Supporting information for

Heterogeneous Dynamics of Protein-RNA Interactions across Transcriptome-Derived Messenger RNA Populations.

Burak Çetin[1], Gary J. Song[2], and Seán E. O'Leary[2*]

1. Graduate Program in Cell, Molecular, and Developmental Biology, University of California, Riverside, Riverside, CA 92521.
2. Department of Biochemistry, University of California, Riverside, Riverside, CA 92521.

*Corresponding author: sean.oleary@ucr.edu

Table of Contents

Supplementary Tables

Yeast growth and total RNA isolation.

A 5 mL starter culture of the yeast strain W303 (a generous gift from Justin Chartron, UC Riverside) was grown in 5 mL YPAD medium at 30 ˚C overnight. The starter culture was then used to inoculate 1 L of YPAD medium, and the culture was grown overnight at 30 ˚C, to stationary phase ($O.D._{600}$ ~ 3.0). Cells were then harvested by centrifugation (at 4,000 × *g* for 10 minutes at 15 ˚C). The resulting cell pellet was washed twice with ddH_2O to remove traces of growth medium, and was then resuspended in 25 mL RNA Lysis Buffer (50 mM Tris-HCl, pH 8.0 at 4C, 2% SDS, 2% β-mercaptoethanol, 10 mM EDTA). Cells were lysed by vortexing with zirconium beads (0.5 mm; 33% (v/v) BioSpec), through five cycles of one-minute vortexing followed by a two-minute pause with incubation on ice. The resulting lysate was centrifuged at 13,000 × *g* for ten minutes at 4 ˚C, and the supernatant was recovered. The supernatant, containing soluble protein and nucleic acids, was then subjected to two successive extractions with 1 volume of acidic phenol-chloroform (pH 4.5) to enrich for RNA. The aqueous phase was recovered, and nucleic acids were precipitated using 2.5 M LiCl. The resulting pellet was washed twice with 80% EtOH and resuspended in ddH_2O (typically ~1 mg /mL based on the 260-nm absorbance in a volume of ~4 mL). The total RNA preparation was first assessed for degradation on a 1% TBE-agarose gel, staining with ethidium bromide. To optimize the density of immobilized mRNA molecules in the ZMW array for single-molecule analysis, the RNA was then concentrated using Vivaspin® 6 centrifugal concentrators (Sartorius, MWCO 10 kDa), and aliquots were stored at –20 ˚C. Quality of RNA preparations was further assessed by BioAnalyzer size distribution analysis (Supplementary Fig. 1).

Protein expression and purification.

Hexahistidine-tagged recombinant *Saccharomyces cerevisiae* eIF4E and eIF4A were purified as reported previously (1). eIF4E was fluorescently labeled by thiol-maleimide chemistry, utilizing a protein variant that contains a single alanine to cysteine mutation (A124C in the standard yeast sequence numbering) as previously described (1). Briefly, *E.coli* cells heterologously overexpressing eIF4E or eIF4A from a pET-28a(+) vector were

grown in LB medium to an $O.D._{600}$ of ~0.6, then protein overexpression was induced by addition of 1 mM IPTG. Overexpression was allowed to proceed at 16 ˚C overnight.

For production of Cy5-labeled eIF4E, the protein was purified by a two-step procedure. Following affinity purification *via* the His_6 tag on Ni-NTA agarose (Qiagen), the protein sample (typically 1.5 mL of ~15 μM) was treated with 1 mg Cy5-maleimide (GE Healthcare Life Sciences). Labeling was allowed to proceed at 4 ˚C overnight, then unreacted dye was removed by gel filtration on a Bio-Rad 10DG desalting column. The protein was then further purified by size-exclusion chromatography on a Superdex 75 Increase column (10/300 GL; GE Healthcare Life Sciences).

Recombinant yeast eIF4A was similarly purified by Ni-NTA chromatography; the eluate was applied to a 5 mL Q HP anion-exchange column (GE Healthcare Life Sciences), which was eluted with a gradient of 0.1 – 1 M KCl. This step removed bound RNA effectively. Purification was then completed with using size exclusion chromatography on a Superdex 75 Increase column. Purified recombinant eIF4A was active as an enzyme in NADH-coupled ATPase assays conducted according to (2).

The purity of both proteins was assessed by 10% SDS-PAGE (Supplementary Fig. 2a); the RNA-dependent ATPase activity of purified eIF4A was assayed with a poly(U) substrate (Supplementary Fig. 2b). k_{cat} and K_M values were 0.026 ± 0.001 and 173 ± 35 μM bases, respectively. The k_{cat} was slightly higher than previously-reported values (3,4), which may reflect differences in assay conditions.

Preparation of Cy3-labeled, biotinylated immobilization oligonucleotide.

For selective immobilization of mRNAs, which contain a poly(A) tail, on the SMRT cell surface, a 5'-biotinylated and 3' amino-modified 45-mer oligo(dT) (Integrated DNA Technologies) was resuspended in 0.1 M sodium bicarbonate (pH ~8.3) to a final concentration of 100 μM. Cyanine 3 NHS ester (Lumiprobe), dissolved in DMSO (8 mM stock), was added to the oligonucleotide, to a final concentration of 2 mM. The conjugation

reaction was allowed to proceed at room temperature for four hours. Unreacted dye was then removed by four successive chloroform extractions, followed by desalting through a Micro-BioSpin P6 size-exclusion column (Bio-Rad). The resulting labeled oligonucleotide was stored in 10 mM Tris-HCl, pH 7.4 at –20 ˚C. The typical labeling efficiency was 50%, as determined from UV/Visible spectrophotometry.

Hybridization of total RNA to biotin-5'-$(dT)_{45}$-3'-Cy3 for single-molecule experiments.

Total RNA (5 mg/mL) was hybridized to 30 nM biotin-5'-$d(T)_{45}$-3'-Cy3 in 50 mM Bis-Tris propane (pH 7.0) and 0.5 M NaCl, by heating to 98 ˚C for three minutes in a thermocycler, followed by slow cooling to room temperature at 0.2 C/sec. The annealed product was used without further purification for RNA immobilization. Hybridization was confirmed by electrophoresing the resulting mixture on a 1% TBE-Agarose gel, then imaging for fluorescence with a Typhoon fluorescence scanner. These conditions resulted in near-quantitative capture of the immobilization oligo by the mRNA population (Supplementary Fig. 1b).

SMRT cell preparation and single-molecule experimental setup.

SMRT cells were purchased from Pacific Biosciences (part number 100-171-800). The chip was prepared as in (5), with minor modifications as follows. Prior to imaging, the chip was first wet with assay buffer (50 mM HEPES-KOH, pH 7.5, 3 mM $Mg(OAc)_2$, 100 mM KOAc) for two minutes, followed by derivatization with 16 μM NeutrAvidin (Invitrogen) for five minutes at room temperature. The chip was washed three times with assay buffer, and the total RNA sample with hybridized Cy3/biotin-$d(T)_{45}$ was immobilized for ten minutes, at a Cy3 concentration of 30 nM. The chip was subsequently washed three times with assay buffer and blocked for 15 minutes to prevent non-specific Cy5-eIF4E binding, using a mixture of BioLipidure 203 and 206 (2.5% (v/v) each; NOF America), 1 μM unlabeled eIF4E, and 1 mg/mL BSA. The chip was next washed once in imaging buffer (assay buffer supplemented with 1 mg/mL BSA, 2 mM PCA/PCD oxygen scavenging

system (6), then 20 µL of imaging buffer was added prior to imaging on the custom RS instrument. Cy5-eIF4E was robotically delivered to the sample in 20 µL imaging buffer, to a final concentration of 50 – 75 nM for replicate experiments. Mg•ATP was added to the delivery mix to produce a final concentration of 2.5 mM during imaging of eIF4E-mRNA binding. For experiments that included eIF4A, a final concentration of 2 µM eIF4A was included. Movies were imaged at 10 frames per second for ten minutes.

Experiments on individual *in vitro* mRNA transcripts were carried out in a similar fashion. Binding rates shown for individual mRNAs were reproducible to within 25–30% between replicate experiments.

Reverse transcription and cDNA library construction.

Total RNA (600 ng) was reverse transcribed using SuperScript™ II Reverse Transcriptase (Invitrogen). The manufacturer's protocol was followed, with modifications developed for single-cell RNA-seq and reported in the Smart-seq2 procedure (7); for analysis of bulk mRNA, single-cell lysate was substituted by bulk input RNA resuspended in ddH_2O. For reverse transcription in the ZMW array (SMRT cell), we utilized a modified Smart-seq2 protocol using a template switching oligo with a locked nucleic acid and a modified oligo d(T)VN primer (AAGCAGTGGTATCAACGCAGAGTACT_{30}VN) to synthesize both the first and the second strand (7). To provide suitable cDNA yields, the SMRT cell was incubated with total RNA::Cy3/biotin-oligo at four times the concentration used for single-molecule imaging, for one hour, and then washed three times in assay buffer. Following the washes, dNTPs and oligo d(T)VN were added to concentrations specified by Smart-seq2 procedure (7), respectively in a reaction volume of 20 µL with 1 U/µL RNAse inhibitor (New England Biolabs) The SMRT cell was overlaid with 20 µL SMRT Cell Oil (Pacific Biosciences, part number 100-171-800), to minimize evaporation during reaction. The cell was then placed into a Petri dish filled with water, floating on a water bath heated to 72 ˚C, to anneal the d(T)VN primer. Following incubation for 10 minutes, the chip was placed on ice for five minutes and the SuperScript™ II Reverse Transcriptase, DTT, template switching oligo, 5× First Strand Buffer, and $MgCl_2$ were added to a volume of 40 µL, along with an additional 25 µL of mineral oil, which was

required to cover the larger liquid surface caused by the increased sample volume. The chip was then returned to the water-filled Petri dish at 42 ˚C, and incubated for two hours. At the end of this incubation, the sample was aspirated from the SMRT cell and placed in a 1.5 mL microcentrifuge tube, then incubated at 72 ˚C for a further 15 minutes to inactivate the reverse transcriptase. After this step, 3 µL aliquots of the reverse transcription mixture were removed for PCR amplification with the KAPA HiFi HotStart ReadyMix polymerase (Roche), using the ISPCR oligonucleotide (AAGCAGTGGTATCAACGCAGAGT) described in (7). as a primer and standard cycling parameters with the number of PCR cycles increased to 30 (7). The resulting PCR mixture was purified with a MinElute PCR purification Kit (Qiagen). The cDNA size distribution was analyzed by Agilent BioAnalyzer. RT-PCR with this cDNA library confirmed the presence of select reference genes, such as *TDH3* (glyceraldehyde 3-phosphate dehydrogenase; GAPDH), and *UBC6* (Supplementary Figure 1c; Supplementary Table 1). These genes were chosen as they show mRNA abundances that vary by ~8 fold in yeast RNA-seq experiments (8).

Single-molecule fluorescence trajectories.

Fluorescent time trajectories from ZMWs were extracted using custom MATLAB code and binding was scored within the trajectories through the appearance of anticorrelated bursts between the donor (Cy3) and acceptor (Cy5) intensity, which is characteristic of FRET. Confirmation of single step photobleaching events to background fluorescence intensity levels in each assigned trace ensured the presence of single molecules, and traces containing unwanted photophysical effects such as blinking were excluded from the analysis.

Single-molecule data analysis.

Event timings from manually-curated traces were analyzed by an in-house MATLAB script. An [$f \times m$] matrix $\boldsymbol{\mathcal{D}}$ was constructed, where f is the number of frames in the movie (typically 5,900) and m is the number of molecules in the dataset. Each matrix element

$d_{f,m}$ was assigned a zero value if no event occurred for that molecule during that frame, or a nonzero value if an event occurred during the frame. A script was then written to count the number, a_m, of contiguous zero elements prior to a nonzero element, for each molecule. This script also counted the number, b_m, of contiguous nonzero elements before the next zero element. The arrival times $\{\tau_{\mathrm{arr}}\}_m$and durations $\{\tau_{\mathrm{dur}}\}_m$ of binding events for each molecule were then computed (in seconds, with a frame rate of 10 fps) as

$$\{\tau_{\mathrm{arr}}\}_m = \left\{\frac{a_m}{10}\right\}, \{\tau_{\mathrm{dur}}\}_m = \left\{\frac{b_m}{10}\right\}.$$

The total number of events in each trace was determined as the number of components n_m of each $\{\tau_{\mathrm{arr}}\}_m$.

Next, the empirical cumulative distribution function of the arrival times for each molecule was constructed from $\{\tau_{\mathrm{arr}}\}_m$. Traces with $n_m < 10$ were excluded from this procedure and the subsequent analysis, to ensure sufficient data were present for distribution fitting. The resulting distributions were then individually fit to the cumulative distribution function

$$P(t) = Ae^{-\lambda_1 t} + (1 - A)e^{-\lambda_2 t}.$$

"Robust" non-linear least squares regression was performed using the Levenberg-Marquardt algorithm with bisquares weighting. Fitted rates were included in the final on-rate distributions only when found to be satisfactory based on visual inspection of the residuals (Supplementary Figure 4a). The fast-phase amplitude, A, was typically greater than 0.8, i.e. this phase accounted for at least 80% of the distribution function; the slow phase was not considered further for data analysis. The slow phase may be a physical attribute of observing binding reactions of this type in zero-mode waveguides, as it was also observed in entirely unrelated studies on binding of a ribosomal protein to nascent ribosomal RNA (9) (Supplementary Discussion D2). To reduce the possibility of mis-assignment of mean arrival rates for each trace due to insufficient fit quality, fits were rejected where the 95% confidence intervals for the fast-phase rate parameter estimates lay further apart than 30% of the fitted value. 30% was selected as the cutoff because it is the typical variation between replicate experiments of this type when different preparations of reagents are utilized for data collection. The resulting association rate (s^{-}

[1]) for each molecule, $\langle \tau_{\mathrm{arr}} \rangle_m = 1/\lambda_1$, was then normalized for eIF4E concentration and labeling efficiency to arrive at the value of the association rate constant ($\mu M^{-1}\ s^{-1}$) referred to in the text. A small number of molecules with $k_{\mathrm{obs,on}} > 3\ s^{-1}$ were excluded in Figure panels 3(a) and 3(b) for clarity.

For analysis of binding event durations, the arithmetic mean was calculated as

$$\langle \tau_{\mathrm{dur}} \rangle_m = \sum_{i=1}^{|\tau_{\mathrm{dur}}|_m} \frac{\tau_i}{|\tau_{\mathrm{dur}}|_m}.$$

Number-of-event distributions for individual *in vitro* transcribed mRNAs.

In vitro transcription templates for the *JJJ1, SSA1, ATP4, GIC1* and *NCE102* mRNAs were prepared by PCR amplification from *S. cerevisiae* genomic DNA (EMD Milipore). These mRNAs were chosen as they show a ~3-fold difference in enrichment of eIF4E binding in an *in vivo* RIP-seq experiment (10). The mRNAs were then *in-vitro* transcribed with in-house purified T7 RNA polymerase. Commercial Vaccinia Capping System (NEB) and *E. coli* poly(A) polymerase (NEB) were used to add a 5' cap and 3' poly(A) tail to the *in-vitro* transcripts, per the manufacturer's instructions. The final capped and tailed mRNAs were purified by acidic phenol-chloroform extraction and LiCl precipitation as described for total RNA. Purified mRNAs were hybridized to the immobilization oligonucleotide through the same procedure used for total RNA, but with approximately 500 nM final mRNA concentration. Single-molecule fluorescence experiments were set up and analyzed identically to those with transcriptome-derived mRNA, in the presence of 70 nM Cy5-eIF4E.

Goodness-of-fit filter for arrival-rate distributions.

To further test whether the minimum, median, and maximum of the eIF4E-mRNA arrival-rate distribution was not biased by fitting parameters, a sub-distribution was created from traces where exponential fitting led to a root-mean-squared error (standard error of the regression) of less than 0.1, i.e. where the fitted curve lay on average no further than probability-0.1 from the data. While this significantly reduced the size of the final, filtered

dataset, each parameter of the distribution remained essentially unchanged (Supplementary Table 2).

Supplementary Discussion

D1. FRET signal as a measure of true eIF4E-mRNA association and dissociation rates.

To establish whether the eIF4E-mRNA FRET signal reliably reports on association and dissociation of eIF4E, rather than on changes in RNA conformation which move cap-bound eIF4E in and out of FRET range to the poly(A) tail, we carried out an experiment where we directly observed eIF4E-mRNA binding through excitation of its attached Cy5 fluorophore. We found that ~84% of eIF4E fluorescence events resulted in FRET across ~3,000 binding events on 100 mRNAs in the population (Supplementary Figure 3a,b). Conservatively, a further ~5–10% of the events lacking FRET can be explained by non-specific eIF4E interaction with the ZMW surface, placing a lower limit of ~89% on the proportion of true binding events that result in FRET. From this we conclude that initial eIF4E-cap binding out of FRET range to the poly(A) tail does not contribute significantly to our association-rate measurements. Once eIF4E is bound, if disappearance of FRET were due to conformational rearrangement, eIF4E-mRNA binding events in this direct-illumination experiment would increase in duration relative to the FRET experiment, and the events would contain FRET and non-FRET segments corresponding to the intervals where eIF4E was within and outside of FRET range to the Cy3 donor on the poly(A) tail. In contrast, we found that the duration of the events measured by FRET and direct illumination was identical within experimental error (Supplementary Figure 3c), and we found no evidence for events with mixed FRET/no-FRET behavior (Supplementary Figure 3a). Thus, conformational rearrangements also do not appear to contribute significantly to our dissociation-rate measurements.

D2. Double-exponential character of eIF4E-mRNA association kinetics.

The slow phase in the eIF4E–mRNA arrival-time distributions could have several origins. A clear possibility is that immobilized mRNAs present an ensemble of conformations with variable access for eIF4E to the cap structure. In this scenario, perturbation of the equilibrium between conformations is predicted alter the amplitudes of the fast- and slow-

binding contributions to the arrival-time cumulative distribution function. However, in experiments where we saturated the mRNA with an RNA-binding protein (eIF4A•(ATP•Mg)), we did not observe a significant change in the slow-phase amplitude (Supplementary Figure 3d,e). Thus, conformational equilibria do not appear to explain the existence of the slow phase.

Indeed, double-exponential association kinetics for protein-RNA interactions in zero-mode waveguides have been observed in unrelated studies of protein-RNA binding (9). Thus, the double-exponential behavior appears to be more a feature of ZMWs than the specific system under study. Since the floor of each zero-mode waveguide is decorated across its area with biotin, mRNAs can immobilize in a range of spatial environments, lying on a spectrum from mRNAs immobilized in the center of the waveguide to those immobilized at the edges. Differential positioning within the zero-mode waveguide may result in a range of diffusional profiles for eIF4E, in addition to hindered eIF4E binding to mRNAs where the cap is sterically occluded due to proximity to the ZMW side wall.

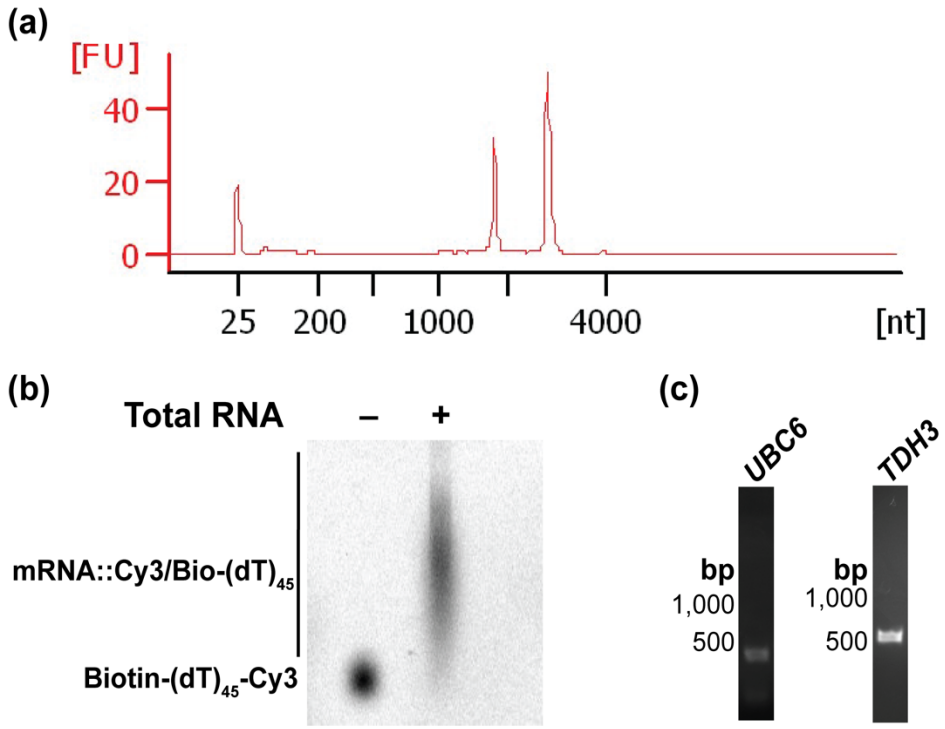
D3. Correlation of mRNA secondary structure with eIF4E-mRNA interaction kinetics.

For our analysis on how structural propensity near the 5' mRNA cap may modulate eIF4E affinity, we mined a published dataset with *in vitro* measurements of mRNA structure (PARS-seq) (11) and a further dataset where the structural propensity for the first thirty nucleotides was calculated (12). We *in-vitro* transcribed two yeast mRNAs that bracket the range of structural propensities near the cap (*SSA1* and *ATP4*): these have PARS scores for the first thirty nucleotides of 45.7 and -54.13, where a higher score indicates higher structural propensity. We found the eIF4E association rates (Supplementary Figure 6) for these mRNAs to be 4 μM^{-1} s^{-1} and 28 μM^{-1} s^{-1}, respectively, i.e. more structure resulted in a slower binding rate (Supplementary Figure 7a). We also prepared and analyzed two mRNAs with experimentally-identified internal ribosome entry site (IRES) elements (*GIC1* and *NCE102*) and found the eIF4E association rates to be 8 μM^{-1} s^{-1} for *GIC1* and 25 μM^{-1} s^{-1} for *NCE102*, which respectively lie above and substantially above the population median.

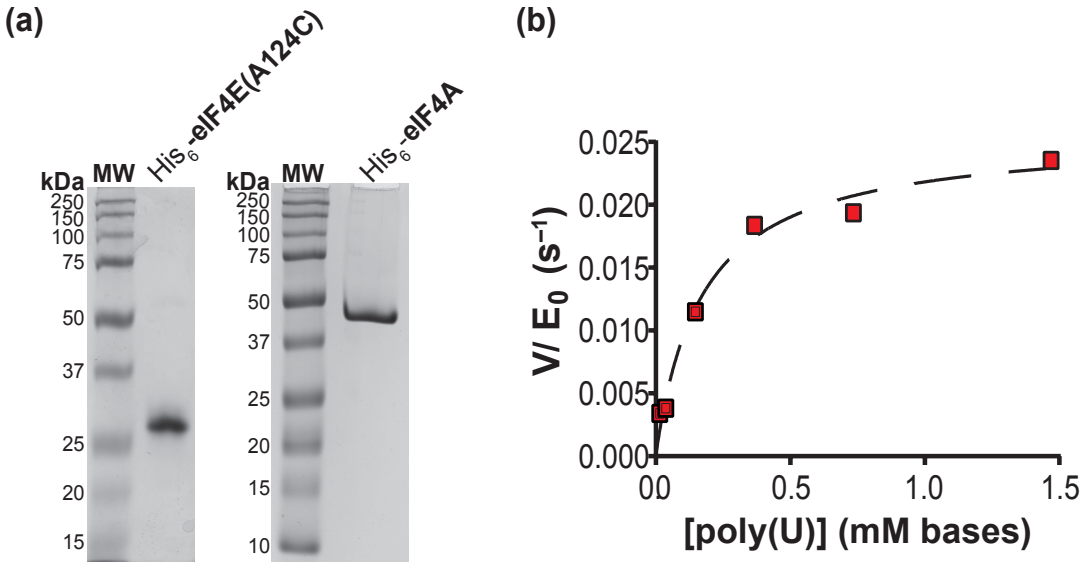
This correlation held for the 5'-proximal 20, 30, and 40 nucleotides (Supplementary Figure 7a). We did not consider the PARS score for the first 10 nucleotides, because the read depth underlying this score is variable across the mRNAs, reducing the reliability of comparisons between them. The rate-structure correlation apparently begins to weaken past 40 nt due to increased structural propensity toward 3' boundaries of the RNA transcript leaders. No correlation was observed between eIF4E association rates and the average PARS score per nucleotide for the full transcripts.

Correlation of the association rates with the $\Delta G_{folding}$ values for the isolated 5' 10-, 20-, 30-, and 40-nucleotide sequences of the mRNAs was undertaken because no experimental measurements of structure exist to our knowledge for these local mRNA elements. The lack of correlation is perhaps unsurprising since in the folded structures of the full-length mRNAs these cap-proximal sequences may base-pair with RNA regions quite remote from the cap.

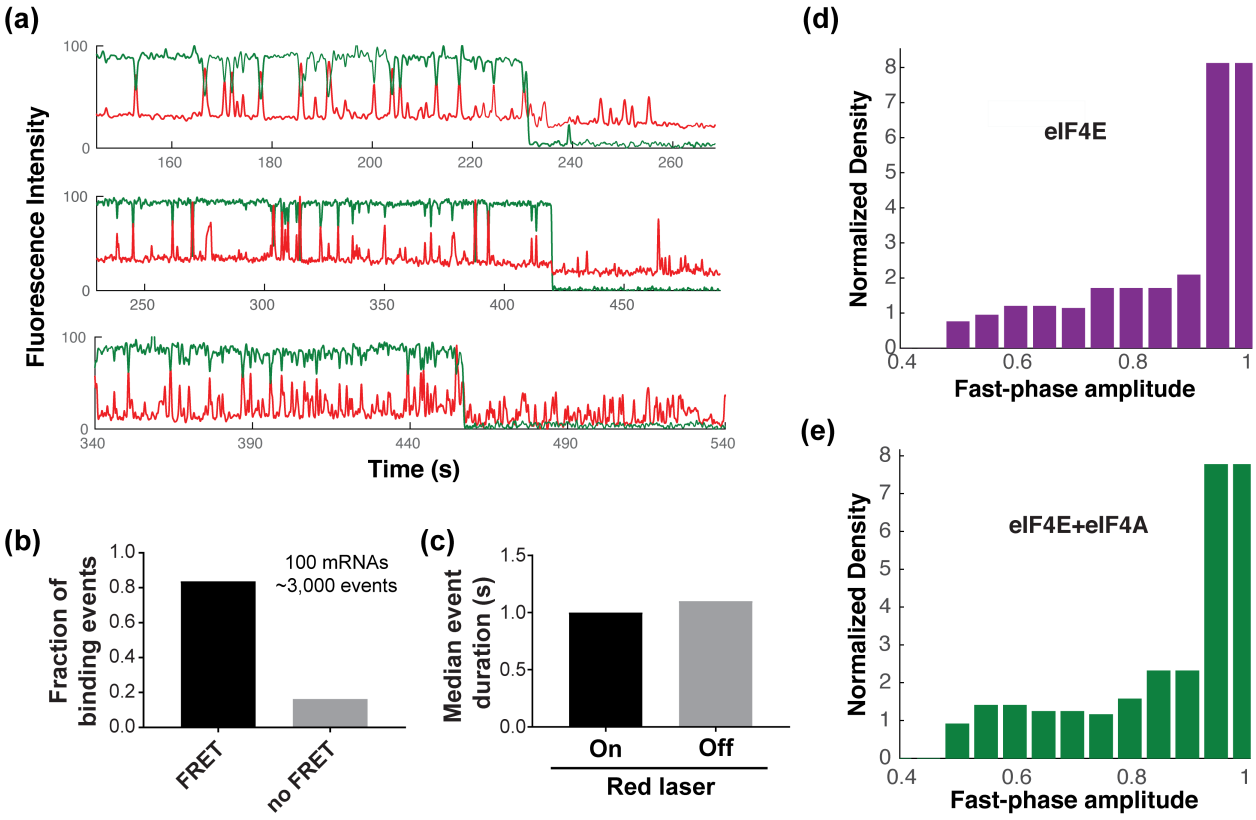
In general, IRES-containing elements tend to be highly structured in their leader sequences, leading to the expectation that the basal eIF4E association rate may be low. However, the *GIC* and *NCE102* mRNAs with experimentally-identified IRES elements actually vary in their structural propensity near the cap. That these mRNAs follow the structure-rate trend for non-IRES mRNAs is consistent with a model where the IRES-containing mRNAs do not leverage their structures to exclude eIF4E binding, but rather to promote direct ribosome recruitment.



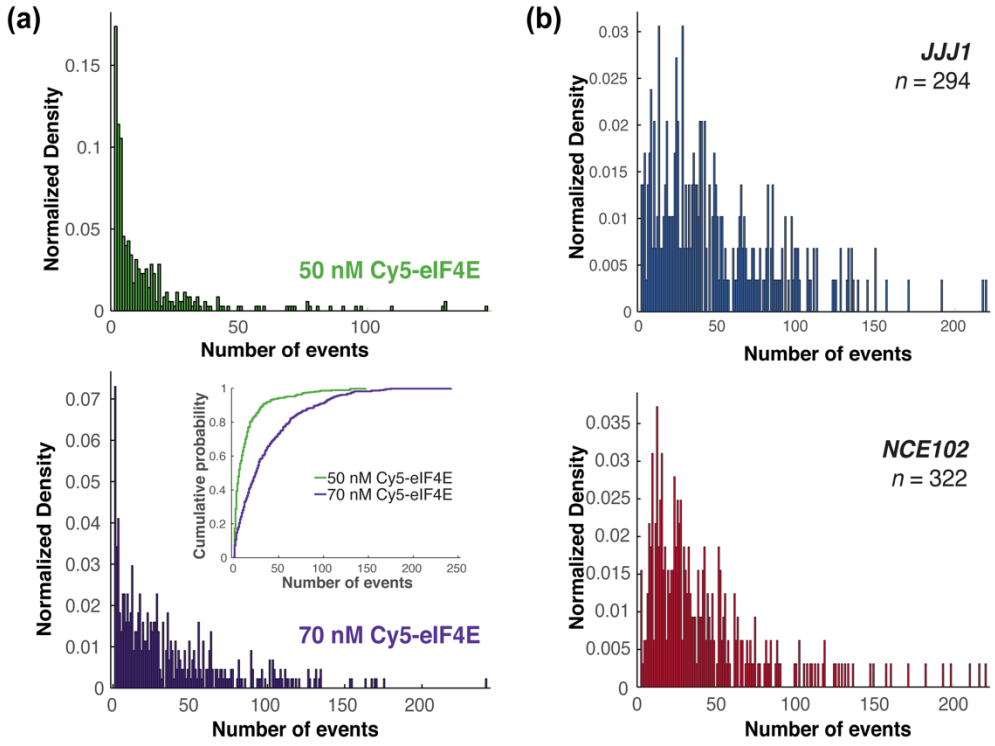
Supplementary Figure 1. Preparation of *Saccharomyces cerevisiae* total RNA. (a) BioAnalyzer analysis of total RNA preparation. The dominant peaks confirm the presence of intact ribosomal RNA. **(b)** Hybridization of immobilization/fluorescent labeling oligo (biotin-5'-(dT)45-3'-Cy3) to total mRNA, assayed by TBE-agarose electrophoresis with imaging for Cy3 oligo fluorescence. **(c)** Single-gene PCR analyses of ZMW-derived cDNA library, confirming the presence of mRNA for highly-expressed *TDH3* (GAPDH) and *UBC6* (~0.13-fold expression relative to GAPDH in publicly available RNA-seq data (7).



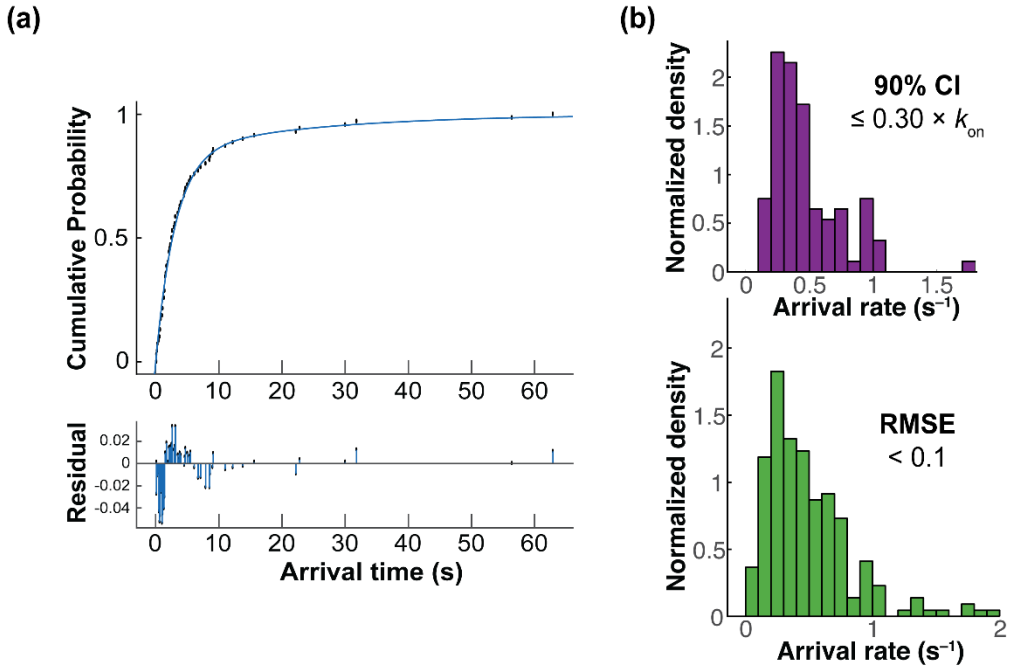
Supplementary Figure 2. Preparation and activity of protein reagents. (a) SDS-PAGE (10%) analysis of purified His_6-eIF4E(A124C) and His_6-eIF4A. **(b)** Michaelis-Menten plot for ATPase activity of recombinant His_6-eIF4A, in the presence of 5 μM eIF4A, saturating (5 mM) ATP and varying concentrations of poly(U) RNA.



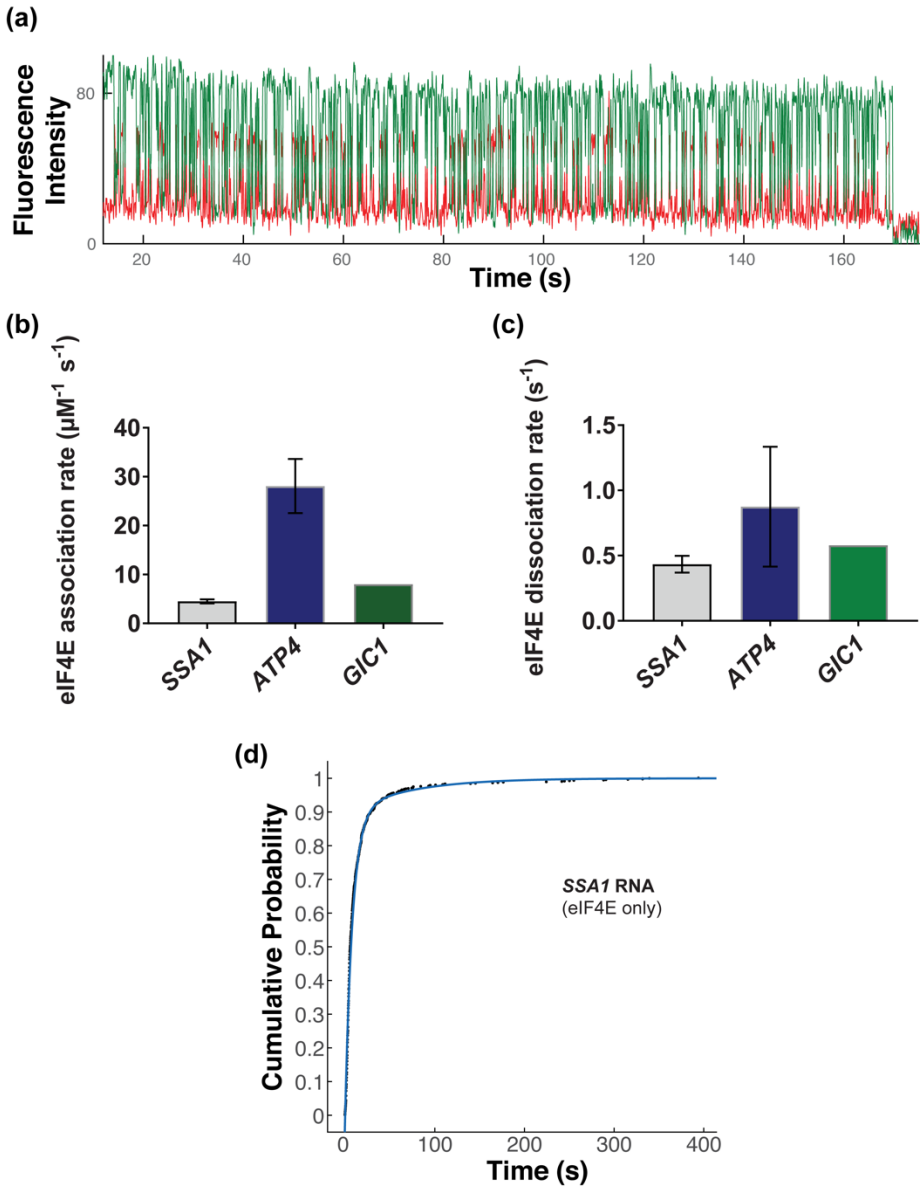
Supplementary Figure 3. Validation of smFRET signal as a reporter for eIF4E•mRNA complex formation. (a) Representative single-molecule trajectories for eIF4E–mRNA interaction observed with direct excitation of the Cy5 fluorophore attached to eIF4E, allowing tracking of all eIF4E binding events rather than those that result in FRET only. **(b)** Fractions of eIF4E-mRNA binding events in the direct-excitation experiment that show FRET (84%) and no FRET (16%) in a 10-minute movie across a population of 100 mRNAs. **(c)** Median lifetime of the eIF4E•mRNA complex compared between the smFRET ("Off") and direct-illumination ("On") experiments. **(d)** Distribution of amplitudes for the fast phase of eIF4E-mRNA association across the mRNA population in the eIF4E-only experiment. **(e)** Fast-phase amplitude distribution in the eIF4A+ATP experiment. Inclusion of eIF4A does not appreciably alter the distribution relative to panel (d).



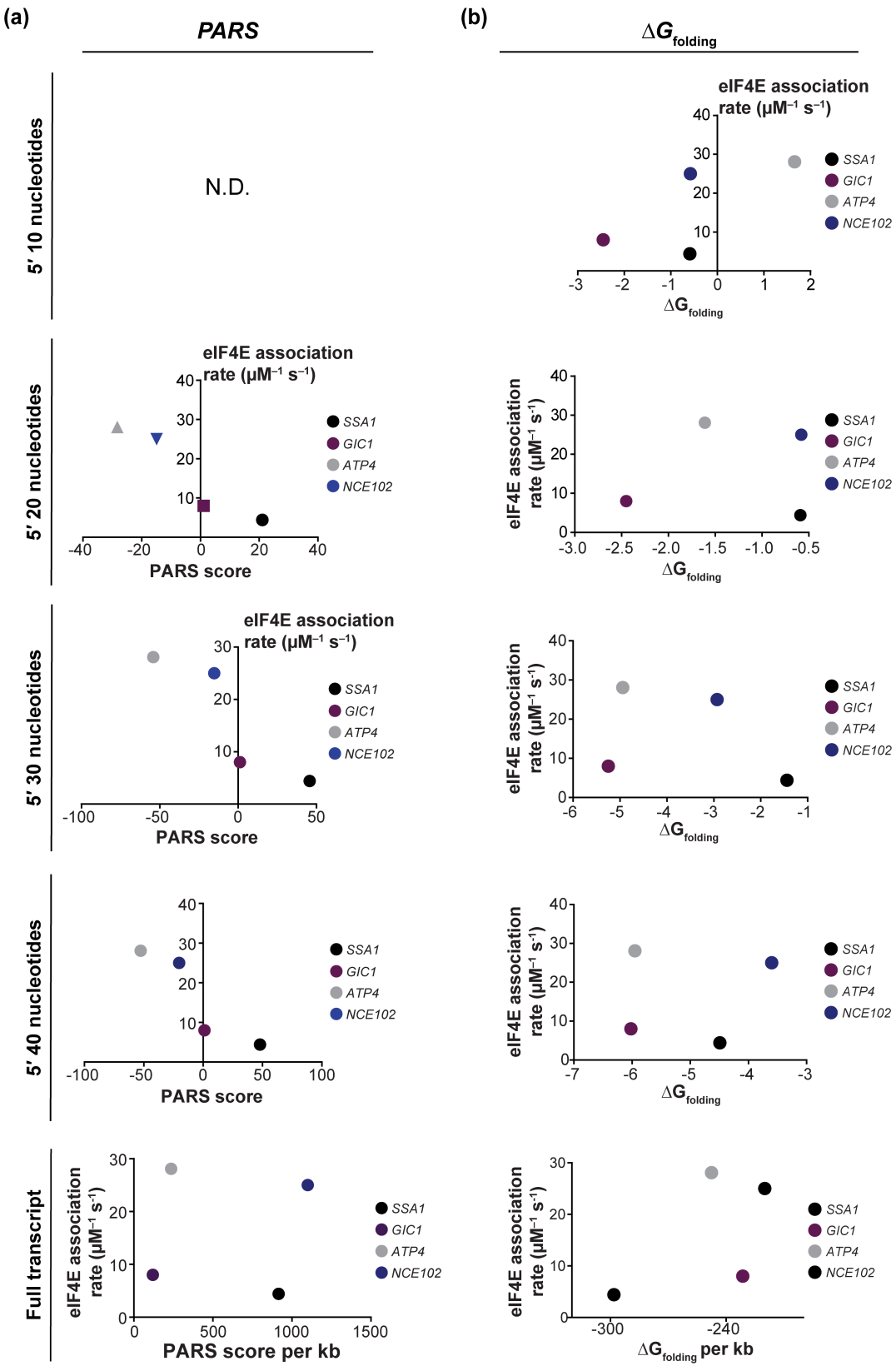
Supplementary Figure 4. Characteristics of kinetic heterogeneity in eIF4E-mRNA binding (a) Distributions of numbers of eIF4E-mRNA binding events on at varying eIF4E concentration. Top – distribution for 50 nM eIF4E binding to 351 mRNAs. Bottom – distribution of events for 70 nM eIF4E binding on 438 immobilized mRNA molecules (as in Fig. 2e, presented here for comparison with top panel). **(b)** Distributions of event numbers for 70 nM eIF4E binding to populations of the *JJJ1* (top) and *NCE102* (bottom) *in-vitro* transcripts.



Supplementary Figure 5. Goodness-of-fit analysis for on-rate parameters. (a) Top – representative exponential fit for eIF4E binding rate on an arbitrarily chosen mRNA with 70 binding events. Bottom – residual plot for the same fit. (b) Effects of increasing stringency of fitting on distributions of fitted arrival rates for 70 nM eIF4E binding to the transcriptome-derived mRNA population, as refined by (top) the 95% confidence intervals of the fitted rate parameter being separated by < 30% of the fitted rate, and (bottom) the root-mean-square error of the fit being <0.1.



Supplementary Figure 6. eIF4E–mRNA interaction kinetics on transcripts with varying structural propensities in their 5' leaders. (a) Sample smFRET trajectory for eIF4E-mRNA interaction. **(b)** eIF4E-association rates from exponential fitting. **(c)** Fitted eIF4E-dissociation rates. **(d)** Representative exponential fit to the eIF4E-mRNA arrival-time distribution for *SSA1* mRNA.



[Legend overleaf]

Supplementary Figure 7. Correlation between mRNA structural properties and eIF4E-mRNA association rate. (a) Correlation plots for eIF4E-mRNA association rate and PARS score for the first 20–40 nucleotides and the entire transcript. Insufficient data were available to calculate the PARS score for the first 10 nucleotides. The *SSA1* and *ATP4* mRNAs bracket the range of available PARS scores transcriptome-wide. **(b)** Correlation plots for eIF4E-mRNA association rate and computed folding free energy change at 30˚C for the first 10–40 nucleotides and the entire transcript.

Supplementary Table 1. Primers for RT-PCR amplification of reference genes from the SMRT-cell-derived cDNA library.

Gene	Direction	Primer Sequence
***TDH3* (GAPDH)**	Forward	5'-TACACTTCTGACTTGAAGATTGTTTCCAACG-3'
***TDH3* (GAPDH)**	Reverse	5'-GCCTTGGCAACGTGTTCAACC-3'
UBC6	Forward	5'-GGACGTTTCAAGCCCAACACACGATTATGC-3'
UBC6	Reverse	5'-CTTGTTCAGCGCGTATTCTGTCTTCAGGG-3'

Supplementary Table 2. 5th, 50th, and 95th percentile values for distribution of eIF4E-mRNA arrival rates filtered by varying goodness-of-fit metrics.

Filter	Percentile		
	5th (s^{-1})	**50th (s^{-1})**	**95th (s^{-1})**
None	0.07	0.3855	1.4648
90% CI	0.1805	0.3866	0.988
RMSE	0.1206	0.4164	1.6630

Supplementary References

1. O'Leary, S. E.; Petrov, A.; Chen, J.; Puglisi, J. D. Dynamic recognition of the mRNA cap by *Saccharomyces cerevisiae* eIF4E. *Structure* **2013**, *21*(12), 2197–2207.
2. Bradley, M. J.; De La Cruz, E. M. Analyzing ATP utilization by DEAD-Box RNA helicases using kinetic and equilibrium methods. *Methods Enzymol.*, **2012,** *511*, 29–63.
3. Andreou, A. Z.; Klostermeier, D. eIF4B and eIF4G jointly stimulate eIF4A ATPase and unwinding activities by modulation of the eIF4A conformational cycle. *J. Mol. Biol.* **2014,** *426*(1), 51–61.
4. Yourik, P.; Aitken, C. E.; Zhou, F.; Gupta, N.; Hinnebusch, A. G.; Lorsch, J. R. Yeast eIF4A enhances recruitment of mRNAs regardless of their structural complexity *Elife*. **2017,** *6*:e31476
5. Chen , J.; Dalal, R. V.; Petrov, A. N.; Tsai, A.; O'Leary, S. E.; Chapin, K.; Cheng, J.; Ewan, M.; Hsiung, P.-L.; Lundquist, P.; Turner, S. W.; Hsu, D. R.; Puglisi, J. D. High-throughput platform for real-time monitoring of biological processes by multicolor single-molecule fluorescence. *Proc. Natl. Acad. Sci. U. S. A*. **2014**, *111*(2), 664–669.
6. Aitken, C. E.; Marshall, R. A.; Puglisi, J. D. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophys. J.* **2008,** *94*(5), 1826–35.
7. Picelli, S.; Faridani, O. R.; Björklund, A. K.; Winberg, G.; Sagasser, S.; Sandberg, R. Full-length RNA-Seq from single cells using Smart-seq2. *Nat. Protoc*. **2014**, *9*(1), 171–181.
8. Lahtvee, P.-J.; Sánchez, B. J.; Smialowska, A.; Kasvandik, S.; Elsemman, I. E.; Gatto, F.; Nielsen, J. Absolute quantification of protein and mRNA abundances demonstrate variability in gene-specific translation efficiency in yeast. *Cell Syst.* **2017**, *4*(5), 495–504.e5.

9. Duss, O.; Stepanyuk, G. A.; Grot, A.; O'Leary, S. E.; Puglisi, J. D.; Williamson, J. R. Real-time assembly of ribonucleoprotein complexes on nascent RNA transcripts. *Nat. Commun*. **2018**, *9*(1), 5087.

10. Costello, J.; Castelli, L. M.; Rowe, W.; Kershaw, C. J.; Talavera, D.; Mohammad-Qureshi, S. S.; Sims, P. F. G.; Grant, C. M.; Pavitt, G. D.; Hubbard, S. J.; Ashe, M. P. Global mRNA selection mechanisms for translation initiation. *Genome Biol.* **2015,** *16*(1), 10.

11. Kertesz, M.; Wan, Y.; Mazor, E.; Rinn, J. L.; Nutter, R. C.; Chang, H. Y.; Segal, E. Genome-wide measurement of RNA secondary structure in yeast. *Nature* **2010**, *467*(7311), 103–107.

12. Sen, N.D.; Zhou, F.; Harris, M. S.; Ingolia, N. T.; Hinnebusch, A. G. eIF4B stimulates translation of long mRNAs with structured 5′ UTRs and low closed-loop potential but weak dependence on eIF4G*. Proc Natl Acad Sci U S A*.. **2016,** *113*(38), 10464–10472.

13. Freeberg, M. A.; Han, T.; Moresco, J. J.; Kong, A.; Yang, Y.-C.; Lu, Z. J.; Yates, J. R.; Kim, J. J. Pervasive and dynamic protein binding sites of the mRNA transcriptome in *Saccharomyces cerevisiae*. *Genome Biol.* **2013,** *14*(2), R13.

14. Ghaemmaghami, S.; Huh, W.-K.; Bower, K.; Howson, R. W.; Belle, A.; Dephoure, N.; O'Shea, E. K.; Weissman, J. S. Global analysis of protein expression in yeast." *Nature* **2003,** *425*(6959), 737–41.