
Alternative transcription start site selection leads to large differences in translation activity in yeast

MARIA F. ROJAS-DURAN and WENDY V. GILBERT¹

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

ABSTRACT

mRNA levels do not accurately predict protein levels in eukaryotic cells. To investigate contributions of 5' untranslated regions (5' UTRs) to mRNA-specific differences in translation, we determined the 5' UTR boundaries of 96 yeast genes for which *in vivo* translational efficiency varied by 80-fold. A total of 25% of genes showed substantial 5' UTR heterogeneity. We compared the capacity of these genes' alternative 5' UTR isoforms for cap-dependent and cap-independent translation using quantitative *in vitro* and *in vivo* translation assays. Six out of nine genes showed mRNA isoform-specific translation activity differences of greater than threefold in at least one condition. For three genes, *in vivo* translation activities of alternative 5' UTR isoforms differed by more than 100-fold. These results show that changing genes' 5' UTR boundaries can produce large changes in protein output without changing the overall amount of mRNA. Because transcription start site (TSS) heterogeneity is common, we suggest that TSS choice is greatly under-appreciated as a quantitatively significant mechanism for regulating protein production.

Keywords: TSS; 5' UTR; translation; cap-independent

INTRODUCTION

The proteins of a cell are the central determinants of cellular form and function. Regulation of the proteome is therefore the primary output of signaling pathways that connect cell physiology to internal and external environmental cues. Although most genome-wide studies of eukaryotic gene expression use mRNA levels as a proxy for protein levels, recent studies of yeast and human cells show that mRNA levels correlate only weakly with protein levels for the majority of genes (Ghaemmaghami et al. 2003; Beyer et al. 2004; Kolkman et al. 2006; Brockmann et al. 2007; Ingolia et al. 2009). Furthermore, based on quantitative genome-scale measurements of mRNA and protein abundance and turnover, it has been suggested that gene-specific differences in translation are at least as important as transcriptional control in determining steady-state protein levels (Schwanhäusser et al. 2011). The molecular mechanisms underlying these widespread differences in translational efficiency are poorly understood.

Translation initiation is thought to be the rate-limiting step in protein synthesis for most genes. For eukaryotic

mRNAs, translation initiation requires the concerted action of many initiation factors (eIFs) to recruit and position a ribosome at the AUG initiation codon (for review, see Jackson et al. 2010). Most eukaryotic mRNAs require a 5' -m⁷GpppN cap structure for efficient translation under normal conditions. Recognition of the cap by a complex of eIF4E, eIF4G, and the DEAD box helicase eIF4A with a cofactor eIF4B stimulates binding of a 43S pre-initiation complex (PIC) comprised of a 40S small ribosomal subunit, a ternary complex of eIF2-GTP-Met-tRNA_i^{Met}, and initiation factors eIF1, eIF1A, eIF5, and eIF3. mRNAs with short unstructured 5' UTRs do not require the cap-binding complex to associate with PICs *in vitro* (Algire et al. 2002). Conversely, some mRNAs with long-structured 5' UTRs require the action of additional DEXD/H-box "helicase" proteins, Ded1 or Dbp1 in yeast, and DXH29 in vertebrates, for efficient translation (Berthelot et al. 2004; Pisareva et al. 2008). 5' UTRs also play active roles in ribosome recruitment; the 5' UTRs of many viral mRNAs contain high-affinity binding sites for one or more host translation factors, permitting efficient cap-independent initiation (Filbin and Kieft 2009). Some cellular 5' UTRs are also capable of cap-independent initiation by mechanisms that are less well understood (Gilbert 2010). 5' UTR features thus influence both the efficiency and the factor requirements for ribosome recruitment to specific mRNAs.

¹Corresponding author

E-mail wgilbert@mit.edu

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.035865.112>.

Most studies of 5' UTR-mediated translation effects have investigated a single 5' UTR isoform for a given gene, yet alternative 5' UTR production is ubiquitous in eukaryotes. Based on sequencing of full-length cDNAs, many moderately to highly expressed yeast genes have multiple distinct 5' UTR variants that are generated by alternative TSS selection (Miura et al. 2006). Human transcription start sites are quite heterogeneous. At least 6000 human genes express mRNAs with alternative first (5' UTR) exons, and alternative TSS use is regulated in a tissue-specific manner (Carninci et al. 2006; Kimura et al. 2006; Wang et al. 2008; Yamashita et al. 2011). Recent work suggests that alternative TSS selection contributes more to mammalian mRNA isoform diversity than alternative splicing in some tissues (Pal et al. 2011). Mammalian genes with alternative 5' UTRs tend to be lowly expressed and are more likely to encode proteins with regulatory functions than genes with invariant 5' UTRs (Resch et al. 2009). Because the majority of alternative TSS choices do not change the mRNA's coding potential (Kimura et al. 2006; Miura et al. 2006; Yamashita et al. 2011), any biological effects must arise from isoform-specific mRNA regulation, e.g., due to differences in translation, stability, and/or localization. Here we describe experiments that examine the quantitative effects of specific yeast 5' UTR sequences on translation. The results demonstrate large effects on both cap-dependent and cap-independent translation activity conferred by alternative 5' UTR mRNA isoforms.

RESULTS AND DISCUSSION

Ribosome footprint profiling in rapidly dividing yeast revealed 100-fold differences between messages in translation activity (average number of ribosomes per mRNA) that have not yet been explained (Ingolia et al. 2009). To investigate whether differences between genes' 5' UTRs could account for this level of variation, we selected 96 genes for in vitro 5' UTR translational reporter studies. The chosen genes sampled the range of translation efficiencies observed in vivo (Fig. 1A). Endogenous 5' UTR boundaries were determined by RNA ligase-mediated rapid amplification of cDNA ends (RLM-RACE), and four independent clones were sequenced for each gene. Many genes' 5' UTR isolates differed by only a few nucleotides (nt) at their 5' ends, while 24 genes had 5' UTRs differing by >50 nt (Fig. 1B; Supplemental Table S1). This trend was also observed in a large-scale yeast cDNA sequencing study (Miura et al. 2006). For genes characterized in both studies, similar 5' UTRs were observed, despite differences in strain background and culture conditions (described in Materials and Methods) (Fig. 1C). The relative ratios of selected alternative 5' UTR mRNAs present in cells were quantified by qRT-PCR using primers specific for the longest mapped 5' UTR isoform compared with ORF primers. Amplicon abundance was related to absolute mRNA isoform abundance by

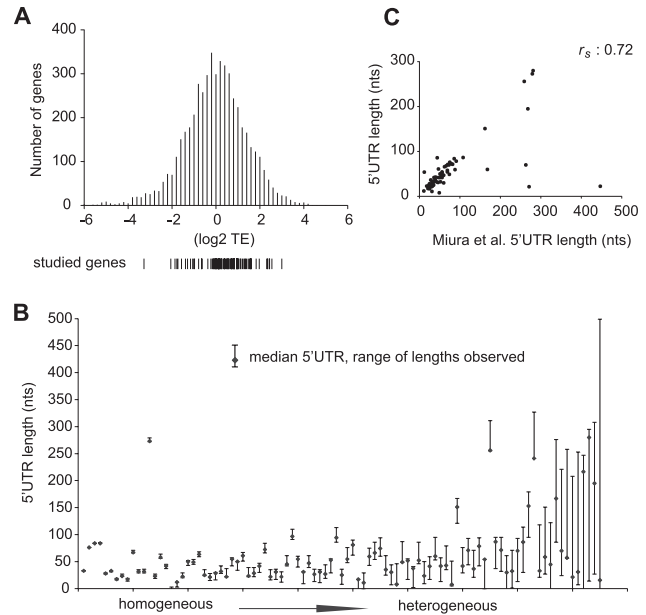


FIGURE 1. Some yeast genes have very heterogeneous 5' UTRs. (A) Histogram of yeast genes' relative translation efficiencies (TE) as determined by Ingolia et al. (2009). TEs for the 96 genes selected for study are indicated below. (B) 5' UTR lengths as determined by 5' RACE for 96 genes. (C) Comparison of 5' RACE results with 5' UTR boundaries from published cDNAs. Each dot represents the median 5' UTR length determined for one gene. A total of 27 out of 96 genes were represented by a single cDNA in Miura et al. (2006) and 32 genes were not represented. r_s indicates the Spearman rank correlation of 5' UTR lengths from 5' RACE clones and cDNAs.

comparison to recombinant in vitro-synthesized RNA standards (Supplemental Fig. S1A). The relative abundance of the longest isoforms detected by 5' RACE ranged from 10% (*FAR7*) to ~100% (*PRE2*) (Supplemental Fig. S1B).

Alternative 5' UTRs confer large differences in cap-dependent translation activity

5' UTR heterogeneity due to alternative TSS selection is common in eukaryotes. In most cases, the impact on protein production is not known. To explore the translational regulatory potential of alternative TSS choice, nine genes with the most heterogeneous 5' UTRs were selected for further characterization (Fig. 2A). None of these genes contained 5' UTR introns, which are rare in *Saccharomyces cerevisiae*. For each gene, reporter constructs were generated containing the longest and shortest 5' UTR variants fused to the Firefly luciferase open reading frame (ORF) under control of the T7 promoter. All translational fusions included 36 nt of endogenous ORF sequence at the N terminus to preserve the AUG context (Fig. 2B). A single guanosine was introduced at the 5' end of yeast 5' UTRs beginning with A/C/U to facilitate transcription initiation by T7 polymerase. Poly(A)₆₂ tails were encoded in the DNA templates. Capped mRNA was synthesized in vitro by run-off

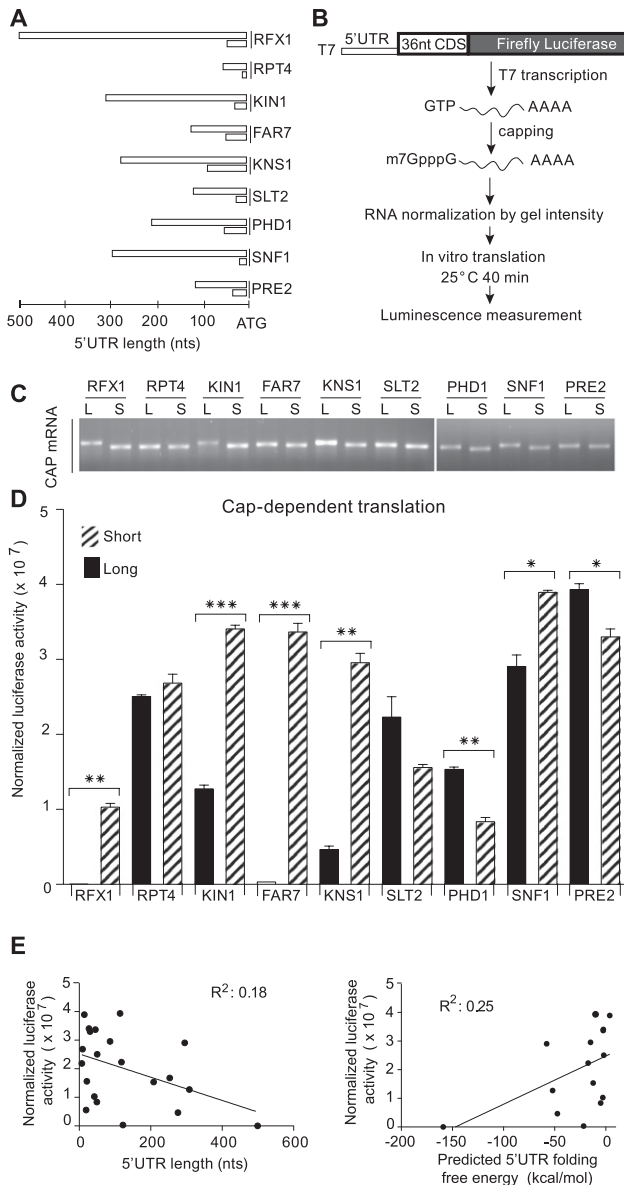


FIGURE 2. Alternative 5' UTR isoforms differ in translation efficiency in vitro. (A) The longest and shortest 5' UTR variants for nine genes with heterogeneous 5' UTRs. (B) Schematic of reporter constructs and translation assay. (C) Representative ethidium bromide stained gels used to normalize luciferase activity to mRNA concentrations. (D) Mean normalized luciferase activity from translation of capped (5'-m⁷GpppG) mRNAs. Error bars represent standard deviations. *P*-values determined by Student's *t*-test (two-tailed) are indicated above the bars; (*) *P* < 0.05; (**) *P* < 0.01; (***) *P* < 0.001. (E) Comparison of mean translation activity to (left) 5' UTR length and (right) predicted stability as determined by free-energy minimization using mfold (Zuker 2003).

transcription with T7 followed by capping with Vaccinia capping enzyme. Purified m⁷G-capped mRNAs were diluted to roughly equal concentrations and quantified by densitometry of agarose gels before use in translation assays (Fig. 2C); precise RNA concentrations as determined by gel analysis were used to normalize translation activity per

mRNA. Translation was performed in wild-type yeast extracts, which are strongly stimulated by m⁷G caps and poly(A) tails (Gilbert et al. 2007).

The 18 5' UTRs tested varied >1000-fold in cap-dependent translation activity (Fig. 2D). Even the most poorly translated, *RFX1* "long" (*RFX1.L*), was 50-fold above background and could be measured reliably. The low-translation activity of *RFX1.L* was not due to poor capping efficiency; long and short isoforms of *RFX1* were equally well capped post-transcriptionally, and similar results were obtained in translation assays of mRNAs prepared by cotranscriptional capping (data not shown). In some cases, long and short 5' UTR isoforms of a single gene showed substantial differences in activity. The shorter variants of *RFX1* and *FAR7* were translated >100-fold better than the long, and seven out of nine genes tested showed significant differences between long and short 5' UTR isoforms. Although the largest differences between isoforms (*RFX1*, *FAR7*, and *KNS1* "long" versus "short") favored the shorter 5' UTR isoforms, length was only weakly anticorrelated with translation activity overall (Fig. 2E, left). Notably, *FAR7.L*, *SLT2.L*, and *PRE2.L* 5' UTRs are approximately the same length, but spanned the full range of observed translation activities. Likewise, predicted RNA secondary structure, which tends to increase with 5' UTR length, was only weakly predictive of poor translation activity (Fig. 2E, right). Furthermore, for three out of nine genes, the longer 5' UTR isoform was more active. Together, these results show that intrinsic differences between 5' UTRs are sufficient to cause large differences in translation activity that are not readily predicted by simple rules.

Next we tested whether alternative 5' UTRs are sufficient to cause large differences in translation activity in vivo. For six genes that showed significant differences in cap-dependent translation in vitro, in vivo reporter constructs were generated containing the longest and shortest 5' UTR variants fused to the Firefly luciferase ORF under control of a modified inducible GAL1 promoter that generates transcripts with a defined 5' end (Fig. 3A; data not shown). Translation activity for each 5' UTR construct was determined by measuring luciferase activity in whole-cell lysates normalized to total protein concentration and reporter mRNA levels as determined by qRT-PCR. The in vivo assays mimicked the effects of alternative 5' UTRs observed in vitro. In every case, the "long" or "short" 5' UTR variant that was better translated in vitro was also better translated in vivo. Moreover, the quantitative differences between variants were comparable in vivo to those observed in vitro for most constructs (cf. Fig. 2D and Fig. 3B). Two 5' UTRs that showed relatively poor translation in vitro were even less active in vivo (*KNS1.L* and *PHD1.S*). A potential explanation for this difference is that these mRNAs may fail to compete for limiting translation factors in vivo in the presence of abundant cellular mRNA. In contrast, the *FAR7.L* mRNA was somewhat better translated in vivo than in vitro; more robust translation in vivo might indicate the presence or

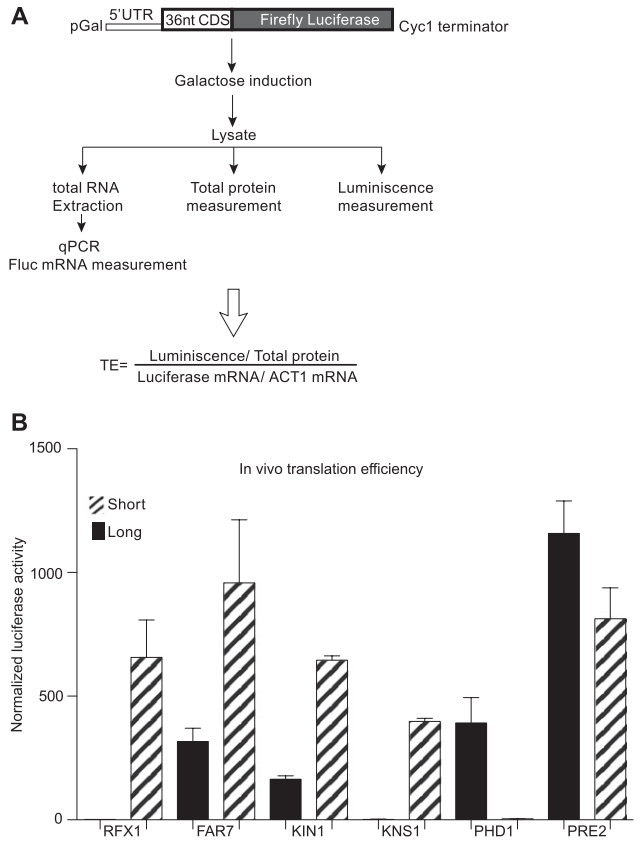


FIGURE 3. Alternative 5' UTR isoforms differ in translation efficiency in vivo. (A) Schematic of reporter constructs and translation assay. (B) Mean normalized luciferase activity from translation of alternative 5' UTR mRNAs. Error bars represent standard deviations.

increased activity of additional factors that mitigate the translational defects of this 5' UTR. Overall, these results show that changing the 5' ends of mRNAs can have dramatic consequences for protein synthesis.

Alternative 5' UTRs differ in their capacity for cap-independent translation

Cap-independent translation is thought to play crucial roles in cellular responses to a variety of stresses including nutrient deprivation, oxidative stress, apoptosis, and viral infection by facilitating continued synthesis of a subset of cellular proteins under conditions of reduced cap-binding activity (for review, see López-Lastra et al. 2005) (but, see also Gilbert 2010). To determine whether alternative 5' UTR variants could contribute differentially to cap-independent protein synthesis, uncapped mRNAs were synthesized and translated in vitro. Translation from all 5' UTRs was greatly reduced in the absence of an m⁷G cap (cf. Figs. 4B and 2D). Short unstructured 5' UTRs are thought to have reduced cap dependence; however, we observed little or no relationship between length/structure and activity (Fig. 4C,D). The “long” variant of *RFX1* showed almost no translation activity in either condition (discussed below).

Intriguingly, *PHD1* and *PRE2*, whose “long” and “short” 5' UTR variants had fairly similar capacities for cap-dependent translation in vitro (Fig. 2D), showed substantial isoform specificity in cap-independent translation (Fig. 4B). For both *PHD1* and *PRE2* the “long” 5' UTR isoform was preferentially translated without a cap (7.8 ± 1.6 - and 4.7 ± 1.1 -fold better than the “short” isoform, respectively). *PRE2* encodes an essential subunit of the 20S proteasome important for maintaining protein homeostasis in response to a variety of stresses. *PHD1* encodes a transcriptional activator required for starvation-induced invasive growth (Jin et al. 2008). Thus, both proteins are likely to be required under conditions of reduced cap-dependent initiation. Notably, both *PRE2* and *PHD1* continue to be translated in glucose-starved yeast, under conditions of globally reduced translation initiation (Arribere et al. 2011); the mRNA isoform(s) responsible for this translation were not determined.

Surprisingly small differences in 5' UTR sequences were sufficient to cause substantial effects on translation. *FAR7*, *KNS1*, *PHD1*, and *PRE2* “long” and “short” 5' UTR variants differ by fewer than 200 nt but showed greater than threefold translation activity differences under cap-dependent and/or cap-independent conditions (Figs. 2D, 4B). Previous estimates of the extent of mammalian gene regulation by

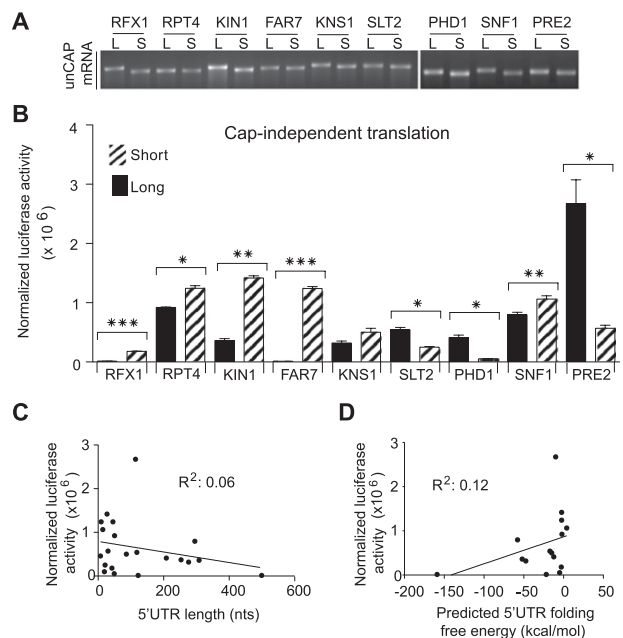


FIGURE 4. Alternative 5' UTR isoforms differ in capacity for cap-independent translation. (A) Representative ethidium bromide stained gels used to normalize luciferase activity to mRNA concentrations. (B) Mean normalized luciferase activity from translation of uncapped (5'-GTP) mRNAs. Error bars represent standard deviations. *P*-values were calculated by Student's *t*-test (two-tailed); (*) *P* < 0.05; (**) *P* < 0.01; (***) *P* < 0.001. (C) Comparison of mean translation activity to 5' UTR length and (D) predicted stability as determined by free-energy minimization using mfold (Zuker 2003).

alternative promoters (leading to alternative 5' UTRs) did not consider TSSs “alternative” if they were separated by fewer than 500 nt (Kimura et al. 2006; Gupta et al. 2011); small differences in TSS are frequently regarded as transcriptional noise. Our results demonstrate that seemingly minor changes in transcription—shifting the site of initiation by 50–200 bases—can have quantitatively significant effects on protein production. It remains to be determined whether such small differences in 5' UTR sequences can cause equally large translational effects in multicellular eukaryotes, but given the high level of conservation of translation factors it seems likely. If so, the functional diversity of mammalian mRNA isoforms is much greater than previously estimated.

“On/Off” switches generated by alternative transcription start site selection

What is the function of very poorly translated mRNAs such as the longer variant of *RFX1*? Notably, this poorly translated species makes up the majority of *RFX1* mRNA under standard growth conditions (~60%, Supplemental Fig. S1B). One possibility is that conditions exist where this mRNA isoform is relatively well translated. The long variant of *RFX1* contains multiple upstream open reading frames (uORFs), which typically reduce translation of downstream ORFs under normal growth conditions (Calvo et al. 2009). Under stress conditions leading to increased eIF2 α phosphorylation, global translation initiation is reduced, while certain uORF-containing transcripts are preferentially translated (Sonenberg and Hinnebusch 2009). Consistent with this possible mode of regulation, *RFX1* is implicated in cellular stress responses. However, mutation of all uAUGs did not improve translation of *RFX1.L* in vitro (data not shown).

Alternatively, the long 5' UTR isoform of *RFX1* may not function to produce protein at all. Transcribing through a promoter region can inhibit transcription initiation at downstream sites through a process known as transcription interference (for review, see Hainer and Martens 2011). The best-characterized example of this type of regulation in yeast involves production of a non-coding RNA (ncRNA), but it is the act of transcription rather than some activity of the ncRNA that leads to repression of the downstream gene (Martens et al. 2004). Our results show that “coding” RNAs are not always translated.

Thus, the potential for gene regulation by transcription interference is not limited to promoters that produce ncRNAs; transcription of a nontranslating mRNA isoform could function similarly to produce an “on/off” regulatory switch by blocking synthesis of a translatable mRNA (Fig. 5). This type of regulation was recently demonstrated for *tco1⁺* in *Schizosaccharomyces pombe* (Sehgal et al. 2008). Regulated TSS switching to produce shorter 5' UTR isoforms was observed in response to mating factor treatment, nitrogen starvation, and osmotic stress in *S. cerevisiae*. Twelve of the identified TSS-regulated genes were poorly translated under conditions where their long 5' UTR isoforms predominated, and they showed coordinated changes in transcript structure and translation efficiency following stress (Law et al. 2005).

The importance of alternative 3' UTR production for post-transcriptional regulation of gene expression is beginning to be appreciated through quantitative genome-wide studies (Mayr and Bartel 2009; Singh et al. 2009; Jan et al. 2011; Wu et al. 2011). Alternative 5' UTR production is equally widespread, but the effects of alternative TSS selection on protein production have not yet been investigated systematically. Our results suggest that for the majority of genes with alternative 5' UTR isoforms, protein synthesis could be regulated under certain conditions by altering transcription start site choice.

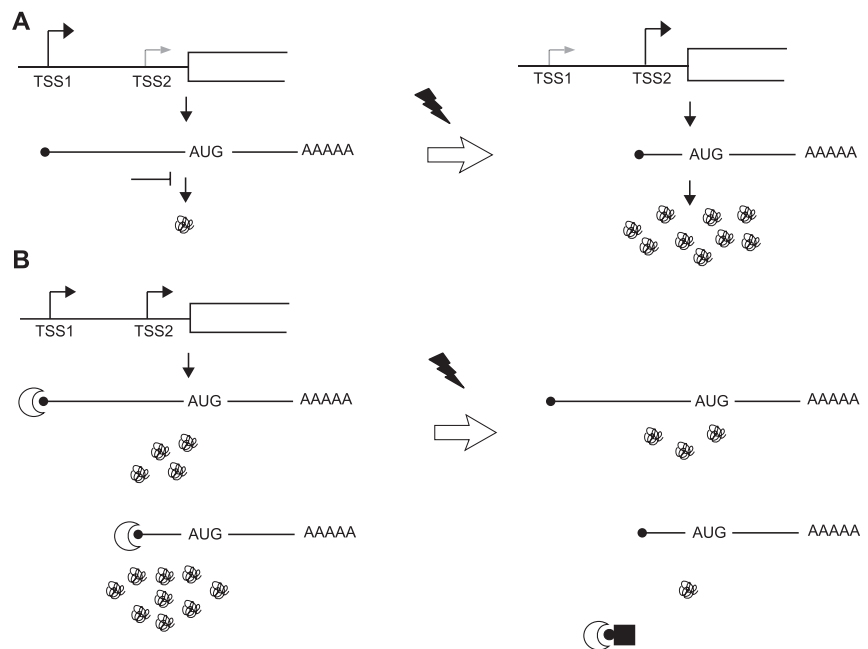


FIGURE 5. Model for the role of transcription start-site selection in translational control of gene expression. (A) Turning genes on and off by switching between poorly translated and efficiently translated 5' UTR isoforms. (B) Heterogeneous transcription start-site selection produces 5' UTR isoforms that respond differently to global changes in translation factor activity. In this example, the “long” 5' UTR variant is better translated than the corresponding “short” isoform under stress conditions, causing reduced cap-dependent initiation due to sequestration of eIF4E.

MATERIALS AND METHODS

Yeast strains and plasmids

S. cerevisiae of genotype MATa his3 trp1 ura3 L-A-o Gal⁺ were used to prepare translation extracts as described (Gilbert et al. 2007). To construct in vitro translation reporters, 5' UTR sequences were amplified from genomic DNA from *S. cerevisiae* strain Sigma 1278b with Phusion High-Fidelity DNA polymerase (NEB). Forward primers introduced a HindIII site followed by the T7 promoter (underlined): 5'-agcaAAGCTTTAATACGACTCACTATA GGG-gene-specific sequence. Reverse primers introduced a unique NcoI site downstream from the first 36 bases of ORF sequence: 5'-agcaCCATGG-gene specific sequence. HindIII–NcoI fragments were cloned in frame with Firefly luciferase. To construct yeast low-copy (LEU2, CEN ARS) plasmids expressing 5' UTR translation reporters in vivo, the same 5' UTR sequences were cloned as BglII–NcoI fragments downstream from the GAL1 promoter and in frame with Firefly luciferase, followed by the CYC1 3' UTR and terminator region. For full primer sequences see Supplemental Table S2. Plasmids are listed in Supplemental Table S3.

5' RACE

5' ends of mRNAs were determined by RNA ligase-mediated rapid amplification of cDNA ends (FirstChoice RLM-RACE, Ambion). cDNAs were synthesized from yeast (Sigma 1278b) RNA from pooled total and polysomal fractions from cells grown in rich media (1% Yeast extract, 2% Peptone, 0.01% Adenine hemisulfate, 2% Dextrose) at 30°C to OD600 1.0 and starved for glucose for 10, 20, or 30 min. The 5' ends mapped from these conditions were compared with published transcript boundaries determined from pooled cDNAs synthesized from yeast strain S288c grown in minimal medium and strain SK1 grown in acetate medium (Miura et al. 2006).

RNA synthesis and capping

RNA was synthesized with T7 by run-off transcription of linearized plasmid templates and purified on Zymo-Clean columns (Zymo Research). Capped RNA was prepared with recombinant Vaccinia capping enzyme (Shuman 1990). Concentrations of purified RNAs were determined by densitometry of ethidium bromide-stained bands compared with an RNA standard of known concentration on denaturing agarose gels using Quantity One software (BioRad).

Translation

Translation in vitro was initiated by adding m⁷G-capped or uncapped RNA template (2.5 nM and 10 nM final concentration, respectively) to micrococcal nuclease-treated yeast extract in Translation Buffer and incubated at 25°C for 40 min. Final concentrations in reactions: 22 mM Hepes-KOH (pH 7.4), 120 mM potassium acetate, 1.5 mM magnesium acetate, 0.75 mM ATP, 0.1 mM GTP, 0.04 mM each amino acid, 1.7 mM DTT, 25 mM creatine phosphate, 5 µg of creatine kinase, 10 U RNasin Plus (Promega), 2 mM phenylmethanesulfonyl fluoride, and 1× protease inhibitor cocktail (Roche). Reactions were stopped by transfer to ice and immediate addition of ice cold Passive Lysis Buffer (Promega). Luciferase activity was measured using the Luciferase Bright-Glo Assay System (Promega) on a Berthold Centro XS Luminometer.

To determine translation efficiency in vivo, wild-type cells transformed with pGAL_5'UTR_FLuc plasmids were grown to log phase in medium containing 2% galactose. Whole-cell lysates were prepared by vortexing with glass beads in 1× PBS with protease inhibitors (2 mM phenylmethanesulfonyl fluoride, and 1× protease inhibitor cocktail (Roche)). Luciferase activity was measured using the Luciferase Bright-Glo Assay System (Promega) on a Berthold Centro XS Luminometer. Fluc mRNA levels were determined by qRT-PCR on total cellular RNA isolated from whole-cell lysates by hot phenol extraction and normalized to *ACT1* mRNA levels.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

We thank the members of the Gilbert lab for helpful discussions and Mary Kay Thompson, Boris Zinshteyn, Kristen Bartoli, and Chris Burge for critical reading of the manuscript. Stewart Shuman provided Vaccinia capping enzyme plasmids. Sarah Walker and Jon Lorsch gave advice on optimizing RNA capping. This work was funded by NIH grant GM081399 to W.G.

Received August 6, 2012; accepted September 17, 2012.

REFERENCES

- Algire MA, Maag D, Savio P, Acker MG, Tarun SZ, Sachs AB, Asano K, Nielsen KH, Olsen DS, Phan L, et al. 2002. Development and characterization of a reconstituted yeast translation initiation system. *RNA* **8**: 382–397.
- Arribere JA, Doudna JA, Gilbert WV. 2011. Reconsidering movement of eukaryotic mRNAs between polysomes and P-bodies. *Mol Cell* **44**: 745–758.
- Berthelot K, Muldoon M, Rajkowsch L, Hughes J, McCarthy JE. 2004. Dynamics and processivity of 40S ribosome scanning on mRNA in yeast. *Mol Microbiol* **51**: 987–1001.
- Beyer A, Hollunder J, Nasheuer HP, Wilhelm T. 2004. Post-transcriptional expression regulation in the yeast *Saccharomyces cerevisiae* on a genomic scale. *Mol Cell Proteomics* **3**: 1083–1092.
- Brockmann R, Beyer A, Heinisch JJ, Wilhelm T. 2007. Posttranscriptional expression regulation: What determines translation rates? *PLoS Comput Biol* **3**: e57. doi: 10.1371/journal.pcbi.0030057.
- Calvo SE, Pagliarini DJ, Mootha VK. 2009. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc Natl Acad Sci* **106**: 7507–7512.
- Carninci P, Sandelin A, Lenhard B, Katayama S, Shimokawa K, Ponjavic J, Semple CA, Taylor MS, Engström PG, Frith MC, et al. 2006. Genome-wide analysis of mammalian promoter architecture and evolution. *Nat Genet* **38**: 626–635.
- Filbin ME, Kieft JS. 2009. Toward a structural understanding of IRES RNA function. *Curr Opin Struct Biol* **19**: 267–276.
- Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS. 2003. Global analysis of protein expression in yeast. *Nature* **425**: 737–741.
- Gilbert WV. 2010. Alternative ways to think about cellular internal ribosome entry. *J Biol Chem* **285**: 29033–29038.
- Gilbert WV, Zhou K, Butler TK, Doudna JA. 2007. Cap-independent translation is required for starvation-induced differentiation in yeast. *Science* **317**: 1224–1227.
- Gupta R, Bhattacharyya A, Agosto-Perez FJ, Wickramasinghe P, Davuluri RV. 2011. MPromDb update 2010: An integrated resource

- for annotation and visualization of mammalian gene promoters and CHIP-seq experimental data. *Nucleic Acids Res* **39**: D92–D97.
- Hainer SJ, Martens JA. 2011. Transcription of ncDNA: Many roads lead to local gene regulation. *Transcription* **2**: 120–123.
- Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS. 2009. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **324**: 218–223.
- Jackson RJ, Hellen CU, Pestova TV. 2010. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* **11**: 113–127.
- Jan CH, Friedman RC, Ruby JG, Bartel DP. 2011. Formation, regulation and evolution of *Caenorhabditis elegans* 3'UTRs. *Nature* **469**: 97–101.
- Jin R, Dobry CJ, McCown PJ, Kumar A. 2008. Large-scale analysis of yeast filamentous growth by systematic gene disruption and overexpression. *Mol Biol Cell* **19**: 284–296.
- Kimura K, Wakamatsu A, Suzuki Y, Ota T, Nishikawa T, Yamashita R, Yamamoto J, Sekine M, Tsuritani K, Wakaguri H, et al. 2006. Diversification of transcriptional modulation: Large-scale identification and characterization of putative alternative promoters of human genes. *Genome Res* **16**: 55–65.
- Kolkman A, Daran-Lapujade P, Fullaondo A, Olsthoorn MM, Pronk JT, Slijper M, Heck AJ. 2006. Proteome analysis of yeast response to various nutrient limitations. *Mol Syst Biol* **2**: 2006.0026. doi: 10.1038/msb4100069.
- Law GL, Bickel KS, MacKay VL, Morris DR. 2005. The under-translated transcriptome reveals widespread translational silencing by alternative 5' transcript leaders. *Genome Biol* **6**: R111. doi: 10.1186/gb-2005-6-13-r111.
- López-Lastra M, Rivas A, Barría MI. 2005. Protein synthesis in eukaryotes: The growing biological relevance of cap-independent translation initiation. *Biol Res* **38**: 121–146.
- Martens JA, Laprade L, Winston F. 2004. Intergenic transcription is required to repress the *Saccharomyces cerevisiae* *SER3* gene. *Nature* **429**: 571–574.
- Mayr C, Bartel DP. 2009. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138**: 673–684.
- Miura F, Kawaguchi N, Sese J, Toyoda A, Hattori M, Morishita S, Ito T. 2006. A large-scale full-length cDNA analysis to explore the budding yeast transcriptome. *Proc Natl Acad Sci* **103**: 17846–17851.
- Pal S, Gupta R, Kim H, Wickramasinghe P, Baubet V, Showe LC, Dahmane N, Davuluri RV. 2011. Alternative transcription exceeds alternative splicing in generating the transcriptome diversity of cerebellar development. *Genome Res* **21**: 1260–1272.
- Pisareva VP, Pisarev AV, Komar AA, Hellen CU, Pestova TV. 2008. Translation initiation on mammalian mRNAs with structured 5'UTRs requires DExH-box protein DHX29. *Cell* **135**: 1237–1250.
- Resch AM, Ogurtsov AY, Rogozin IB, Shabalina SA, Koonin EV. 2009. Evolution of alternative and constitutive regions of mammalian 5'UTRs. *BMC Genomics* **10**: 162. doi: 10.1186/1471-2164-10-162.
- Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M. 2011. Global quantification of mammalian gene expression control. *Nature* **473**: 337–342.
- Sehgal A, Hughes BT, Espenshade PJ. 2008. Oxygen-dependent, alternative promoter controls translation of *tc1⁺* in fission yeast. *Nucleic Acids Res* **36**: 2024–2031.
- Shuman S. 1990. Catalytic activity of vaccinia mRNA capping enzyme subunits coexpressed in *Escherichia coli*. *J Biol Chem* **265**: 11960–11966.
- Singh P, Alley TL, Wright SM, Kamdar S, Schott W, Wilpan RY, Mills KD, Graber JH. 2009. Global changes in processing of mRNA 3' untranslated regions characterize clinically distinct cancer subtypes. *Cancer Res* **69**: 9422–9430.
- Sonenberg N, Hinnebusch AG. 2009. Regulation of translation initiation in eukaryotes: Mechanisms and biological targets. *Cell* **136**: 731–745.
- Wang ET, Sandberg R, Luo S, Khrebtkova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**: 470–476.
- Wu X, Liu M, Downie B, Liang C, Ji G, Li QQ, Hunt AG. 2011. Genome-wide landscape of polyadenylation in *Arabidopsis* provides evidence for extensive alternative polyadenylation. *Proc Natl Acad Sci* **108**: 12533–12538.
- Yamashita R, Sathira NP, Kanai A, Tanimoto K, Arauchi T, Tanaka Y, Hashimoto S, Sugano S, Nakai K, Suzuki Y. 2011. Genome-wide characterization of transcriptional start sites in humans by integrative transcriptome analysis. *Genome Res* **21**: 775–789.
- Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**: 3406–3415.