HHS Public Access

Author manuscript

Nat Chem Biol. Author manuscript; available in PMC 2014 May 01.

Published in final edited form as:

Nat Chem Biol. 2013 November ; 9(11): 671–673. doi:10.1038/nchembio.1334.

Global profiling of stimulus-induced polyadenylation in cells using a poly(A) trap

Dusica Curanovic[1], **Michael Cohen**[1,2], **Irtisha Singh**[3,*], **Christopher E. Slagle**[4,*], **Christina S. Leslie**[3], and **Samie R. Jaffrey**[1,§]

[1]Department of Pharmacology, Weill Medical College, Cornell University, New York, NY 10065, USA

[2]Department of Pharmacology, Oregon Health Sciences University, Portland, OR, 97239-3098, USA

[3]Computational Biology Program, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10065

[4]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599

Abstract

Polyadenylation of mRNA leads to increased protein expression in response to diverse stimuli, but it is difficult to identify mRNAs that become polyadenylated in living cells. Here we describe a click chemistry-compatible nucleoside analog that is selectively incorporated into poly(A) tails of transcripts in cells. Next-generation sequencing of labeled mRNAs enables a transcriptome-wide profile of polyadenylation and provides insights into the mRNA sequence elements that are correlated with polyadenylation.

A major mechanism by which extracellular signaling molecules exert their effects in cells is by inducing the polyadenylation of specific transcripts, enabling their translation[1]. Up to 30% of cytosolic transcripts in mammalian cells exist in a deadenylated state, and may therefore be subject to this form of regulation[2,3]. However, identification of the specific transcripts that are polyadenylated in response to a given stimulus remains a challenge.

Current methods for identifying transcripts targeted for polyadenylation rely on the fractionation of RNAs based on the length of their poly(A) tails[4–6]. However, these methods do not distinguish between transcripts with pre-existing poly(A) tails, and those that become polyadenylated in response to a signal. Selective labeling of transcripts as they become

§Correspondence should be addressed to: S.R.J. (srj2003@med.cornell.edu).
*These authors contributed equally to this work.

AUTHOR CONTRIBUTIONS

D.C. and S.R.J. designed experiments, analyzed the data and wrote the manuscript, D.C. performed the experiments, C.E.S. performed oocyte injection experiments, I.S. and C.L. performed bioinformatics analyses, M.S.C. designed and synthesized EA and EAMP, and wrote portions of the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

polyadenylated in cells, followed by their recovery, can enable the identification of the mRNAs that are selectively polyadenylated in response to any signaling pathway.

To selectively label mRNAs during polyadenylation, we sought to develop an adenosine analog that contains a click chemistry-compatible tag, and which can be utilized by cellular poly(A) polymerases. To identify the position least likely to perturb enzymatic function, we examined the crystal structure of the mammalian poly(A) polymerase complexed with 3′-deoxyadenosine triphosphate[7]. The crystal structure revealed that the C2 atom of adenosine is distant from the enzyme's catalytic triad and does not make contacts with the surrounding residues (Supplementary Results, Supplementary Fig. 1a). Additionally, previous studies demonstrated that adenosine analogs containing halogen modifications of the C2 position do not inhibit poly(A) polymerases[8], suggesting that an alkyne moiety at the C2 position would be tolerated by the enzyme[9]. Importantly, poly(A) polymerases exhibit high conservation across paralogs and orthologs in the ATP binding pocket (Supplementary Fig. 1b), suggesting broad tolerance of the C2 alkyne modification. Therefore, we chose to append an alkyne moiety to the C2 atom of adenosine, yielding 2-ethynyl adenosine (EA) (Fig. 1a, Supplementary Note).

We next determined whether EA can be used as a substrate in polyadenylation reactions. We performed in vitro polyadenylation of an oligo$(A)_{12}$ oligonucleotide with adenosine 5′-triphosphate (ATP) and increasing concentrations of EA triphosphate (EATP). The products of the reactions were separated on a denaturing polyacrylamide gel and imaged using an RNA dye. Reactions containing ATP or EATP yielded long poly(A) chains, while the control reactions that contained chain-terminating cordycepin triphosphate[10] yielded no polyadenylation products (Supplementary Fig. 2a). To verify that EA is incorporated into poly(A) tails, the polyadenylation products were clicked with biotin-azide, and the presence of biotin was assessed by streptavidin-AP blotting. The signal increased with increasing EATP concentrations and product length, suggesting that the molecule is incorporated throughout the poly(A) chains (Supplementary Fig. 2b). Therefore, EA can be used to label polyadenylation products.

We next asked if EA is incorporated into RNA in living cells. As a control, we used 5-ethynyl uridine (EU), an alkyne-modified uridine analogue that is incorporated into cellular RNA by RNA polymerases[11]. We incubated rat E18 cortical neuron cultures (Fig. 1b) or HEK293 cells (Supplementary Fig. 3) with EU or EA, isolated total cellular RNA and performed the click reaction with biotin-azide to label the incorporated nucleosides. Both EU and EA were readily incorporated into RNA, as measured by streptavidin-AP blotting (Fig. 1b). Inhibition of cellular transcription with actinomycin D resulted in a complete loss of EU signal. However, actinomycin D reduced, but did not block, the incorporation of EA into RNA, indicating that EA is incorporated into RNA in both a transcription-dependent and transcription-independent manner (Fig. 1b). To test if EA is incorporated into RNA by polyadenylation, we inhibited cellular polyadenylation with cordycepin. This treatment substantially reduced EA incorporation into RNA (Fig. 1b). Simultaneous application of cordycepin and actinomycin D prevented EA incorporation (Fig. 1b). Therefore, EA is incorporated into cellular RNA by transcription as well as polyadenylation.

To confirm that EA is incorporated into poly(A) tails, we removed poly(A) tails by oligo(dT)/RNaseH treatment. The EU signal was unaffected by the removal of poly(A) tails. However, EA signal in RNA from actinomycin D-treated cells was lost (Fig. 1c). These data indicate that EA is metabolically incorporated into poly(A) tails in living cells.

We next wanted to define the set of mRNAs that are polyadenylated during cellular signaling. We investigated polyadenylation in *Xenopus laevis* that occurs during oocyte maturation. Application of progesterone to immature oocytes induces the polyadenylation of cytoplasmically stored, deadenylated mRNAs, which activate cellular pathways that orchestrate oocyte maturation[12]. We treated immature oocytes with EA in the presence or absence of progesterone. Transcription was inhibited by treatment with actinomycin D, which does not affect oocyte maturation[13]. Extracted RNA was clicked with biotin-azide and EA incorporation was analyzed by streptavidin blotting. The biotin signal from progesterone-treated oocytes was 5-fold greater than from non-treated cells, indicating that progesterone stimulation resulted in increased EA incorporation into RNA in a transcription-independent manner (Supplementary Fig. 4a).

We next wanted to confirm that EA is incorporated into newly polyadenylated transcripts. We used streptavidin beads to isolate EA-labeled transcripts, and performed RT-PCR to determine if transcripts known to undergo polyadenylation during oocyte maturation were recovered (Supplementary Fig. 4b). Each of the known polyadenylation targets was retained on the beads that contained RNA isolated from progesterone-treated oocytes (Fig. 2a, Supplementary Figure 5a), unlike the transcripts known to undergo deadenylation during oocyte maturation[14] (Supplementary Fig. 5b). These data show that EA labels transcripts undergoing stimulus-induced polyadenylation in living cells.

We sought to identify all mRNAs that are polyadenylated in oocytes in response to progesterone treatment. We subjected the pre-pulldown RNAs and the pulldown EA-labeled RNAs to deep sequencing. By comparing these samples, we identified 1737 transcripts that were significantly enriched by streptavidin pulldown (FDR-corrected $p < 0.1$), and which we designated “EA-trapped transcripts”. To enable bioinformatics comparisons of transcripts that exhibit different polyadenylation behaviors, we also defined RNAs that are not enriched by pulldown, which we termed “background transcripts”. The group of transcripts most underrepresented in the pulldown fraction (FDR-corrected $p < 0.1$) was termed “EA-depleted” (Supplementary Fig. 6, Supplementary Data Set 1).

We first asked if known targets of polyadenylation were present in the EA-trapped fraction. The well-established targets of progesterone-induced polyadenylation[15–19] were significantly enriched in the EA-trapped fraction (Supplementary Data Set 1). We next biochemically confirmed that other transcripts identified in this screen were polyadenylated during oocyte maturation. Validation was performed using anchored RT-PCR, which measures the overall poly(A) tail length of transcripts[20]. Randomly selected transcripts from our dataset exhibited increases in poly(A) tail lengths of 100–120 nucleotides following progesterone treatment (Fig. 2b). Prior to progesterone treatment, the poly(A) tails of these transcripts were adenylated to different extents, indicating that EA can be used to label newly polyadenylated transcripts regardless of their initial tail length.

For transcripts that are known to not undergo polyadenylation during oocyte maturation[16], no increase in tail length was seen (Supplementary Fig. 7). Importantly, among the EA-trapped transcripts that we tested biochemically, 91% exhibited progesterone-induced poly(A) tail elongation (Supplementary Table 1). Therefore, the EA-trapped transcript list reliably identifies mRNAs that become polyadenylated in response to progesterone.

Gene Ontology analysis revealed that transcripts targeted for polyadenylation participate in diverse cellular processes (Supplementary Fig. 8, Supplementary Tables 3–18). To confirm that the transcripts we identified are biologically relevant in oocyte maturation, we investigated lysine methyltransferase *setd8a*, which is highly enriched in our dataset. We determined that the *setd8a* transcript is subject to regulated polyadenylation and deadenylation during oocyte maturation (Supplementary Fig. 9a, 9b) and cell cycle resumption in human cells (Supplementary Fig. 9c). Knockdown of *setd8a* (Supplementary Fig. 10a, 10b) in *Xenopus* oocytes resulted in slower maturation kinetics and fewer oocytes undergoing GVBD compared to control oocytes (Supplementary Fig. 10c). These data show that *setd8a* is required for efficient oocyte maturation. Together, our findings reveal the diversity of polyadenylation-regulated processes and suggest that the targets we identified may be regulated by polyadenylation in other species and signaling contexts.

We next sought to investigate the 3′ UTR regulatory elements associated with polyadenylation. We first asked whether the motifs known to regulate polyadenylation were present in the 3′ UTRs of EA-trapped transcripts. The cytoplasmic polyadenylation element (CPE) is a conserved cis-regulatory element known to direct poly(A) tail elongation during progesterone-stimulated oocyte maturation, as well as in other signaling paradigms across species[3]. Indeed, the known CPE variants[21] were significantly enriched in EA-trapped transcripts compared to background RNA, and significantly depleted in EA-depleted RNA (Fig. 3). Additional 3′ UTR elements (MBE, PBE and ARE) known to regulate the stability and translational capacity of transcripts during oocyte maturation were also significantly enriched in EA-trapped transcripts, and depleted in EA-depleted transcripts (Fig. 3). These findings further support the physiological relevance of the polyadenylation targets identified using the EA-enabled poly(A) trapping method.

To identify new 3′ UTR elements that may regulate progesterone-stimulated polyadenylation, we performed a de novo search for all possible 8-mer sequences that were significantly enriched in the 3′ UTRs of EA-trapped transcripts compared to background RNA. The 249 most highly-enriched 8-mers (FDR-corrected $p < 0.0001$) were U-rich, and 69% contained at least 4 consecutive U residues. In addition to the known CPE sequences, among these 8-mers were numerous U-rich elements (Supplementary Table 2). The CPEs and UREs were enriched in 100 nucleotides proximal to the 3′ end in EA-trapped transcripts over background RNA (CPE, $p < 1.72 \times 10^{-6}$; URE, $p < 1.77 \times 10^{-7}$) (Supplementary Figs. 11, 12). These findings reveal that the 3′ UTRs of transcripts that undergo progesterone-induced polyadenylation are characterized by a pronounced enrichment of CPEs as well as novel UREs within 100 nucleotides proximal to the 3′ end of the transcripts. Using synthetic RNA reporter constructs, we determined that the novel UREs support polyadenylation to different extents in a CPEB-dependent manner (Supplementary Fig. 13), and therefore likely represent novel CPE variants.

Taken together, these data show that selective labeling of cellular mRNAs by endogenous poly(A) polymerases allows transcriptome-wide profiling of polyadenylation. Subsequent bioinformatics analysis of polyadenylation targets may provide novel insights into the mechanisms underlying the polyadenylation response in different signaling contexts. We expect that EA will facilitate the identification of novel polyadenylation events that mediate the effects of diverse signaling pathways in cells.

ONLINE METHODS

In vitro EA phosphorylation

To generate the triphosphate form of EA, we enzymatically phosphorylated 100 mmol 2-ethynyl adenosine monophosphate (EAMP, see Supplementary Note). In brief, EAMP was converted to 2-ethynyl adenosine diphosphate using 10 units of myokinase (Sigma Aldrich) and 100 mmol inosine triphosphate as a phosphate donor. The reaction was performed in poly(A) polymerase buffer (Affymetrix) at 37°C for 1 h. The product was then converted to 2-ethynyl adenosine triphosphate by adding 10 units of nucleotide diphosphate kinase to the reaction for an additional hour at 37°C. In parallel, in vitro phosphorylation of an equal amount of AMP (Sigma Aldrich) was performed.

In vitro polyadenylation

One microgram of poly$(A)_{12}$ was used as a substrate for polyadenylation reactions. Five units of yeast poly(A) polymerase were used in each polyadenylation reaction, supplemented with poly(A) polymerase buffer (Affymetrix). The reactions contained 10 mM ATP, substituted with increasing concentrations of the in vitro generated 2-ethynyl adenosine triphosphate or adenosine triphosphate. Polyadenylation reactions were performed at 37°C for 45 minutes. The products were then clicked with biotin-azide, purified, and equal amounts resolved on denaturing polyacrylamide gels.

Cell cultures

All reagents were from Invitrogen, unless stated otherwise. Cortical neurons were dissected from embryonic day 18 Sprague-Dawley rats, dissociated as described previously [22], and plated at a density of 10,000 cells per cm^2 in poly-L-lysine and laminin-coated dishes. The neurons were cultured for two weeks in Neurobasal medium supplemented with 1x B-27, 2 mM Glutamax, penicillin and streptomycin. The cultures were treated with 10 μM 5-fluorodeoxyuridine on day two to eliminate mitotic cells. HEK293 cells were cultured in DMEM supplemented with 10% FBS, penicillin and streptomycin. To induce G0 arrest, HEK293 cells were switched to serum-free medium for 3 days[23]. To resume the cell cycle, serum-starved HEK293 cells were treated with 20% FBS for 4 hours[23], at which point RNA and protein analyses were performed.

Ovaries extracted from *Xenopus laevis* were obtained from Nasco. The oocytes were defolliculated by incubating ovary pieces with gentle agitation in 60% Leibovitz L-15 media containing 2 mg/mL collagenase type I for 2 hours at 25°C. The oocytes were washed extensively and allowed to recover overnight at 16°C in 60% Leibovitz L-15 media

supplemented with 1% BSA, penicillin and streptomycin. *Xenopus* oocyte cultures were maintained at 16°C.

Antibodies

Antibody against activated caspase 3 from Santa Cruz Biotechnology was used for determining cell viability in immunofluorescence assays using standard protocols. Antibody against *Xenopus* CPEB that was used in oocyte injections was a gift from Joel Richter[24].

Drug treatments

Transcription by RNA polymerases I and II was inhibited in cortical neurons by incubating the cultures with 2 μM actinomycin D for 1 h. To inhibit polyadenylation, cordycepin (Sigma) was used at 20 μM for 1 h. The drugs were kept in the media for the duration of the experiment. To assess EA or EU incorporation into cellular RNA, 100 μM EA or EU was added to drug- or vehicle-treated cultures for 1 h.

Prior to all experiments with *Xenopus* oocytes, the cells were pre-treated with 39 μM actinomycin D for 2 h, which was kept in the media for the duration of the experiment. To induce oocyte maturation, 2 μM progesterone was added to the cultures. Oocytes were treated with 300 μM EA during progesterone or vehicle stimulation.

RNA isolation and click reaction

RNA was harvested using Trizol Reagent (Invitrogen) according to the manufacturer's instructions. Equal amounts of RNA were used for the click reaction with 0.5 mM biotin-azide. The click reaction was performed using the Click-iT Nascent RNA Capture kit as recommended by the manufacturer (Invitrogen), with the exception that the total volume of RNA in DEPC-treated water was 5 μL, and the reactions were brought up to 50 μL volume with DMSO. The reactions were purified using RNeasy spin columns (QIAGEN) as recommended by the manufacturer, with 3 additional washes with buffer RPE, and eluted in 15 μL of DEPC-treated water.

Anti-biotin blotting

All reagents were from Ambion. Equal amounts of RNA of each sample, ranging from 500 ng to 2 μg in each experiment, were denatured by heating at 50°C for 1 h in glyoxal loading dye. The samples were separated in a 1% denaturing agarose gel prepared with NorthernMax-Gly buffer. RNA was transferred onto positively charged nylon membranes using standard downward capillary transfer methods. In vitro polyadenylation products were separated on a denaturing 6% TBE-polycarylamide gel, and transferred using semi-dry transfer apparatus. Biotin was detected with streptavidin-AP using the BrightStar Biodetect kit.

Isolation of EA-labeled RNA and RT-PCR

Two micrograms of RNA clicked with 0.5 mM biotin-azide were subjected to streptavidin bead pulldown using the Click-iT Nascent RNA Capture kit, as recommended by the manufacturer. The RNA bound to 0.5 mg streptavidin beads was reverse-transcribed by

resuspending the beads in a 50 μL RT reaction containing Superscript III, RT buffer, dNTPs, RNaseOUT and a 1:1 mix of oligo(dT) and random hexamer primers. Reverse transcription was carried out at 45°C for 1 h, with shaking to prevent the beads from settling. The samples were then treated with RNaseH at 37°C for 20 min. After heating the samples at 80°C for 5 min, the beads were magnetically immobilized, and the reactions collected. For subsequent PCRs, 1 μL of the cDNA preparation was used per reaction. qPCR was performed using iQ SYBR Green Master Mix (Life Technologies). qPCR data were normalized to a spiked-in in vitro transcribed RNA (*hlb4*-derived RNA probe, see Synthesis of RNA Probes), which was biotinylated during transcription with biotin-14-CTP.

Anchored RT-PCR

Changes in transcript poly(A) tail length were determined by performing anchored RT-PCR, as described previously [20]. Briefly, 1 μg of each RNA sample was ligated to 0.1 μg of oligo (5′-P-GGTCACCTTGATCTGAAGC-NH_2-3′) using T4 RNA ligase. The oligo is phosphorylated at the 5′ end and blocked at the 3′ end by an amino modification that prevents concatemeric ligation. The ligation reaction was performed at 17°C overnight. Reverse transcription was performed using the entire ligation reaction, 0.1 μg of universal reverse primer (5′-GCTTCAGATCAAGGTGACCTTTTT-3′) and Superscript III, supplemented with reaction buffer, dNTPs and DTT. The resulting cDNA was diluted 3x with water, and 1 μL was used for subsequent PCR analysis.

Transcript-specific forward primers were 24-mers designed approximately 100 bp upstream of the polyadenylation signal. Universal reverse primer was used in each reaction, and PCR was performed using Platinum Taq DNA Polymerase. The products were resolved in 3% agarose or 6% TBE-polyacrylamide gels, and visualized with SYBR Gold.

Synthesis of RNA probes

RNA probes used to assess the function of different 3′ UTR elements were synthesized in vitro using the mMESSAGE mMACHINE T7 kit (Life Technologies). The DNA oligonucleotides that were used in transcription reactions contained a T7 promoter upstream of the modified *hlb4* 3′ UTR sequence, and a short poly(A) tail (5′-TTTTTTTTTTTTAAGAAAAAAAAAACAATTACTTACTTTATTTATAGAATTAAACATTAAAAACCAGATTGTACAGAGTCCAGAGTTGTAGACGCCCTATAGTGAGTCGTATTACTACTA-3′). The *hlb4* 3′ UTR sequence was altered to eliminate a U-rich stretch found at positions 879 to 884 of *hlb4*, and to include a unique primer annealing site (5′-CGTCTACAACTCTGGACTCTGTAC-3′) that was used in subsequent anchored RT-PCR assays as the gene-specific forward priming site. The endogenous CPE (5′-TTTTTAAT-3′), at positions 899 to 906 was replaced with a mutant CPE sequence (5′-TATGGAAT-3′) or a URE sequence.

Oocyte injections

Stage VI immature oocytes were injected according to standard procedures[25]. In knockdown experiments, the oocytes were injected with 10 nL droplets containing 5 ng of thioate end-blocked *setd8a* antisense oligonucleotide (5′-TACACCTCGCCCTTTCCCAG-3′), scrambled oligonucleotide (5′-ATACTCCTACGCTCCTCGCC-3′) or water. In experiments

testing polyadenylation of RNA probes, 0.1 fmol of each RNA probe was injected per oocyte. In some experiments, 10 nL containing 40 ng of anti-CPEB antibody or pre-immune rabbit serum was injected 12 hours before the injection of RNA probes.

Statistical analysis of experimental data

Means of 3 biological replicates are plotted in all qPCR graphs, and error bars represent standard deviation. Statistical significance between different conditions was determined by unpaired 2-tailed *t*-test.

Next-generation sequencing and processing

We isolated EA-labeled RNA and subjected it to deep sequencing. cDNA for deep sequencing was prepared using the Illumina TruSeq RNA Sample Preparation Kit with the following modifications: RNA from progesterone-treated oocytes was clicked with biotin-azide and pulled down with streptavidin beads as described above. The bead-bound RNA was introduced into the TruSeq protocol at step 12 of the subheading “Make RFP”. At this step, the RNA was fragmented and cDNA synthesis primed as described in the TruSeq protocol. The cDNA libraries were sequenced on an Illumina HiSeq 2000 instrument to generate 50 bp reads. The raw data were processed using CASAVA 1.8 pipeline (Illumina).

The sequencing reads from two biological replicates of the EA-pulldown cDNA library, and the total RNA library, were mapped with Bowtie[26] to the Xenopus Gene Collection cDNA library of 11,516 full-length clones. We used the DEseq Bioconductor package[27] to compute the differential enrichment of transcripts in the EA-pulldown samples relative to total RNA. After scaling read counts based on library size, DEseq estimates the significance of the difference in log mean read counts mapping to full-length transcripts in two conditions based on a negative binomial noise model. The dispersion function for this noise model was estimated from the two EA-pulldown replicates. We used a transcriptome-wide FDR-corrected *p*-value threshold of $p < 0.1$ to identify transcripts that were significantly enriched or depleted in the EA-pulldown.

The enriched (EA-trapped) and depleted (EA-depleted) datasets were subjected to GO analysis using the DAVID Functional Annotation Tool[28,29] (NCBI). Annotation terms were derived for 55% of the transcripts from the Swiss-Prot and Protein Information Resource database, and the EASE score was determined for each enriched term. A background *Xenopus laevis* gene file was used in enrichment analysis.

Computational and statistical analyses of large transcript sets

We parsed 3′UTR sequences from 10,585 transcripts using annotated sequence files downloaded from NCBI GenBank. We used these 3′UTR sequences to compute motif enrichments (or depletions) for transcripts that were enriched (EA-trapped) or depleted (EA-depleted) after pulldown. We first assessed the enrichment of known CPE motifs, MBE, PBE and ARE in the 3′UTRs of EA-trapped transcripts (FDR-corrected $p < 0.1$). We used Fisher’s exact test to compute the overrepresentation of each element in the EA-trapped transcripts relative to a background set of transcripts that are neither enriched nor depleted after pulldown. We also determined the depletion of these motifs in EA-depleted transcripts,

again computed against a background of transcripts that are neither enriched nor depleted in the pulldown.

To identify novel motifs overrepresented in the EA-trapped transcripts, we computed the enrichment of each possible 8-mer sequence in the EA-trapped set against the background of transcripts that are neither EA-trapped nor EA-depleted. The 249 most-enriched 8-mers were identified by Fisher's exact test (FDR-corrected $p < 0.0001$). Of these, 69% had stretches of 4 Us or longer, and were described as U-rich. A one-sided Kolmogorov-Smirnov test was performed to compare the U-rich element and CPE positions in the 3′ UTRs in each dataset.

We used UNAFold software[30] for secondary structure analysis. We established the number of secondary structures for each 3′ UTR in each transcript set (EA-trapped, background and EA-depleted), and determined the number of times each nucleotide appears single stranded, i.e. is not base paired. These per-nucleotide counts were normalized by the number of UNAFold-generated secondary structures to give an estimated probability that the nucleotide is unpaired. The distribution of per-nucleotide single-strandedness probabilities over the 3′UTR describes the overall propensity to remain single-stranded or to form secondary structures. To identify the positions in 3′ UTRs most likely to contain secondary structures, the single-strandedness probability of a given base was plotted against its position in the 3′ UTR.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank members of the Jaffrey lab for helpful comments and suggestions. We thank Joel Richter for the anti-CPEB antibody. We gratefully acknowledge Clay Bracken and Nina Svensen for assistance with NMR spectroscopy. This work was supported by NIH grants NIDA T32DA007274 (D.C. and M.S.C) and NINDS NS56306 (S.R.J.).

References

1. Weill L, Belloc E, Bava FA, Mendez R. Nat Struct Mol Biol. 2012; 19:577–85. [PubMed: 22664985]
2. Meijer HA, et al. Nucleic Acids Res. 2007; 35:e132. [PubMed: 17933768]
3. Novoa I, Gallego J, Ferreira PG, Mendez R. Nat Cell Biol. 2010; 12:447–56. [PubMed: 20364142]
4. Graindorge A, Thuret R, Pollet N, Osborne HB, Audic Y. Nucleic Acids Res. 2006; 34:986–95. [PubMed: 16464828]
5. Beilharz TH, Preiss T. Methods. 2009; 48:294–300. [PubMed: 19233282]
6. Meijer HA, de Moor CH. Methods Mol Biol. 2011; 703:123–35. [PubMed: 21125487]
7. Martin G, Keller W, Doublie S. Embo J. 2000; 19:4193–203. [PubMed: 10944102]
8. Chen LS, Sheppard TL. J Biol Chem. 2004; 279:40405–11. [PubMed: 15265873]
9. Martin G, Moglich A, Keller W, Doublie S. J Mol Biol. 2004; 341:911–25. [PubMed: 15328606]
10. Horowitz B, Goldfinger BA, Marmur J. Arch Biochem Biophys. 1976; 172:143–8. [PubMed: 766704]
11. Jao CY, Salic A. Proc Natl Acad Sci U S A. 2008; 105:15779–84. [PubMed: 18840688]

12. Radford HE, Meijer HA, de Moor CH. Biochim Biophys Acta. 2008; 1779:217–29. [PubMed: 18316045]
13. Schorderet-Slatkine S. Cell Differ. 1972; 1:179–89. [PubMed: 4670882]
14. de Moor CH, Meijer H, Lissenden S. Semin Cell Dev Biol. 2005; 16:49–58. [PubMed: 15659339]
15. de Moor CH, Richter JD. Mol Cell Biol. 1997; 17:6419–26. [PubMed: 9343404]
16. Sheets MD, Fox CA, Hunt T, Vande Woude G, Wickens M. Genes Dev. 1994; 8:926–38. [PubMed: 7926777]
17. Charlesworth A, Cox LL, MacNicol AM. J Biol Chem. 2004; 279:17650–9. [PubMed: 14752101]
18. Paris J, Philippe M. Dev Biol. 1990; 140:221–4. [PubMed: 2358121]
19. Culp PA, Musci TJ. Dev Biol. 1998; 193:63–76. [PubMed: 9466888]
20. Rassa JC, Wilson GM, Brewer GA, Parks GD. Virology. 2000; 274:438–49. [PubMed: 10964786]
21. MacNicol MC, MacNicol AM. Mol Reprod Dev. 2010; 77:662–9. [PubMed: 20652998]
22. Cohen MS, Bas Orth C, Kim HJ, Jeon NL, Jaffrey SR. Proc Natl Acad Sci U S A. 2011; 108:11246–51. [PubMed: 21690335]
23. Krek W, DeCaprio JA. Methods Enzymol. 1995; 254:114–24. [PubMed: 8531680]
24. Groisman I, et al. Cell. 2000; 103:435–47. [PubMed: 11081630]
25. Cohen S, Au S, Pante N. J Vis Exp. 2009
26. Jiang H, Wong WH. Bioinformatics. 2008; 24:2395–6. [PubMed: 18697769]
27. Anders S, Huber W. Genome Biol. 2010; 11:R106. [PubMed: 20979621]
28. Huang da W, Sherman BT, Lempicki RA. Nat Protoc. 2009; 4:44–57. [PubMed: 19131956]
29. Huang da W, Sherman BT, Lempicki RA. Nucleic Acids Res. 2009; 37:1–13. [PubMed: 19033363]
30. Markham NR, Zuker M. Methods Mol Biol. 2008; 453:3–31. [PubMed: 18712296]
31. Belloc E, Mendez R. Nature. 2008; 452:1017–21. [PubMed: 18385675]

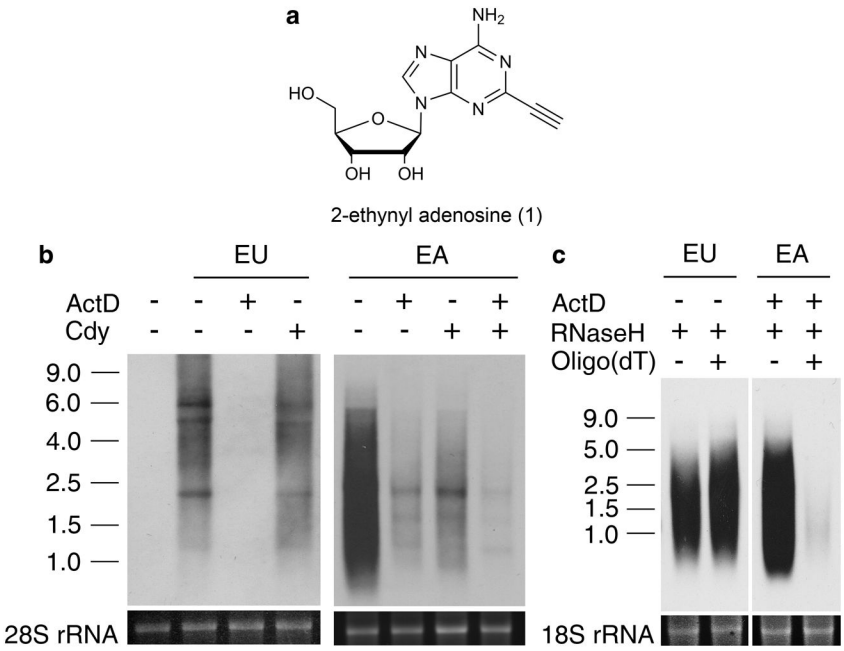

Figure 1. 2-Ethynyl adenosine (EA) synthesis and incorporation

(**a**) Structure of 2-ethynyl adenosine (**1**). (**b**) EA can be incorporated into mRNA during transcription and polyadenylation. Streptavidin blot showing EU or EA incorporation into RNA in cortical neuron cultures pre-treated with actinomycin D (ActD) or cordycepin (Cdy). (**c**) EA is incorporated into poly(A) tails in cells. Streptavidin blot of EU or EA-labeled cellular RNA after treatment with RNaseH in the presence or absence of oligo(dT). rRNA stained with ethidium bromide is shown as a loading control. Full blots and gel images are presented in Supplementary Figure 14.

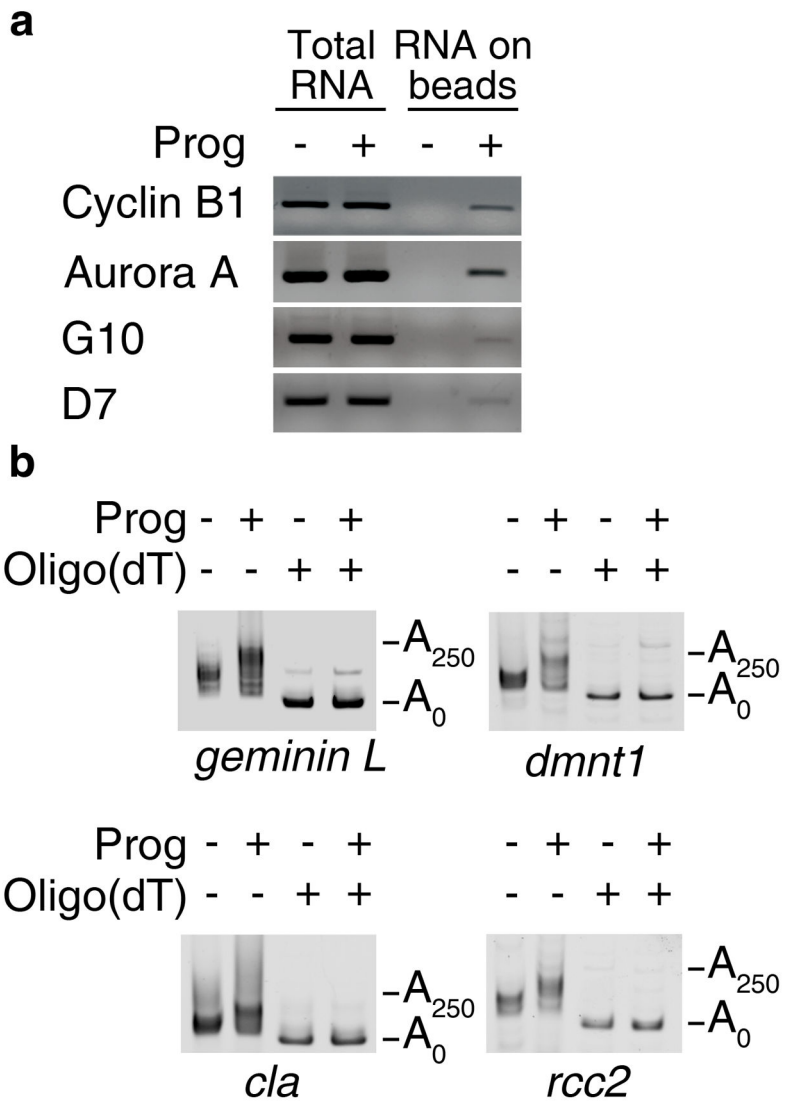

Figure 2. Using EA to isolate newly polyadenylated transcripts

(**a**) Known targets of polyadenylation are labeled by EA following progesterone treatment of *Xenopus* oocytes. EA-labeled RNA was clicked with biotin-azide and recovered on streptavidin beads. Specific transcripts were detected by RT-PCR. (**b**) Validation of polyadenylation state changes upon progesterone stimulation. RNA was isolated from oocytes incubated with progesterone or vehicle control, and treated with RNaseH in the presence or absence of oligo(dT). Anchored RT-PCR was used to measure poly(A) tail length. Full gel images are presented in Supplementary Figures 15 and 16.

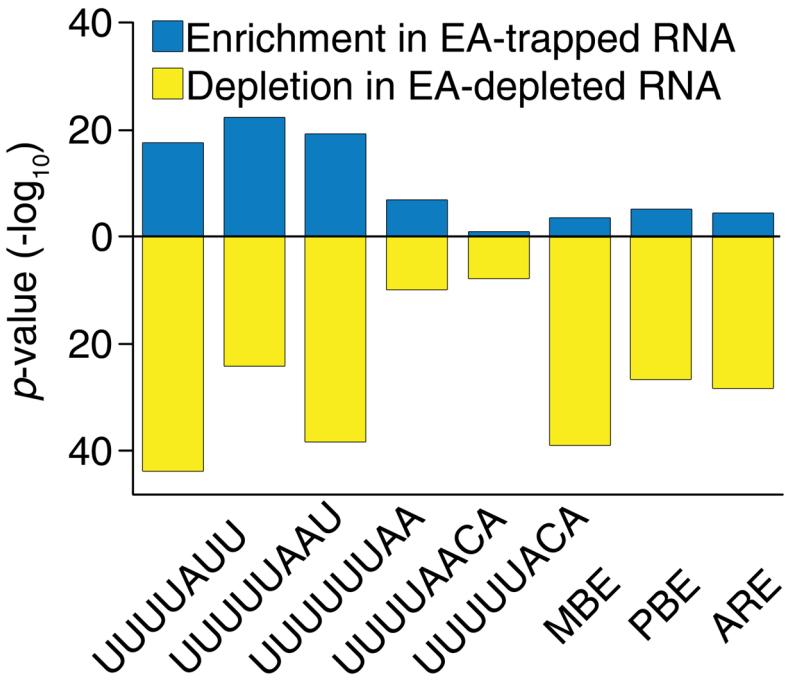

Figure 3. Enrichment analysis of regulatory motifs in 3′ UTRs of EA-trapped transcripts

The known CPE variants and other regulatory motifs are significantly enriched in the 3′UTRs of EA-trapped transcripts (blue bars), and depleted in 3′ UTRs of EA-depleted transcripts (yellow bars), relative to background transcripts.