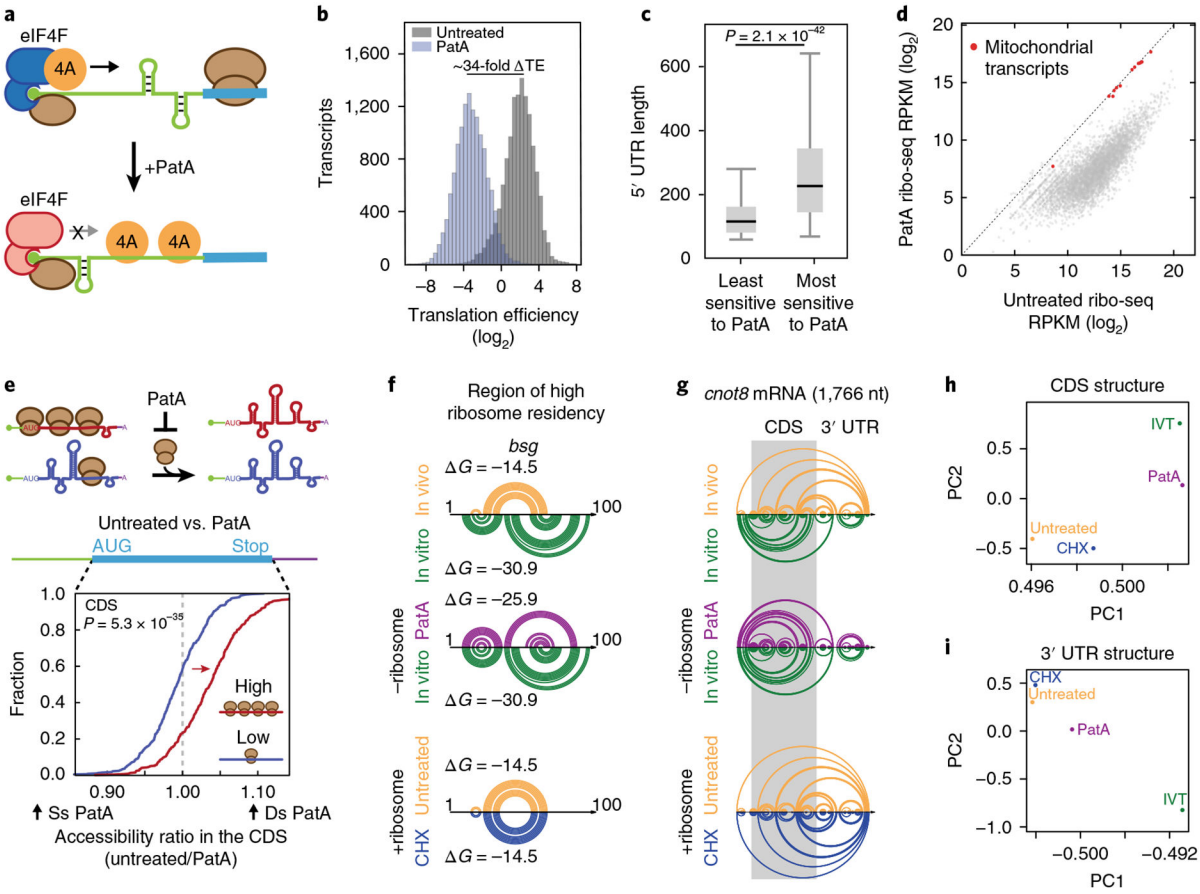

Fig. 3 |. I Ribosomes promote alternative mRNA structures in the early embryo.

a, Schematic view of the PatA translation-inhibition mechanism. PatA decreases the levels of functional eIF4F initiation complex, which is essential for cap-dependent translation, by trapping eIF4A (4A) on mRNAs and ectopically enhancing its RNA helicase activity. **b**, Distribution of translation efficiencies of PatA-treated (blue) and untreated (gray) embryos. **c**, Correlation between 5′ -UTR length and sensitivity to PatA treatment. The mRNA subgroups least and most sensitive to PatA correspond to the top (n = 394 transcripts) and bottom (n = 394 transcripts) quintiles of the change in translation efficiency (PatA - untreated), respectively. P value was computed with a one-sided Mann–Whitney U test. Box, first to last quartiles; whiskers, 1.5x interquartile range; center line, median. **d**, Scatter plot of the per-transcript ribosome footprints (ribo-seq ribosome profiling) in PatA-treated versus untreated embryos. Results for mitochondrial genes (red), which were unaffected by PatA treatment, remained unchanged. RPKM, reads per kilobase per million reads. **e**, Schematic representation depicting the effect of PatA, along with its corresponding cumulative distribution of CDS accessibility ratios (untreated/PatA). Accessibility (red arrow) was lower in highly translated mRNAs (red, n = 392) in PatA-treated than untreated embryos, an effect not observed for weakly translated mRNAs (blue, n = 392). P value was computed with a one-sided Mann–Whitney U test. **f**, Arc plots of DMS-seq-guided RNA secondary-structure predictions of a high-ribosome-occupancy region found in the *bsg* (*basigin*) gene. Each arc represents a base-pair interaction **g**, Arc plots of DMS-seq-guided RNA secondary-

structure predictions of the full-length *cnot8* mRNA from SeqFold. The CDS region is shadowed in gray. **h**,**i**, Principal component (PC)-analysis biplots of the SeqFold mRNA structure predictions of the four tested conditions (untreated, in vitro, PatA treated and CHX treated) for CDS (**h**, $n = 1{,}130$) and 3′ -UTR (**i**, $n = 1{,}085$) regions. Source data for **h** and **i** are available online.

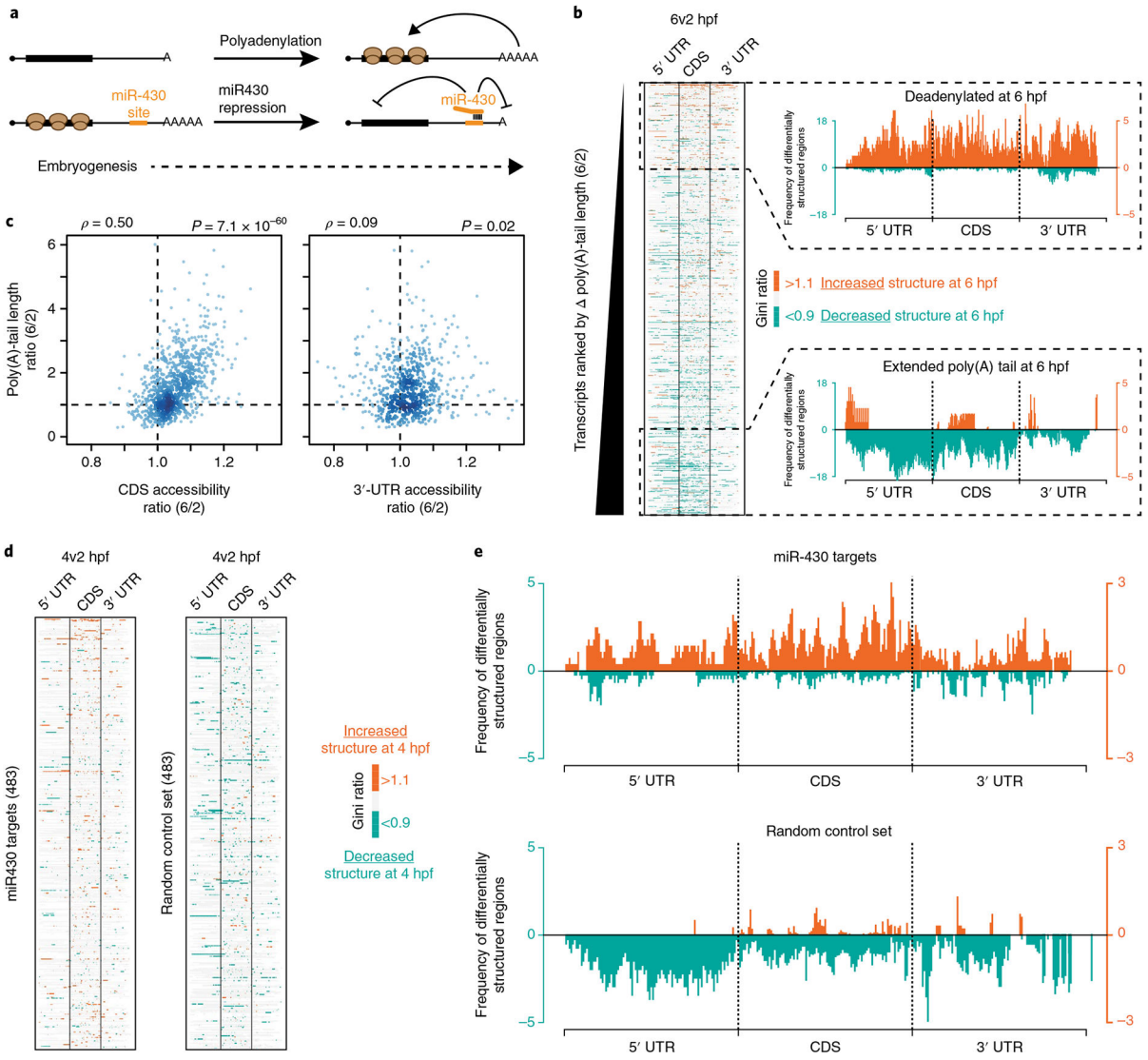

Fig. 4 |. Polyadenylation and miR-430 activity influence mRNA structure.

a, Schematic view of the effects of poly(A)-tail length and miRNA-mediated repression on translation during early embryogenesis. **b**, Left, structure changes along each transcript (total of 965 mRNAs with sufficient coverage in DMS-seq and poly(A)-tail-length profiling experiments at both 2 and 6 hpf), ranked by changes in poly(A)-tail length (6/2 hpf). Structure changes were identified by computation of the Gini ratio for each 100-nt sliding window (6/2 hpf). Right, cumulative distribution of differentially structured windows along the transcripts, for both the top 20% (top) and bottom 20% (bottom) transcripts, binned by changes in poly(A)-tail length. **c**, Correlation between changes in poly(A)-tail length and accessibility (6/2 hpf), for both CDS (left; $n = 939$) and 3′ UTR (right, $n = 724$). Spearman correlation coefficients (ρ) and the corresponding asymptotic P values are shown. **d**, mRNA structure changes along each of the miR-430 targets ($n = 483$) or a random set of transcripts ($n = 483$) taken from the non-miR-430 targets. Regions with structure changes were identified by computation of the Gini ratio for each 100-nt sliding window (4/2 hpf). **e**,

Metaplot of the differentially structured regions along the transcripts, comparing 2 and 4 hpf (4v2), for both miR-430 targets (n = 483) and non-miR-430 targets (n = 1,495).

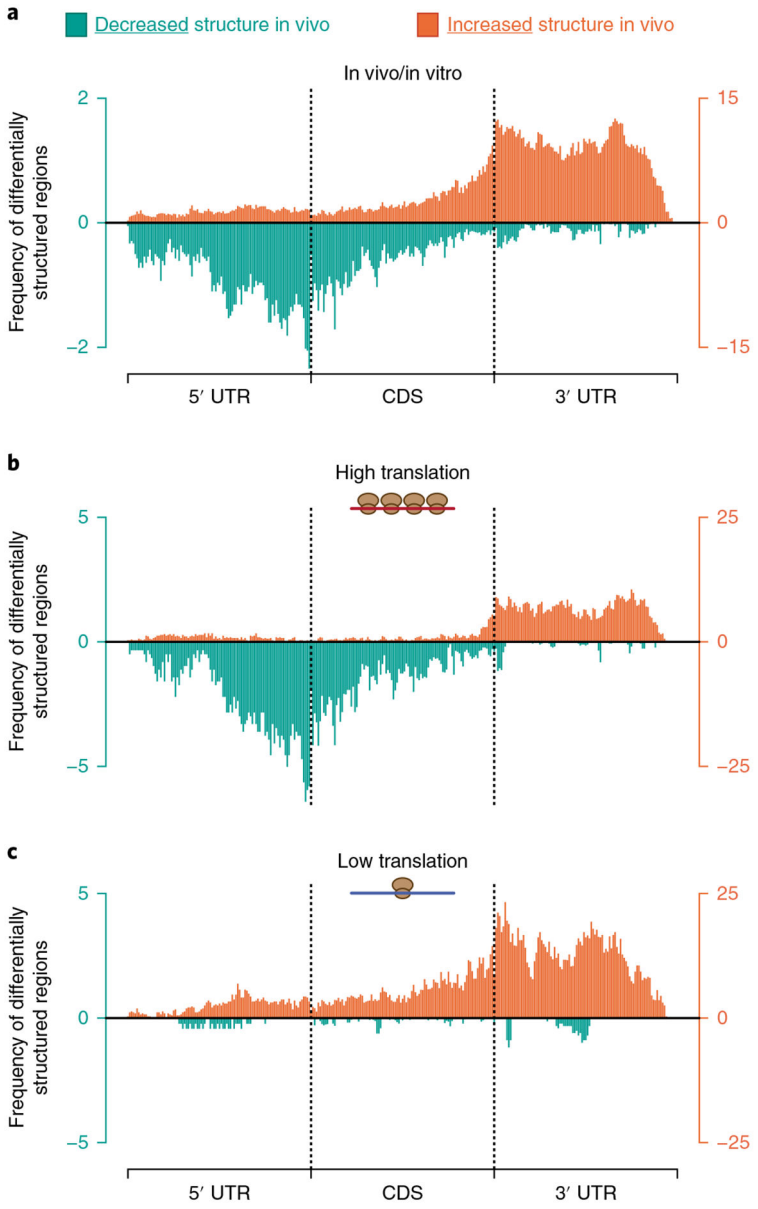

Fig. 5 |. 3′-UTR regions are more structured in the cell than in vitro.

a, Distribution of differentially structured windows along the transcripts, comparing in vivo and in vitro conditions. Regions with increased structure in vivo are orange, whereas those with decreased structure in vivo are turquoise. **b,c**, Distribution of differentially structured windows, comparing in vivo and in vitro conditions, for transcripts with either high translation efficiency (top 20%; **b)** or low translation efficiency (bottom 20%; **c).**

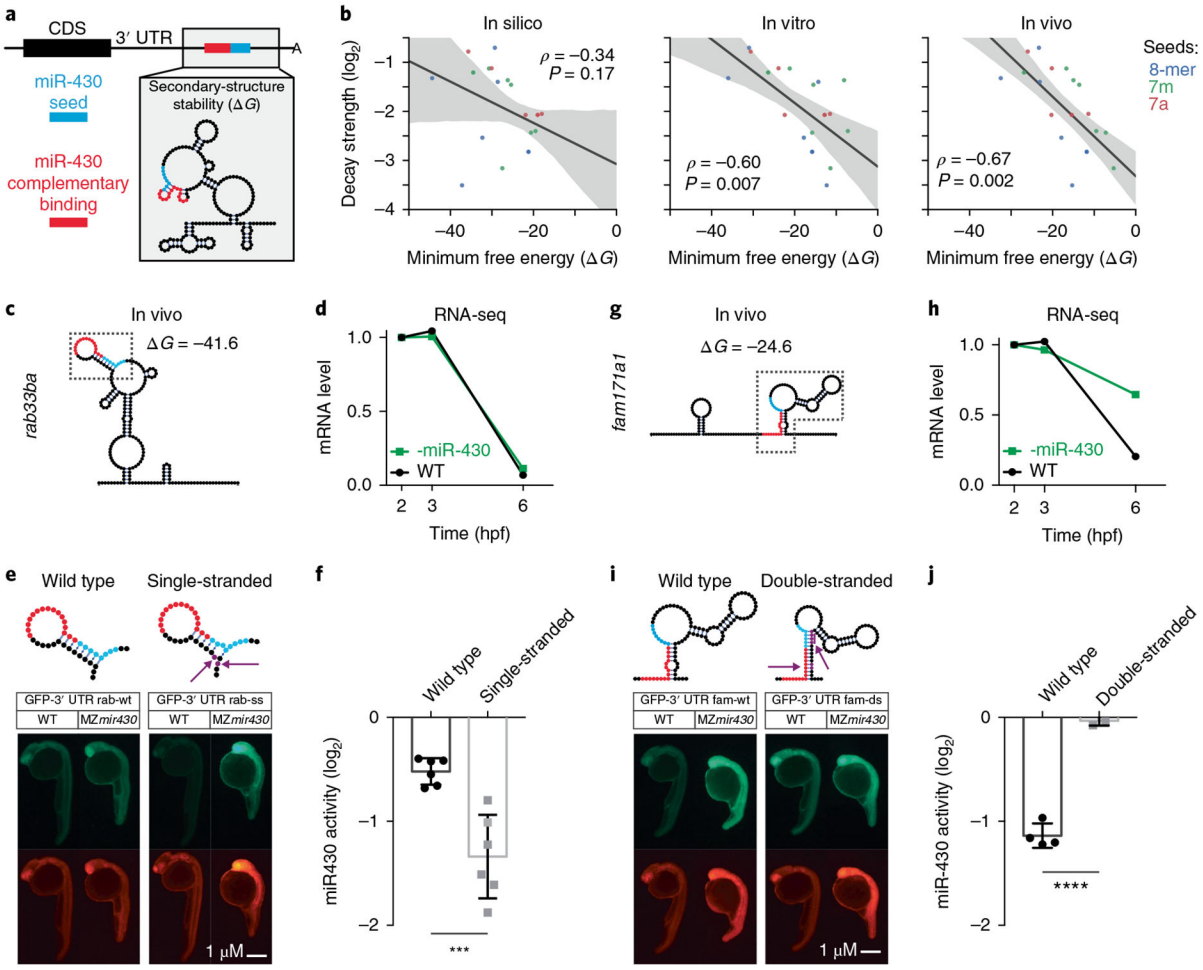

Fig. 6 |. Cellular-specific 3′-UTR structures modulate miR-430 activity and gene expression.

a, Cartoon representation of the RNA secondary structure found at a miR-430 target site. **b**, Correlation between the decay strength and the stability of the RNA structure of 18 endogenous miR-430-binding sites (colored according to their seep type), where the RNA secondary structure was predicted with an in silico approach (left) or DMS-seq guided by in vitro (middle) or in vivo structural data (right). Spearman correlation coefficients (ρ), asymptotic *P* values and regression models (shaded region) are shown ($n = 18$ different miR-430-binding sites). **c**,**g**, In vivo predicted secondary structure and stability (**Δ*G***) of the miR430-target site found in the *rab33ba* 3′ UTR (**c**) and in the *fam171a1* 3′ UTR (**g**). **d**,**h**, Changes in mRNA abundance during the MZT of the endogenous *rab33ba* (**d**) and *fam171a1* (**h**) transcripts in wild-type (WT) conditions (black, average of two independent RNA-seq replicates) or in conditions in which miR-430 activity is inhibited (green, one RNA-seq replicate). **e**,**i**, Fluorescence microscopy at 24 hpf of GFP reporters containing either a wild-type sequence of the *rab33ba* (rab-wt) or *fam171a1* (fam-wt) target site, compared with a destabilized version of *rab33ba* (*rab33ba* single-stranded, rab-ss (**e**)) or a stabilized version of *fam171a1* (*fam171a1* double-stranded, fam-ds, (**i**)) target site in the 3′ UTR. The specific mutations of the rab-ss and fam-ds constructs are highlighted in purple and with arrows. DsRed mRNA was co-injected as a control. Reporters were injected in both

the wild type and the maternal-zygotic mutant of miR-430 (MZ*mir430*) lacking the entire miR-430 locus. **f**,**j**, miR-430 activity calculated from the fluorescence of the GFP normalized to DsRed and the GFP/DsRed ratio of the MZ*mir430* mutants for each construct. Data are represented as mean ± s.d. (**f**: wt, $n = 6$; ss, $n = 6$; **j**: wt, $n = 4$; ds, $n = 3$ independent replicates). Student *t*-test *P* values (two tailed) are indicated as ***$P < 0.001$ and ****$P < 0.0001$. Source data for **b**, **d**, **h**, **f** and **j** are provided online.

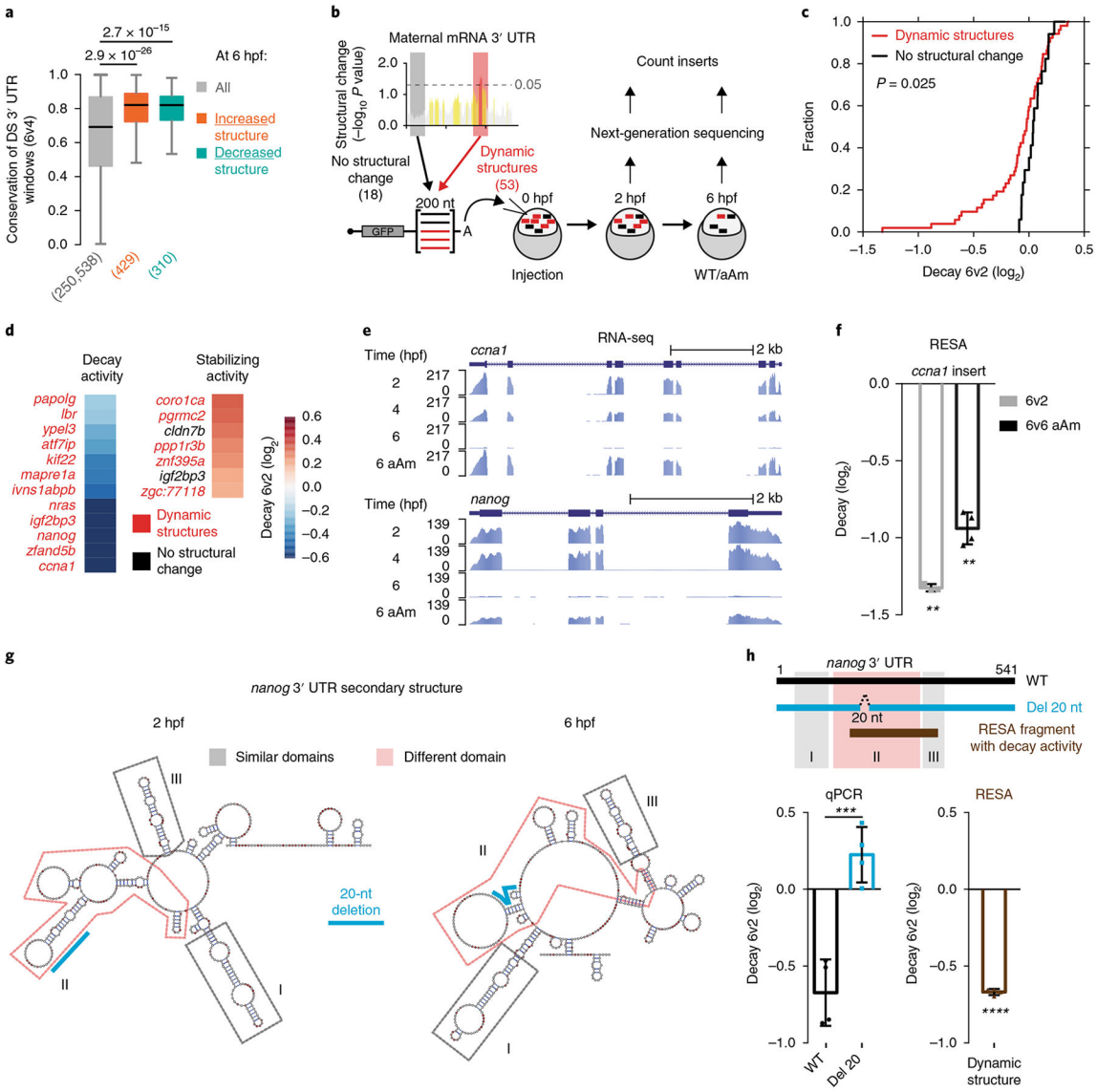

Fig. 7 |. Dynamic 3′-UTR structures are enriched in decay regulatory elements.

a, Sequence conservation of differentially structured 3′ -UTR regions between 4 and 6 hpf (6v4), compared with all 3′ -UTR regions analyzed (gray). *P* values were computed with one-sided Mann–Whitney *U* tests. Box, first to last quartiles; whiskers, 1.5× interquartile range; center line, median. **b**, Cartoon representation of the RESA experiment used to validate the regulatory activity of 3′ -UTR regions that were identified as changing (red) or not changing (gray) in their RNA structure during the MZT. **c**, Cumulative distribution of the decay activity measured by the RESA validation experiment for 3′ -UTR regions with dynamic structures ($n = 53$, red) or with no structural change ($n = 18$, black) during the MZT. *P* value computed by one-sided Mann–Whitney *U* test. **d**, RESA-validated 3′ -UTR sequences with identified decay or stabilizing elements. **e**, Endogenous mRNA expression of *ccna1* and *nanog* at 2, 4 and 6 hpf, as well as for α-amanitin (aAm)-treated embryos collected at 6 hpf. **f**, Quantification of the decay activity of *ccna1* measured with RESA

(mean ± s.d., 6v2 $n = 3$; 6v6 aAm $n = 4$ independent replicates), focusing on reporter-level changes between 2 and 6 hpf (6v2, light gray), and between 6 hpf untreated and aAm-treated samples (6v6 aAm, dark gray). *P* values were calculated with Student two-tailed *t* test; $^{**}P < 0.01$. **g**, Predicted RNA secondary structures of the 542-nt *nanog* 3′ UTR at 2 and 6 hpf. Structural domains that are similar at both developmental stages (I and III) are boxed in gray, whereas the one changing (II) is boxed in red. Cyan lines highlight the 20-nt deletion that disrupts a stem region in both structures. **h**, Location of the 20-nt deletion and the 200-nt RESA fragment with decay activity within the *nanog* 3′ UTR (top). Decay activity of the *nanog* wild-type 3′ UTR (black) and with the 20-nt deletion (cyan), quantified by qPCR (6v2), compared with the activity of the 200-nt fragment from the RESA experiment (brown). Data are represented as mean ± s.d. ($n = 4$ independent replicates for qPCR and RESA). $^{**}P < 0.01$, $^{***}P < 0.001$ and $^{****}P < 0.0001$ (Student two-tailed *t* test). Source data for **c**, **d** and **f** and are available online.