A Multi-Level Approach to Understanding the Regulation of Translation Initiation

by

Heather R. Keys

B.S. Biology
University of New Mexico

SUBMITTED TO THE DEPARTMENT OF BIOLOGY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN BIOLOGY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

SEPTEMBER 2016

Signature of Author: Signature redacted

Department of Biology
June 30, 2016

Certified by: Signature redacted

David M. Sabatini
Professor of Biology
Member, Whitehead Institute
Thesis Advisor

Accepted by: Signature redacted

Michael T. Hemann
Associate Professor of Biology
Co-Chair, Biology Graduate Committee

A Multi-Level Approach to Understanding the Regulation of Translation Initiation

by

Heather R. Keys

Submitted to the Department of Biology on 30 June 2016 in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biology

ABSTRACT

mRNA translation is an extremely complex process required for life. Translation consumes vast amounts of cellular resources, and organisms have evolved tight regulatory mechanisms to control this process, which are often deregulated in cancer and other disease states. Initiation, as the rate-limiting step in translation, is particularly well regulated. Two kinase pathways that respond to cellular stresses, the GCN2 and mTORC1 pathways, sense amino acid insufficiency to inhibit translation initiation at distinct points. GCN2 is activated in response to amino acid deprivation and inhibits formation of the ternary complex, comprising eIF2, GTP, and the initiator methionyl-tRNA, which is required for recognition of the start codon. Although translation of most mRNAs is greatly suppressed when GCN2 is activated, mRNAs with certain *cis* elements escape inhibition. In contrast, the mTORC1 pathway is inhibited by the lack of amino acids, which ultimately results in the disruption of eIF4F, a multiprotein initiation factor complex that coordinates the recruitment of the small ribosomal subunit to the 5' end of mRNA. Like a decrease in the amount of ternary complex, disruption of eIF4F also suppresses translation of most mRNAs; however, the translation of a subset of mRNAs harboring a 5'TOP motif is even more dramatically reduced when mTORC1 is inhibited. Here we describe the translational program downstream of amino acid insufficiency, and present evidence of a novel uORF in murine ATF4 whose ribosome occupancy is regulated by the presence of amino acids. We identify the 4EBPs as the mTORC1 substrates that mediate the major effects of mTORC1 inhibition on translation of mRNAs both globally and on 5'TOP mRNAs specifically. Although we cannot mechanistically explain the dependence of 5'TOP mRNA translation on mTORC1 activity, we uncover a surprising role of the cap-proximal sequence in eIF4E recruitment. We systematically assess how the juxtacap sequence modulates eIF4E binding and translation, and present a model whereby the juxtacap sequence dictates the cap-proximal RNA secondary structure in an mRNA-context-dependent manner.

Thesis Advisor: David M. Sabatini
Professor of Biology
Member, Whitehead Institute

ACKNOWLEDGEMENTS

First and foremost I want to thank my advisor, David Sabatini. Getting to this point in my career has been a long and difficult road, with many twists and turns, and he has been there for me throughout it all. Our relationship has evolved tremendously throughout the years. When I first joined the lab in 2007, he encouraged my abundant enthusiasm, always took the time to discuss my (in retrospect incredibly naïve and simple) ideas, and somehow managed to deal with my superhuman stubbornness. He has supported me through so many failures that I've lost count. He has allowed me the freedom to follow my interests, wherever they lead. And he has taught me how to be a true scientist. David, I can never thank you enough.

I would also like to thank my thesis and defense committees. A big thank you goes out to Sean Ryder for taking the time to be a part of my thesis defense committee. Sean, I really appreciate it. In the early days, Iain Cheeseman served on my committee, and although the focus of my work changed and I absolved him of his thesis committee responsibilities, he has still given me some great advice about life and my career. Thank you so much, Iain. And a huge thank you goes out to Wendy Gilbert and Mike Hemann. I find you both so incredibly inspiring, both scientifically and in life. Thank you both so much for being such great role models, and for guiding me throughout my time at MIT.

During my many years in the lab, I spent most of my waking hours with two of the most amazing baymates anyone could ask for. Peggy was there for the beginning. The late nights, the silliness, trying to navigate our way through science and figure out who we were as people. She welcomed me into her space, always had time to talk (often against her wishes), and she was a great example of what a young scientist should aspire to. Her curiosity for all things, big and small, from science to Michael Jackson to what I had for lunch, has always been a source of inspiration, and continues to be to this day. Peggy, thank you for all the years.

When Peggy left, bench space was scarce, people were vying for prime real estate, and I wanted to get ahead of the curve. So, without knowing him all that well, I made a judgment call and lobbied for and recruited Larry to join the bay. It is one of the best decisions I've ever made. As all baymates do, Larry and I have had our ups and downs. I've said it before: you get to choose your friends, but you're sort of stuck with family...and baymates are like family. I was fortunate enough to also be stuck with a friend. Larry is a tremendous scientist; it never ceases to amaze me how quickly he gets to the heart of an idea, and then immediately jumps several steps beyond it. This is a characteristic I've seen in all great scientists, and I can only hope that being exposed to it 24/7 has rubbed off on me a bit. Larry is also a fantastic human being. He literally kept offering to help so often that he wore down my stubborn refusal to accept help from anyone. He has also brought me baked goods on many occasions. This sounds trivial, but one should not discount the ability of a surprise baked good to turn a bad day into something better. Larry, thanks for your awesomeness.

I would also like to single out a few former and current labmates who have all had an impact on me at different times over the years. Yasemin, Tim, and Shomit – you embodied the spirit of the "independent grad student" that has defined the way I approach science. Dave G. and Jan, do you guys remember when we used to go out dancing? I do. Doug, you and I had some pretty late/fun nights together in lab, and some pretty deep life conversations. Brian, Naama, Alejo, Nora, Maria, Sam, and Edie, you guys have all been great friends. Thanks for that.

I also want to single out Carson for a huge "thank you." He almost single-handedly resurrected my interest in science at a time when I was the most discouraged, and he has been a wonderful friend and collaborator ever since. He is one of the most thoughtful, creative, and amazing people I know, and I am extremely fortunate to know him. Carson, thank you so much.

I have made so many great friendships during my PhD, without which I would never have made it to this point. I was really fortunate to be a part of a wonderful class, but in particular I want to mention Corbin, Luke, LL, Justin, and Michelle – I miss you guys all the time. I also want to mention Andy, Daisy, Scarborough, and Mike G. – you guys and the Muddy have given me a great perspective on life, and I'm so glad we're friends.

I would literally not be at MIT if it weren't for Alex (I would probably be at UCSF or Berkeley). He knows all my flaws as a person, and yet has always supported and encouraged me, and been proud of me, and been there for me when life has thrown curves. I am lucky to have several people in my life worthy of the title of "fantastic human being," and he is one of them. Alex, thank you so much for being my friend.

I also want to thank my two best friends for their support throughout the entire PhD process. I have known Jenny since our first day of Biology Lab in undergrad in New Mexico, and she has always encouraged me in whatever endeavor I have chosen to do. She has consistently reminded me of what's important in life, and she has been a source of inspiration to me time and again. She has also been my adventure buddy on more than one occasion, although perhaps I shouldn't thank her for that, since one of us usually gets injured. Jenny, thanks for being you! The other person I want to thank wholeheartedly is Nina. She has been my Massachusetts partner-in-crime for more years than I care to think about, and she has been a true friend throughout. She has always been accepting of me, no matter what, and she deals with my stubbornness in a saint-like manner. Nina, I appreciate you. Thanks for always being there for me when I need it most.

Finally, I want to thank my father. He has been my "best" best friend for many years now, and has tried to offer guidance and advice (unsuccessfully), tried to entice me to move to small-town Illinois (also unsuccessfully), listened to my worries and struggles, made me martinis...and above all has been an overall great source of inspiration and unwavering support throughout my life. He taught me that everyone has flaws, that you should approach life with a sense of humor and lightheartedness, and that you should not back down from what you believe in. I have known throughout this process that he would love me no matter what, and that I'd always have a place to call home. Dad, this one's for you.

TABLE OF CONTENTS

Chapter 1

Introduction

Heather R. Keys[1-2]

[1]Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA

[2]Department of Biology, MIT, Cambridge, MA 02139, USA

Importance of Translation

mRNA translation is a highly complex and fundamental process necessary for organismal development and cellular function. Translation consumes a vast amount of cellular energy not only during the initiation and elongation of each individual peptide chain, but also to create the machinery necessary for protein synthesis. Translation in eukaryotes involves multiple types of RNA, as well as numerous proteins (Figure 1). The eukaryotic ribosome contains around 80 proteins (Nakao, Yoshihama, & Kenmochi, 2004), with over 200 proteins involved in ribosome assembly (Henras et al., 2008). There are over a dozen translation initiation factors, several of which are comprised of multiple subunits, as well as a handful of elongation and termination factors. In yeast, transcription of ribosomal RNA (rRNA) accounts for approximately 60% of the transcriptional activity of the cell, with ribosomes comprising about 30% of the cytoplasmic volume (Warner, 1999). Furthermore, around 30% of the mRNAs in a yeast cell encode ribosomal proteins (Warner, 1999). In rapidly dividing mammalian cells, rRNA comprises about 80% of the total RNA, with tRNA accounting for another 15% (Lodish, 2000), and a secretory cell may use up to half of its energy for translation (Pannevis & Houlihan, 1992). Regardless of the precise details, each cell devotes a significant fraction of its resources to the process of translation, and organisms have unsurprisingly evolved several mechanisms to tightly regulate this fundamental activity.

Translation Initiation in Eukaryotes

Translation occurs in three stages: initiation, elongation, and termination. Initiation is typically the rate-limiting step in mammals, and we will focus on initiation here (Golini et al., 1976; Walden, Godefroy-Colburn, & Thach, 1981). During initiation, a ternary complex comprised of the initiation factor eIF2, GTP, and the initiator methionine (Met-$tRNA_i^{Met}$) forms. The ternary complex is recruited to the 40S small ribosomal subunit in a process involving initiation factors eIF1, eIF1A, eIF3, and eIF5 to form the 43S pre-initiation complex (43S PIC) (Benne & Hershey, 1978; Chaudhuri, Chowdhury, & Maitra, 1999; Jackson, Hellen, & Pestova, 2010; Majumdar, Bandyopadhyay, & Maitra, 2003; Peterson, Safer, & Merrick, 1979; Thomas, Goumans, Voorma, & Benne, 1980; Thomas, Spaan, van Steeg, Voorma, & Benne, 1980; Trachsel, Erni, Schreier, & Staehelin, 1977). Interaction with the initiation factor complex eIF4F, which recognizes the 5' 7-methylguanosine cap of mRNA, recruits the 43S PIC to the 5' end of an mRNA.

In the canonical scanning model of translation initiation, after attachment to the mRNA, the small ribosomal subunit with associated initiation factors (48S initiation complex; 48S IC) scans the 5' UTR until it encounters the first AUG codon (Kozak, 1978). A component of eIF4F, eIF4A, possesses helicase activity and unwinds secondary structure in the 5' UTR of mRNAs to facilitate 43S PIC binding and the scanning process (Pestova & Kolupaeva, 2002; Rogers, Richter, & Merrick, 1999; Svitkin et al., 2001). After recruitment of the 43S PIC to the mRNA, eIF5 acts as a GAP for eIF2 to promote hydrolysis of the GTP in the ternary complex, although the

committed step of this reaction, inorganic phosphate (P_i) release, does not occur immediately (Algire, Maag, & Lorsch, 2005).

Once the 48S IC recognizes the initiation codon, primarily through its interaction with the Met-$tRNA_i^{Met}$, but with some contribution by the sequence context surrounding the start codon (Kozak sequence), P_i is released and the 48S IC is committed to assembly of the ribosomal complex at the chosen site (Algire et al., 2005). Subsequently, the initiation factors associated with the 48S IC are displaced in a step involving eIF5B, and the 60S large ribosomal subunit joins (Pestova et al., 2000). Finally, eIF5B must hydrolyze GTP to dissociate from the 80S ribosome to complete translation initiation (Antoun, Pavlov, Andersson, Tenson, & Ehrenberg, 2003; Pestova et al., 2000). Although control can occur at other stages in translation, as the rate-limiting step, initiation is regulated particularly well.

Regulation of Translation Initiation in Eukaryotes

Formation of the ternary complex is one major target of translational regulation. The protein component of the ternary complex, eIF2, is a GTPase comprised of three subunits: alpha, beta, and delta. The nucleotide-bound state of eIF2 is regulated by both a GAP, eIF5, and a GEF, eIF2B. In its GDP-bound form, eIF2 cannot participate in the ternary complex, and thus eIF2B facilitates ternary complex formation (Wortham & Proud, 2015). In response to various cellular stresses, the alpha subunit of eIF2 is phosphorylated at serine 51, which inhibits the GEF activity of eIF2B and ternary complex formation (Kimball, 1999).

Attenuation of initiation via eIF2α phosphorylation has both general and specific consequences. A reduction in ternary complex unsurprisingly attenuates translation initiation on most mRNAs. Somewhat unexpectedly, however, translation of particular mRNAs is increased when ternary complex formation is inhibited (Dever et al., 1992; Harding et al., 2000). These mRNAs contain short upstream open reading frames (uORFs), which typically inhibit translation of the downstream protein-coding ORF (Wethmar, 2014). When the availability of ternary complex is decreased, ribosomes can reinitiate translation at the downstream coding ORF, resulting in upregulated translation of these messages under conditions of stress, as is the case for ATF4 (Lu, Harding, & Ron, 2004; Vattem & Wek, 2004).

Translation initiation is also regulated at the point of recruitment of the 43S PIC to the mRNA. The 43S PIC is recruited to mRNA through the interaction of eIF3, a component of the 43S PIC, with eIF4G, a component of eIF4F. The assembly of eIF4F is regulated by the mechanistic target of rapamycin complex 1 (mTORC1), and will be discussed below.

Figure 1: Overview of Eukaryotic Translation Initiation.

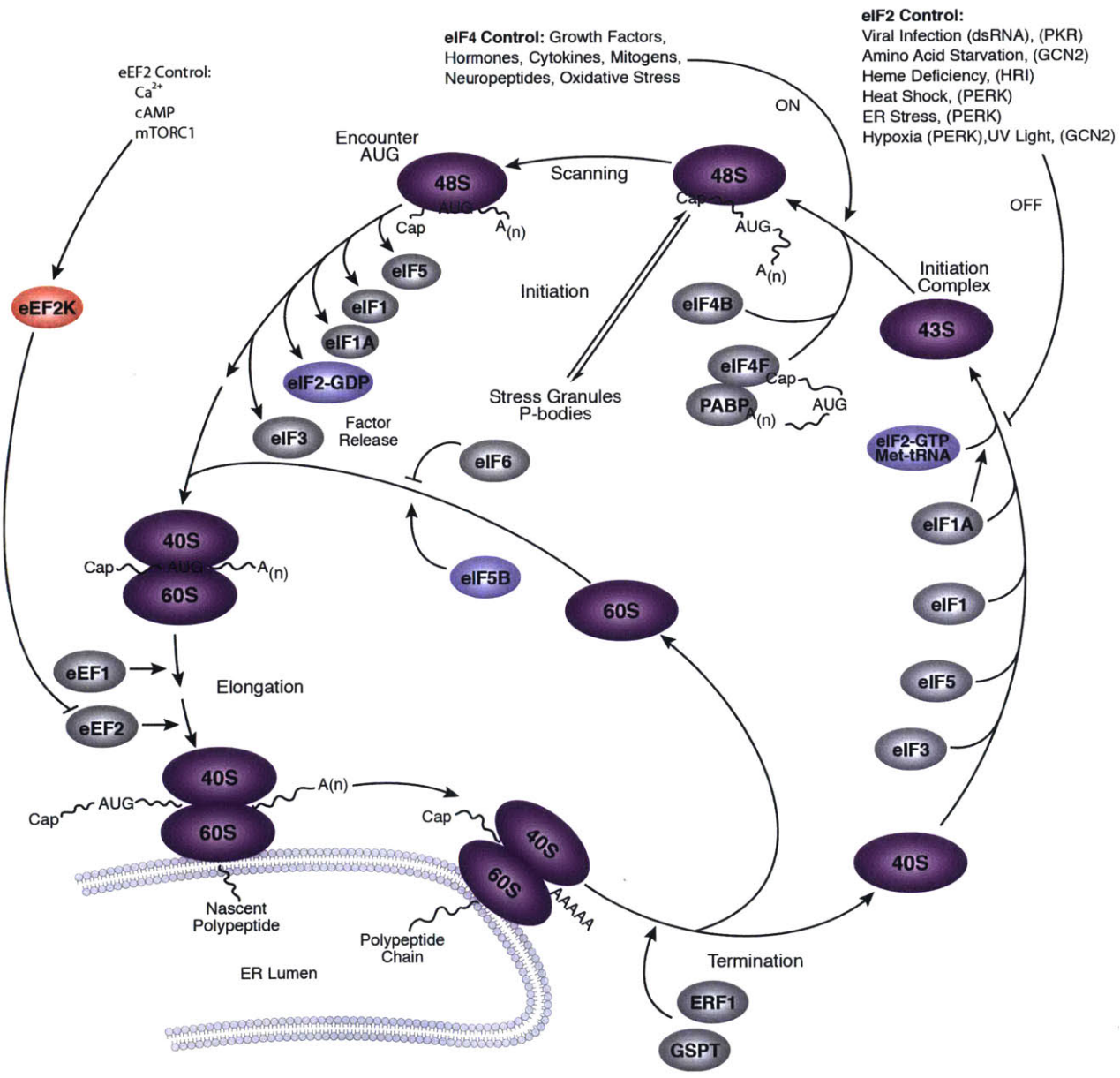


Figure 1: Overview of eukaryotic translation.

Ribosomal subunits are shown in magenta, GTPases in purple, kinases in peach, and all other proteins in gray. Major points of control and the proteins involved are indicated. Illustration adapted courtesy of Cell Signaling Technology, Inc; www.cellsignal.com.

mTORC1 Overview

mTORC1 is a highly conserved multiprotein complex that is important for regulating a number of energetically costly processes that lead to cell growth and proliferation. An intricate network of proteins exists that responds to numerous extracellular and intracellular signals to coordinately regulate mTORC1. The mTORC1 pathway integrates organismal and cellular nutritional cues such as growth factors, amino acids, glucose, oxygen availability, cellular energy levels, and cellular stress signals to promote mRNA translation, ribosome and mitochondrial biogenesis, nucleotide and lipid biosynthesis and to suppress autophagy (Laplante & Sabatini, 2012) (Figure 2).

Figure 2: Overview of mTORC1 Signaling

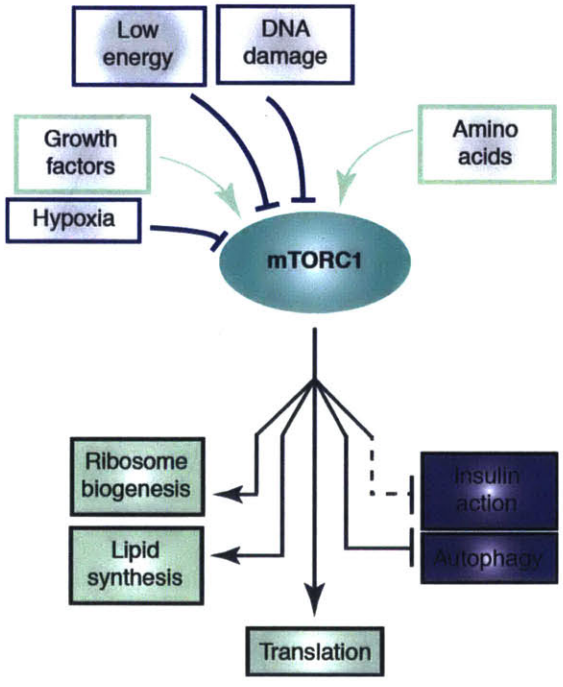


Figure 2: Overview of mTORC1 signaling

Shown are the general cellular processes regulated by mTORC1, and the signals mTORC1 integrates. Adapted from (Efeyan, Zoncu, & Sabatini, 2012).

In the simplest sense, mTORC1 is a coincidence detector regulated by two sets of GTPases. The small GTPase Rheb (Ras homolog enriched in brain) resides at the late endosomal and lysosomal membranes, and its activity is regulated in response to growth factor signals through inhibition of its GTPase activating protein (GAP), TSC (Inoki, Li, Xu, & Guan, 2003; Li, Inoki, & Guan, 2004; Sancak et al., 2008; Tee, Manning, Roux, Cantley, & Blenis, 2003; Zhang et al., 2003). The heterodimeric Rag GTPase complex (comprised of RagA or B with RagC or D) also resides at the lysosome, and its activity is regulated in response to lysosomal and cytoplasmic amino acid signals through several mechanisms that converge on Ragulator, its guanine nucleotide exchange factor (GEF), and GATOR1, its GAP (Bar-Peled et al., 2013; Bar-Peled, Schweitzer, Zoncu, & Sabatini, 2012; Chantranupong et al., 2016; E. Kim, Goraksha-Hicks, Li, Neufeld, & Guan, 2008; Sancak et al., 2010; Sancak et al., 2008; Saxton et al., 2016; S. Wang et al., 2015; Wolfson et al., 2016). In the presence of sufficient amino acids, the Rag complex recruits mTORC1 to the lysosomal surface, where Rheb activates it in the presence of sufficient growth factor signals.

mTORC1 in Cancer

mTOR is a member of the phosphatidylinositol-3-kinase-related kinase (PIKK) family of protein kinases, and is itself a part of the oncogenic PI3K signaling pathway. Activating mutations in PI3K, as well as amplifications of Akt, a downstream effector of PI3K that positively regulates mTORC1, are observed in many cancers; additionally, the tumor suppressor PTEN, a lipid phosphatase that counteracts the lipid kinase activity of

PI3K, is frequently mutated (Lim, Crowe, & Yang, 2015). The mTORC1 pathway is activated in a variety of cancers and overgrowth syndromes, such as Tuberous Sclerosis, Birt-Hogg-Dube, Peutz-Jeghers, Cowden, and Neurofibromatosis (Efeyan et al., 2012) (Figure 3). Activating mutations have also been identified in mTOR itself (Grabiner et al., 2014). Analogs of the small molecule rapamycin, a partial inhibitor of some functions of mTORC1, are currently under investigation for treating a variety of cancers, and are currently in use in some cases (Zaytseva, Valentino, Gulhati, & Evers, 2012). It remains an open question which downstream targets of mTORC1 are most relevant in cancer.

Figure 3: Selected Syndromes and Cancers Related to the mTORC1 Pathway

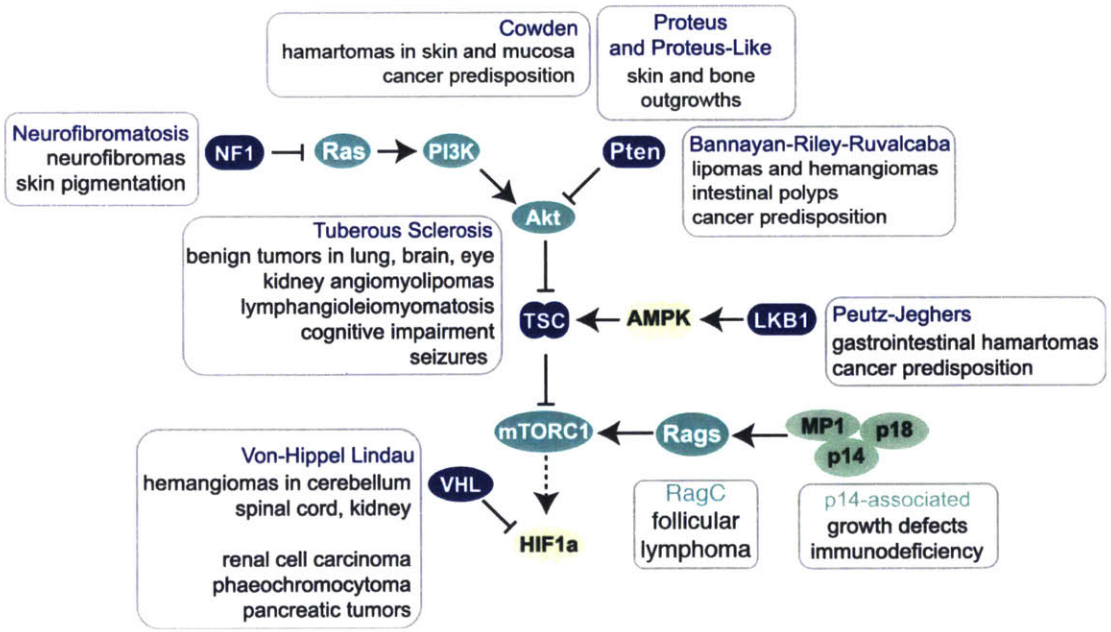


Figure 3: Selected syndromes and cancers related to the mTORC1 pathway.

NF1, Pten, TSC, LKB1, and VHL are all tumor suppressors. Activating mutations in Ras and PI3K frequently occur in various cancers, as does Akt overexpression. A mutation was recently identified in RagC that activates mTOR in follicular lymphoma (Okosun et al., 2016). Adapted from (Zoncu, Efeyan, & Sabatini, 2011).

Downstream of mTORC1

The precise control of mTORC1 is so critical and is so often deregulated, because mTORC1, through phosphorylation of myriad substrates, promotes a number of anabolic processes and inhibits a number of catabolic processes that ultimately lead to cell growth. mTORC1 negatively regulates autophagy, a pro-survival catabolic process that degrades cellular components under conditions of stress, through phosphorylation of its substrate Ulk1 (Y. C. Kim & Guan, 2015). mTORC1 was also linked to regulation of SREBP and CAD to regulate lipid and nucleotide synthesis, and YY1 and PGC1α to regulate mitochondrial biogenesis (Laplante & Sabatini, 2012). Ribosome biogenesis is regulated through Maf1 phosphorylation and activation of TIF1A, an RNA PolI-associated initiation factor (Iadevaia, Liu, & Proud, 2014). mTORC1 also regulates mRNA translation through direct and indirect phosphorylation of several substrates, including ribosomal protein S6 kinases (S6Ks), eIF4E binding proteins (4EBPs), Eef2 kinase (Eef2K), and the initiation factors eIF4B and eIF4G (Laplante & Sabatini, 2012).

mTORC1 in Translation

S6Ks

mTORC1 phosphorylates a variety of substrates with roles in translation; the S6Ks are among the best characterized of these. The phosphorylation of S6K1 and S6K2 is potently induced by insulin and growth factors, and their phosphorylation is completely inhibited by acute rapamcyin treatment, unlike that of some other mTORC1 substrates (Dufner & Thomas, 1999; Jeno, Ballou, Novak-Hofer, & Thomas, 1988; Thoreen et al., 2009). S6K1 and S6K2 mediate the effects of mTORC1 on translation and cell growth through phosphorylation of myriad substrates, such as ribosomal protein S6 (rpS6), eIF4B, PDCD4, eEF2K, CBP80, and SKAR (Dorrello et al., 2006; Ferrari, Bandi, Hofsteenge, Bussian, & Thomas, 1991; Jastrzebski, Hannan, Tchoubrieva, Hannan, & Pearson, 2007; Krieg, Hofsteenge, & Thomas, 1988; Raught et al., 2004; Richardson et al., 2004; X. Wang et al., 2001; Wilson, Wu, & Cerione, 2000). S6K1 also regulates ribosome biogenesis (Jastrzebski et al., 2007). Although S6Ks are clearly important regulators of translation, other mTORC1 substrates play a large role as well.

4EBPs

4EBP1, 2, and 3 are part of a family of proteins that disrupt the eIF4F complex (described below), which is important for cap-dependent translation. Hypophosphorylated 4EBPs bind the initiation factor eIF4E and inhibit its interaction with eIF4G, resulting in a severe suppression of translation (Haghighat, Mader, Pause, & Sonenberg, 1995; Marcotrigiano, Gingras, Sonenberg, & Burley, 1999). When

mTORC1 is active, it phosphorylates the 4EBPs at multiple sites, rendering them inactive for eIF4E binding (Marcotrigiano et al., 1999; Richter & Sonenberg, 2005). Mechanistically, this occurs via a disorder-to-order transition of 4EBP upon phosphorylation; phosphorylation stabilizes free 4EBP (Peter et al., 2015; Tait et al., 2010). The ratio of 4EBPs to eIF4E is under homeostatic control and has been linked to the response of cells to mTOR inhibitors (Bah et al., 2015). Thus, the 4EBPs are important translational regulators downstream of mTORC1.

eIF4F

The recruitment of the 43S PIC to the mRNA is one of two main points at which translation initiation is regulated, and the eIF4F complex is the target of this regulatory mechanism. eIF4F comprises three core members: eIF4E, eIF4G, and eIF4A (Merrick, 2015). eIF4E binds the 5' ^{me7}G cap to localize eIF4F to the 5' end of mRNA (Sonenberg, Morgan, Merrick, & Shatkin, 1978). eIF4G is a large scaffolding molecule that facilitates a number of interactions important for 43S PIC recruitment and translation: it links the 5' and 3' ends of mRNA via interactions with eIF4E and PABP; it binds eIF3, a component of the 43S PIC, to recruit the 43S PIC to the mRNA; it binds RNA directly, which stabilizes the interaction of eIF4E with the 5' cap; and it binds eIF4A, the third component of eIF4F (Prevot, Darlix, & Ohlmann, 2003). eIF4A is a helicase that resolves secondary structure in the 5'UTR of mRNAs, which promotes scanning of the small ribosomal subunit (Svitkin et al., 2001).

eIF4F assembly is disrupted by mTORC1 inhibition; specifically, the interaction between eIF4E and eIF4G is disrupted through competitive binding of hypophosphorylated 4EBPs to eIF4E, ultimately leading to decreased recruitment of the 43S PIC to mRNA and decreased translation initiation (Brunn et al., 1997; Marcotrigiano et al., 1999; Pause et al., 1994). Interestingly, the disruption of eIF4F by mTORC1 inhibition causes a dramatic and selective decrease in the translation of a subset of mRNAs that encode the main components of the translational apparatus, such as ribosomal proteins and initiation factors (Hsieh et al., 2012; Jefferies, Reinhard, Kozma, & Thomas, 1994; Thoreen et al., 2012). mRNAs in this class possess a variable length stretch of 5' cytidines and uridines, typically directly adjacent to the 5' cap. Hence, these mRNAs are called 5'TOP (Terminal OligoPyrimidine) mRNAs.

Translational Regulation of 5'TOP mRNAs

Following the discovery that mTORC1 regulates the translation of the translational machinery itself, much effort has been focused on characterizing the molecular mechanism of this process. One early model hypothesized that rpS6 mediates translational regulation of this class of mRNAs, in response to its phosphorylation by the S6Ks (Meyuhas, 2000). However, studies performed in knock-in mice whose rpS6 contained serine-to-alanine substitutions at all five phosphorylatable residues argue against this model (Ruvinsky et al., 2005). Other evidence also pointed to S6Ks as the mediator of 5'TOP mRNA translational regulation downstream of mTORC1, but the translation of these mRNAs in S6K1/2 double knockout MEFs and ES cells was still

regulated, suggesting that a different mTORC1 substrate is responsible (Pende et al., 2004). More recently, several groups have argued for a variety of protein factors, and even a microRNA, as regulators of 5'TOP mRNA translation (Biberman & Meyuhas, 1999; Damgaard & Lykke-Andersen, 2011; Kakegawa et al., 2007; Orom, Nielsen, & Lund, 2008; Tcherkezian et al., 2014), but it remains to be determined which, if any, of these factors is the primary determinant of mTORC1-mediated translational control of 5'TOP mRNAs.

Thesis Overview

The general focus of this thesis is the process of translation initiation. Specifically, in Chapter 2 we describe the translational programs downstream of a cellular stress known to regulate initiation, amino acid deprivation. We find that the short-term response to amino acid deprivation is largely translational, while the long-term response is primarily transcriptional, and we find that translational upregulation following acute amino acid deprivation is likely uORF-independent. In Chapter 3, we explore the mechanism of translational regulation of an important class of mRNAs that is regulated at the initiation step, 5'TOP mRNAs. We show that regulation of these mRNAs by acute mTOR inhibition is dependent on 4EBPs, and present evidence that eIF4E affinity differences for mRNAs do not explain their translational regulation. We also uncover a role of the cap-proximal sequence in the recruitment of the initiation factor eIF4E to mRNA. Finally, we systematically dissect the role of the juxtacap sequence in the recruitment of eIF4E to mRNA, its effect on translation, and its modulatory mechanism.

REFERENCES

Algire, M. A., Maag, D., & Lorsch, J. R. (2005). Pi release from eIF2, not GTP hydrolysis, is the step controlled by start-site selection during eukaryotic translation initiation. *Mol Cell, 20*(2), 251-262. doi: 10.1016/j.molcel.2005.09.008

Antoun, A., Pavlov, M. Y., Andersson, K., Tenson, T., & Ehrenberg, M. (2003). The roles of initiation factor 2 and guanosine triphosphate in initiation of protein synthesis. *EMBO J, 22*(20), 5593-5601. doi: 10.1093/emboj/cdg525

Bah, A., Vernon, R. M., Siddiqui, Z., Krzeminski, M., Muhandiram, R., Zhao, C., . . . Forman-Kay, J. D. (2015). Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature, 519*(7541), 106-109. doi: 10.1038/nature13999

Bar-Peled, L., Chantranupong, L., Cherniack, A. D., Chen, W. W., Ottina, K. A., Grabiner, B. C., . . . Sabatini, D. M. (2013). A Tumor suppressor complex with GAP activity for the Rag GTPases that signal amino acid sufficiency to mTORC1. *Science, 340*(6136), 1100-1106. doi: 10.1126/science.1232044

Bar-Peled, L., Schweitzer, L. D., Zoncu, R., & Sabatini, D. M. (2012). Ragulator is a GEF for the rag GTPases that signal amino acid levels to mTORC1. *Cell, 150*(6), 1196-1208. doi: 10.1016/j.cell.2012.07.032

Benne, R., & Hershey, J. W. (1978). The mechanism of action of protein synthesis initiation factors from rabbit reticulocytes. *J Biol Chem, 253*(9), 3078-3087.

Biberman, Y., & Meyuhas, O. (1999). TOP mRNAs are translationally inhibited by a titratable repressor in both wheat germ extract and reticulocyte lysate. *FEBS Lett, 456*(3), 357-360.

Brunn, G. J., Hudson, C. C., Sekulic, A., Williams, J. M., Hosoi, H., Houghton, P. J., . . . Abraham, R. T. (1997). Phosphorylation of the translational repressor PHAS-I by the mammalian target of rapamycin. *Science, 277*(5322), 99-101.

Chantranupong, L., Scaria, S. M., Saxton, R. A., Gygi, M. P., Shen, K., Wyant, G. A., . . . Sabatini, D. M. (2016). The CASTOR Proteins Are Arginine Sensors for the mTORC1 Pathway. *Cell, 165*(1), 153-164. doi: 10.1016/j.cell.2016.02.035

Chaudhuri, J., Chowdhury, D., & Maitra, U. (1999). Distinct functions of eukaryotic translation initiation factors eIF1A and eIF3 in the formation of the 40 S ribosomal preinitiation complex. *J Biol Chem, 274*(25), 17975-17980.

Damgaard, C. K., & Lykke-Andersen, J. (2011). Translational coregulation of 5'TOP mRNAs by TIA-1 and TIAR. *Genes Dev, 25*(19), 2057-2068. doi: 10.1101/gad.17355911

Dever, T. E., Feng, L., Wek, R. C., Cigan, A. M., Donahue, T. F., & Hinnebusch, A. G. (1992). Phosphorylation of initiation factor 2 alpha by protein kinase GCN2 mediates gene-specific translational control of GCN4 in yeast. *Cell, 68*(3), 585-596.

Dorrello, N. V., Peschiaroli, A., Guardavaccaro, D., Colburn, N. H., Sherman, N. E., & Pagano, M. (2006). S6K1- and betaTRCP-mediated degradation of PDCD4 promotes protein translation and cell growth. *Science, 314*(5798), 467-471. doi: 10.1126/science.1130276

Dufner, A., & Thomas, G. (1999). Ribosomal S6 kinase signaling and the control of translation. *Exp Cell Res, 253*(1), 100-109. doi: 10.1006/excr.1999.4683

Efeyan, A., Zoncu, R., & Sabatini, D. M. (2012). Amino acids and mTORC1: from lysosomes to disease. *Trends Mol Med, 18*(9), 524-533. doi: 10.1016/j.molmed.2012.05.007

Ferrari, S., Bandi, H. R., Hofsteenge, J., Bussian, B. M., & Thomas, G. (1991). Mitogen-activated 70K S6 kinase. Identification of in vitro 40 S ribosomal S6 phosphorylation sites. *J Biol Chem, 266*(33), 22770-22775.

Golini, F., Thach, S. S., Birge, C. H., Safer, B., Merrick, W. C., & Thach, R. E. (1976). Competition between cellular and viral mRNAs in vitro is regulated by a messenger discriminatory initiation factor. *Proc Natl Acad Sci U S A, 73*(9), 3040-3044.

Grabiner, B. C., Nardi, V., Birsoy, K., Possemato, R., Shen, K., Sinha, S., . . . Sabatini, D. M. (2014). A diverse array of cancer-associated MTOR mutations are hyperactivating and can predict rapamycin sensitivity. *Cancer Discov, 4*(5), 554-563. doi: 10.1158/2159-8290.CD-13-0929

Haghighat, A., Mader, S., Pause, A., & Sonenberg, N. (1995). Repression of cap-dependent translation by 4E-binding protein 1: competition with p220 for binding to eukaryotic initiation factor-4E. *EMBO J, 14*(22), 5701-5709.

Harding, H. P., Novoa, I., Zhang, Y., Zeng, H., Wek, R., Schapira, M., & Ron, D. (2000). Regulated translation initiation controls stress-induced gene expression in mammalian cells. *Mol Cell, 6*(5), 1099-1108.

Henras, A. K., Soudet, J., Gerus, M., Lebaron, S., Caizergues-Ferrer, M., Mougin, A., & Henry, Y. (2008). The post-transcriptional steps of eukaryotic ribosome biogenesis. *Cell Mol Life Sci, 65*(15), 2334-2359. doi: 10.1007/s00018-008-8027-0

Hsieh, A. C., Liu, Y., Edlind, M. P., Ingolia, N. T., Janes, M. R., Sher, A., . . . Ruggero, D. (2012). The translational landscape of mTOR signalling steers cancer initiation and metastasis. *Nature, 485*(7396), 55-61. doi: 10.1038/nature10912

Iadevaia, V., Liu, R., & Proud, C. G. (2014). mTORC1 signaling controls multiple steps in ribosome biogenesis. *Semin Cell Dev Biol, 36*, 113-120. doi: 10.1016/j.semcdb.2014.08.004

Inoki, K., Li, Y., Xu, T., & Guan, K. L. (2003). Rheb GTPase is a direct target of TSC2 GAP activity and regulates mTOR signaling. *Genes Dev, 17*(15), 1829-1834. doi: 10.1101/gad.1110003

Jackson, R. J., Hellen, C. U., & Pestova, T. V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol, 11*(2), 113-127. doi: 10.1038/nrm2838

Jastrzebski, K., Hannan, K. M., Tchoubrieva, E. B., Hannan, R. D., & Pearson, R. B. (2007). Coordinate regulation of ribosome biogenesis and function by the ribosomal protein S6 kinase, a key mediator of mTOR function. *Growth Factors, 25*(4), 209-226. doi: 10.1080/08977190701779101

Jefferies, H. B., Reinhard, C., Kozma, S. C., & Thomas, G. (1994). Rapamycin selectively represses translation of the "polypyrimidine tract" mRNA family. *Proc Natl Acad Sci U S A, 91*(10), 4441-4445.

Jeno, P., Ballou, L. M., Novak-Hofer, I., & Thomas, G. (1988). Identification and characterization of a mitogen-activated S6 kinase. *Proc Natl Acad Sci U S A, 85*(2), 406-410.

Kakegawa, T., Ohuchi, N., Hayakawa, A., Hirata, S., Matsuda, M., Kogure, K., . . . Kaspar, R. L. (2007). Identification of AUF1 as a rapamycin-responsive binding protein to the 5'-terminal oligopyrimidine element of mRNAs. *Arch Biochem Biophys, 465*(1), 274-281. doi: 10.1016/j.abb.2007.06.001

Kim, E., Goraksha-Hicks, P., Li, L., Neufeld, T. P., & Guan, K. L. (2008). Regulation of TORC1 by Rag GTPases in nutrient response. *Nat Cell Biol, 10*(8), 935-945. doi: 10.1038/ncb1753

Kim, Y. C., & Guan, K. L. (2015). mTOR: a pharmacologic target for autophagy regulation. *J Clin Invest, 125*(1), 25-32. doi: 10.1172/JCI73939

Kimball, S. R. (1999). Eukaryotic initiation factor eIF2. *Int J Biochem Cell Biol, 31*(1), 25-29.

Kozak, M. (1978). How do eucaryotic ribosomes select initiation regions in messenger RNA? *Cell, 15*(4), 1109-1123.

Krieg, J., Hofsteenge, J., & Thomas, G. (1988). Identification of the 40 S ribosomal protein S6 phosphorylation sites induced by cycloheximide. *J Biol Chem, 263*(23), 11473-11477.

Laplante, M., & Sabatini, D. M. (2012). mTOR signaling in growth control and disease. *Cell, 149*(2), 274-293. doi: 10.1016/j.cell.2012.03.017

Li, Y., Inoki, K., & Guan, K. L. (2004). Biochemical and functional characterizations of small GTPase Rheb and TSC2 GAP activity. *Mol Cell Biol, 24*(18), 7965-7975. doi: 10.1128/MCB.24.18.7965-7975.2004

Lim, H. J., Crowe, P., & Yang, J. L. (2015). Current clinical regulation of PI3K/PTEN/Akt/mTOR signalling in treatment of human cancer. *J Cancer Res Clin Oncol, 141*(4), 671-689. doi: 10.1007/s00432-014-1803-3

Lodish, H. F. (2000). *Molecular cell biology* (4th ed.). New York: W.H. Freeman.

Lu, P. D., Harding, H. P., & Ron, D. (2004). Translation reinitiation at alternative open reading frames regulates gene expression in an integrated stress response. *J Cell Biol, 167*(1), 27-33. doi: 10.1083/jcb.200408003

Majumdar, R., Bandyopadhyay, A., & Maitra, U. (2003). Mammalian translation initiation factor eIF1 functions with eIF1A and eIF3 in the formation of a stable 40 S preinitiation complex. *J Biol Chem, 278*(8), 6580-6587. doi: 10.1074/jbc.M210357200

Marcotrigiano, J., Gingras, A. C., Sonenberg, N., & Burley, S. K. (1999). Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. *Mol Cell, 3*(6), 707-716.

Merrick, W. C. (2015). eIF4F: a retrospective. *J Biol Chem, 290*(40), 24091-24099. doi: 10.1074/jbc.R115.675280

Meyuhas, O. (2000). Synthesis of the translational apparatus is regulated at the translational level. *Eur J Biochem, 267*(21), 6321-6330.

Nakao, A., Yoshihama, M., & Kenmochi, N. (2004). RPG: the Ribosomal Protein Gene database. *Nucleic Acids Res, 32*(Database issue), D168-170. doi: 10.1093/nar/gkh004

Okosun, J., Wolfson, R. L., Wang, J., Araf, S., Wilkins, L., Castellano, B. M., . . . Fitzgibbon, J. (2016). Recurrent mTORC1-activating RRAGC mutations in follicular lymphoma. *Nat Genet, 48*(2), 183-188. doi: 10.1038/ng.3473

Orom, U. A., Nielsen, F. C., & Lund, A. H. (2008). MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Mol Cell, 30*(4), 460-471. doi: 10.1016/j.molcel.2008.05.001

Pannevis, M. C., & Houlihan, D. F. (1992). The energetic cost of protein synthesis in isolated hepatocytes of rainbow trout (Oncorhynchus mykiss). *J Comp Physiol B, 162*(5), 393-400.

Pause, A., Belsham, G. J., Gingras, A. C., Donze, O., Lin, T. A., Lawrence, J. C., Jr., & Sonenberg, N. (1994). Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function. *Nature, 371*(6500), 762-767. doi: 10.1038/371762a0

Pende, M., Um, S. H., Mieulet, V., Sticker, M., Goss, V. L., Mestan, J., . . . Thomas, G. (2004). S6K1(-/-)/S6K2(-/-) mice exhibit perinatal lethality and rapamycin-sensitive 5'-terminal oligopyrimidine mRNA translation and reveal a mitogen-activated protein kinase-dependent S6 kinase pathway. *Mol Cell Biol, 24*(8), 3112-3124.

Pestova, T. V., & Kolupaeva, V. G. (2002). The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes Dev, 16*(22), 2906-2922. doi: 10.1101/gad.1020902

Pestova, T. V., Lomakin, I. B., Lee, J. H., Choi, S. K., Dever, T. E., & Hellen, C. U. (2000). The joining of ribosomal subunits in eukaryotes requires eIF5B. *Nature, 403*(6767), 332-335. doi: 10.1038/35002118

Peter, D., Igreja, C., Weber, R., Wohlbold, L., Weiler, C., Ebertsch, L., . . . Izaurralde, E. (2015). Molecular architecture of 4E-BP translational inhibitors bound to eIF4E. *Mol Cell, 57*(6), 1074-1087. doi: 10.1016/j.molcel.2015.01.017

Peterson, D. T., Safer, B., & Merrick, W. C. (1979). Role of eukaryotic initiation factor 5 in the formation of 80 S initiation complexes. *J Biol Chem, 254*(16), 7730-7735.

Prevot, D., Darlix, J. L., & Ohlmann, T. (2003). Conducting the initiation of protein synthesis: the role of eIF4G. *Biol Cell, 95*(3-4), 141-156.

Raught, B., Peiretti, F., Gingras, A. C., Livingstone, M., Shahbazian, D., Mayeur, G. L., . . . Hershey, J. W. (2004). Phosphorylation of eucaryotic translation initiation factor 4B Ser422 is modulated by S6 kinases. *EMBO J, 23*(8), 1761-1769. doi: 10.1038/sj.emboj.7600193

Richardson, C. J., Broenstrup, M., Fingar, D. C., Julich, K., Ballif, B. A., Gygi, S., & Blenis, J. (2004). SKAR is a specific target of S6 kinase 1 in cell growth control. *Curr Biol, 14*(17), 1540-1549. doi: 10.1016/j.cub.2004.08.061

Richter, J. D., & Sonenberg, N. (2005). Regulation of cap-dependent translation by eIF4E inhibitory proteins. *Nature, 433*(7025), 477-480. doi: 10.1038/nature03205

Rogers, G. W., Jr., Richter, N. J., & Merrick, W. C. (1999). Biochemical and kinetic characterization of the RNA helicase activity of eukaryotic initiation factor 4A. *J Biol Chem, 274*(18), 12236-12244.

Ruvinsky, I., Sharon, N., Lerer, T., Cohen, H., Stolovich-Rain, M., Nir, T., . . . Meyuhas, O. (2005). Ribosomal protein S6 phosphorylation is a determinant of cell size and glucose homeostasis. *Genes Dev, 19*(18), 2199-2211. doi: 10.1101/gad.351605

Sancak, Y., Bar-Peled, L., Zoncu, R., Markhard, A. L., Nada, S., & Sabatini, D. M. (2010). Ragulator-Rag complex targets mTORC1 to the lysosomal surface and is necessary for its activation by amino acids. *Cell, 141*(2), 290-303. doi: 10.1016/j.cell.2010.02.024

Sancak, Y., Peterson, T. R., Shaul, Y. D., Lindquist, R. A., Thoreen, C. C., Bar-Peled, L., & Sabatini, D. M. (2008). The Rag GTPases bind raptor and mediate amino acid signaling to mTORC1. *Science, 320*(5882), 1496-1501. doi: 10.1126/science.1157535

Saxton, R. A., Knockenhauer, K. E., Wolfson, R. L., Chantranupong, L., Pacold, M. E., Wang, T., . . . Sabatini, D. M. (2016). Structural basis for leucine sensing by the Sestrin2-mTORC1 pathway. *Science, 351*(6268), 53-58. doi: 10.1126/science.aad2087

Sonenberg, N., Morgan, M. A., Merrick, W. C., & Shatkin, A. J. (1978). A polypeptide in eukaryotic initiation factors that crosslinks specifically to the 5'-terminal cap in mRNA. *Proc Natl Acad Sci U S A, 75*(10), 4843-4847.

Svitkin, Y. V., Pause, A., Haghighat, A., Pyronnet, S., Witherell, G., Belsham, G. J., & Sonenberg, N. (2001). The requirement for eukaryotic initiation factor 4A (eIF4A) in translation is in direct proportion to the degree of mRNA 5' secondary structure. *RNA, 7*(3), 382-394.

Tait, S., Dutta, K., Cowburn, D., Warwicker, J., Doig, A. J., & McCarthy, J. E. (2010). Local control of a disorder-order transition in 4E-BP1 underpins regulation of translation via eIF4E. *Proc Natl Acad Sci U S A, 107*(41), 17627-17632. doi: 10.1073/pnas.1008242107

Tcherkezian, J., Cargnello, M., Romeo, Y., Huttlin, E. L., Lavoie, G., Gygi, S. P., & Roux, P. P. (2014). Proteomic analysis of cap-dependent translation identifies LARP1 as a key regulator of 5'TOP mRNA translation. *Genes Dev, 28*(4), 357-371. doi: 10.1101/gad.231407.113

Tee, A. R., Manning, B. D., Roux, P. P., Cantley, L. C., & Blenis, J. (2003). Tuberous sclerosis complex gene products, Tuberin and Hamartin, control mTOR signaling by acting as a GTPase-activating protein complex toward Rheb. *Curr Biol, 13*(15), 1259-1268.

Thomas, A., Goumans, H., Voorma, H. O., & Benne, R. (1980). The mechanism of action of eukaryotic initiation factor 4C in protein synthesis. *Eur J Biochem, 107*(1), 39-45.

Thomas, A., Spaan, W., van Steeg, H., Voorma, H. O., & Benne, R. (1980). Mode of action of protein synthesis initiation factor eIF-1 from rabbit reticulocytes. *FEBS Lett, 116*(1), 67-71.

Thoreen, C. C., Chantranupong, L., Keys, H. R., Wang, T., Gray, N. S., & Sabatini, D. M. (2012). A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature, 485*(7396), 109-113. doi: 10.1038/nature11083

Thoreen, C. C., Kang, S. A., Chang, J. W., Liu, Q., Zhang, J., Gao, Y., . . . Gray, N. S. (2009). An ATP-competitive mammalian target of rapamycin inhibitor reveals rapamycin-resistant functions of mTORC1. *J Biol Chem, 284*(12), 8023-8032. doi: 10.1074/jbc.M900301200

Trachsel, H., Erni, B., Schreier, M. H., & Staehelin, T. (1977). Initiation of mammalian protein synthesis. II. The assembly of the initiation complex with purified initiation factors. *J Mol Biol, 116*(4), 755-767.

Vattem, K. M., & Wek, R. C. (2004). Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proc Natl Acad Sci U S A, 101*(31), 11269-11274. doi: 10.1073/pnas.0400541101

Walden, W. E., Godefroy-Colburn, T., & Thach, R. E. (1981). The role of mRNA competition in regulating translation. I. Demonstration of competition in vivo. *J Biol Chem, 256*(22), 11739-11746.

Wang, S., Tsun, Z. Y., Wolfson, R. L., Shen, K., Wyant, G. A., Plovanich, M. E., . . . Sabatini, D. M. (2015). Metabolism. Lysosomal amino acid transporter SLC38A9 signals arginine sufficiency to mTORC1. *Science, 347*(6218), 188-194. doi: 10.1126/science.1257132

Wang, X., Li, W., Williams, M., Terada, N., Alessi, D. R., & Proud, C. G. (2001). Regulation of elongation factor 2 kinase by p90(RSK1) and p70 S6 kinase. *EMBO J, 20*(16), 4370-4379. doi: 10.1093/emboj/20.16.4370

Warner, J. R. (1999). The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci, 24*(11), 437-440.

Wethmar, K. (2014). The regulatory potential of upstream open reading frames in eukaryotic gene expression. *Wiley Interdiscip Rev RNA, 5*(6), 765-778. doi: 10.1002/wrna.1245

Wilson, K. F., Wu, W. J., & Cerione, R. A. (2000). Cdc42 stimulates RNA splicing via the S6 kinase and a novel S6 kinase target, the nuclear cap-binding complex. *J Biol Chem, 275*(48), 37307-37310. doi: 10.1074/jbc.C000482200

Wolfson, R. L., Chantranupong, L., Saxton, R. A., Shen, K., Scaria, S. M., Cantor, J. R., & Sabatini, D. M. (2016). Sestrin2 is a leucine sensor for the mTORC1 pathway. *Science, 351*(6268), 43-48. doi: 10.1126/science.aab2674

Wortham, N. C., & Proud, C. G. (2015). eIF2B: recent structural and functional insights into a key regulator of translation. *Biochem Soc Trans, 43*(6), 1234-1240. doi: 10.1042/BST20150164

Zaytseva, Y. Y., Valentino, J. D., Gulhati, P., & Evers, B. M. (2012). mTOR inhibitors in cancer therapy. *Cancer Lett, 319*(1), 1-7. doi: 10.1016/j.canlet.2012.01.005

Zhang, Y., Gao, X., Saucedo, L. J., Ru, B., Edgar, B. A., & Pan, D. (2003). Rheb is a direct target of the tuberous sclerosis tumour suppressor proteins. *Nat Cell Biol, 5*(6), 578-581. doi: 10.1038/ncb999

Zoncu, R., Efeyan, A., & Sabatini, D. M. (2011). mTOR: from growth signal integration to cancer, diabetes and ageing. *Nat Rev Mol Cell Biol, 12*(1), 21-35. doi: 10.1038/nrm3025

Chapter 2

Regulation of mRNA Translation by Amino Acids

Heather R. Keys[1-5] and David M. Sabatini[1-5]

[1]Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA

[2]Department of Biology, MIT, Cambridge, MA 02139, USA

[3]Howard Hughes Medical Institute, MIT, Cambridge, MA 02139, USA

[4]Broad Institute, Cambridge, MA 02142, USA

[5]The David H. Koch Institute for Integrative Cancer Research, Cambridge, MA 02139, USA

SUMMARY

Translation is a highly regulated cellular process that is especially important for proliferative cells, and is often dysregulated in disease states such as cancer. Several lines of evidence suggest that both the mTORC1 and GCN2 pathways, which regulate translation initiation, play a role in tumor growth. Amino acid deprivation, a cancer-associated cellular stress, activates the GCN2 pathway while inhibiting the mTORC1 pathway, causing decreased translation initiation through multiple distinct mechanisms. However, the translational programs downstream of these regulatory pathways are not fully understood. Here, we characterize the translational program of cells deprived of amino acids both acutely and long-term using ribosome footprinting. The response to short-term amino acid starvation is mostly translational, whereas many mRNAs are found at higher levels after long-term deprivation, suggesting that the prolonged response relies heavily on transcription or mRNA stabilization. We confirm increased protein levels for several translationally regulated mRNAs, but also observe decreased protein levels for others, underscoring the importance of validating ribosome footprinting results via orthogonal methods. Translationally upregulated mRNAs were not enriched in upstream open reading frames, suggesting that alternative mechanisms exist to increase translation initiation under amino acid stress. We also identify two novel putative uORFs in the ATF4 5'UTR whose translation is amino acid responsive, suggesting that translational regulation of murine ATF4 is more complex than previously appreciated.

INTRODUCTION

The translation of a molecule of mRNA into a protein that can carry out a cellular function is required for the survival of all organisms, and is particularly important in proliferating cells that must duplicate their entire contents every cell cycle. Even acellular organisms, such as viruses, must commandeer this process in a host cell in order to reproduce. Like other energetically costly cellular processes, translation is quite tightly regulated. A well-studied point of regulation occurs at the initiation step: under certain conditions, such as nutrient stress, both initiator methionyl-tRNA (Met-tRNA$_i^{Met}$) recruitment to the 40S ribosomal subunit and the recruitment of the 43S pre-initiation complex (43S PIC) to the 5' end of the mRNA are inhibited, effectively reducing the translation of most mRNAs (Sonenberg & Hinnebusch, 2009; Thoreen, 2013). However, the cell must also produce new proteins to adapt to the stress, and mRNAs containing particular *cis* elements can escape this global inhibition (Wethmar, 2014).

Amino acid deprivation is a specific stress that has been studied in the context of the whole organism (Kilberg, Balasubramanian, Fu, & Shan, 2012; Mortimore & Poso, 1987), and has more recently been investigated in the context of cancer (Baracos & Mackenzie, 2006). Asparagine, arginine, and serine have all been shown to play roles in cancer or cancer cell proliferation (Emadi, Zokaee, & Sausville, 2014; Labuschagne, van den Broek, Mackay, Vousden, & Maddocks, 2014; Patil, Bhaumik, Babykutty, Banerjee, & Fukumura, 2016; Zhang et al., 2014), and glutamine is consumed more highly by cancer cells (Coles & Johnstone, 1962; DeBerardinis et al., 2007; Reitzer,

Wice, & Kennell, 1979). Although amino acids are generally studied in cancer in the context of cancer cell metabolism (Mayers & Vander Heiden, 2015), amino acid levels could play a more direct role in certain cancers that are auxotrophic for particular amino acids (Agrawal, Alpini, Stone, Frenkel, & Frankel, 2012; Kuo, Savaraj, & Feun, 2010). Amino acid stress could also play a more indirect role in certain cancers, by modulating the translational program through the GCN2 and mechanistic target of rapamycin complex 1 (mTORC1) pathways (Kimball, 2002).

Translational control in general has also been studied in many different contexts, and is also beginning to be investigated in cancer (Bhat et al., 2015). It is well appreciated that tumors experience nutrient and oxygen stress during their growth, both of which can regulate translation (Ackerman & Simon, 2014; Holcik & Sonenberg, 2005). There are a variety of links between the translational machinery and cancer, including enlarged nucleoli and overexpression of initiation factors, as well as translational deregulation (Bhat et al., 2015; Orsolic et al., 2016). mTORC1 is a kinase complex that responds to growth factors, nutrients, cellular energy levels, and hypoxia to regulate processes important for cell growth, including translation (Laplante & Sabatini, 2012). Upstream regulators or mTORC1 are mutated in many cancers (Laplante & Sabatini, 2012), and mutations are found in mTOR itself (Grabiner et al., 2014), underscoring the important role of this pathway in this disease. mTOR inhibitors have been approved for clinical therapy for certain malignancies, and many trials remain ongoing (Zaytseva, Valentino, Gulhati, & Evers, 2012); however, it is not entirely clear which substrates and functions downstream of mTOR play the greatest role in disease. mTORC1 promotes

translation by inhibiting the 4EBPs, a class of proteins that inhibits translation initiation by disrupting the interaction between eIF4E and eIF4G (Brunn et al., 1997; Marcotrigiano, Gingras, Sonenberg, & Burley, 1999). An increase in phosphorylated, inactive 4EBP1 is observed in a mouse thymic lymphoma model driven by the expression of the oncogene Akt2, and expression of a non-phosphorylatable form of 4EBP1 inhibits tumor growth (Hsieh et al., 2010). Pharmacological inhibition of mTOR also inhibits tumor growth in this model (Hsieh et al., 2010). Small molecule inhibition of the interaction between eIF4E and eIF4G, which mimics disruption of this interaction by hypophosphorylated 4EBPs, reduces growth of xenografts (Chen et al., 2012). Furthermore, inhibition of eIF4A, the helicase component of eIF4F, can inhibit tumor growth (Cencic et al., 2009), and its cellular inhibitor, PDCD4, is a tumor suppressor (Yang et al., 2003). These data suggest that translational control downstream of mTORC1 is particularly relevant in cancer.

There is also evidence that specific stress responses programs regulated through 4EBP-independent mechanisms, such as control of ATF4 translation mediated by the initiation inhibitor GCN2, can be advantageous for tumor growth (Wang et al., 2013; Ye et al., 2010). However, a small molecule that causes phosphorylation of eIF2α can conversely inhibit tumor growth (Chen et al., 2012), suggesting that the role of this translational program in cancer is not straightforward. Regardless, these studies support the notion that translational regulation plays a role in cancer progression, and that it may be therapeutically advantageous to identify the translational program that helps drive this disease.

Attempts have been made to identify translationally regulated mRNAs by hybridizing RNA isolated from polysome fractions of sucrose gradients to microarrays (Zong, Schummer, Hood, & Morris, 1999). However, this method measures changes in the distribution of mRNAs within a sucrose gradient, which does not allow resolution of small changes in translation. Furthermore, it provides no information about the position of ribosomes on the mRNA, which can give insight into the regulatory mechanism. More recently, the ribosome footprinting method was developed to address these shortcomings using Illumina sequencing to quantitatively measure ribosome occupancy (Ingolia, Ghaemmaghami, Newman, & Weissman, 2009).

Despite the importance of translation in basic cell survival, cellular processes, and in disease states, there is still a limited molecular understanding of the global consequences of modulation of translation under stress conditions. Here we have taken a transcriptome-wide approach to analyzing cellular adaptation and translational regulation in response to one stress, amino acid deprivation, which regulates both mTORC1 and GCN2.

RESULTS

Characterization of the Global Translational Response to Short- and Long-Term Amino Acid Deprivation

Prior to performing ribosome footprinting analysis, we characterized the response of SV40 large T antigen-immortalized MEFs (Jiang et al., 2003) to short-term and long-term amino acid deprivation. We chose these MEFs because they are matched controls for GCN2-/- MEFs (Jiang et al., 2003), on which we planned to also perform ribosome footprinting. We starved cells of all exogenous amino acids for either two or 18 hours, and compared the global translational response to control cells by separation of ribosomes on a sucrose gradient, followed by detection of RNA by absorbance at 254nm (Figure 1A). Control cells contained few free 40S and 60S ribosomal subunits, with most of the ribosomes present in polysomes (Figure 1A), indicating that translation initiation is quite efficient in these cells. In contrast, cells deprived of amino acids for two hours showed a striking reduction in polysomes and a concomitant increase in the 80S monosome peak (Figure 1A), representing a severe inhibition of initiation. After 18 hours of amino acid starvation, mRNAs were once again present in polysomes, although there was still an increase in monosomes and free 60S ribosomal subunits relative to amino acid-replete cells (Figure 1A). This result suggests that cells can adapt their translational program to prolonged periods of amino acid deprivation, although likely not indefinitely. It is worth noting that equal protein quantities were loaded onto each sucrose gradient, and that many more cells were required to obtain the same amount of protein from the cells starved of amino acids for 18 hours (data not shown). Thus, although mRNAs do

return to polysomes after long-term amino acid deprivation (Figure 1A), in absolute terms the amount of translation that occurs is greatly reduced.

We also examined the activity of two signaling pathways known to regulate translation in response to amino acid levels, the mTORC1 pathway and the GCN2 pathway. After two hours of amino acid deprivation, we observed a decrease in mTORC1 activity, as indicated by decreased phosphorylation of the mTORC1 substrate S6K (Figure 1B). Another mTORC1 substrate, 4EBP1, also showed decreased phosphorylation after two hours of amino acid starvation, which was detected by the appearance of lower-molecular-weight species (Figure 1B). Autophagy is also activated downstream of mTORC1 inhibition (Laplante & Sabatini, 2012). Induction of autophagy can be detected as an initial increase in lipidated LC3b (LC3b II) and reduction in cytoplasmic LC3b (LC3b I), followed by a reduction in both forms of LC3b as autophagy proceeds (Tanida, Ueno, & Kominami, 2008). At two hours post-amino acid starvation, we observed a reduction in both forms of LC3b, which we interpreted as indicating the induction of autophagy (Figure 1B). Levels of mTOR, as well as raptor and rictor, two proteins that are part of the mTORC1 and mTORC2 complexes, respectively, remain constant (Figure 1B). Levels of S6K and eIF2α remain constant as well (data not shown). At this time point, the GCN2 pathway was also active, as evidenced by a modest increase in eIF2α phosphorylation (Figure 1B). Together, these data suggest that both the mTORC1 and GCN2 pathways contribute to the inhibition of translation initiation after amino acid deprivation in these cells.

Figure 1: Amino Acid Deprivation Affects Translation and Signaling in MEFs

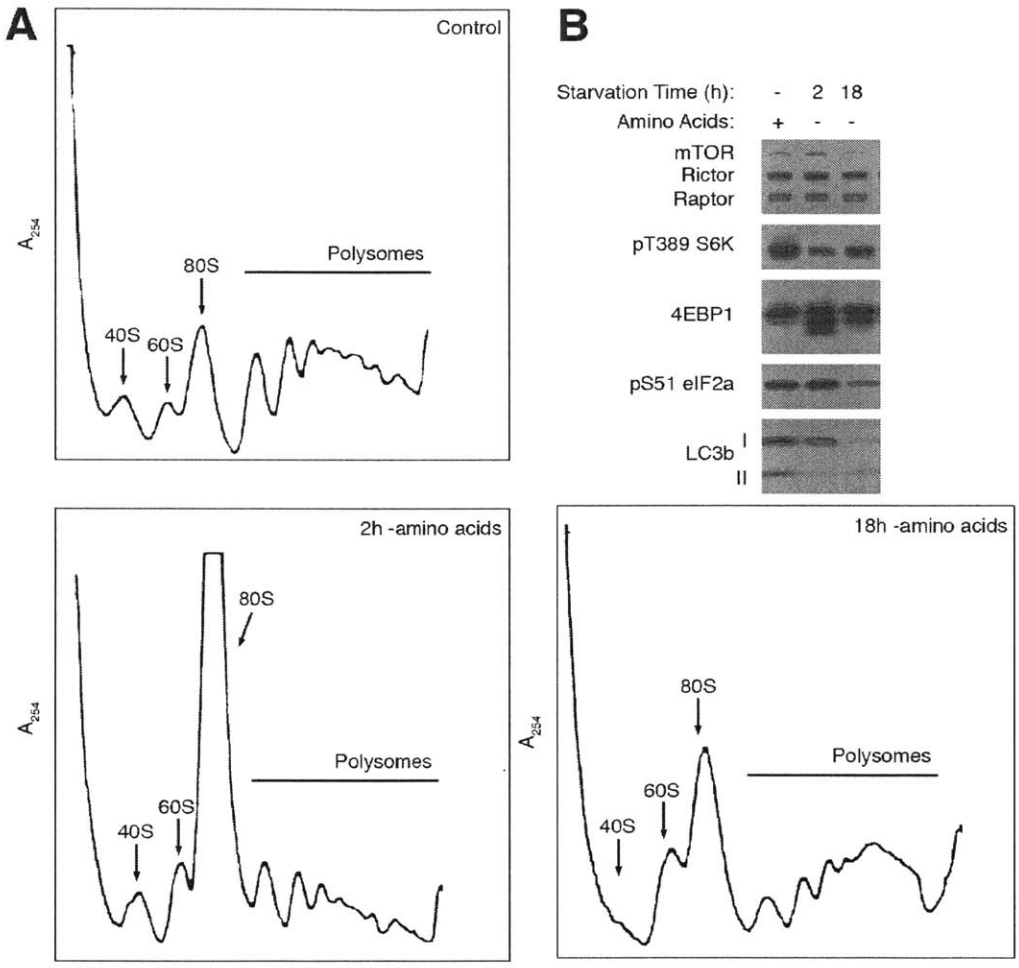


Figure 1: Amino acid deprivation affects translation and signaling in MEFs

A. Amino acid deprivation changes translation in MEFs. MEFs were grown in full media or deprived of all amino acids for 2 or 18 hours. Lysates were fractionated on a sucrose gradient and polysome profiles were detected by absorbance at 254nm.

B. Amino acid deprivation changes signaling in MEFs. MEFs were grown in full media or deprived of all amino acids for 2 or 18 hours. Signaling was assessed by immunoblot.

After 18 hours of amino acid starvation, a time point when mRNAs had redistributed back to polysomes (Figure 1A), signaling through the mTORC1 and GCN2 pathways had also recovered somewhat, relative to two hours of amino acid starvation. We observed an increase in S6K phosphorylation at 18 hours, although S6K phosphorylation did not reach the level seen in amino acid-replete cells (Figure 1B). Similarly, phosphorylation of 4EBP1 recovered incompletely (Figure 1B). Consistent with some level of sustained inhibition the mTORC1 pathway, autophagy appeared to continue, as evidenced by very low levels of LC3b (Figure 1B). Interestingly, eIF2α phosphorylation was actually reduced below the level seen in amino acid-replete cells (Figure 1B), despite equivalent total levels of eIF2α (data not shown), which could indicate a basal level of stress that impinges on eIF2α in amino acid-replete cells (such as ER stress) that is relieved when translation is reduced. These data suggest that translation and signaling remain coupled, even after long-term amino acid deprivation.

Ribosome Footprinting Analysis

To assess the translational consequences of amino acid deprivation on a transcriptome-wide scale, we performed ribosome footprinting (Ingolia et al., 2009). Briefly, the ribosome footprinting method captures ribosome-protected fragments (footprints) after RNAse digestion, which are subsequently processed in parallel with total cellular mRNA to generate an Illumina sequencing library. Sequencing reads from ribosome footprints can be compared to those from total mRNA for each mRNA transcript to calculate the ribosome occupancy of each mRNA, which is an indication of

how highly-translated an mRNA is. We performed ribosome footprinting of MEFs under amino acid-replete conditions or after total amino acid deprivation for two hours or 18 hours. We chose this relatively early time point in order to identify genes which play a master role in adapting the cell to amino acid stress, whereas the eighteen hour time point was chosen to give a snapshot of the downstream changes the cell undergoes in order to survive a lack of amino acids.

Ribosome footprints consisted primarily of 33-nucleotide fragments (Figure 2A). This length is two nucleotides longer than previously reported for mammalian footprinting studies (Ingolia, Lareau, & Weissman, 2011), and could reflect different RNAse digestion conditions (Ingolia, Brar, Rouskin, McGeachy, & Weissman, 2012), or a different ribosomal conformation (Lareau, Hite, Hogan, & Brown, 2014). We also observed a preference in the reading frame for footprints but not total mRNA fragments (Figure 2B), although this preference was less marked than has been previously observed for yeast footprints (Ingolia et al., 2009), but is similar to mammalian footprints (Ingolia et al., 2011). Finally, we saw greater ribosome footprint coverage of the 5'UTR than the 3'UTR (Figure 2C), whereas coverage was similar between the 5'UTR and 3'UTR in the total mRNA sample (Figure 2C). This result is also expected, based on other ribosome footprinting studies (Ingolia et al., 2009; Ingolia et al., 2011). Altogether, the quality of our ribosome footprinting data was generally similar to previous datasets.

Figure 2: Ribosome Footprinting QC

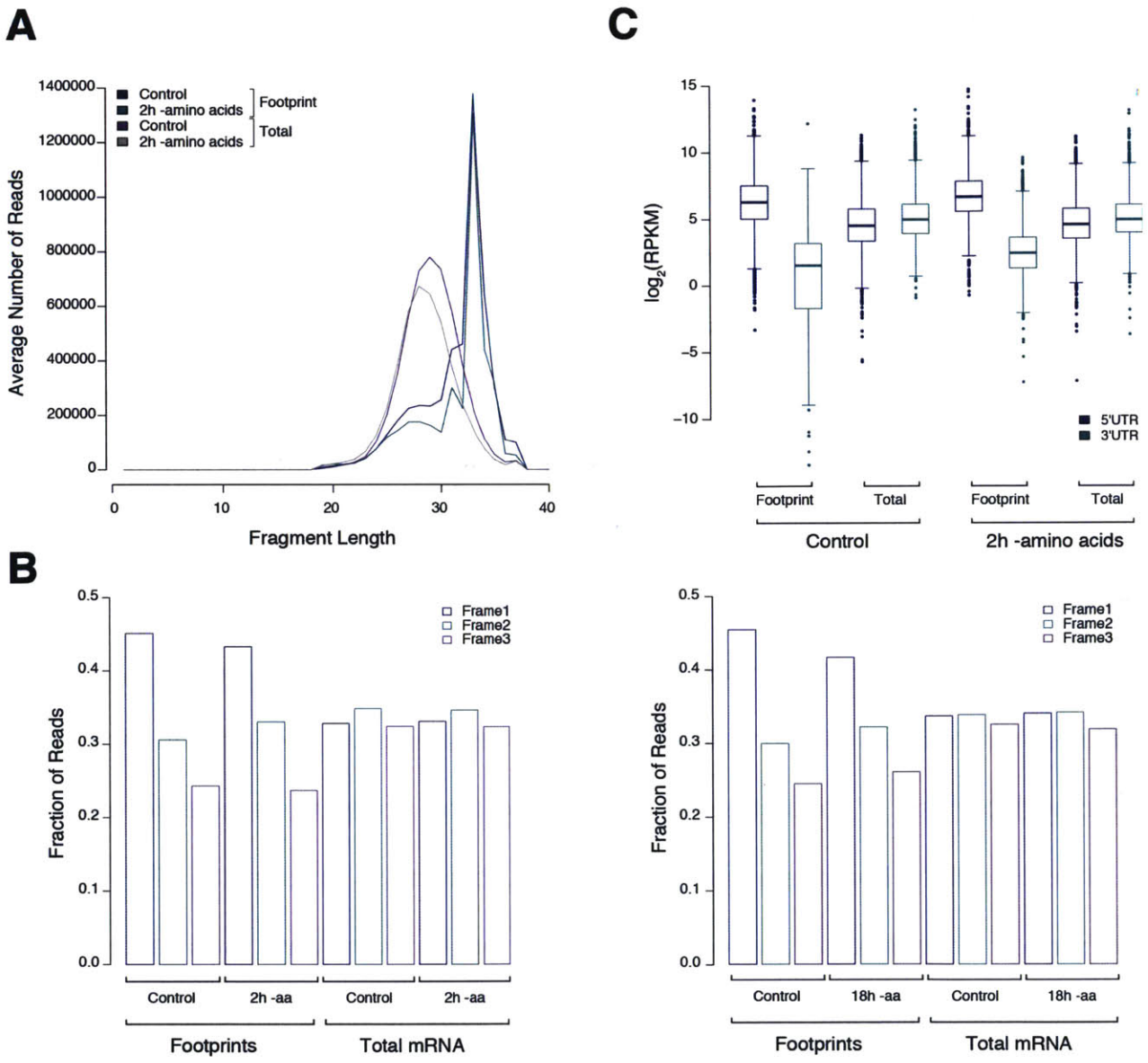


Figure 2: Ribosome Footprinting QC

A. Ribosome footprints are relatively uniform. Fragment length is plotted against average number of mapped reads for each fragment length.

B. Reading frame preference is observed for ribosome footprints but not total mRNA. Reads mapping to the region -13 to +148 around the start codon of all transcripts were averaged for each position, then the fraction of reads in each reading frame throughout the region was calculated. Each condition is indicated.

C. Ribosome footprints are depleted for reads mapping the 3'UTR. The reads mapping to either the 5'UTR or 3'UTR were normalized for feature length and for the number of mapped reads and plotted. Ribosome footprint and total mRNA are indicated, as is the treatment condition.

mRNAs are Translationally Regulated in Response to Amino Acid Deprivation

After two hours of amino acid starvation, the fold change in ribosome footprints was overall quite similar to the change in total mRNA levels (Figure 3A; Pearson $R^2 = 0.82$, $p < 0.0001$). Despite this correlation between mRNA levels and footprints, 116 mRNAs exhibited a statistically significant increase in ribosome footprints, as determined by EdgeR analysis (Figure 3A, all colors). Of these mRNAs, 65 also showed a significant increase in total levels (Figure 3A, "Total" and "Total & Occupancy"), suggesting that the increased ribosome footprints observed for these mRNAs were not solely due to an increase in translation initiation efficiency. Of the remaining 51 mRNAs with increased ribosome footprints, 21 had changes in their ribosome occupancies greater than two standard deviations from the mean (Figure 3A, "Occupancy"). Importantly, ATF4, which is known to be translationally upregulated after amino acid withdrawal, was a member of this group, suggesting that these mRNAs may be subject to similar translational control. Three of the mRNAs with increased total levels also had changes in their ribosome occupancies greater than two standard deviations from the mean (Figure 3A, "Total & Occupancy"). The remaining 30 mRNAs with increased ribosome footprints (Figure 3A, "Footprint Only") could also be translationally upregulated, although perhaps not as dramatically, or through a different mechanism.

The comparison of ribosome footprints with total mRNA from amino acid-replete cells and cells starved for amino acids for 18 hours revealed an even higher correlation between footprints and mRNA levels (Figure 3B; Pearson R^2 = 0.8571, $p < 0.0001$), suggesting that protein synthesis at this time point is primarily driven by changes in mRNA levels. Additionally, the levels of many more mRNAs were significantly higher than at the two-hour time point (Figure 3B), implying that transcriptional or mRNA stability alterations play a major role in the cellular adaptation to prolonged amino acid deprivation. Consistent with transcriptional alterations being particularly important, known ATF4 targets were upregulated at 18 hours, such as ASNS, CEBPγ, and ATF3 (Harding et al., 2003). Of the 439 mRNAs with significantly increased ribosome footprints (Figure 3B, all colors), 273 also exhibited an increase at the mRNA level (Figure 3B, "Total" and "Total & Occupancy"). Four mRNAs with increased total levels also exhibited an increase in ribosome occupancy greater than two standard deviations from the mean (Figure 3B, "Total & Occupancy"). Only 24 of the remaining mRNAs with increased footprints showed an increase in ribosome occupancy greater than two standard deviations from the mean (Figure 3B, "Occupancy"), which further supports the notion that adaptation to long-term amino acid stress is largely dependent on changes in mRNA levels.

We reasoned that mRNAs that are particularly important for the cellular response to amino acid deprivation would be more highly translated after both acute and long-term deprivation. Of the 116 mRNAs with significantly increased ribosome footprints two hours post-amino acid deprivation, 55 of these continued to show elevated ribosome

footprints after 18 hours. When we looked at functional annotation for these 55 mRNAs using the DAVID database (Dennis et al., 2003), three of the top four most highly enriched categories involved transcriptional regulation or the nucleus (data not shown). These data further suggest that the adaptation to long-term amino acid stress is largely transcriptional.

Figure 3: Ribosome Occupancy After Amino Acid Deprivation

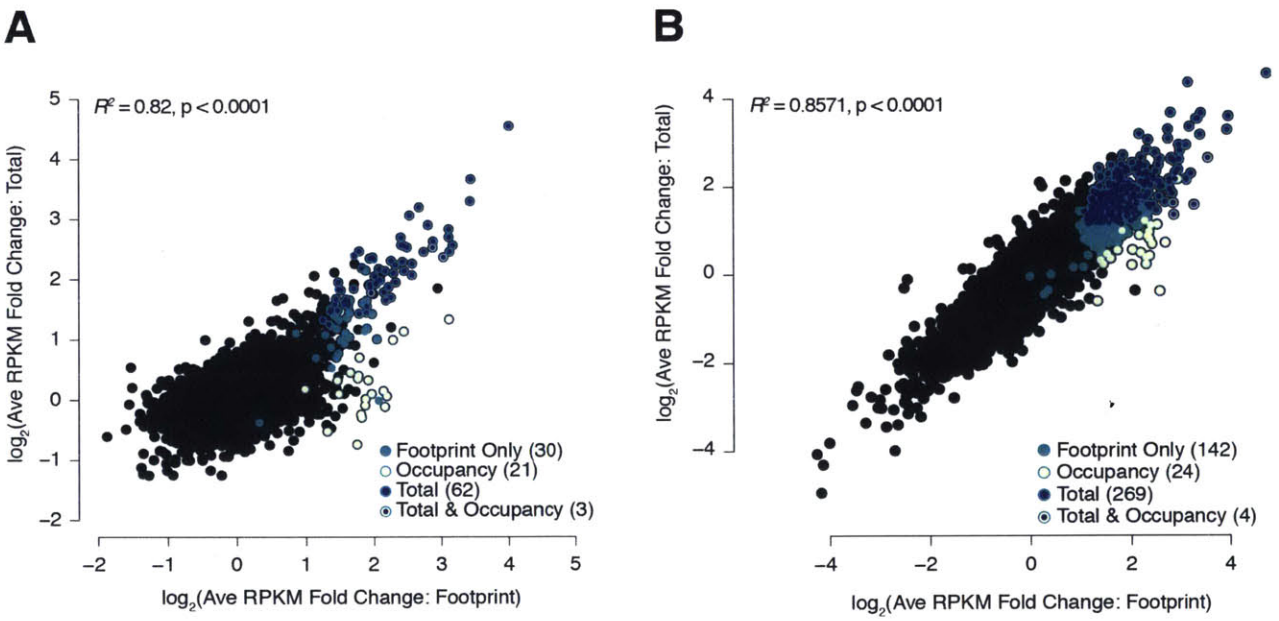


Figure 3: Ribosome occupancy after amino acid deprivation

A. Amino acid deprivation for two hours results in changes in total levels, footprint levels, and ribosome occupancy of mRNAs. The log_2-transformed average fold change in read density of ribosome footprints after two hours of amino acid deprivation is plotted against the log_2-transformed average fold change in read density of total mRNA fragments. Both sets of values are corrected for number of mapped reads. All colored points represent mRNAs whose footprints were statistically significantly increased with amino acid deprivation, as determined by EdgeR analysis. Purple indicates mRNAs whose total levels were also statistically significantly increased, and beige indicates mRNAs whose average ratio of footprint-to-total reads was greater than two standard deviations from the mean.

B. Amino acid deprivation for 18 hours results in changes in total levels, footprint levels, and ribosome occupancy of mRNAs. The average fold change in read density of ribosome footprints after 18 hours of amino acid deprivation is plotted against the average fold change in read density of total mRNA fragments. Both sets of values are corrected for number of mapped reads. All colored points represent mRNAs whose footprints were statistically significantly increased with amino acid deprivation, as determined by EdgeR analysis. Red indicates mRNAs whose total levels were also statistically significantly increased, and yellow indicates mRNAs whose average ratio of footprint-to-total reads was greater than two standard deviations from the mean.

An increase in the number of ribosome footprints for an mRNA between conditions could indicate that the mRNA is translated more efficiently, in which case the increased reads would be fairly evenly distributed throughout the open reading frame (ORF). Alternatively, the increased reads could show positional differences, which could indicate ribosomal stalling or mRNA features that can affect translation, such as upstream open reading frames (uORFs). Of the 7 mRNAs with increased ribosome footprints but not increased total levels after two hours and 18 hours of amino acid deprivation (Figure 4A), we selected three for further analysis, plus ATF4 as a control (Figure 4A and 4B). We plotted the position of each read against the number of reads at each position, normalized to the number of mapped reads in each sample for the two hour time point (Figure 4B). For both footprint and total mRNA reads, we set the position of the annotated start codon at zero. For each mRNA, footprinting reads increased throughout the ORF after two hours of amino acid deprivation (Figure 4B), indicating that translation of these mRNAs is increased. Total levels of these mRNAs did not significantly increase at this time point, although we did notice a slight increase in levels of ATF4 and Mafg (Figure 4B).

Interestingly, we also observed positional differences in ribosome footprints (Figure 4B). The ATF4 transcript encodes two uORFs, which regulate its translation by controlling the level of translation initiation that occurs at the start codon of the downstream coding ORF (Lu, Harding, & Ron, 2004; Vattem & Wek, 2004). We observed a striking increase in ribosome footprints at two uORFs upon amino acid deprivation (Figure 4B), one of which has been previously described (Lu et al., 2004; Vattem & Wek, 2004). However, we did not observe a significant change in occupancy at the other canonical ATF4 uORF, which could reflect either a lack of coverage of this region, or that an alternative mechanism is used for translational regulation of ATF4 during amino acid stress. We also observed utilization of a uORF in the Mafg mRNA. Similarly to ATF4, we saw an increase in the ribosome footprints at the uORF, as well as in the coding ORF (Figure 4B), suggesting that these uORFs promote translation reinitiation at the downstream coding ORF under these conditions. Finally, we saw greatly increased ribosome footprints at certain positions throughout the Ahnak ORF (Figure 4B). Although we observed increased footprints throughout the Ahnak ORF under amino acid deprivation (Figure 4B), suggesting increased translation of this mRNA, the much greater increases in ribosome footprints at particular positions throughout the mRNA suggest ribosomes could be stalled at those sites. To verify increased translation of this mRNA, further experiments would be necessary.

Figure 4: Ribosome Occupancy of Selected mRNAs

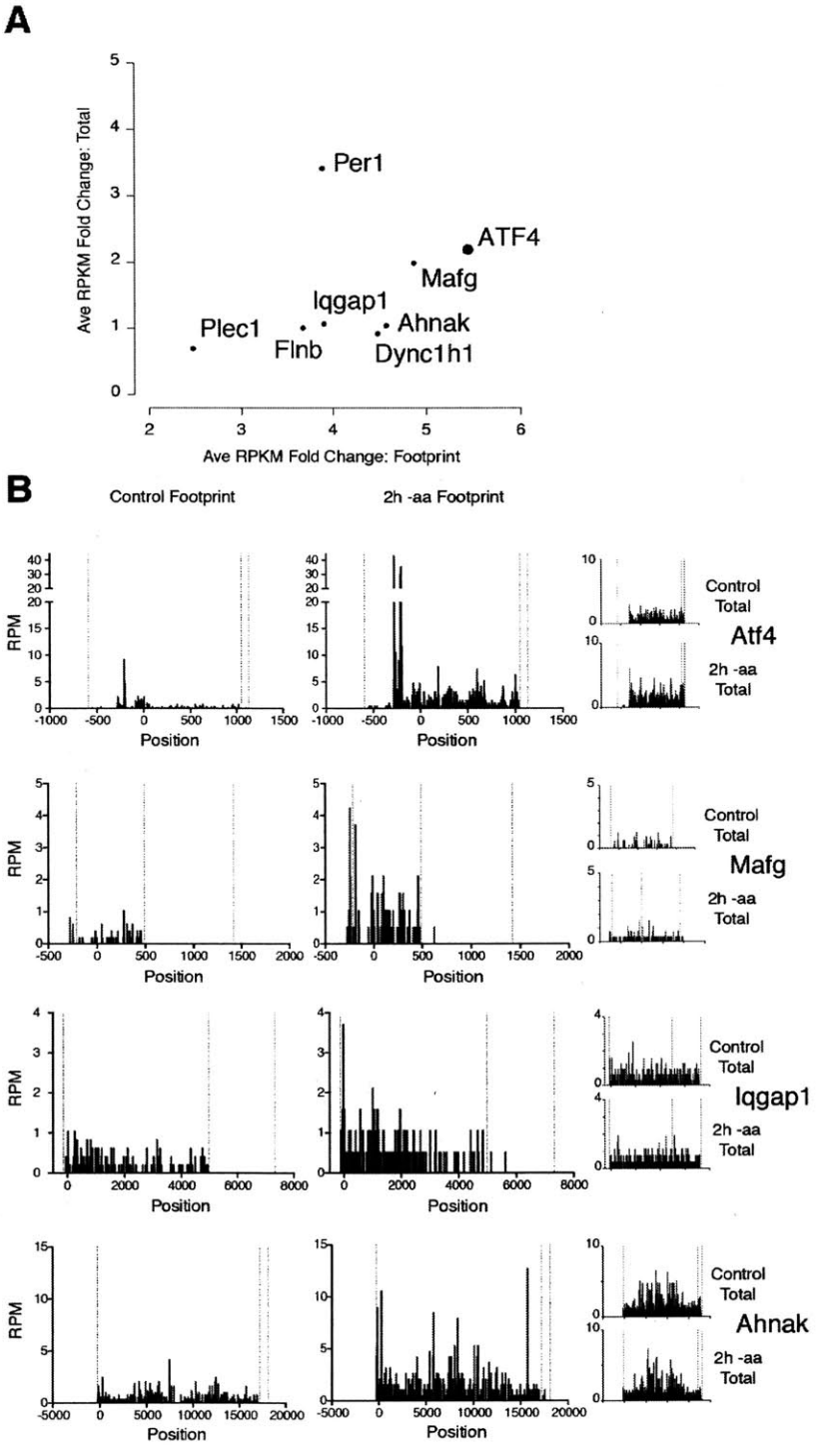


Figure 4: Ribosome occupancy of selected mRNAs

A. Selected translationally upregulated mRNAs. The average fold change in ribosome footprint after two hours of amino acid starvation is plotted against the average fold change in total levels for mRNAs that met the following criteria: they were statistically significantly increased in footprints after two hours and 18 hours of amino acid deprivation; their average ratio of footprint-to-total reads was greater than two standard deviations from the mean after two hours and 18 hours of amino acid deprivation. ATF4, a control mRNA, is also indicated.

B. Positional coverage of selected mRNAs reveals increased coverage after amino acid deprivation, novel uORF usage, and possible ribosome stalling. Reads mapping to the indicated mRNAs were normalized to the number of total mapped reads and plotted against mRNA position, with the start codon set to zero. Ribosome footprint and total mRNA coverage is indicated. The three vertical dashed lines indicate, in order: the annotated RefSeq transcription start site, the annotated stop codon, and the annotated transcript end.

Validation of Candidate Translationally Upregulated mRNAs

To validate candidate translationally regulated mRNAs, we utilized a polysome gradient-qPCR approach. Briefly, we separated RNA by sucrose gradient fractionation and isolated mRNAs that are occupied by one, 2-3, 4-7, or greater than 7 ribosomes; the region of the gradient containing greater than 7 ribosomes was divided into two fractions (Figure 5A). We performed qPCR on each fraction to detect levels of candidate mRNAs. If an mRNA is well translated, it should be found in the fractions containing more ribosomes. Conversely, if an mRNA is poorly translated, it should be observed in the less dense fractions. Therefore, mRNAs translated more efficiently due to increased initiation under amino acid deprivation should shift their distribution toward fractions containing more ribosomes.

An mRNA with a change in ribosome footprints near the mean of the change observed for all mRNAs was used as a control. Mapk3 was associated with fewer ribosomes upon amino acid deprivation (Figure 5B). This is expected; since we observed many fewer polysomes in general when cells were starved for amino acids (Figure 1A), we expected that most mRNAs would redistribute to fractions with fewer polysomes. Conversely, mRNAs that were likely to be translationally upregulated according to our ribosome footprinting results were associated with more ribosomes upon amino acid deprivation (Figure 5C), although to varying degrees. Two mRNAs with both increased footprints and increased total levels showed a slight redistribution toward fractions containing more ribosome (Figure 5D), suggesting that these mRNAs may be translationally regulated, in addition to exhibiting regulated transcription or stability. Finally, two mRNAs whose footprints were significantly increased without an increase in total levels exhibited different behaviors (Figure 5E). Ccnl1 showed a quite dramatic shift to fractions with greater numbers of ribosomes (Figure 5E), similarly to some of the mRNAs with greatly increased ribosome occupancy (Figure 5B). Cdc6, on the other hand, showed only a modest increase in association with greater than 7 ribosomes (Figure 5E). It is unclear whether this increase reflects an increase in association with ribosomes or a decrease, as the fractions with fewer ribosomes and the fraction with more ribosomes were both decreased (Figure 5E). It is possible that two populations of this mRNA exist, and that they are differentially regulated upon amino acid deprivation. However, it is likely that polysome gradient-qPCR method simply cannot resolve the numbers of ribosomes well enough to truly measure the distribution for some mRNAs.

Figure 5: Most mRNAs Validate by Polysome Gradient-qPCR Analysis.

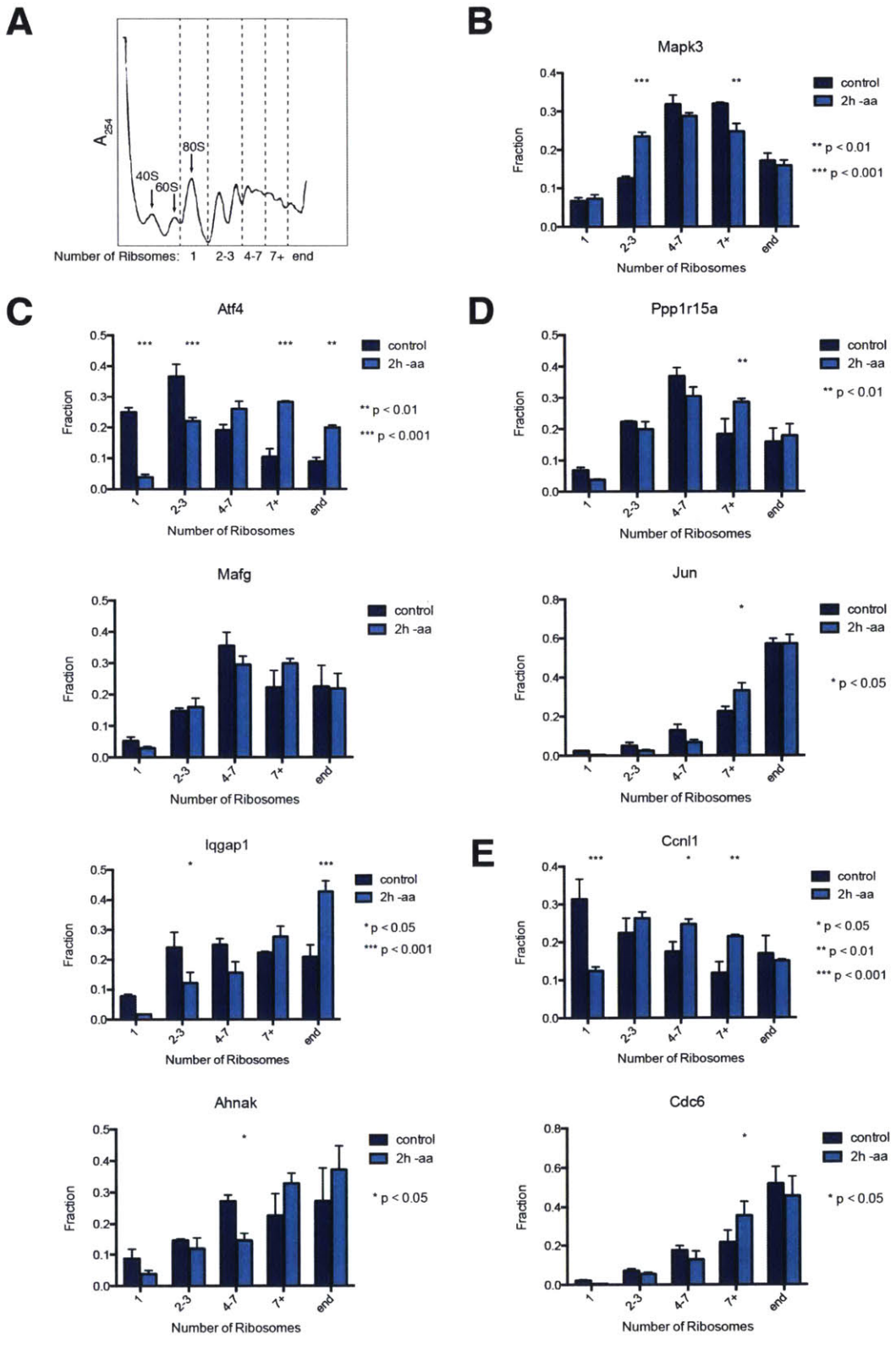


Figure 5: Most mRNAs validate by polysome gradient-qPCR analysis

A. Schematic of fractionation. Lysates were fractionated on a sucrose gradient and polysomes were detected by absorbance at 254nm. Fractions containing the indicated numbers of ribosomes were collected.

B. A control mRNA shows decreased polysome association after amino acid deprivation. The fraction of Mapk3 associated with the indicated numbers of ribosomes is plotted for the indicated conditions. Data are expressed as the mean ± s.e.m. (n = 3). Significance was determined using two-way ANOVA comparing the mean of each treatment for each fraction with Bonferroni correction for multiple comparisons. P-values are indicated.

C. mRNAs with increased ribosome footprints and greatly increased ribosome occupancy associate with heavier polysomes after amino acid deprivation. The fraction of each mRNA associated with the indicated numbers of ribosomes is plotted for each condition. Data are expressed as the mean ± s.e.m. (n = 3). Significance was determined using two-way ANOVA comparing the mean of each treatment for each fraction with Bonferroni correction for multiple comparisons. P-values are indicated.

D. mRNAs with increased ribosome footprints and increased total levels after amino acid deprivation show a modest increase in polysome association. The fraction of each mRNA associated with the indicated numbers of ribosomes is plotted for the each condition. Data are expressed as the mean ± s.e.m. (n = 3). Significance was determined using two-way ANOVA comparing the mean of each treatment for each fraction with Bonferroni correction for multiple comparisons. P-values are indicated.

E. mRNAs with increased footprints but not greatly increased ribosome occupancy exhibit both increased and decreased association with polysomes after amino acid deprivation. The fraction of each mRNA associated with the indicated numbers of ribosomes is plotted for the each condition. Data are expressed as the mean ± s.e.m. (n = 3). Significance was determined using two-way ANOVA comparing the mean of each treatment for each fraction with Bonferroni correction for multiple comparisons. P-values are indicated.

Although most mRNAs that showed increased ribosome occupancy by ribosome footprinting were also associated with more ribosomes by polysome gradient-qPCR, we

sought to further validate our results by immunoblot. ATF4, which is known to be translationally upregulated, which was confirmed by our ribosome footprinting results, was also upregulated at the protein level when cells were starved for amino acids, whereas mTOR, rictor, and raptor levels were unchanged (Figure 6A). Iqgap1 is another mRNA that showed increased ribosome footprints and occupancy in our experiment, and we observed an increase in the protein upon amino acid deprivation, although not to as great an extent as ATF4 (Figure 6A).

We also compared mRNAs with increased ribosome footprints, but not greatly increased occupancy, with mRNAs whose ribosome footprints and total levels both increase. Cdc6, for which the qPCR validation results were unclear (Figure 5E), showed a striking reduction in protein level upon amino acid deprivation (Figure 6B). In contrast, Jun, which exhibited an increase in both ribosome occupancy and total mRNA, showed an increase at the protein level as well (Figure 6B). Although some of our candidate mRNAs validated, these results underscore the importance of verifying ribosome footprinting results at the protein level.

We reasoned that mRNAs with increased ribosome footprints and increased mRNA levels upon both short term and long term amino acid deprivation would be particularly important for the adaptive response to lack of amino acids, and would be more likely to also exhibit increased protein levels, as was observed for Jun. This hypothesis was true for Sqstm1, although the increase in protein was very modest (Figure 6C). Myc, however, was reduced when amino acids were removed (Figure 6C). This reduction could be reversed by proteasomal inhibition (data not shown), suggesting that for some

Figure 6: Not All mRNAs that Show Increased Ribosome Occupancy Are Also Increased at the Protein Level.

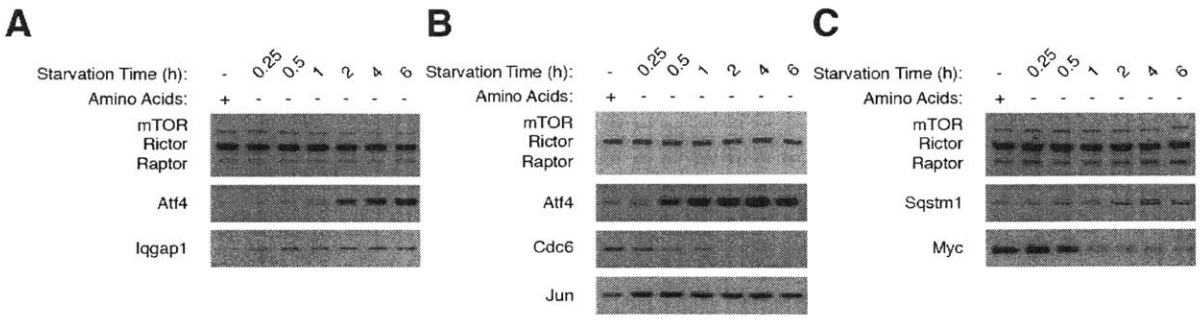


Figure 6: Not all mRNAs that show increased ribosome occupancy are also increased at the protein level

A. mRNAs with increased ribosome footprints and highly increased ribosome occupancy are increased at the protein level after amino acid deprivation. Lysates from cells starved for amino acids for the indicated times were analyzed by immunoblotting for the indicated proteins.

B. An increase in ribosome footprints after amino acid deprivation does not necessarily predict an increase at the protein level. Lysates from cells starved for amino acids for the indicated times were analyzed by immunoblotting for the indicated proteins.
C. An increase in ribosome footprints and total mRNA does not necessarily predict an increase at the protein level. Lysates from cells starved for amino acids for the indicated times were analyzed by immunoblotting for the indicated proteins.

proteins, synthesis and degradation are regulated disparately under certain conditions. This result further indicates that ribosome footprinting results should be verified at the protein level.

The most well characterized mechanism for increased translation following amino acid deprivation is through the presence of an inhibitory upstream open reading frame (uORF), which can be bypassed under certain stress conditions to result in increased

translation of the downstream protein-coding ORF. The canonical regulation of ATF4 is an example of this regulatory mechanism (Jackson, Hellen, & Pestova, 2010; Lu et al., 2004; Vattem & Wek, 2004). To see if uORFs were more frequently found in mRNAs whose translation is increased by amino acid deprivation, we compared our ribosome footprinting data from control and amino acid-deprived cells with published uORF data from mouse ES cells (Ingolia et al., 2011) (Figure 7A). We saw no enrichment of uORFs in mRNAs with increased footprints and ribosome occupancy after two hours of amino acid deprivation (data not shown). Surprisingly, we even observed depletion of uORFs in mRNAs with increased ribosome footprints after two hours of amino acid starvation (two-sided Fisher's Exact Test, $p = 2.937 \times 10^{-10}$). Furthermore, the cumulative distributions of ribosome occupancy fold change for all mRNAs and for mRNAs harboring uORFs were almost identical (Figure 7B), suggesting that the presence of a uORF does not generally confer translational upregulation in response to amino acid stress. The distribution of mRNAs containing uORFs was significantly shifted toward higher ribosome occupancy ($p \sim 7.702 \times 10^{-5}$); however, this shift appeared to be primarily due mRNAs that are translationally repressed following amino acid deprivation (Figure 7B). In other words, the presence of a uORF is generally associated with a milder translational repression in response to amino acid stress, rather than with translational upregulation as is typically associated with the uORFs of ATF4.

ATF4 is the stereotypical mRNA whose translation is increased in response to amino acid stress, and acted as a positive control for translationally upregulated mRNAs in our experiment. Although yeast GCN4 mRNA harbors four uORFs (Hinnebusch, 1997),

ATF4 is thought to possess only two uORFs in mammals, the second of which overlaps the coding ORF to inhibit its translation under normal conditions (Lu et al., 2004; Vattem & Wek, 2004). We looked closely at our footprinting data for ATF4, and we observed a peak 12-14 nucleotides upstream of the canonical first uORF, which is located 196 nucleotides upstream of the start of the coding ORF (Figure 7C). The position of the observed footprinting peak relative to uORF1 is expected, as it is approximately the distance from the edge of the ribosome to the A site of the ribosome. Furthermore, we observed footprint coverage along the subsequent 9 nucleotides, suggesting that this 3-amino-acid-long ORF is indeed translated (Figure 7C). We observed very low coverage 12-14 nucleotides upstream of the canonical second uORF, and hence could not reliably determine initiation at this uORF (Figure 7C). However, we did observe a low level of footprint coverage downstream of the canonical second uORF and upstream of the coding ORF start, suggesting that this uORF may be translated at a low level (Figure 7C).

When ternary complex formation is inhibited, such as by amino acid deprivation, the second canonical uORF is bypassed and initiation at the downstream coding ORF can proceed (Lu et al., 2004; Vattem & Wek, 2004). While we did observe a general increase in ribosome footprint coverage across the entire coding ORF two hours after amino acid deprivation (Figure 4B), we surprisingly observed a slight increase in coverage along canonical uORF2, although we still did not observe a peak 12-14 nucleotides upstream of the canonical uORF2 AUG (Figure 7B). Even more surprising, however, was the dramatic increase in footprints approximately 80 nucleotides

Figure 7: uORFs in Translational Regulation Following Amino Acid Deprivation

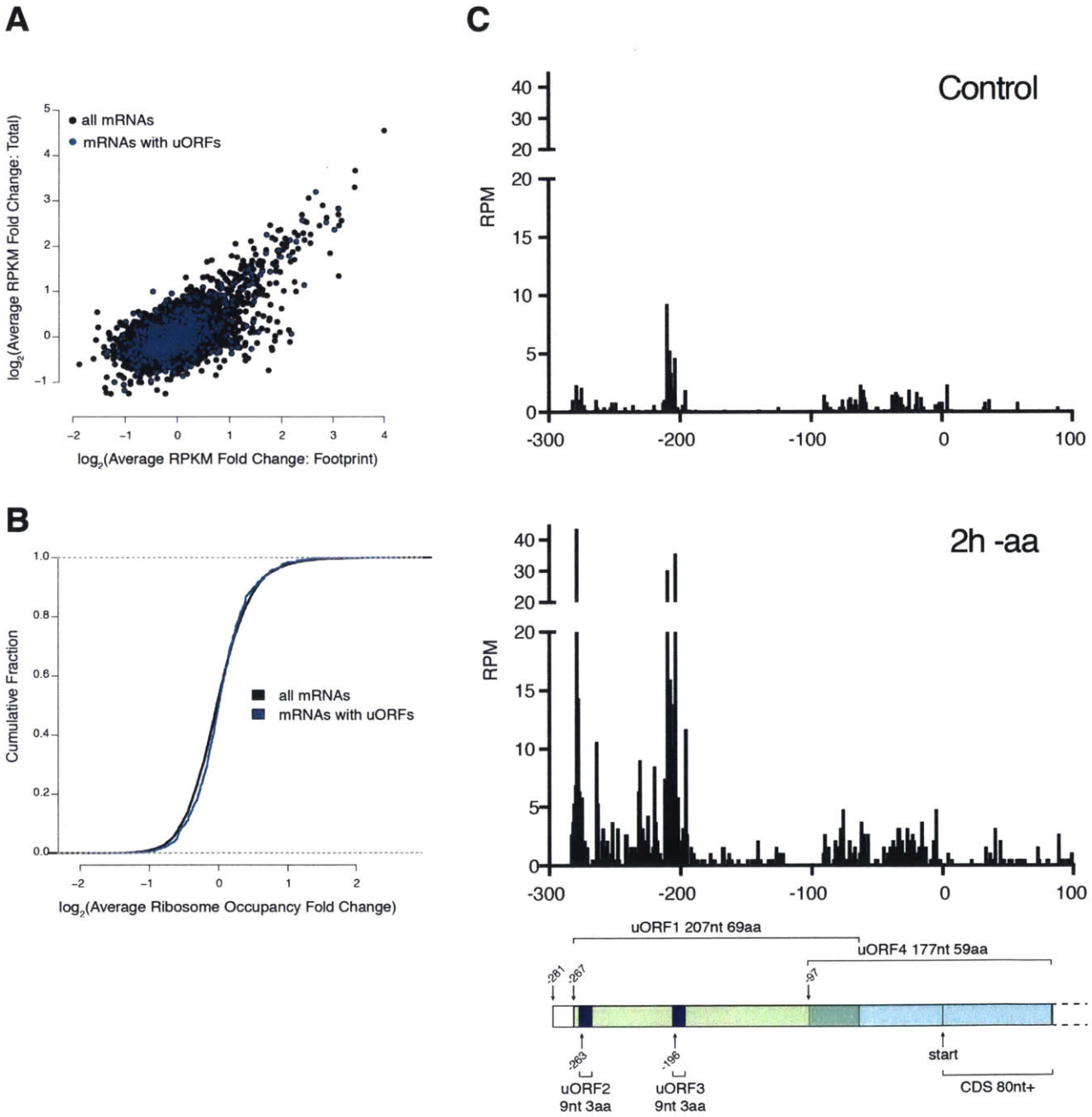


Figure 7: uORFs in Translational Regulation Following Amino Acid Deprivation

A. uORFs are not enriched in mRNAs whose translation is sensitive to amino acid deprivation. The $\log_2$ of the average ribosome footprint fold change upon amino acid starvation for two hours is plotted against the $\log_2$ average fold change in total mRNA for all mRNAs (indicated in black). mRNAs containing uORFs are indicated in blue.

B. The presence of a uORF is associated with milder translational repression after acute amino acid deprivation, but not with translational upregulation. The cumulative distribution of the $\log_2$-transformed average ribosome occupancy fold change upon amino acid starvation for two hours is plotted for all mRNAs (black) and for mRNAs containing uORFs (blue).

C. ATF4 contains two putative novel CUG uORFs, at least one of which is translated upon amino acid deprivation. Footprint reads mapping to the indicated region of ATF4 were normalized by number of total mapped reads and plotted against mRNA position, with the start codon set to zero. Each condition is indicated. The architecture of the mouse ATF4 mRNA is diagrammed. The two newly-identified uORFs are designated "uORF1" and "uORF2," while the canonical uORFs are designated "uORF3" and "uORF4."

upstream of the first canonical uORF (Figure 7C). These footprints were also observed in the amino acid replete condition, although at a much lower frequency (Figure 7C). We did not find any AUG codons in the mRNA sequence in this region, but we noticed two CUG codons spaced four nucleotides apart, located 12 and 16 nucleotides downstream of the position with dramatically increased footprint coverage. CUG codons are one of the most frequently used non-AUG start codons (Asano, 2014), supporting the notion that one or both of these uORFs are actually translated. The first putative new uORF is predicted to encode a protein 69 amino acids in length, whereas putative uORF2 encodes a three-amino-acid-long peptide. We cannot distinguish between the use of these two putative uORFs, but our data indicate that ribosomes initiate translation at

one or both of these positions, and that this initiation is amino acid-sensitive (Figure 7C). We propose that one or both of these uORFs are involved in the translational response of ATF4 to amino acid stress.

DISCUSSION

Translational regulation in response to stress has been well studied in many organisms. However, few mRNAs have been definitively shown to be translationally upregulated in response to stress (Le & Maizel, 1997; Lee, Cevallos, & Jan, 2009; Lu et al., 2004; Palam, Baird, & Wek, 2011; Vattem & Wek, 2004; Zhou et al., 2008), suggesting either a very limited translational response program, or an insufficiency of the methods used to identify new translationally regulated mRNAs. Understanding the translational response to a stress such as lack of amino acids is important for identifying candidates downstream of the translational regulators mTORC1 and GCN2 that could be targeted in disease. Using ribosome footprinting, we have identified mRNAs that translationally respond to amino acid stress, providing a snapshot of both acute and long-term translational regulation that was previously lacking.

It will be critical to determine which translationally upregulated mRNAs are functionally important for the response to stress, both in cells and *in vivo*. It is tantalizing to think that mRNAs which are upregulated in response to acute amino acid starvation, and which continue to be translated when cells are exposed to long-term amino acid stress, are more likely to be functionally relevant. However, it will be important to test the effect of loss-of-function of these candidates for proliferation and survival under amino acid stress *in vitro*, and for their contribution to tumor growth *in vivo*.

It will also be important to determine the relative contributions of the mTORC1 and GCN2 pathways to translational regulation under stress. The major downstream target of the GCN2 pathway, ATF4, was shown to contribute to tumor survival in a mouse

model (Ye et al., 2010), while the mTORC1 pathway is active in a variety of cancers (Lim, Crowe, & Yang, 2015); thus, inhibition of translation initiation may have opposing roles in cancer, depending on the regulatory pathway involved. Candidate mRNAs translationally regulated via mTORC1 could be sensitive to pharmacological inhibition of mTOR, and could serve as biomarkers for responsive tumors. There is currently no pharmacological inhibitor of the GCN2 pathway approved for use in the clinic, although several investigational compounds exist (Joshi, Kulkarni, & Pal, 2013). Future work could combine individual pharmacological inhibition of the mTORC1 and GCN2 pathways with ribosome footprinting analysis, to disentangle the roles of each pathway in translational control under stress. Alternatively, footprinting from GCN2 loss-of-function cells, such as GCN2-/- MEFs (Jiang et al., 2003), could be compared with footprinting from cells in which the mTORC1 pathway has been rendered resistant to amino acid starvation, such as $RagA^{GTP/GTP}$ knock-in MEFs (Efeyan et al., 2013). Preliminary experiments using these genetic models suggest that the GCN2 pathway is required for the major acute translational response to amino acid deprivation, but that the ability of mTORC1 to respond to amino acids is important for shaping global translation, even in amino acid replete conditions (data not shown).

Although we identified mRNAs that were also increased at the protein level upon amino acid deprivation, we suspect that many of the mRNAs translationally upregulated according to ribosome footprinting measurements actually harbor stalled ribosomes. Ribosome footprints could increase due to increased translation, or they could increase due to decreased elongation rates. Although amino acid deprivation is known to

dramatically decrease translation initiation rates (Kimball, 2002), we hypothesize that the limited availability of charged tRNAs dramatically decreases elongation rates as well. To distinguish the effects of amino acid deprivation on elongation from its effects on initiation, we could perform ribosome footprinting in the presence of harringtonine, a small molecule that blocks the first cycles of elongation, but allows elongation on polysomes to proceed (Fresno, Jimenez, & Vazquez, 1977). Comparison of footprinting results from harringtonine-treated cells deprived of amino acids with our current data would allow us to distinguish between mRNAs whose increased ribosome footprint is due to stalled ribosomes or decreased elongation rate, and mRNAs whose increased footprint is due to increased initiation.

Originally, ribosome footprinting measurements were compared with measurements of the proteome by mass spectrometry to validate the ribosome footprinting method (Ingolia et al., 2009). However, the correlation of ribosome occupancy with protein abundance for a single cell type under a single condition does not necessarily indicate congruity in these measurements across all cell types or conditions. In particular, we emphasize that validation of ribosome footprinting data at the protein level is particularly important for conditions that perturb translation at points other than initiation.

One goal of the ribosome footprinting performed herein was to identify mechanisms of translational regulation in response to amino acid deprivation. One potential mechanism for translational regulation is through a cap-independent mechanism, for example via the presence of an internal ribosome entry site (IRES), which could be utilized more efficiently when cap-dependent translation is dramatically decreased (Jackson, 2013;

Licursi, Komatsu, Pongnopparat, & Hirasawa, 2012). An IRES would be difficult to identify by sequence analysis alone, but could be identified by functional assays. Another mechanism is via the canonical uORF-dependent mechanism (Jackson, 2013). We have no evidence that the mRNAs translationally upregulated following amino acid deprivation are enriched in uORFs (Figure 7A and data not shown), suggesting that the effect of ternary complex depletion on translation of uORFs and downstream coding ORFs may be quite mRNA-specific. Furthermore, the positions of the ribosome footprints in the 5' UTR of ATF4 (Figure 7C) lead us to question the current model for ATF4 translational upregulation in mammals. At the very least, functional assays will be necessary to determine the effect of the newly identified putative uORFs on ATF4 regulation.

It has become clear that the translational stress response program plays a role in cancer, and it likely plays an important role in other disease states as well. Here, we present a transcriptome-wide snapshot of mRNAs that are translationally regulated in response to short- and long-term amino acid deprivation, with a particular emphasis on mRNAs that are translationally upregulated. This work extends our understanding of the translational program beyond the well-studied control of the transcription factor ATF4, and suggests that alternative, although currently unexplored, mechanisms exist beyond the regulation by uORFs for influencing translation rates under stress.

MATERIALS AND METHODS

Reagents

Primary antibodies were purchased from Cell Signaling Technology (mTOR #2972, Rictor #2140, pT389 S6K #9234, 4EBP1 #9452, pS51 eIF2a #9722, LC3b #2775, Iqgap1 #2293, Sqstm1 #5114, Myc #5605, Jun #9165, Cdc6 #3387), EMDMillipore (Raptor #09-217), and Santa Cruz Biotechnology (CREB-2 (ATF4) #SC-200). HRP-conjugated econdary antibodies were purchased from Santa Cruz Biotechnology. Enzymes were purchased from New England Biolabs, except where noted. Primers were obtained from Integrated DNA Technologies. Antibiotics and chemicals were purchased from Sigma, except where noted. TBE-Urea gels, TBE gels, and sample loading dyes were purchased from ThermoFisher. RPMI and fetal bovine serum (FBS) were purchased from US Biologicals.

Amino Acid Starvation

MEFs immortalized with SV40 large T antigen were seeded at a density of 3, 4, or $6x10^6$ cells (corresponding to samples for no amino acid starvation, two hours amino acid starvation, or 18 hours amino acid starvation) in 15cm dishes in 25ml complete RPMI supplemented with 10% FBS and penicillin/streptomycin (day 0). The following day (day 1), the 18-hour time point cells were washed three times with 20ml warm PBS, and RPMI without amino acids supplemented with 10% dialyzed FBS and penicillin/streptomycin was added. On day 2, the two-hour time point cells were washed three times with 20ml warm PBS, and RPMI without amino acids supplemented with

10% dialyzed FBS and penicillin/streptomycin was added. The media on the control cells was removed and replaced with complete RPMI supplemented with 10% FBS and penicillin/streptomycin for two hours. Prior to harvest, cells were treated with 100 µg/ml cycloheximide in ethanol for 5 minutes at 37C in batches spaced 15 minutes apart. Each condition was harvested on ice in a batch every 15 minutes.

Cell Lysis for Polysome Analysis and Ribosome Footprinting

Cells were washed with 10ml ice-cold PBS containing 100µg/ml cycloheximide, and lysed in 750µl Polysome Lysis Buffer (15 mM HEPES pH 7.4, 7.5 mM $MgCl_2$, 100 mM NaCl, 1% TritonX-100, 2 mM DTT (ThermoFisher), one EDTA-free protease inhibitor tablet per 10ml buffer (Roche), and 100 µg/ml cycloheximide). Cells were incubated for 10min at 4C with end-over-end rotation, and lysate was cleared by centrifugation for 5 minutes at 13,300xg at 4C. Soluble material was transferred to new tubes, and protein concentrations were normalized between conditions.

Sucrose Gradient Preparation and Polysome Analysis

10%-50% sucrose gradients were prepared the day before use by underlaying 5.5ml 10% Polysome Gradient Buffer (15 mM HEPES pH 7.4, 7.5 mM $MgCl_2$, 100 mM NaCl, 100 µg/ml cycloheximide, 2 mM DTT, 20U/ml SuperaseIn (ThermoFisher), and 10% sucrose) with 5.5ml 50% Polysome Gradient Buffer (same as 10% Polysome Gradient Buffer, but containing 50% sucrose). Tubes were parafilmed and placed horizontally for 4-6 hours at 4C, then carefully turned upright and kept at 4C overnight. The following

day, 750µl-850µl of cell lysate was loaded onto each gradient and ultracentrifuged for two hours at 36,000*xg* at 4C. Polysomes were detected by monitoring absorbance at 254 nm, and fractionated into fractions of approximately equal volume.

Polysome Gradient-qPCR

For polysome gradient-qPCR analysis, 500µl of each fraction was transferred to a new tube containing 5 ng of firefly luciferase (Promega) and 250µl Polysome Lysis Buffer. SDS was added to 0.5% and Proteinase K (Roche) was added to 200 µg/ml, then samples were incubated for 45 minutes at 50C. RNA was extracted twice with 750µl acid phenol/chloroform (ThermoFisher), once for 5 minutes at 65C and once for 5 minutes at room temperature. The aqueous layer was extracted with 750µl chloroform, and precipitated with NaOAc and isopropanol for 30 minutes at -20C. RNA was recovered by centrifugation for 30 minutes at 13,300*xg*, pellets were washed once with ice-cold 80% ethanol, and resuspended in 10µl 10 mM Tris pH 8. cDNA synthesis was performed with Superscript III (ThermoFisher) according to manufacturer's instructions using an oligodT_{20} primer for reverse transcription. qPCR was performed with 2.5µl cDNA per 10µl reaction using 2xSYBR Green master mix (Roche) according to manufacturer's instructions, with primers present at a concentration of 5 µM. All values for mRNAs assayed from polysome gradient fractions were normalized to the spiked in firefly luciferase control, and to the total level of the mRNA (see total mRNA preparation, below).

Total RNA Preparation

Total mRNA was prepared using RNA-STAT 60 (Amsbio). Briefly, media was removed from cells seeded and treated in parallel with those used for polysome analysis, and 1ml of RNA-STAT 60 was added. Cells were scraped and homogenized by pipetting. Homogenate was incubated for 5 minutes at room temperature and transferred to tubes. 200µl of chloroform was added, and samples were vigorously mixed for 30 seconds at room temperature. Samples were incubated for 2 minutes at room temperature, followed by centrifugation at 12,000*xg* for 15 minutes at 4C. The aqueous layer was transferred to new tubes and RNA was precipitated with NaOAc and isopropanol as described above. RNA was resuspended in 20µl 10 mM Tris pH 8. 2 µg RNA was used per 20µl cDNA reaction with Superscript III according to manufacturer's instructions. qPCR was performed as described above. Each mRNA assayed by qPCR was normalized to the control mRNA Rplp0.

Ribosome Footprint and Sequencing Library Preparation

Ribosome footprints were prepared using a modified version of the protocol described by Ingolia *et al.* (Ingolia et al., 2009). Briefly, 600µl lysate was mixed with 700 mM NaCl Polysome Lysis Buffer (15 mM HEPES pH 7.4, 7.5 mM $MgCl_2$, 700 mM NaCl, 1% TritonX-100, and 2 mM DTT), and RNAse I (ThermoFisher) was added to 0.1U/µl. Samples were digested for 45 minutes at room temperature with mixing. Samples were proteinase K digested, acid phenol/chloroform extracted, and NaOAc/isopropanol precipitated as described above.

All other steps for footprint and total mRNA processing for sequencing library preparation were performed as described (Ingolia et al., 2009).

Data Analysis

Raw sequencing data was processed as described previously (Ingolia et al., 2009), using custom Python scripts. Statistically significant changes in total mRNA and footprints were determined using the EdgeR package (Robinson, McCarthy, & Smyth, 2010). All other data manipulation was performed in R.

uORF Analysis

mRNAs containing uORFs were extracted from published mouse ES cell data (Ingolia et al., 2011).

Primers

Mapk3_F: ACAGATGAGCCAGTGGCCGAG
Mapk3_R: TGTTAGGGGCCCTCTGGCGC
ATF4_F: GGC CTA AGC CAT GGC GCT CTT C
ATF4_R: CCC CCA AAC CCG ACT GGT CG
Mafg_F: GCTCACTGTGTCCCCCGGGTTA
Mafg_R: CCTGGCTCCCGCTTCACCTTTAAG
Iqgap1_F: CTACGGTCGCCCGCGTCTTC
Iqgap1_R: CCAGGACGGAGCCATAGTGCG
Ahnak_F: GCTCTGAAGTGGTTCTGAGCGGG
Ahnak_R: TCCTCCGACCTCAGGCGTGG
Ppp1r15a_F: GCTCGGAAGGTACACTTCGCTGAGA
Ppp1r15a_R: ATGCGTCGAGCAAAGCGGCT
Jun_F: CCCGCGTGAAGTGACGGACC
Jun_R: AGTTCTTGGCGCGGAGGTGC
Ccnl1_F: CCCTGGTTCAAACTGCCTGGAAT
Ccnl1_R: ACCAATGGGGCCGAGTTGGC

Cdc6_F: GTGTCCAGAGAGTTGTGTGGCCAT
Cdc6_R: ACGTCCACACACGGAGTGGTG
Rplp0_F: GTG TAC TCA GTC TCC ACA GA
Rplp0_R: GGA GCC ATG GAT TGC ACA TT

ACKNOWLEDGEMENTS

We would like to especially thank Carson Thoreen for sharing his knowledge of ribosome footprinting, for generously sharing his custom Python scripts for data analysis, and overall for his collegial, helpful, and inspirational attitude. We would like to thank Peggy Hsu and Larry Schweitzer for helpful discussions and suggestions. We would also like to thank the Whitehead Institute Genome Technology Center for sequencing. This work was supported by grants from the NIH (R01 CA129105, R01 CA103866, R37 AI047389) to D.M.S. D.M.S. is also an investigator of the Howard Hughes Medical Institute.

REFERENCES

Ackerman, D., & Simon, M. C. (2014). Hypoxia, lipids, and cancer: surviving the harsh tumor microenvironment. *Trends Cell Biol, 24*(8), 472-478. doi: 10.1016/j.tcb.2014.06.001

Agrawal, V., Alpini, S. E., Stone, E. M., Frenkel, E. P., & Frankel, A. E. (2012). Targeting methionine auxotrophy in cancer: discovery & exploration. *Expert Opin Biol Ther, 12*(1), 53-61. doi: 10.1517/14712598.2012.636349

Asano, K. (2014). Why is start codon selection so precise in eukaryotes? *Translation (Austin), 2*(1), e28387. doi: 10.4161/trla.28387

Baracos, V. E., & Mackenzie, M. L. (2006). Investigations of branched-chain amino acids and their metabolites in animal models of cancer. *J Nutr, 136*(1 Suppl), 237S-242S.

Bhat, M., Robichaud, N., Hulea, L., Sonenberg, N., Pelletier, J., & Topisirovic, I. (2015). Targeting the translation machinery in cancer. *Nat Rev Drug Discov, 14*(4), 261-278. doi: 10.1038/nrd4505

Brunn, G. J., Hudson, C. C., Sekulic, A., Williams, J. M., Hosoi, H., Houghton, P. J., . . . Abraham, R. T. (1997). Phosphorylation of the translational repressor PHAS-I by the mammalian target of rapamycin. *Science, 277*(5322), 99-101.

Cencic, R., Carrier, M., Galicia-Vazquez, G., Bordeleau, M. E., Sukarieh, R., Bourdeau, A., . . . Pelletier, J. (2009). Antitumor activity and mechanism of action of the cyclopenta[b]benzofuran, silvestrol. *PLoS One, 4*(4), e5223. doi: 10.1371/journal.pone.0005223

Chen, L., Aktas, B. H., Wang, Y., He, X., Sahoo, R., Zhang, N., . . . Halperin, J. A. (2012). Tumor suppression by small molecule inhibitors of translation initiation. *Oncotarget, 3*(8), 869-881. doi: 10.18632/oncotarget.598

Coles, N. W., & Johnstone, R. M. (1962). Glutamine metabolism in Ehrlich ascites-carcinoma cells. *Biochem J, 83*, 284-291.

DeBerardinis, R. J., Mancuso, A., Daikhin, E., Nissim, I., Yudkoff, M., Wehrli, S., & Thompson, C. B. (2007). Beyond aerobic glycolysis: transformed cells can engage in glutamine metabolism that exceeds the requirement for protein and nucleotide synthesis. *Proc Natl Acad Sci U S A, 104*(49), 19345-19350. doi: 10.1073/pnas.0709747104

Dennis, G., Jr., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., & Lempicki, R. A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol, 4*(5), P3.

Efeyan, A., Zoncu, R., Chang, S., Gumper, I., Snitkin, H., Wolfson, R. L., . . . Sabatini, D. M. (2013). Regulation of mTORC1 by the Rag GTPases is necessary for neonatal autophagy and survival. *Nature, 493*(7434), 679-683. doi: 10.1038/nature11745

Emadi, A., Zokaee, H., & Sausville, E. A. (2014). Asparaginase in the treatment of non-ALL hematologic malignancies. *Cancer Chemother Pharmacol, 73*(5), 875-883. doi: 10.1007/s00280-014-2402-3

Fresno, M., Jimenez, A., & Vazquez, D. (1977). Inhibition of translation in eukaryotic systems by harringtonine. *Eur J Biochem, 72*(2), 323-330.

Grabiner, B. C., Nardi, V., Birsoy, K., Possemato, R., Shen, K., Sinha, S., . . . Sabatini, D. M. (2014). A diverse array of cancer-associated MTOR mutations are hyperactivating and can predict rapamycin sensitivity. *Cancer Discov, 4*(5), 554-563. doi: 10.1158/2159-8290.CD-13-0929

Harding, H. P., Zhang, Y., Zeng, H., Novoa, I., Lu, P. D., Calfon, M., . . . Ron, D. (2003). An integrated stress response regulates amino acid metabolism and resistance to oxidative stress. *Mol Cell, 11*(3), 619-633.

Hinnebusch, A. G. (1997). Translational regulation of yeast GCN4. A window on factors that control initiator-trna binding to the ribosome. *J Biol Chem, 272*(35), 21661-21664.

Holcik, M., & Sonenberg, N. (2005). Translational control in stress and apoptosis. *Nat Rev Mol Cell Biol, 6*(4), 318-327. doi: 10.1038/nrm1618

Hsieh, A. C., Costa, M., Zollo, O., Davis, C., Feldman, M. E., Testa, J. R., . . . Ruggero, D. (2010). Genetic dissection of the oncogenic mTOR pathway reveals druggable addiction to translational control via 4EBP-eIF4E. *Cancer Cell, 17*(3), 249-261. doi: 10.1016/j.ccr.2010.01.021

Ingolia, N. T., Brar, G. A., Rouskin, S., McGeachy, A. M., & Weissman, J. S. (2012). The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat Protoc, 7*(8), 1534-1550. doi: 10.1038/nprot.2012.086

Ingolia, N. T., Ghaemmaghami, S., Newman, J. R., & Weissman, J. S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science, 324*(5924), 218-223. doi: 10.1126/science.1168978

Ingolia, N. T., Lareau, L. F., & Weissman, J. S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell, 147*(4), 789-802. doi: 10.1016/j.cell.2011.10.002

Jackson, R. J. (2013). The current status of vertebrate cellular mRNA IRESs. *Cold Spring Harb Perspect Biol, 5*(2). doi: 10.1101/cshperspect.a011569

Jackson, R. J., Hellen, C. U., & Pestova, T. V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol, 11*(2), 113-127. doi: 10.1038/nrm2838

Jiang, H. Y., Wek, S. A., McGrath, B. C., Scheuner, D., Kaufman, R. J., Cavener, D. R., & Wek, R. C. (2003). Phosphorylation of the alpha subunit of eukaryotic initiation factor 2 is required for activation of NF-kappaB in response to diverse cellular stresses. *Mol Cell Biol, 23*(16), 5651-5663.

Joshi, M., Kulkarni, A., & Pal, J. K. (2013). Small molecule modulators of eukaryotic initiation factor 2alpha kinases, the key regulators of protein synthesis. *Biochimie, 95*(11), 1980-1990. doi: 10.1016/j.biochi.2013.07.030

Kilberg, M. S., Balasubramanian, M., Fu, L., & Shan, J. (2012). The transcription factor network associated with the amino acid response in mammalian cells. *Adv Nutr, 3*(3), 295-306. doi: 10.3945/an.112.001891

Kimball, S. R. (2002). Regulation of global and specific mRNA translation by amino acids. *J Nutr, 132*(5), 883-886.

Kuo, M. T., Savaraj, N., & Feun, L. G. (2010). Targeted cellular metabolism for cancer chemotherapy with recombinant arginine-degrading enzymes. *Oncotarget, 1*(4), 246-251. doi: 10.18632/oncotarget.135

Labuschagne, C. F., van den Broek, N. J., Mackay, G. M., Vousden, K. H., & Maddocks, O. D. (2014). Serine, but not glycine, supports one-carbon metabolism and proliferation of cancer cells. *Cell Rep, 7*(4), 1248-1258. doi: 10.1016/j.celrep.2014.04.045

Laplante, M., & Sabatini, D. M. (2012). mTOR signaling in growth control and disease. *Cell, 149*(2), 274-293. doi: 10.1016/j.cell.2012.03.017

Lareau, L. F., Hite, D. H., Hogan, G. J., & Brown, P. O. (2014). Distinct stages of the translation elongation cycle revealed by sequencing ribosome-protected mRNA fragments. *Elife, 3*, e01257. doi: 10.7554/eLife.01257

Le, S. Y., & Maizel, J. V., Jr. (1997). A common RNA structural motif involved in the internal initiation of translation of cellular mRNAs. *Nucleic Acids Res, 25*(2), 362-369.

Lee, Y. Y., Cevallos, R. C., & Jan, E. (2009). An upstream open reading frame regulates translation of GADD34 during cellular stresses that induce eIF2alpha phosphorylation. *J Biol Chem, 284*(11), 6661-6673. doi: 10.1074/jbc.M806735200

Licursi, M., Komatsu, Y., Pongnopparat, T., & Hirasawa, K. (2012). Promotion of viral internal ribosomal entry site-mediated translation under amino acid starvation. *J Gen Virol, 93*(Pt 5), 951-962. doi: 10.1099/vir.0.040386-0

Lim, H. J., Crowe, P., & Yang, J. L. (2015). Current clinical regulation of PI3K/PTEN/Akt/mTOR signalling in treatment of human cancer. *J Cancer Res Clin Oncol, 141*(4), 671-689. doi: 10.1007/s00432-014-1803-3

Lu, P. D., Harding, H. P., & Ron, D. (2004). Translation reinitiation at alternative open reading frames regulates gene expression in an integrated stress response. *J Cell Biol, 167*(1), 27-33. doi: 10.1083/jcb.200408003

Marcotrigiano, J., Gingras, A. C., Sonenberg, N., & Burley, S. K. (1999). Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. *Mol Cell, 3*(6), 707-716.

Mayers, J. R., & Vander Heiden, M. G. (2015). Famine versus feast: understanding the metabolism of tumors in vivo. *Trends Biochem Sci, 40*(3), 130-140. doi: 10.1016/j.tibs.2015.01.004

Mortimore, G. E., & Poso, A. R. (1987). Intracellular protein catabolism and its control during nutrient deprivation and supply. *Annu Rev Nutr, 7*, 539-564. doi: 10.1146/annurev.nu.07.070187.002543

Orsolic, I., Jurada, D., Pullen, N., Oren, M., Eliopoulos, A. G., & Volarevic, S. (2016). The relationship between the nucleolus and cancer: Current evidence and emerging paradigms. *Semin Cancer Biol, 37-38*, 36-50. doi: 10.1016/j.semcancer.2015.12.004

Palam, L. R., Baird, T. D., & Wek, R. C. (2011). Phosphorylation of eIF2 facilitates ribosomal bypass of an inhibitory upstream ORF to enhance CHOP translation. *J Biol Chem, 286*(13), 10939-10949. doi: 10.1074/jbc.M110.216093

Patil, M. D., Bhaumik, J., Babykutty, S., Banerjee, U. C., & Fukumura, D. (2016). Arginine dependence of tumor cells: targeting a chink in cancer's armor. *Oncogene.* doi: 10.1038/onc.2016.37

Reitzer, L. J., Wice, B. M., & Kennell, D. (1979). Evidence that glutamine, not sugar, is the major energy source for cultured HeLa cells. *J Biol Chem, 254*(8), 2669-2676.

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics, 26*(1), 139-140. doi: 10.1093/bioinformatics/btp616

Sonenberg, N., & Hinnebusch, A. G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell, 136*(4), 731-745. doi: 10.1016/j.cell.2009.01.042

Tanida, I., Ueno, T., & Kominami, E. (2008). LC3 and Autophagy. *Methods Mol Biol, 445*, 77-88. doi: 10.1007/978-1-59745-157-4_4

Thoreen, C. C. (2013). Many roads from mTOR to eIF4F. *Biochem Soc Trans, 41*(4), 913-916. doi: 10.1042/BST20130082

Vattem, K. M., & Wek, R. C. (2004). Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proc Natl Acad Sci U S A, 101*(31), 11269-11274. doi: 10.1073/pnas.0400541101

Wang, Y., Ning, Y., Alam, G. N., Jankowski, B. M., Dong, Z., Nor, J. E., & Polverini, P. J. (2013). Amino acid deprivation promotes tumor angiogenesis through the GCN2/ATF4 pathway. *Neoplasia, 15*(8), 989-997.

Wethmar, K. (2014). The regulatory potential of upstream open reading frames in eukaryotic gene expression. *Wiley Interdiscip Rev RNA, 5*(6), 765-778. doi: 10.1002/wrna.1245

Yang, H. S., Jansen, A. P., Komar, A. A., Zheng, X., Merrick, W. C., Costes, S., . . . Colburn, N. H. (2003). The transformation suppressor Pdcd4 is a novel eukaryotic translation initiation factor 4A binding protein that inhibits translation. *Mol Cell Biol, 23*(1), 26-37.

Ye, J., Kumanova, M., Hart, L. S., Sloane, K., Zhang, H., De Panis, D. N., . . . Koumenis, C. (2010). The GCN2-ATF4 pathway is critical for tumour cell survival and proliferation in response to nutrient deprivation. *EMBO J, 29*(12), 2082-2096. doi: 10.1038/emboj.2010.81

Zaytseva, Y. Y., Valentino, J. D., Gulhati, P., & Evers, B. M. (2012). mTOR inhibitors in cancer therapy. *Cancer Lett, 319*(1), 1-7. doi: 10.1016/j.canlet.2012.01.005

Zhang, J., Fan, J., Venneti, S., Cross, J. R., Takagi, T., Bhinder, B., . . . Thompson, C. B. (2014). Asparagine plays a critical role in regulating cellular adaptation to glutamine depletion. *Mol Cell, 56*(2), 205-218. doi: 10.1016/j.molcel.2014.08.018

Zhou, D., Palam, L. R., Jiang, L., Narasimhan, J., Staschke, K. A., & Wek, R. C. (2008). Phosphorylation of eIF2 directs ATF5 translational control in response to diverse stress conditions. *J Biol Chem, 283*(11), 7064-7073. doi: 10.1074/jbc.M708530200

Zong, Q., Schummer, M., Hood, L., & Morris, D. R. (1999). Messenger RNA translation state: the second dimension of high-throughput expression screening. *Proc Natl Acad Sci U S A, 96*(19), 10632-10636.

Chapter 3

Investigating the Role of eIF4E in the Translational Regulation of 5'TOP mRNAs

Heather R. Keys[1-5], Carson C. Thoreen[6], Lynne Chantranupong[1-5], Tim Wang[1-5], and David M. Sabatini[1-5]

[1]Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA

[2]Department of Biology, MIT, Cambridge, MA 02139, USA

[3]Howard Hughes Medical Institute, MIT, Cambridge, MA 02139, USA

[4]Broad Institute, Cambridge, MA 02142, USA

[5]The David H. Koch Institute for Integrative Cancer Research, Cambridge, MA 02139, USA

[6]Department of Cellular and Molecular Physiology, Yale University, New Haven, CT 06510, USA

A portion of this work has been published in:
A unifying model for mTORC1-mediated regulation of mRNA translation. Thoreen CC, Chantranupong L, Keys HR, Wang T, Gray NS, and Sabatini DM. *Nature*. 2012 May 3;485(7396):109-13.

All experiments and analysis were performed by HRK, except Figure 1A-1D and 1F. Figures 1A-1D and 1F were performed and analyzed primarily by CCT, with experimental assistance from LC and HRK and analysis assistance by TW.

SUMMARY

The translation of nearly all eukaryotic mRNAs is facilitated by a cytoplasmic protein complex called eIF4F that recognizes the 5' terminal cap structure. Deregulation of eIF4F frequently occurs in common human diseases such as cancer, and is capable of driving tumorigenesis, altering energy metabolism, and causing a spectrum of autism-like disorders in experimental systems. An important regulator of eIF4F is the 4EBP family of translational repressors that are regulated by the mTORC1 signaling pathway. We show that 4EBPs selectively inhibit the translation of mRNAs with a 5' terminal oligopyrimidine (5'TOP) motif, which encode many proteins required for growth, including nearly all components of the translation machinery. 5'TOP mRNAs are preferentially lost from eIF4E, but not the other eIF4F components eIF4G and eIF4A, when mTOR is inhibited, and the extent of loss of mRNAs from eIF4E correlates with the degree of translational inhibition. Moreover, capped 5'UTRs with pyrimidine-rich sequences near the 5' terminus show a markedly lower affinity for eIF4F *in vitro*, with varied effects on both the on-rate and off-rate of eIF4E and eIF4G. Furthermore, recombinant eIF4E exhibits binding differences for 5'UTRs with pyrimidine- and purine-rich cap-proximal sequences, indicating that eIF4E drives the affinity differences observed with eIF4F. Our results suggest that 5'TOP mRNAs are dependent on high concentrations of eIF4F for efficient translation, explaining why their translation is selectively impaired by perturbations that decrease eIF4F stability, such as activation of 4EBPs by mTORC1. However, preliminary data refutes this model, as eIF4E binds a reporter mRNA harboring a true 5'TOP motif similarly to a reporter mRNA containing a mutant 5'TOP motif *in vitro*, despite differences in the degree of translational repression

of these mRNAs when mTOR is inhibited. We also present preliminary evidence that some endogenous 5'TOP mRNAs, which show decreased eIF4E binding and translational repression upon mTOR inhibition, actually have a higher affinity for eIF4E than mRNAs that do not contain a 5'TOP motif. In summary, although eIF4F and eIF4E exhibit binding preferences for 5'UTRs and endogenous mRNAs with different cap-proximal sequences, these binding differences do not explain the preferential translational repression of 5'TOP mRNAs observed with mTOR inhibition.

INTRODUCTION

The translation of messenger RNAs into proteins is a tightly regulated and essential cellular process, and is important not only for cellular homeostasis, but also for a myriad of processes from embryonic development to adaptation to stress. It is therefore unsurprising that mRNA translation is deregulated in a variety of diseases, including cancer (Bhat et al., 2015). Translation in eukaryotes is regulated at multiple steps, namely the initiation and elongation phases, and is even regulated through multiple mechanisms at the initiation step (Fraser, 2015; Sonenberg & Hinnebusch, 2009). One mechanism for regulation of translation initiation is through the cap binding complex eIF4F. eIF4F is comprised of three initiation factors: eIF4E, which directly binds the 5' cap of mRNA; eIF4G, a scaffolding molecule which binds RNA, eIF4E, polyA binding protein, and other initiation factors associated with the 40S ribosomal subunit, thereby bridging the mRNA and the ribosome; and eIF4A, an RNA helicase that associates with eIF4G and is thought to participate in secondary structure resolution in the 5'UTR of mRNAs initiating translation (Merrick, 2015; Prevot, Darlix, & Ohlmann, 2003).

One mechanism of eIF4F regulation is through eIF4E-binding proteins (4EBPs). 4EBPs are phosphorylated by the protein complex mTORC1 (mechanistic target of rapamycin complex 1) under conditions of nutrient and growth factor sufficiency, thereby maintaining 4EBPs in an inactive form that is unable to interact with eIF4E (Bah et al., 2015; Brunn et al., 1997; Pause et al., 1994). Upon inhibition of mTORC1 through perturbations such as amino acid deprivation, serum starvation, or treatment with mTORC1 inhibitors like rapamycin or Torin1, 4EBPs are dephosphorylated, bind to

eIF4E, and block the interaction of eIF4E with eIF4G (Marcotrigiano, Gingras, Sonenberg, & Burley, 1999).

For more than two decades it has been known that inhibition of mTORC1 causes selective translational repression of a subset of mRNAs harboring a stretch of pyrimidines directly downstream of the 5' 7-methylguanosine cap (Jefferies, Reinhard, Kozma, & Thomas, 1994). These mRNAs containing a 5' terminal oligopyrimidine tract (5'TOP mRNAs) are particularly important, because they encode many proteins of the translational machinery, including ribosomal proteins and initiation and elongation factors. Furthermore, mRNAs encoding ribosomal proteins comprise a substantial percentage of the cellular mRNA pool: in yeast, this estimate is around 30% (Warner, 1999). The translational regulation of 5'TOP mRNAs is a crucial link between cell state cues, such as nutrient sufficiency, and the production of biomass for cell growth and division. Unsurprisingly, there has been great interest in uncovering the most proximal regulator of 5'TOP mRNA translation. There is evidence that this factor acts in *trans*, and several proteins and even a microRNA have been proposed to mediate both positive and negative translational regulation of these mRNAs in different cellular contexts (Biberman & Meyuhas, 1999; Damgaard & Lykke-Andersen, 2011; Kakegawa et al., 2007; Orom, Nielsen, & Lund, 2008; Tcherkezian et al., 2014). However, the search for a universal regulator of 5'TOP mRNA translation still continues.

Here we show that translational repression of 5'TOP mRNAs by mTORC1 inhibition is almost completely dependent on 4EBPs. Additionally, we show a correlation between the loss of mRNA from eIF4E and the degree of translational repression upon mTORC1

inhibition. Capped 5'UTRs with cap-proximal pyrimidine-rich sequences bind eIF4F and eIF4E poorly compared with purine-rich sequences *in vitro*, suggesting that eIF4E may interact differently with mRNAs and perhaps the larger eIF4F complex in the context of mRNAs with different cap-proximal sequences, and that 5'TOP mRNAs may rely more heavily on high levels of eIF4F for their translation. Although eIF4E shows a preference for purine-rich cap-proximal sequences *in vitro*, preliminary evidence indicates that eIF4E does not prefer non-TOP mRNAs to mRNAs containing a true 5'TOP motif. Although eIF4F affinity for mRNAs with different cap-proximal nucleotide compositions could still contribute to translational regulation upon mTORC1 inhibition, we conclude that eIF4E affinity does not solely determine this phenomenon.

RESULTS

Acute mTOR Inhibition Causes Translational Repression of 5'TOP mRNAs Primarily Through the 4EBPs

We characterized the global translational response of cells to acute inhibition of mTOR by polysome profiling. We treated p53-/- mouse embryonic fibroblasts (MEFs) for two hours with the mTOR inhibitor Torin1, and visualized the polysome profiles of control and treated cells by separating ribosomes on a sucrose gradient, followed by detection of RNA by absorbance at 254 nm. mTOR inhibition caused a severe defect in translation initiation, as evidenced by a dramatic increase in the 80S monosome peak, with a simultaneous reduction in polysomes (Figure 1A).

To measure the translation of individual mRNAs on a transcriptome-wide scale, we performed ribosome footprinting on MEFs treated with vehicle or the ATP-competitive inhibitor Torin1 for two hours (Thoreen et al., 2009). The ribosome footprinting method uses deep sequencing to quantify the number of ribosome-protected fragments on each individual mRNA (ribosome footprints). Ribosome footprints measurements are compared with measurements of the total level of each mRNA in the sample to calculate the ribosome occupancy, an indicator of how highly-translated an mRNA is.

Ribosome footprinting revealed that 5'TOP mRNAs were among the mRNAs whose translation was most inhibited by Torin1 treatment (Figure 1B). 5'TOP mRNAs encode mainly components of the translational machinery; hence, their translation is particularly important for sustaining the entire process of translation. We also isolated mRNAs from vehicle- or Torin1-treated MEFs by sucrose gradient fractionation, and subjected them

to qPCR analysis. Although an mRNA that lacks a 5'TOP motif, Actb, showed a shift from heavier to lighter polysomes in the Torin1-treated sample, indicating a decrease in translation (Figure 1C), the decrease was quite modest compared with the dramatic shift from polysomes to monosomes seen for mRNAs containing a 5'TOP motif (Figure 1C). Furthermore, even mRNAs containing an oligopyrimidine stretch downstream of a single purine (TOP-like mRNAs), such as Ybx1, showed this striking shift from polysomes (Figure 1C). These data indicate a difference in the regulation of mRNA translation between 5'TOP and non-TOP mRNAs downstream of mTOR inhibition.

A number of mTORC1 substrates play roles in translation, and it was unclear which of these contributed most highly to the observed translational repression of 5'TOP mRNAs. We performed ribosome footprinting on p53-null MEFs also lacking both 4EBP1 and 4EBP2 (4EBP1/2 double knockout MEFs; 4EBP1/2 DKO MEFs) treated with vehicle or Torin1 for two hours to determine their influence on the translational response to mTOR inhibition. It is worth noting that these cells were reported to lack expression of the third canonical 4EBP family member, 4EBP3, and thus effectively lack all three proteins (Dowling et al., 2010). In the absence of 4EBPs, Torin1 has very little effect on global translation (Figure 1D). Similarly, translation of 5'TOP mRNAs is almost unaffected by mTOR inhibition in 4EBP1/2 DKO MEFs (Figure 1D), suggesting that the translational regulation of 5'TOP mRNAs downstream of mTOR inhibition occurs primarily through 4EBPs.

When mTORC1 is inhibited, the 4EBPs disrupt the eIF4F complex by competing with eIF4G for eIF4E binding (Marcotrigiano et al., 1999). Since 4EBPs were required for the

Figure 1: mTOR Translationally Regulates 5'TOP mRNAs

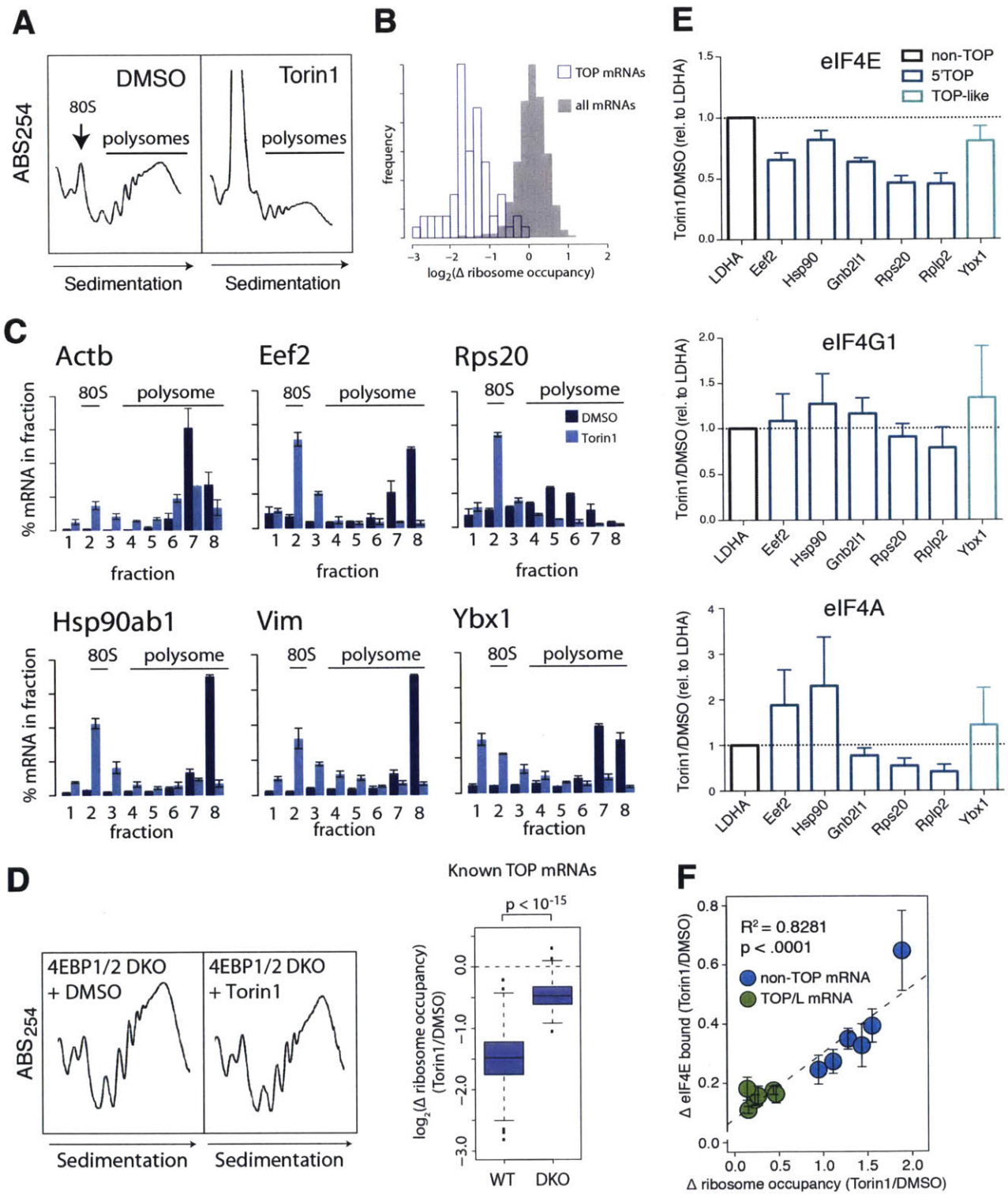


Figure 1: mTOR translationally regulates 5'TOP mRNAs

A. mTOR inhibition represses translation in MEFs. MEFs were treated with vehicle or 250 nM Torin1 for 2 hours. Lysates were fractionated on a sucrose gradient and polysome profiles were detected by absorbance at 254nm.

B. mTOR inhibition especially decreases translation of 5'TOP mRNAs. Ribosome footprints reads were normalized to total mRNA levels for each mRNA for vehicle- and Torin1-treated samples to calculate the ribosome occupany. The $\log_2$ fold change in ribosome occupancy was plotted for all mRNAs and for 5'TOP mRNAs.

C. mTOR inhibition greatly inhibits the association of 5'TOP mRNAs with polysomes. The fraction of each mRNA associated with the indicated gradient fraction is plotted for each condition. Data are expressed as the mean ± s.e.m. (n = 4).

D. Translational repression caused by mTOR inhibition is mostly 4EBP-dependent. 4EBP1/2 DKO MEFs were treated with vehicle or 250 nM Torin1 for 2 hours. Lysates were fractionated on a sucrose gradient and polysome profiles were detected by absorbance at 254 nm. The $\log_2$ fold change in ribosome occupancy for 5'TOP mRNAs after treatment with Torin1 is plotted for wildtype and 4EBP1/2 DKO MEFs.

E. 5'TOP mRNAs are preferentially lost from eIF4E following mTOR inhibition. qPCR was performed on RNA extracted from immunoprecipitates of the indicated proteins stably expressed in HEK-293T cells following vehicle or 250 nM Torin1 treatment for two hours. Values were normalized to a spiked in luciferase control and to total levels of each indicated mRNA, then the fold difference relative to vehicle treatment was calculated. It should be noted that the fold difference after Torin1 treatment was decreased for binding to all eIF4F components, especially eIF4A and eIF4G1. The fold difference was then normalized to the fold difference for LDHA, which was set to one. Data are expressed as the mean ± s.e.m. (n = 3).

F. Ribosome occupancy changes correlate with mRNA loss from eIF4E after mTOR inhibition. qPCR was performed on RNA extracted from immunoprecipitates of Flag-eIF4E from MEFs following vehicle or 250 nM Torin1 treatment for two hours. Values were normalized to a spiked in luciferase control and to total levels of each indicated mRNA, then the fold difference relative to vehicle treatment was calculated. Changes in eIF4E binding were plotted against changes in ribosome occupancy. eIF4E binding data represent the mean ± s.e.m. (n = 4).

preferential translational repression of 5'TOP mRNAs, we hypothesized that mTOR inhibition might affect binding of eIF4F to 5'TOP mRNAs more than to mRNAs lacking a 5'TOP motif. We immunoprecipitated eIF4E, eIF4G1, or eIF4A from HEK-293T cells stably expressing Flag epitope-tagged versions of each protein treated for two hours with vehicle or Torin1. We isolated mRNA from each immunoprecipitate in parallel with total mRNA, and performed qPCR for individual mRNAs with and without 5'TOP motifs. As expected, all mRNAs were lost from all eIF4F components, and to a greater extent from eIF4G1 and eIF4A (data not shown). However, of all three eIF4F components, only eIF4E showed preferential loss of 5'TOP and TOP-like mRNAs when mTOR was inhibited, relative to a non-TOP mRNA (Figure 1E). Furthermore, the degree of translational repression observed correlated with the degree of mRNA loss from eIF4E (Figure 1F), suggesting that eIF4E binds 5'TOP and TOP-like mRNAs differently than non-TOP mRNAs and that these differences affect translation.

TOP-like 5'UTRs Recruit eIF4E and eIF4G1 Differentially from Cell Lysates

To directly test the effects of the 5'TOP motif on eIF4F binding, we developed an in vitro binding system. We modified the T7 RNA Polymerase consensus sequence to encode only a single guanosine at position +1 from the 5' cap (Dunn & Studier, 1983), instead of the canonical guanosines at positions +1, +2, and +3 (Figure 2A). We cloned the 5'UTRs of several murine 5'TOP mRNAs downstream of the modified promoter, such that the 5'TOP sequence began at position +2 (Figure 2A). For each of these 5'UTRs, we mutated each 5'TOP motif nucleotide to the complementary purine, yielding

matched pyrimidine-rich, TOP-like and purine-rich, non-TOP pairs (Figure 2A). We transcribed and cotranscriptionally capped each 5'UTR *in vitro* in the presence of biotinylated UTP (Figure 2B). We added the biotinylated UTP at a concentration such that, on average, each 5'UTR would contain only a single labeled UTP. We mixed each 5'UTR with HeLa cell lysate, and isolated the 5'UTRs and their interacting proteins by streptavidin isolation using magnetic beads (Figure 2C). We isolated the bound proteins, and detected them by immunoblot. We observed more binding of the eIF4F components eIF4E and eIF4G1 to the purine-rich 5'UTRs as compared to the 5'UTRs containing the TOP-like cap proximal sequences for all 5'UTR pairs (Figure 2C). We were unable to detect eIF4A, the third component of eIF4F (data not shown).

Figure 2: Pyrimidine-rich TOP-like 5'UTRs Recruit eIF4E and eIF4G1 from Cell Lysates More Poorly than Purine-rich 5'UTRs

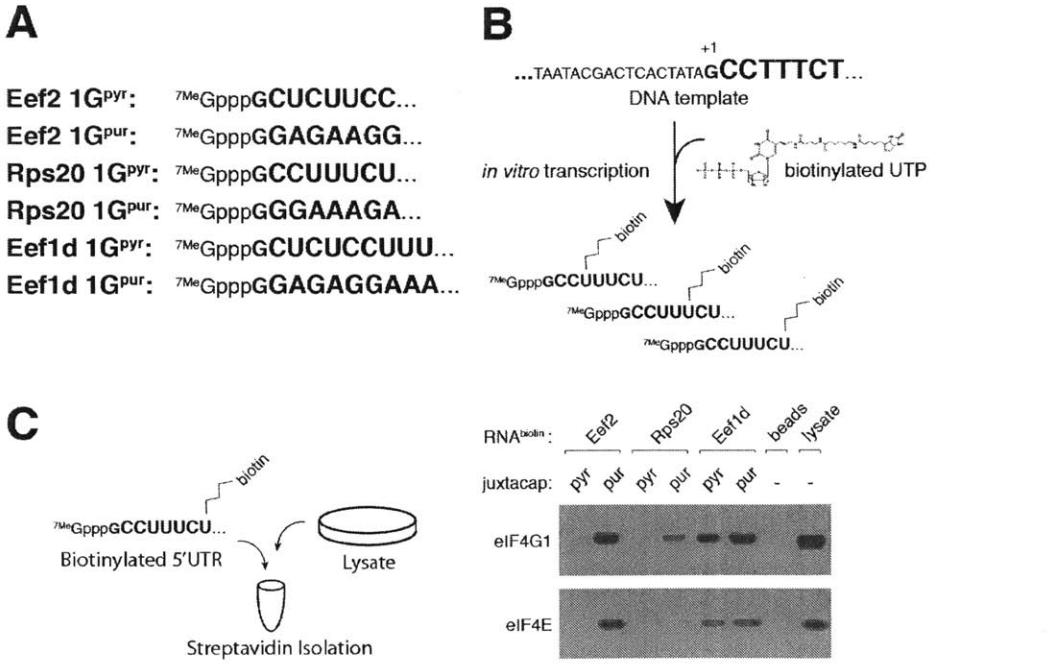


Figure 2: Pyrimidine-rich TOP-like 5'UTRs recruit eIF4E and eIF4G1 from cell lysates more poorly than purine-rich 5'UTRs

A. Sequences of pyrimidine-rich and purine-rich 5'UTRs tested. For each pair, the pyrimidine-rich version encodes the indicated 5'UTR harboring a 5'TOP motif with a single guanosine upstream. The purine-rich version encodes the same 5'UTR, but with the 5'TOP motif mutated to the complementary purine at each position.

B. Schematic of 5'UTR biotinylation. Biotin-labeled UTP was incorporated at a random position once on average into each 5'UTR during transcription *in vitro*.

C. Pyrimidine-rich TOP-like 5'UTRs recruit eIF4E and eIF4G1 less well from cell lysates than purine-rich 5'UTRs. Each indicated biotinylated 5'UTR was mixed with HeLa cell lysate, and bound proteins were recovered by streptavidin isolation. Bound proteins were detected by immunoblotting with the indicated antibodies.

We sought to test the dynamics of the 5'UTR interaction with eIF4E and eIF4G from lysates. We incubated HeLa cell lysate with biotinylated 5'TOP-like or purine-rich 5'UTRs pre-bound to beads for different amounts of time, washed away unbound proteins, and assessed binding of eIF4E and eIF4G1 by immunoblot. Both proteins bound more quickly to purine-rich 5'UTRs (Figure 3A and Figure 3B), suggesting that the cap-proximal sequence can affect the on-rate of these proteins. We also tested the off-rate of eIF4E and eIF4G from these 5'UTRs. We incubated HeLa cell lysate with each biotinylated 5'UTR containing the pyrimidine- or purine-rich cap-proximal sequence and streptavidin beads, removed unbound proteins through a series of washes, and diluted the samples in an excess of fresh binding buffer to minimize rebinding of a protein to the 5'UTR after dissociation. We monitored loss of eIF4E and eIF4G1 from the 5'UTRs by removing aliquots of the binding reaction at different time points and rewashing the beads, then isolating the bound proteins and detecting them

Figure 3: TOP-like 5'UTRs Have Different Affinities for eIF4E and eIF4G1 from Lysates

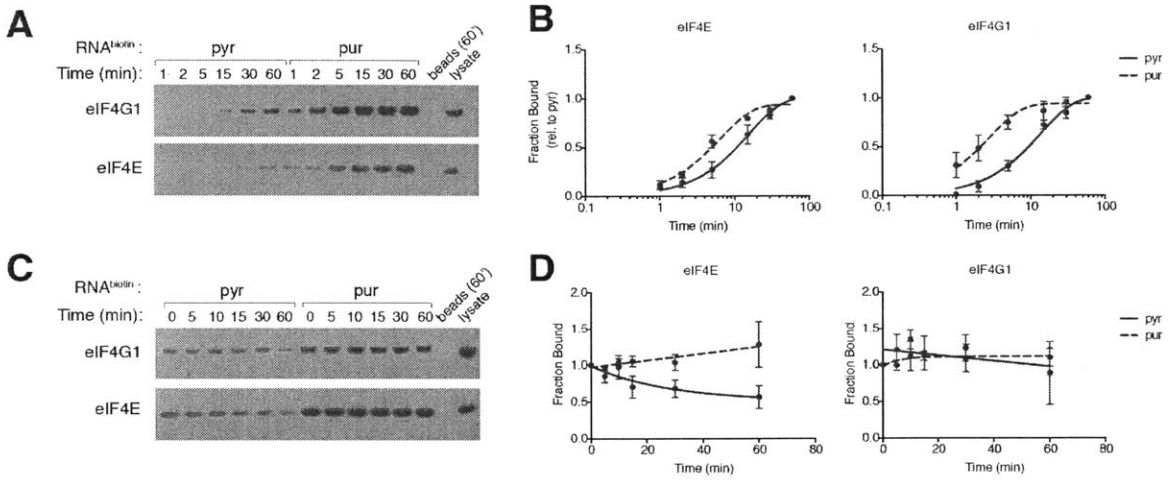


Figure 3: TOP-like 5'UTRs have different affinities for eIF4E and eIF4G1 from lysates

A. eIF4E and eIF4G1 from lysates associate more slowly with pyrimidine-rich TOP-like 5'UTRs. Each biotinylated Eef2 5'UTR was mixed with HeLa cell lysate for the indicated times, and bound proteins were recovered by streptavidin isolation. Bound proteins were detected by immunoblotting with the indicated antibodies.

B. Non-linear curve fit of on-rate of eIF4E and eIF4G1 from lysates. Densitometry was performed on immunoblots from experiments performed as in Figure 3A. Points are plotted as the mean ± s.e.m. (n = 3).

C. The dissociation of eIF4E from lysates is slower for pyrimidine-rich TOP-like 5'UTRs. Each biotinylated Eef2 5'UTR was mixed with HeLa cell lysate, and bound proteins were recovered by streptavidin isolation. Beads were diluted for the indicated times and subjected to a second round of streptavidin isolation. Bound proteins were detected by immunoblotting with the indicated antibodies.

D. Quantification of off-rate of eIF4E and eIF4G1 from lysates. Densitometry was performed on immunoblots from experiments performed as in Figure 3C. Points are plotted as the mean ± s.e.m. (n = 3).

by immunoblot. We observed almost no dissociation of eIF4G1 from either 5'UTR, and no dissociation of eIF4E from the 5'UTR with a purine-rich cap-proximal sequence (Figure 3C and Figure 3D). However, eIF4E dissociated more quickly from the TOP-like 5'UTR, suggesting that the cap-proximal sequence can also differentially affect the off-rate of eIF4E (Figure 3C and Figure 3D). Together, these results indicate that purine-rich cap-proximal sequences can retain eIF4E and also recruit eIF4E and eIF4G1 better than pyrimidine-rich sequences.

eIF4F Has a Higher Affinity for Purine-Rich Cap-Proximal Sequences

Cell lysate contains many other protein and RNA factors that could influence eIF4E and eIF4G1 binding, so we sought to more directly test whether eIF4F exhibits binding preferences for the cap-proximal sequence. We co-transfected HEK-293T cells with HA-eIF4G1 and either Flag-eIF4E or Flag-GFP, and purified the eIF4E/eIF4G1 complex or GFP using anti-Flag agarose, followed by elution with Flag peptide. Although these preparations were not 100% pure, as determined by Coomassie staining, they were greatly enriched for eIF4E and eIF4G1 relative to total cell lysate (data not shown). We mixed increasing concentrations of the purified complex with a fixed amount of either the pyrimidine- or purine-rich Eef2 5'UTR, and assessed binding by streptavidin isolation and Western blot (Figure 4A). Both eIF4E and eIF4G1 bound more highly to the purine-rich 5'UTR (Figure 4A), suggesting that the complex has a higher affinity for purine-rich cap-proximal sequences. Importantly, GFP did not bind to either 5'UTR, nor did GFP or the eIF4E/eIF4G1 complex bind to the beads in the absence of the biotinylated 5'UTR (Figure 4A).

To quantitatively measure the affinity of the eIF4E/eIF4G1 complex for each 5'UTR, we performed an electrophoretic mobility shift assay (EMSA) using purified eIF4E/eIF4G1 complex and radiolabeled pyrimidine-rich and purine-rich Rps20 5'UTRs. We transcribed each 5'UTR with unlabeled UTP, and capped each 5'UTR in vitro in the presence of α^{32}P-labeled GTP. This resulted in the ability to detect each capped 5'UTR by autoradiography, as the guanosine of the 5' cap was radiolabeled. We mixed increasing concentrations of the eIF4E/eIF4G1 complex with a fixed amount of each 5'UTR, and quantified the reduction in free RNA (Figure 4B). The purified eIF4E/eIF4G1 complex had a higher affinity for the purine-rich Rps20 5'UTR (K_d = 22.7±6.3 nM) than for the pyrimidine-rich 5'UTR (K_d = 76.6±11.4 nM), whereas the GFP control did not bind RNA (Figure 4B). We observed two bands of low mobility (Figure 4B), which we believed corresponded to the eIF4E/eIF4G1/RNA complex, as we detected eIF4G1 as a doublet by Western blotting (data not shown). Furthermore, we observed an additional band that migrated slightly more slowly than free RNA (Figure 4B), which could have been eIF4E bound to RNA, without eIF4G1. Since we purified the eIF4E/eIF4G1 complex via the Flag-tag fused to eIF4E, the preparation likely contained some amount free eIF4E. This putative eIF4E/RNA complex was more prominent in the binding reactions containing the pyrimidine-rich 5'UTR (Figure 4B), suggesting that the eIF4E/eIF4G1 complex bound the TOP-like 5'UTR with low enough affinity that free eIF4E was able to compete for binding.

Figure 4: Pyrimidine-rich TOP-like 5'UTRs Have a Lower Affinity for Purified eIF4E/eIF4G1 Independently of the eIF4G1 RNA-binding Domains

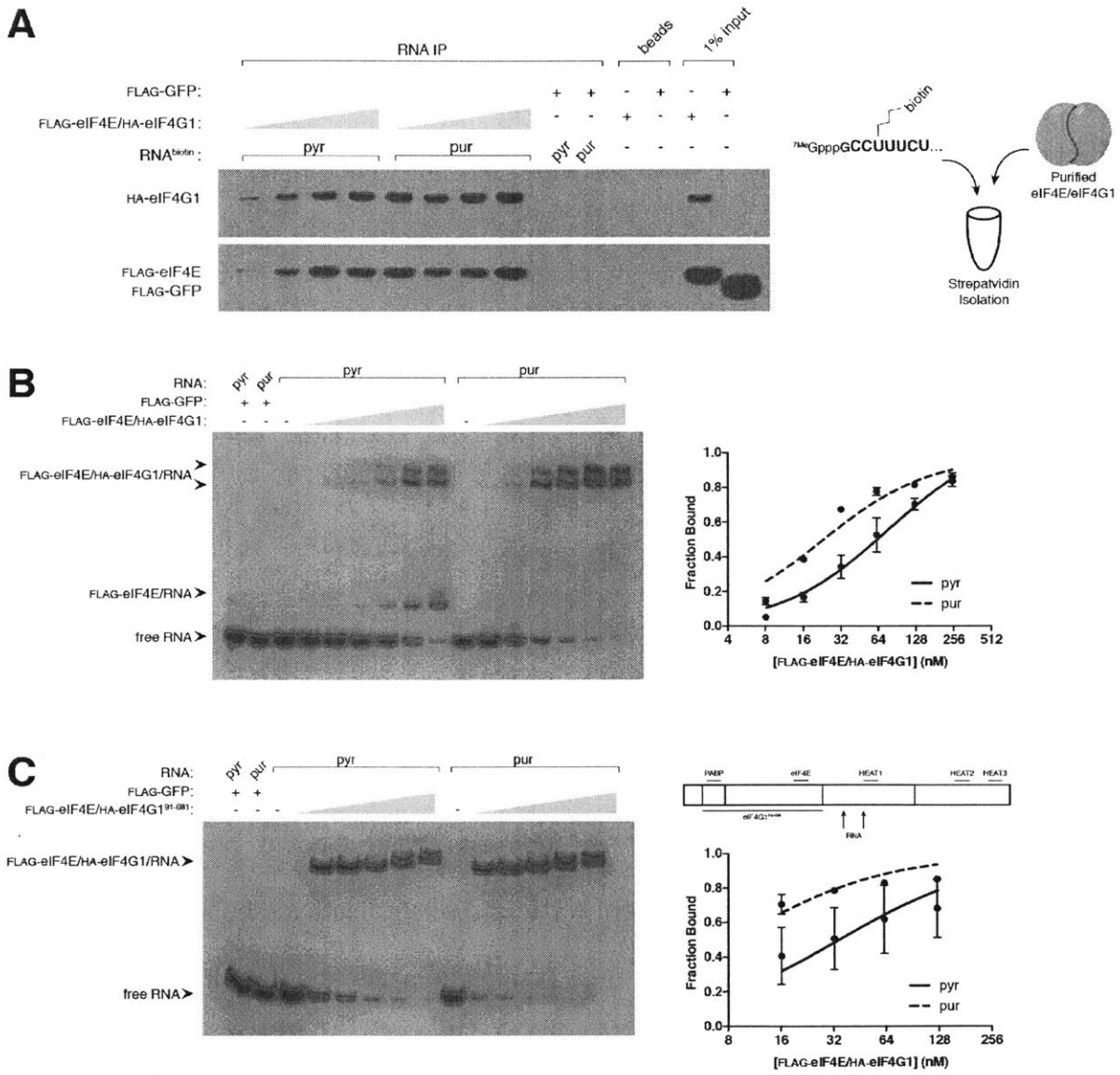


Figure 4: Pyrimidine-rich TOP-like 5'UTRs have a lower affinity for purified eIF4E/eIF4G1 independently of the eIF4G1 RNA-binding domains

A. Pyrimidine-rich TOP-like 5'UTRs recruit purified eIF4E/eIF4G1 less well than purine-rich 5'UTRs. Each biotinylated Eef2 5'UTR was mixed with increasing concentrations of eIF4E/eIF4G1 complex or GFP purified by Flag immunoprecipitation, and bound proteins were recovered by streptavidin isolation. Bound proteins were detected by immunoblotting.

B. Pyrimidine-rich TOP-like 5'UTRs recruit purified eIF4E/eIF4G1 less well than purine-rich 5'UTRs. Non-biotinylated radiolabeled Rps20 5'UTRs were mixed with increasing concentrations of eIF4E/eIF4G1 complex or GFP purified as in Figure 4A, and bound proteins were visualized by EMSA followed by autoradiography. The estimated concentration of eIF4E/eIF4G1 complex is plotted against the fraction of RNA bound, which was calculated by subtracting the free RNA signal from the total RNA signal, as quantified by phosphorimaging. Points are plotted as the mean ± s.e.m. (n = 3).

C. The RNA-binding domains are not required for differential binding of purified eIF4E/eIF4G1 to pyrimidine-rich and purine-rich 5'UTRs. Non-biotinylated radiolabeled Rps20 5'UTRs were mixed with increasing concentrations of eIF4E/eIF4G1$^{91\text{-}681}$ complex or GFP purified as in Figure 4A, which lacks the RNA-, eIF4A-, eIF3-, and MNK1/2-binding regions of eIF4G1. Bound proteins were visualized by EMSA followed by autoradiography. The concentration of eIF4E/eIF4G1$^{91\text{-}681}$ complex is plotted against the fraction of RNA bound, which was calculated by subtracting the free RNA signal from the total RNA signal, as quantified by phosphorimaging. Points are plotted as the mean ± s.e.m. (n = 3).

We hypothesized that the RNA binding domains of eIF4G1 directly interact with the cap-proximal nucleotides, and that these domains have binding preferences that influence the affinity of eIF4F for different 5'UTRs. To test this hypothesis, we made a truncation mutant of eIF4G1 lacking all RNA binding domains and the eIF4A, eIF3, and Mnk1/2 binding sites, but retaining the eIF4E and PABP binding sites (Figure 4C), based on a truncation mutant that has been previously studied by NMR (Marintchev et al., 2009). We co-transfected this HA-eIF4G1$^{91\text{-}681}$ with Flag-eIF4E into HEK-293T cells,

and purified the eIF4E/eIF4G1$^{91-681}$ complex. EMSA using this complex and pyrimidine- or purine-rich Rps20 5'UTRs surprisingly revealed a higher affinity for the purine-rich 5'UTR (Figure 4C; K_d = 8.5±1.5 nM compared to 34.1±5.5 nM for the TOP-like 5'UTR), indicating that the RNA binding domains of eIF4G1 do not greatly influence the binding preferences of eIF4F, and suggesting that eIF4E itself may possess cap-proximal nucleotide preferences.

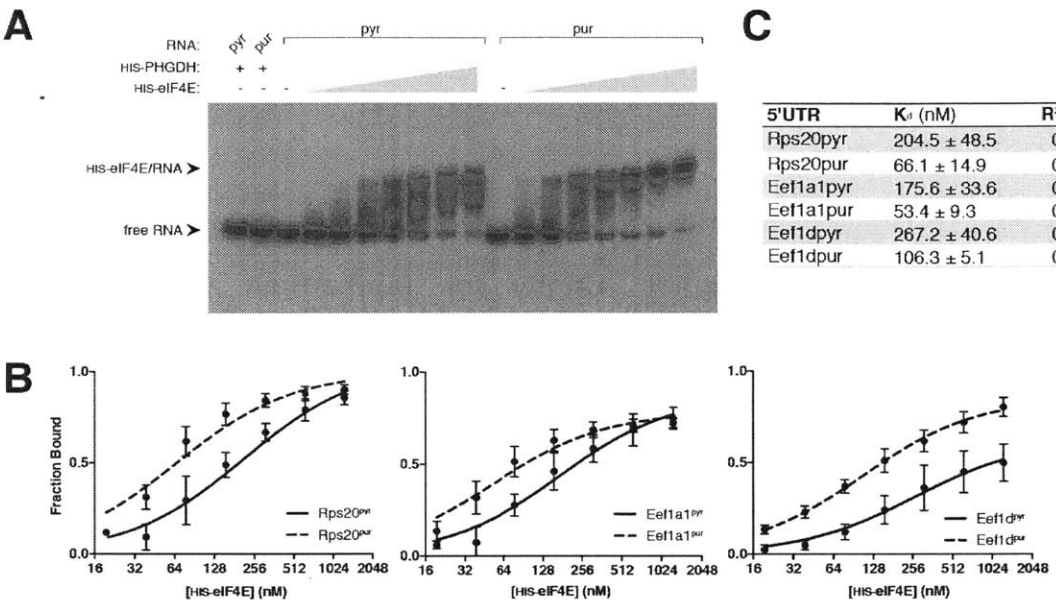
Recombinant eIF4E Has a Higher Affinity for Purine-Rich Cap-Proximal Sequences

To test whether eIF4E has an intrinsic binding preference for purine-rich cap-proximal nucleotides, we performed EMSA with His-eIF4E purified from bacteria and pyrimidine- and purine-rich Rps20, Eef1a1, and Eef1d 5'UTR pairs. In all cases, eIF4E bound more highly to the purine-rich 5'UTR, although the K_d difference for each TOP-like and purine-rich 5'UTR pair varied (Figure 5A, Figure 5B, and Figure 5C). In no case did the control protein, recombinant His-PHGDH, bind RNA (Figure 4A and data not shown). Taken together, our data suggest that eIF4E, not eIF4G1, drives the cap-proximal binding preferences of eIF4F.

Affinity of eIF4E for non-TOP mRNAs is not Higher than for True 5'TOP mRNAs *in vitro*

Our results suggested that TOP-like 5'UTRs have a lower affinity for eIF4E than purine-rich 5'UTRs, consistent with the model that 5'TOP mRNAs are more susceptible

Figure 5: Pyrimidine- and Purine-rich 5'UTRs Have Different Affinities for Recombinant eIF4E



C

5'UTR	K_d (nM)	R^2	p value
Rps20pyr	204.5 ± 48.5	0.9749	0.0027
Rps20pur	66.1 ± 14.9	0.9685	
Eef1a1pyr	175.6 ± 33.6	0.9818	0.0003
Eef1a1pur	53.4 ± 9.3	0.9798	
Eef1dpyr	267.2 ± 40.6	0.9901	<0.0001
Eef1dpur	106.3 ± 5.1	0.9985	

Figure 5: Pyrimidine- and purine-rich 5'UTRs have different affinities for recombinant eIF4E

A. Pyrimidine-rich TOP-like 5'UTRs recruit recombinant eIF4E less well than purine-rich 5'UTRs. Non-biotinylated radiolabeled 5'UTRs were mixed with increasing concentrations of His-tagged eIF4E or a control protein purified from bacteria, and bound proteins were visualized by EMSA followed by autoradiography.

B. Multiple TOP-like 5'UTRs have lower affinity for recombinant eIF4E. The concentration of eIF4E is plotted against the fraction of the indicated RNA bound, which was calculated by subtracting the free RNA signal from the total RNA signal, as quantified by phosphorimaging. Experiments were performed as in Figure 5A. Points are plotted as the mean ± s.e.m. (n = 3).

C. Summary of K_d differences between pyrimidine-rich and purine-rich 5'UTRs. K_ds were obtained by fitting an exponential one-phase association curve to the data from Figure 5B. P-values are reported for the extra sum-of-squares F test comparing the K_ds between members of each pair. The R^2 values for the goodness-of-fit are also reported.

to translational suppression caused by mTOR inhibition because they do not compete for limiting amounts of eIF4F as well as other mRNAs. However, the 5'UTRs we used in our *in vitro* system were not perfectly analogous to true 5'TOP mRNAs, because T7 RNA Polymerase requires a guanosine in the +1 position. To test our hypothesis that 5'TOP mRNAs have a lower affinity for eIF4E than non-TOP mRNAs, we used an alternative *in vitro* binding approach. We transfected HEK-293T cells with a construct encoding *Renilla* luciferase and a polyA tail downstream of the CMV promoter and wildtype murine Eef2 5'UTR, which begins with a cytidine, or a mutant Eef2 5'UTR in which each nucleotide of the 5'TOP motif has been mutated to the complimentary purine. We then isolated and polyA-selected the RNA from cells individually expressing the wildtype or mutant Eef2-*Renilla* mRNAs. We mixed 1 μM streptavidin binding peptide (SBP)-tagged eIF4E purified from bacteria with 50 nM polyA-selected RNA, in the presence of increasing concentrations of ^{me7}GTP, a competitive inhibitor of 5' cap binding to eIF4E, or GTP, which does not compete for cap binding. We then isolated eIF4E and mRNAs bound to eIF4E using streptavidin-coated magnetic beads, extracted the mRNAs, and performed qPCR from each set of samples using primers against *Renilla* luciferase or endogenous mRNAs. Preliminary results indicate that the IC_{50} for ^{me7}GTP is the same for eIF4E binding to wildtype and mutant Eef2-*Renilla* mRNA (Figure 6A), suggesting that there is no difference in affinity between a true 5'TOP and non-TOP mRNA.

In a similar experiment, we mixed 50 nM polyA-selected mRNA with 1 μM SBP-eIF4E in the presence of various ^{me7}GTP concentrations, and performed qPCR on the bound

mRNAs using primers against endogenous 5'TOP and non-TOP mRNAs. Preliminary results suggest that eIF4E does possess different affinities for some endogenous mRNAs, although the affinity for 5/7 mRNAs was similar (Figure 6B and data not shown). However, the two mRNAs most resistant to competition with ^{me7}GTP for eIF4E binding were both 5'TOP mRNAs (Figure 6B). Hsp90ab1, a 5'TOP mRNA, is both more translationally repressed and less well bound to stably expressed Flag-eIF4E after mTOR inhibition than Myc, a non-TOP mRNA (Figure 1F), suggesting that differences in eIF4E affinity cannot explain the differences in response to mTOR inhibition.

Figure 6: eIF4E Binding to True 5'TOP and non-TOP mRNAs Does not Correlate with Their Sensitivity to mTOR Inhibition

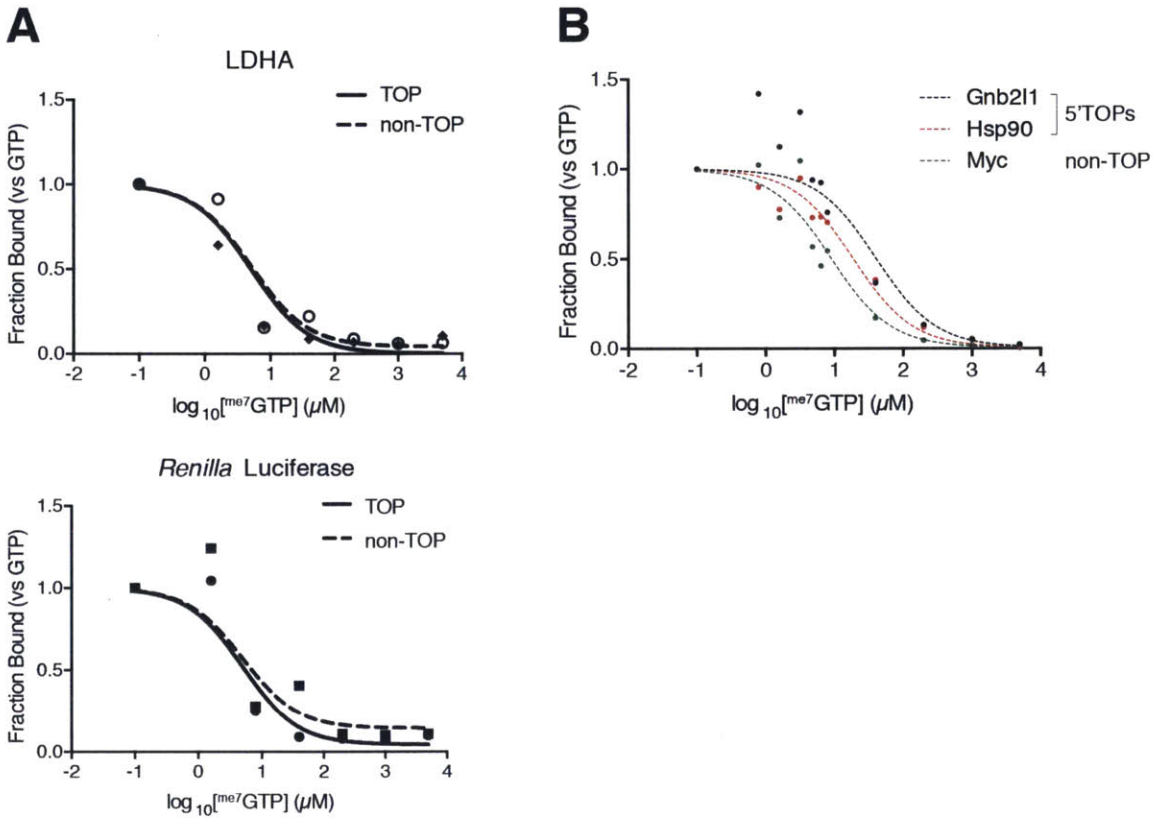


Figure 6: eIF4E binding to true 5'TOP and non-TOP mRNAs does not correlate with their sensitivity to mTOR inhibition

A. True 5'TOP and 5'TOP mutant reporter mRNAs have the same affinity for recombinant eIF4E. PolyA+ mRNA was isolated from HEK-293T cells expressing a construct encoding *Renilla* luciferase downstream of the true Eef2 5'UTR (without a +1 guanosine) or the Eef2 5'UTR in which the 5'TOP motif was mutated to the complimentary purine at each position. A fixed concentration of polyA-selected mRNA was mixed with a single concentration of SBP-tagged eIF4E purified from bacteria, in the presence of increasing concentrations of ^{me7}GTP, a competitive inhibitor of cap binding, or GTP, which does not compete. Bound mRNA was recovered by streptavidin isolation, and qPCR for the indicated mRNAs was performed. qPCR measurements were normalized to a spiked in firefly luciferase control, and then to the GTP-containing sample, which was set at one. Curves were fit using a one-site logIC_{50} model. Values reported are from a single experiment.

B. Endogenous mRNAs that are more sensitive to mTOR inhibition have a higher affinity for eIF4E. PolyA+ mRNA was isolated from HEK-293T cells, and a fixed concentration of polyA-selected mRNA was mixed with a single concentration of SBP-tagged eIF4E purified from bacteria, in the presence of increasing concentrations of ^{me7}GTP or GTP. Bound mRNA was recovered by streptavidin isolation, and qPCR for the indicated endogenous mRNAs was performed. qPCR measurements were normalized to a spiked in firefly luciferase control, and then to the GTP-containing sample, which was set at one. Curves were fit using a one-site logIC_{50} model. Values reported are from a single experiment. 5'TOP and non-TOP mRNAs are indicated.

DISCUSSION

It is well established that the mTORC1 pathway plays a role in cancer and aging, as well as a variety of other diseases and disorders. mTORC1 regulates a wide array of processes to ultimately promote cell growth and proliferation, but it remains unclear which of its downstream effectors are most important in disease. mRNA translation is often deregulated in cancer, and the translational program downstream of mTORC1 likely plays a role. Understanding the mechanistic details of translational regulation in general will facilitate the discovery of new ways to target the production of individual proteins that are relevant to human health and disease. In particular, understanding how the mTORC1 pathway regulates translation is especially important for deciphering its role in cancer and metabolic disorders.

Although our preliminary data indicate that the affinity of eIF4E for mRNAs is not sufficient to explain the dramatic repression of 5'TOP mRNA translation by inhibition of mTOR, the link between 4EBPs, eIF4E binding, and sensitivity to translational regulation downstream of mTOR remains. It is well established that hypophosphorylated 4EBPs competitively inhibit eIF4G binding to eIF4E (Marcotrigiano et al., 1999). However, the dynamics of this process are not understood and could contribute to translational selectivity downstream of mTOR inhibition. Our data suggest that the dynamics of the eIF4F interaction can differ on RNAs of different sequence, as eIF4E is lost more quickly from a TOP-like 5'UTR than from a purine-rich 5'UTR in vitro, and both eIF4E and eIF4G bind more quickly to the 5'UTR with a purine-rich cap-proximal sequence. Several models for how the dynamics of the eIF4F/4EBP/RNA interaction

could be differentially affected by mTORC1 inhibition are consistent with our data. One model is that 4EBPs can target eIF4F bound to an mRNA, and that the eIF4E-eIF4G interaction is more easily disrupted on mRNAs possessing a 5'TOP motif. A second possibility is that 4EBPs can only act on free eIF4F, and that release of the mRNA from eIF4F is necessary for inhibition by 4EBPs to occur. In this case, it could be that sensitivity to mTOR inhibition is dictated primarily by the off-rate of eIF4F from mRNAs, which could be increased by the presence of a 5'TOP motif. eIF4E in complex with a 4EBP can still bind to the 5' cap, although with a lower affinity than when bound to eIF4G (von der Haar, Gross, Wagner, & McCarthy, 2004). It could be that 4EBPs can disrupt both free and mRNA-bound eIF4F, but that the eIF4E-4EBP complex has a preference for mRNAs lacking a 5'TOP motif. This model is consistent with our results that 5'TOP mRNAs are selectively lost from eIF4E in cells, but not other eIF4F complex members, following mTOR inhibition. Although mRNAs bound to eIF4E-4EBP complexes should not be translated, greater stability of eIF4E on an mRNA could allow more opportunity for eIF4G to displace 4EBP and for translation to resume. Therefore, mRNAs that retain eIF4E-4EBP would have a higher probability of being translated. Thorough characterization of the dynamics of 4EBP-mediated inhibition of translation could help uncover the mechanism for 5'TOP mRNA selectivity.

Another enticing possibility is that a factor unrelated to eIF4F binds 5'TOP mRNAs with higher affinity than non-TOPs, and competes with eIF4E for binding. According to this model, disruption of eIF4F by 4EBPs and subsequent loss of eIF4E from mRNA could occur similarly for all mRNAs, but this disruption would allow another protein

factor with a preference for 5'TOP mRNAs the opportunity to outcompete eIF4E for mRNA binding. Although no single protein has been found to regulate 5'TOP mRNAs in this manner under all conditions, this mechanism has been described for the proteins TIA-1 and TIAR when cells are starved for amino acids (Damgaard & Lykke-Andersen, 2011). Although TIA-1 and TIAR did not play a role in translational regulation downstream of mTOR in our study (data not shown), we cannot rule out that a factor or combination of factors may bind 5'TOP mRNAs preferentially when mTOR is inhibited.

A mechanism that incorporates differences in eIF4E or eIF4F affinity or dynamics is very attractive, because it could explain the observation that many factors affect 5'TOP mRNA translation under many different conditions (Biberman & Meyuhas, 1999; Damgaard & Lykke-Andersen, 2011; Kakegawa et al., 2007; Orom et al., 2008; Tcherkezian et al., 2014). If eIF4F has a lower affinity for 5'TOP mRNAs, then any factor that can bind near the 5' end of mRNA and block eIF4F binding would necessarily affect 5'TOP mRNAs more strongly than other mRNAs.

Although we have not elucidated the mechanism whereby mTORC1 inhibition selectively represses translation of 5'TOP mRNAs, we have uncovered a new principle in the regulation of eIF4F recruitment to mRNAs; namely, that mRNAs with different cap-proximal sequences can recruit eIF4F differently. There are very few examples of sequence-dependent differential recruitment of translation initiation factors in the literature, although the model that translation factors affinity differences for mRNAs could contribute to their translational regulation is not new (Lodish, 1974). We propose that general translation initiation factor affinity could be a more common mechanism

used to regulate the translation of particular subsets of mRNAs. Global profiling of mRNA sequence preferences for the initiation machinery could yield important insights into the workings of the fundamental process of protein synthesis.

MATERIALS AND METHODS

Reagents

Primary antibodies were purchased from Cell Signaling Technology (eIF4G #2469, eIF4E #2067, DYKDDDDK Tag #2368). Enzymes were purchased from New England Biolabs, except where noted. Primers were obtained from Integrated DNA Technologies. Antibiotics, Flag-M2 agarose, and chemicals were purchased from Sigma, except where noted. TBE-Urea gels, TBE gels, and sample loading dyes were purchased from ThermoFisher. DMEM and inactivated fetal bovine serum (IFS) were purchased from US Biologicals. pLJM60-Flag-eIF4G1, pCT3-EEF2-TOP and pCT3-EEF2-nonTOP were gifts from C. Thoreen (Yale University).

Cell Lysis, Polysome Analysis, and Gradient-qPCR Analysis in MEFs

4EBP1/2+/+ p53-/- or 4EBP1/2-/- p53-/- MEFs (Petroulakis et al., 2009) were seeded at a density of 5 $x10^6$ cells per 15cm dish in 25ml complete DMEM supplemented with 10% FBS and penicillin/streptomycin. The following day, cells were treated for two hours with either DMSO or 250 nM Torin1 ((Thoreen et al., 2009); available from Tocris). Five minutes prior to lysis, cells were treated with 100ug/ml cycloheximide in ethanol.

Cells were washed with 10ml ice-cold PBS containing 100ug/ml cycloheximide, and lysed in 750µl Polysome Lysis Buffer (15 mM HEPES pH 7.4, 7.5 mM $MgCl_2$, 100 mM KCl, 1% TritonX-100, 2 mM DTT (ThermoFisher), one EDTA-free protease inhibitor tablet per 10ml buffer (Roche), and 100 µg/ml cycloheximide). Lysates were normalized

by protein content, and split into aliquots to run on sucrose gradients or for total mRNA extraction.

10%-50% sucrose gradients were prepared the day before use by underlaying 5.5ml 10% Polysome Gradient Buffer (15 mM HEPES pH 7.4, 7.5 mM $MgCl_2$, 100 mM KCl, 100 µg/ml cycloheximide, 2 mM DTT, 20U/ml SuperaseIn (ThermoFisher), and 10% sucrose) with 5.5ml 50% Polysome Gradient Buffer (same as 10% Polysome Gradient Buffer, but containing 50% sucrose). Tubes were parafilmed and placed horizontally for 4-6 hours at 4C, then carefully turned upright and kept at 4C overnight. The following day, 750µl-850µl of cell lysate was loaded onto each gradient and ultracentrifuged for two hours at 32,000*xg* at 4C. Polysomes were detected by monitoring absorbance at 254 nm, and 1ml fractions were collected.

For polysome gradient-qPCR analysis from MEFs, 5 ng of firefly luciferase (Promega) was added to each fraction and samples were incubated for 5 minutes at 65C. SDS was added to 0.5% and Proteinase K (Roche) was added to 200 µg/ml, then samples were incubated for 45 minutes at 50C. Samples were diluted 1:1 with RNAse-free water, and RNA was extracted twice with 750µl acid phenol/chloroform (ThermoFisher), once for 5 minutes at 65C and once for 5 minutes at room temperature. The aqueous layer was extracted with 750µl chloroform, and precipitated with NaOAc and isopropanol for 30 minutes at -20C. RNA was recovered by centrifugation for 30 minutes at 13,300*xg*, pellets were washed once with ice-cold 80% ethanol, and resuspended in 10µl 10 mM Tris pH 8. cDNA synthesis was performed with Superscript III (ThermoFisher) according to manufacturer's instructions using a random hexamer primer for reverse transcription.

qPCR was performed with 2.5µl cDNA per 10µl reaction using 2X SYBR Green master mix (Roche) according to manufacturer's instructions, with primers present at 5 µM. All values for mRNAs assayed from polysome gradient fractions were normalized to the spiked in firefly luciferase control.

Total RNA preparation

Total mRNA was prepared from cells seeded and treated in parallel as those used for polysome analysis, but lysed in RNA Lysis Buffer (15 mM HEPES pH 7.4, 15 mM $MgCl_2$, 300 mM KCl, 1% SDS). Lysates were homogenized using a 21G needle and incubated at 65C for 5 minutes. RNA was extracted using the hot acid phenol method and precipitated with NaOAc and isopropanol as described above. RNA was resuspended in 200µl 10 mM Tris pH 8 and polyA-selected using oligodT_{25} Dynabeads (ThermoFisher) according to manufacturer's instructions. RNA was resuspended in 20µl 10 mM Tris pH 8 and fragmented according as described previously (Ingolia, Ghaemmaghami, Newman, & Weissman, 2009). Fragments were prepared for small RNA Illumina sequencing in parallel with ribosome footprints as described below.

Ribosome footprint and sequencing library preparation

Ribosome footprints were prepared similarly to the method of Ingolia *et al.* (Ingolia et al., 2009). Briefly, cells were seeded and treated as described above for polysome analysis, then lysed in Footprint Lysis Buffer (15 mM HEPES pH 7.4, 7.5 mM $MgCl_2$, 300 mM KCl, 2 mM DTT, one EDTA-free protease inhibitor tablet per 10ml buffer, and

100 µg/ml cycloheximide). Cleared lysates were digested with 1U/µl RNAseI (ThermoFisher) for 45 minutes at room temperature with constant mixing. Digested samples were layered onto 10-50% sucrose gradients (15 mM HEPES pH 7.4, 7.5 mM $MgCl_2$, 300 mM KCl, 100 µg/ml cycloheximide, 2 mM DTT, 20U/ml SuperaseIn, and 10-50% sucrose) and centrifuged for 2.5 hours at 36,000*xg* at 4C. Gradients were fractionated and monosome fractions were collected. SDS was added to 0.5%, digested with Proteinase K, and hot acid phenol extracted and precipitated as described above. RNA was resuspended in 500µl 10 mM Tris pH 8 with 2.5µl SuperaseIn and centrifuged for 28 minutes in a YM100 microconcentrator (Millipore) to remove large RNA fragments. RNA was precipitated and purified on a 15% TBE-Urea gel and extracted. RNA was prepared for small RNA Illumina Sequencing as previously described (Ingolia et al., 2009).

Data analysis

Raw sequencing data was processed as described previously (Ingolia et al., 2009) using custom Python scripts.

Cloning, expression, and protein purification

eIF4E was amplified by PCR from a plasmid encoding eIF4E (Addgene #38239) and cloned into the pET302/NT-His bacterial expression vector (ThermoFisher) downstream of a 6xHis tag using XhoI and BamHI restriction sites.

meIF4E_Nterm_F: AATTCGCTCGAGatggcgactgtcgaaccg

heIF4E_Nterm_R: GCAGCCGGATCCttaaacaacaaacctatt

A streptavidin-binding peptide (SBP) tag was introduced in place of the 6xHis tag using a geneblock (IDT) encoding SBP with a 5' NdeI site and a 3' XhoI site.

SBP_geneblock_NdeI_XhoI:

GATATACATATGGACGAAAAAACGACCGGGTGGCGTGGCGGCCATGTCGTGGAGG

GCCTGGCAGGCGAGCTGGAACAACTGCGCGCACGTCTGGAACACCATCCTCAGGG

ACAGCGCGAACCAGTGAATTCGCTCGAGatggcg

pLJM1-EGFP is as described (Sancak et al., 2008). pLJM60-Flag-eIF4E (Addgene #38239) and pLJM60-Flag-eIF4G1 (a gift of C. Thoreen) were used directly for stable cell line generation. For transfection experiments, eIF4E and eIF4G1 were subcloned into a modified pRK5 vector containing an N-terminal Flag or HA tag (Bar-Peled, Schweitzer, Zoncu, & Sabatini, 2012; Sancak et al., 2008) using SalI and NotI restriction sites. eIF4G1$^{91\text{-}681}$ was cloned by PCR amplification from pLJM60-Flag-eIF4G1 with primers containing an N-terminal SalI site and a C-terminal NotI site:

eIF4G1_91-681_F: CAAGTCgtcgacGatg**GGA**tcccaagtaatgatgatcccttcc

eIF4G1_91-681_R: CATGATgcggccgcCTAtta**GCC**aaggttggcaaaggatggagtgaa

The resulting product was then cloned back into pLJM60-Flag and also into pRK5-HA using SalI and NotI restriction sites. Human eIF4A was cloned from HEK-293T cDNA using primers containing an N-terminal SalI site and a C-terminal NotI site:

eIF4A_SalI_F: CAAGTCgtcgacGatgtctgcgagccaggattcccgatcc

eIF4A_NotI_R: CATGATgcggccgcCTAtcagatgaggtcagcaacattgagggg

The PCR product was cloned into pLJM60-Flag using SalI and NotI restriction sites.

For transient overexpression, HEK-293T cells were seeded at $4x10^6$ cells per 15cm dish in 20ml DMEM supplemented with 10% IFS and penicillin/streptomycin. The following day, cells were transfected with 2µg each of Flag-eIF4E or Flag-GFP with HA-eIF4G1 and 6ug of empty pRK5 vector using 30µl PEI (1 mg/ml) in 800µl DMEM without serum or antibiotics. Media was refreshed 24 hours post-transfection, and again two hours before lysis at 48 hours post-transfection. Cells were lysed in 700µl Triton Lysis Buffer with 1 mM DTT, lysates were cleared by centrifugation, and protein concentrations were normalized. Lysates were incubated with Flag-M2 agarose for 90 minutes at 4C with mixing, and beads were washed three times for 5 minutes at 4C with Triton Lysis Buffer. Proteins were eluted in 20 mM HEPES pH 7.4, 100 mM KCl, 1 mM DTT, and 100 µg/ml 3xFlag peptide for 30 minutes at 4C.

RNA immunoprecipitation-qPCR in HEK-293Ts

For virus production, HEK-293T cells were seeded at $7x10^5$ cells in 6cm dishes in 4ml DMEM supplemented with 10% IFS and penicillin/streptomycin. Cells were transfected the following day with the following plasmids: 1 µg cDNA, 900 ng delta-VPR, 100 ng VSV-G using 6µl Fugene (Roche) in 200 µl serum- and antibiotic-free DMEM. Virus was harvested at 48 and 72 hours post-transfection and stored at -80C.

Stable cell lines were produced by seeding $2x10^5$ HEK-293T cells in 6-well plates in 2ml DMEM supplemented with 10% IFS and penicillin/streptomycin. The following day,

1ml virus and 12µl polybrene (2 mg/ml) were added to cells. 24 hours post-infection, cells were selected with 1 µg/ml puromycin. Puromycin was removed when cells were passaged.

MEFs or HEK-293Ts stably expressing the indicated Flag-tagged constructs were seeded at $2x10^6$ or $10x10^6$ cells per plate in 10cm or 15cm dishes, respectively and grown overnight. Cells were treated with DMSO or 250 nM Torin1 for two hours. MEFs were washed once with ice-cold PBS and lysed in Polysome Lysis Buffer with 40U/ml SuperaseIn. HEK-293Ts were washed three times with ice-cold PBS and lysed in RNA IP Lysis Buffer (15 mM HEPES pH 7.4, 100 mM NaCl, 2 mM EDTA, 1% TritonX-100, 2 mM DTT, and one EDTA-free protease inhibitor tablet per 25ml buffer). Insoluble material was removed by centrifugation, and lysates were normalized by protein concentration. Aliquots were reserved for total RNA extraction, and Flag-M2 agarose was added to the remaining sample. Samples were immunoprecipitated for two hours at 4C with rotation. Immunoprecipitations from MEFs were washed 6 times with IP Lysis Buffer (50 mM HEPES pH 7.4, 40 mM NaCl, 2mM EDTA, 10 mM pyrophosphate, 10 mM β-glycerophosphate, 1% TritonX-100, and one EDTA-free protease inhibitor tablet), twice with Polysome Lysis Buffer, and eluted with 100 µg/ml 3xFlag peptide in Polysome Lysis Buffer for 10 minutes at 37C. Immunoprecipitations from HEK-293Ts were washed 6 times with RNA IP Lysis Buffer and eluted on ice for 15 minutes with Elution Buffer (15 mM HEPES pH 7.4, 500 mM NaCl, 7.5 mM $MgCl_2$, 2 mM DTT, and 100 µg/ml 3xFlag peptide). Aliquots of the eluate were reserved for immunoblotting.

Total RNA and immunoprecipitated RNA were extracted similarly. SDS was added to 0.5%, as well as 1 ng or 10 ng firefly luciferase, 1 µg yeast tRNA (ThermoFisher), and 200 µg/ml Proteinase K. Samples were incubated at 50C for 45 minutes. RNA was extracted twice with acid phenol/chloroform (once at 65C for 5 minutes, once at room temperature for 5 minutes), once with chloroform, and precipitated with NaOAc and isopropanol. RNA was resuspended in 10 mM Tris pH 8 and reverse transcribed using oligodT_{20} primers and Superscript III Reverse Transcriptase according to manufacturer's instructions. qPCR was performed using SYBR Green PCR Mix (Applied Biosystems) or FastStart Universal SYBR Master Mix (Roche). Individual mRNA measurements were normalized to firefly luciferase measurements, as well as total mRNA abundance and control protein measurements.

5'UTR Cloning

The murine Rps20, Eef2, Eef1d, Eef1a1, and Actb 5'UTRs were cloned from mouse embryonic fibroblast cDNA into pIS1-Eef2-5UTR-renilla (Addgene plasmid #38235) by PCR amplification with the following 5'UTR-specific primers containing a 5' EcoRI restriction site and modified T7 RNA Polymerase promoter sequence (containing only a single guanosine at position +1), and a 3' NheI restriction site. The 5'TOP sequence was subsequently mutated by PCR amplification of each cloned 5'UTR with the same reverse primer and a forward primer harboring the complimentary purine stretch after the +1 guanosine.

Rps20_TOP_1G_F: ttggcgaattcTAATACGACTCACTATAGcCTTTCTGAG
Rps20_5U_R: tggtggctagcGGCGCGGCTTCCTGACCG

Rps20_nT_1G_F: ttggcgaattcTAATACGACTCACTATAGggaaagaGAGCCCCGGCGG

Eef2_TOP_1G_F: ttggcgaattcTAATACGACTCACTATAGctcttccGCCGCAGCCGCC
Eef2_5U_R: tggtggctagcGGATGGCGACGGATTCTC
Eef2_nT_1G_F: ttggcgaattcTAATACGACTCACTATAGgagaaggGCCGCAGCCGCC

Eef1d_TOP_1G_F:
ttggcgaattcTAATACGACTCACTATAGctctcctttgcagtcctgagctcgcagttc
Eef1d_5U_R: tggtggctagcTTTGGCCTCACGCCAGCC
Eef1d_nT_1G_F:
ttggcgaattcTAATACGACTCACTATAGgagaggaaagcagtcctgagctcgcagttc

Eef1a1_TOP_1G_F:
ttggcgaattcTAATACGACTCACTATAGttcccgagggtgggggagaacggtatataa
Eef1a1_5U_R: tggtggctagcTTTGAATTAGCGGTGGTT
Eef1a1_nT_1G_F:
ttggcgaattcTAATACGACTCACTATAGaAGGGgagggtgggggagaacggtatataa

Actb_TOP_1G_F: ttggcgaattcTAATACGACTCACTATAGaCTGTCGAGT
Actb_5U_R: tggtggctagcCTGGTGGCGGGTGTGGACCG

In vitro transcription, RNA processing, and purification

The DNA template for *in vitro* transcription was generated by linearizing the 5'UTR constructs with NheI and transcribed *in vitro* using T7 RNA Polymerase in a reaction containing 1X NEB T7 RNA Polymerase buffer, 100 mM DTT, 0.6U SuperaseIn, 2 mM ATP, 2 mM UTP, 2 mM CTP, 2 mM GTP, 1 µg randomized DNA template, and 2.5U T7 RNA Polymerase per 100µl reaction volume. Transcripts were either capped co-transcriptionally using ^{me7}GpppG cap analog (New England Biolabs) or capped *in vitro* after transcription using the ScriptCap m^{7}G Capping System (Cellscript) according to manufacturer's directions. Biotinylated 5'UTRs were transcribed using biotin-16-UTP (Epicentre) at a concentration such that, on average, only one biotin-labeled UTP

should be incorporated per 5'UTR. All 5'UTRs were acid phenol/chloroform extracted and gel purified.

pCT3-Eef2-TOP, pCT3-Eef2-nonTOP were kind gifts from Carson Thoreen (Yale University). They are based on the pLJM1 lentiviral expression plasmid (Sancak et al., 2008), and encode the murine Eef2 5'UTR (or Eef2 5'UTR in which the 5'TOP motif has been mutated to the complimentary purines) downstream of a CMV promoter and upstream of the *Renilla* luciferase coding sequence and the SV40 polyadenylation signal.

Biotinylated 5'UTR Immunoprecipitiations

HeLa cells were seeded at $6x10^6$ cells per 15cm dish in 20ml DMEM supplemented with 10% IFS and penicillin/streptomycin. Cells were washed once with ice-cold PBS and lysed in 700ul Triton Lysis Buffer (40 mM HEPES pH 7.4, 40 mM KCl, 2 mM EDTA, 10 mM pyrophosphate, 10 mM β-glycerophosphate, 1% TritonX-100, and one EDTA-free protease inhibitor tablet per 25ml buffer). Soluble material was normalized by protein concentration, and 500µl at 1 mg/ml lysate was added to 5µl biotinylated 5'UTR at 25 ng/µl. Binding proceeded for 30 minutes at 4C with mixing, and 10µl Streptavidin M-280 magnetic beads washed three times with 1ml PBS and once with 1ml Triton Lysis Buffer were added to each binding reaction. Reactions were incubated 30 minutes at room temperature with mixing, washed three times with 500µl Triton Lysis Buffer, and were resuspended in 25µl TBE-Urea load buffer (ThermoFisher) to assess RNA levels or protein load buffer for immunoblotting.

EMSA

Uncapped 5'UTRs prepared as described above were radiolabeled *in vitro* using the ScriptCap m^7G Capping System similarly to manufacturer's instructions with α-^{32}P-GTP instead of GTP. Briefly, 10µl of 50 ng/ul 5'UTR was denatured 10 minutes at 75C. Meanwhile, a 10µl reaction containing 3µl water, 2µl 10X ScriptCap Capping Buffer, 1.5µl 3000Ci/mmol α-^{32}P-GTP, 2µl fresh 2mM SAM, and 0.5µl ScriptGuard RNAse Inhibitor was prepared. The denatured RNA was placed on ice, and 1µl Capping enzyme per 10µl reaction was added to the mix just before adding to the chilled 5'UTR. Capping reactions were incubated for 2 hours at 37C, gel purified, and counted in a scintillation counter. Equal counts were used for each 5'UTR in each experiment.

EMSA was performed similarly to the method of Gagnon and Maxwell (Gagnon & Maxwell, 2011). Briefly, reactions were carried out in a volume of 20µl containing a final concentration of 50 mM HEPES pH 7.4, 50 mM KCl, 1% TritonX-100, 20U SuperaseIn, 10 µg yeast tRNA, 80,000 cpm experimental 5'UTR, and protein. Protein concentrations were determined by comparison with a standard curve of BSA using Bradford Assay (Biorad). Mixes were made without protein and aliquotted to tubes, whereupon 10µl protein at different concentrations was added. Reactions were incubated for 30 minutes at 23C at 750rpm in a Thermomixer (Eppendorf). A 6% phosphate gel was made by mixing 6ml 40% acrylamide:bisacrylamide (19:1) (Amresco), 4ml 10X Phosphate Buffer (0.25 M potassium phosphate pH 7), 1.6ml 50% glycerol, 28.4ml water, and 320µl freshly-made 10% ammonium persulfate. 40µl TEMED was added just prior to pouring

the gel. Reactions run on a gel for 2 hours at 150V at 4C. Gels were dried and exposed to film or a phosphorimager screen.

^{me7}GTP Competition-qPCR

HEK-293Ts were seeded at 6x10^6 cells per 20ml in 15cm dishes. The following day, cells were transfected with 1ug pCT3-Eef2-TOP or pCT3-Eef2-nonTOP. Briefly, 1 µg pCT3 and 9 µg empty pRK5 vector was mixed with 30ul PEI (1 mg/ml) and 670µl serum- and antibiotic-free DMEM for 5 minutes at room temperature. The mixture was added drop-wise to cells, and total RNA was harvested 48 hours later according to the RNA-STAT 60 reagent procedure described above. PolyA RNA was selected using oligodT_{25} Dynabeads (ThermoFisher). 375µl of RNA at 2 µg/ul was mixed with 2X Binding Buffer (40 mM Tris pH 7, 2 M LiCl, 4 mM EDTA), incubated for 2 min at 80C, and placed on ice. 3.75µl SuperaseIn was added. 750µl oligodT_{25} Dynabeads were washed twice with 750ul 1X Binding Buffer and resuspended in the RNA/buffer mixture. Binding was allowed to proceed for 5 min at room temperature with mixing. Beads were washed twice with Wash Buffer B (10 mM Tris pH 7, 150 mM LiCl, and 1 mM EDTA), resuspended in 15µl 10mM Tris pH 8, and eluted by heating at 80C for 2 minutes. Eluate was removed to a new tube on ice. Beads were regenerated according to manufacturer's instructions.

Binding reactions were performed in a 20µl reaction containing 1X qPCR Binding Buffer (20 mM HEPES pH 7.4, 150 mM KCl, 2.5 mM $MgCl_2$, 1% TritonX-100, 0.5 mM DTT, 15 µg/ml polyI RNA (Sigma), and 20U SuperaseIn), various concentrations of

^{me7}GTP or GTP, 1 µg polyA RNA (~50nM, calculated based on average mammalian mRNA size), and 1 µM SBP-eIF4E. Mixes were made without RNA or me7GTP or GTP and aliquotted to tubes. ^{me7}GTP or GTP was added and reactions were incubated at 23C for 30 minutes. RNA was added and reactions were incubated an additional hour at 23C at 750rpm in a Thermomixer. 25µl Streptavidin M-280 beads per reaction were washed in a batch three times with 1ml 1X qPCR Binding Buffer. Beads were resuspended in 1X qPCR Binding Buffer and aliquotted to tubes. Binding reactions were added to the beads and incubated one hour at 23C at 750rpm. Beads were washed once with 1ml qPCR Wash Buffer (20 mM HEPES pH 7.4, 150 mM KCl, 0.5 mM EDTA, 1% TritonX-100), and resuspended in 200µl qPCR Elution Buffer (20 mM HEPES pH 7.4, 1 mM EDTA, 1% SDS). Reactions were heated 10 minutes at 70C, and 5 ng firefly luciferase was added. Reactions were extracted with acid phenol/chloroform and NaOAc/isopropanol precipitated. cDNA synthesis and qPCR were performed as described above, except that experimental RNAs were normalized to the spiked in control firefly luciferase, then to the amount of that RNA in the input mixture. The GTP control sample was set as 100% binding.

qPCR Primers

Mouse:

Eef2_F: GAGAATCCGTCGCCATCCGCCA
Eef2_R: CGGGCTTGATGCGTTCAGCGA
Rps20_F: TGACTCACCGCTGTTCGCTCC
Rps20_R: GAGTCGCTTGTGGATCCTCATCTGG
Hsp90ab1_F: GCCGTGCGAGTCGGACTTGGT
Hsp90ab1_R: CCGACACCAAACTGCCCGATCA

Pabpc1_F: CGCTGGACTGCTCAGGGTGC
Pabpc1_R: GGGGGCGCAGATGCCAACAT
Ybx1_F: GGGGTCCTCCACGCAATTACC
Ybx1_R: CGGCGATACCGACGTTGAGGT
Vim_F: ACTGCTGCCCTGCGTGATGT
Vim_R: TCTCACGCATCTGGCGCTCC
Actb_F: TCGTTGCCGGTCCACACCCG
Actb_R: CTCCTCAGGGGCCACACGCAG
Mrpl22_F: TCTGGGCAACGCAGACGCTG
Mrpl22_R: GCCAAAGCGACCTCGGCCAT
Ccnd1_F: GCCCGAGGAGCTGCTGCAAA
Ccnd1_R: GCCTTGCATCGCAGCCACCA
Slc2a1_F: CTGGCATGGCAGGCTGTGCT
Slc2a1_R: CGCCCCCAGAGGGTGGAAGA
Gabarapl1_F: AGCCCCCAAAGCTCGGATAGGA
Gabarapl1_R: GGTGTTCCTGGTACAGCTGACCC
Myc_F: GCCAGCCCTGAGCCCCTAGT
Myc_R: GGGTGCGGCGTAGTTGTGCT

Human:

LDHA_F: GGCTACACATCCTGGGCTAT
LDHA_R: CAGCTCCTTTTGGATCCCCC
Eef2_F: CAGAGCAAGGACGGTGCCGG
Eef2_R: TCTGCACGCACACGCCTGAC
Gnb2l1_F: TGGGATCTCACAACGGGCACCA
Gnb2l1_R: CCGGTTGTCAGAGGAGAAGGCCA
Rps20_F: CCAGTTCGAATGCCTACCAAGACTT
Rps20_R: ACTTCCACCTCAACTCCTGGCTCA
Hsp90ab1_F: CCGGCGCAGTGTTGGGACTG
Hsp90ab1_R: GGTGCACTTCCTCAGGCATCTTG
Rplp2_F: CGACCGGCTCAACAAGGTTA
Rplp2_R: GGCTTTATTTGCAGGGGAGC
Ybx1_F: GGTCATCGCAACGAAGGTTTTGGG
Ybx1_R: CACTGCGAAGGTACTTCCTGGGG
Myc_F: CGTCCTCGGATTCTCTGCTC
Myc_R: GCTGCGTAGTTGTGCTGATG
Sqstm1_F: CCTTGTACCCACATCTCCCG
Sqstm1_R: GCCGCTCCGATGTCATAGTT

firefly luciferase_F: ATCCGGAAGCGACCAACGCC
firefly luciferase_R: GTCGGGAAGACCTGCCACGC
Renilla luciferase_F: TGGCCTCGTGAAATCCCGTTAGT
Renilla luciferase_R: TTTCATCAGGTGCATCTTCTTGCGA

ACKNOWLEDGEMENTS

We would like to thank members of the Sabatini Lab for valuable discussions and input, especially Ellen Edenberg and Larry Schweitzer. We would also like to thank Ellen Edenberg for technical assistance. This work was supported by grants from the NIH (R01 CA129105, R01 CA103866, R37 AI047389) to D.M.S. D.M.S. is also an investigator of the Howard Hughes Medical Institute.

REFERENCES

Bah, A., Vernon, R. M., Siddiqui, Z., Krzeminski, M., Muhandiram, R., Zhao, C., . . . Forman-Kay, J. D. (2015). Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature, 519*(7541), 106-109. doi: 10.1038/nature13999

Bar-Peled, L., Schweitzer, L. D., Zoncu, R., & Sabatini, D. M. (2012). Ragulator is a GEF for the rag GTPases that signal amino acid levels to mTORC1. *Cell, 150*(6), 1196-1208. doi: 10.1016/j.cell.2012.07.032

Bhat, M., Robichaud, N., Hulea, L., Sonenberg, N., Pelletier, J., & Topisirovic, I. (2015). Targeting the translation machinery in cancer. *Nat Rev Drug Discov, 14*(4), 261-278. doi: 10.1038/nrd4505

Biberman, Y., & Meyuhas, O. (1999). TOP mRNAs are translationally inhibited by a titratable repressor in both wheat germ extract and reticulocyte lysate. *FEBS Lett, 456*(3), 357-360.

Brunn, G. J., Hudson, C. C., Sekulic, A., Williams, J. M., Hosoi, H., Houghton, P. J., . . . Abraham, R. T. (1997). Phosphorylation of the translational repressor PHAS-I by the mammalian target of rapamycin. *Science, 277*(5322), 99-101.

Damgaard, C. K., & Lykke-Andersen, J. (2011). Translational coregulation of 5'TOP mRNAs by TIA-1 and TIAR. *Genes Dev, 25*(19), 2057-2068. doi: 10.1101/gad.17355911

Dowling, R. J., Topisirovic, I., Alain, T., Bidinosti, M., Fonseca, B. D., Petroulakis, E., . . . Sonenberg, N. (2010). mTORC1-mediated cell proliferation, but not cell growth, controlled by the 4E-BPs. *Science, 328*(5982), 1172-1176. doi: 10.1126/science.1187532

Dunn, J. J., & Studier, F. W. (1983). Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J Mol Biol, 166*(4), 477-535.

Fraser, C. S. (2015). Quantitative studies of mRNA recruitment to the eukaryotic ribosome. *Biochimie, 114*, 58-71. doi: 10.1016/j.biochi.2015.02.017

Gagnon, K. T., & Maxwell, E. S. (2011). Electrophoretic mobility shift assay for characterizing RNA-protein interaction. *Methods Mol Biol, 703*, 275-291. doi: 10.1007/978-1-59745-248-9_19

Ingolia, N. T., Ghaemmaghami, S., Newman, J. R., & Weissman, J. S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science, 324*(5924), 218-223. doi: 10.1126/science.1168978

Jefferies, H. B., Reinhard, C., Kozma, S. C., & Thomas, G. (1994). Rapamycin selectively represses translation of the "polypyrimidine tract" mRNA family. *Proc Natl Acad Sci U S A, 91*(10), 4441-4445.

Kakegawa, T., Ohuchi, N., Hayakawa, A., Hirata, S., Matsuda, M., Kogure, K., . . . Kaspar, R. L. (2007). Identification of AUF1 as a rapamycin-responsive binding protein to the 5'-terminal oligopyrimidine element of mRNAs. *Arch Biochem Biophys, 465*(1), 274-281. doi: 10.1016/j.abb.2007.06.001

Lodish, H. F. (1974). Model for the regulation of mRNA translation applied to haemoglobin synthesis. *Nature, 251*(5474), 385-388.

Marcotrigiano, J., Gingras, A. C., Sonenberg, N., & Burley, S. K. (1999). Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. *Mol Cell, 3*(6), 707-716.

Marintchev, A., Edmonds, K. A., Marintcheva, B., Hendrickson, E., Oberer, M., Suzuki, C., . . . Wagner, G. (2009). Topology and regulation of the human eIF4A/4G/4H helicase complex in translation initiation. *Cell, 136*(3), 447-460. doi: 10.1016/j.cell.2009.01.014

Merrick, W. C. (2015). eIF4F: a retrospective. *J Biol Chem, 290*(40), 24091-24099. doi: 10.1074/jbc.R115.675280

Orom, U. A., Nielsen, F. C., & Lund, A. H. (2008). MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Mol Cell, 30*(4), 460-471. doi: 10.1016/j.molcel.2008.05.001

Pause, A., Belsham, G. J., Gingras, A. C., Donze, O., Lin, T. A., Lawrence, J. C., Jr., & Sonenberg, N. (1994). Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function. *Nature, 371*(6500), 762-767. doi: 10.1038/371762a0

Petroulakis, E., Parsyan, A., Dowling, R. J., LeBacquer, O., Martineau, Y., Bidinosti, M., . . . Sonenberg, N. (2009). p53-dependent translational control of senescence and transformation via 4E-BPs. *Cancer Cell, 16*(5), 439-446. doi: 10.1016/j.ccr.2009.09.025

Prevot, D., Darlix, J. L., & Ohlmann, T. (2003). Conducting the initiation of protein synthesis: the role of eIF4G. *Biol Cell, 95*(3-4), 141-156.

Sancak, Y., Peterson, T. R., Shaul, Y. D., Lindquist, R. A., Thoreen, C. C., Bar-Peled, L., & Sabatini, D. M. (2008). The Rag GTPases bind raptor and mediate amino acid signaling to mTORC1. *Science, 320*(5882), 1496-1501. doi: 10.1126/science.1157535

Sonenberg, N., & Hinnebusch, A. G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell, 136*(4), 731-745. doi: 10.1016/j.cell.2009.01.042

Tcherkezian, J., Cargnello, M., Romeo, Y., Huttlin, E. L., Lavoie, G., Gygi, S. P., & Roux, P. P. (2014). Proteomic analysis of cap-dependent translation identifies LARP1 as a key regulator of 5'TOP mRNA translation. *Genes Dev, 28*(4), 357-371. doi: 10.1101/gad.231407.113

Thoreen, C. C., Kang, S. A., Chang, J. W., Liu, Q., Zhang, J., Gao, Y., . . . Gray, N. S. (2009). An ATP-competitive mammalian target of rapamycin inhibitor reveals rapamycin-resistant functions of mTORC1. *J Biol Chem, 284*(12), 8023-8032. doi: 10.1074/jbc.M900301200

von der Haar, T., Gross, J. D., Wagner, G., & McCarthy, J. E. (2004). The mRNA cap-binding protein eIF4E in post-transcriptional gene expression. *Nat Struct Mol Biol, 11*(6), 503-511. doi: 10.1038/nsmb779

Warner, J. R. (1999). The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci, 24*(11), 437-440.

Chapter 4

Juxtacap Nucleotide Sequence Modulates eIF4E Binding and Translation

Heather R. Keys[1-2] and David M. Sabatini[1-5]

[1]Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA

[2]Department of Biology, MIT, Cambridge, MA 02139, USA

[3]Howard Hughes Medical Institute, MIT, Cambridge, MA 02139, USA

[4]Broad Institute, Cambridge, MA 02142, USA

[5]The David H. Koch Institute for Integrative Cancer Research, Cambridge, MA 02139, USA

SUMMARY

mRNA translation is an energetically costly activity required for almost all biological processes. The multiprotein complex eIF4F, which bridges the 5' cap and the polyA tail through eIF4E and eIF4G, respectively, is necessary for efficient translation initiation of most mRNAs and is an important target of translational control. Previous work suggests that cap-proximal nucleotides can modulate eIF4E binding to mRNAs, but the effect of specific cap-proximal nucleotide sequences on eIF4E recruitment and the ultimate consequences for translation remain unknown. Using RNA Bind-n-Seq on a model 5' UTR, we systematically identify eIF4E-intrinsic cap-proximal nucleotide binding preferences. mRNAs with highly-bound motifs are translated well in a cell-free system, whereas those with low-eIF4E-binding motifs are not. However, eIF4E juxtacap motif preferences do not dictate the ribosome occupancy of endogenous mRNAs in cells, suggesting that the effect of juxtacap sequence on eIF4E binding and translation is mRNA-context-dependent. Accordingly, a single downstream point mutation that disrupts a predicted base pair with a preferred juxtacap nucleotide increases translation. The juxtacap sequence is a previously unappreciated determinant of eIF4E recruitment to mRNAs, and we propose that differences in mRNA 5' end accessibility defined by the juxtacap sequence are important for establishing translational efficiency.

INTRODUCTION

Protein synthesis is a fundamental process essential for almost all biological phenomena and as such is highly regulated. As the rate-limiting step of protein synthesis under normal conditions (Golini et al., 1976; Walden, Godefroy-Colburn, & Thach, 1981), translation initiation is particularly well regulated. Translation initiation can be dynamically regulated in a global and general manner through modulation of core translation factors (Fraser, 2015; Sonenberg & Hinnebusch, 2009). Additionally, the translation of specific mRNAs can be temporally and spatially controlled in diverse cell types or at specific developmental stages (Fraser, 2015; Sonenberg & Hinnebusch, 2009).

Two main regulatory arms impact translation initiation in eukaryotes. One point of regulation occurs at the formation of the ternary complex, which comprises the initiation factor eIF2, the initiator methionyl-tRNA, and GTP, and which is required for recognition of the start codon (Sonenberg & Hinnebusch, 2009). The other major point of regulation is at the recruitment of the 43S preinitiation complex (43S PIC) to the mRNA (Fraser, 2015). The 43S PIC includes the 40S small ribosomal subunit, the ternary complex, and initiation factors eIF1, eIF1A, eIF3, and eIF5 (Fraser, 2015; Jackson, Hellen, & Pestova, 2010). The 43S PIC is recruited to the mRNA via an interaction between eIF3 and the multiprotein complex eIF4F (Lamphear, Kirchweger, Skern, & Rhoads, 1995). The mechanistic target of rapamycin complex 1 (mTORC1), a kinase complex that integrates nutrient and growth factor inputs to promote several processes important for cell growth, regulates the assembly of eIF4F to impact translation initiation (Thoreen, 2013).

eIF4F comprises the cap-binding protein eIF4E, the scaffolding protein eIF4G, and the helicase eIF4A (Merrick, 2015). eIF4E performs the key first step of eIF4F function by recognizing the 5' end of the mRNA via its interaction with the 7-methylguanosine (^{me7}G) cap, thereby positioning eIF4F and the 43S PIC at the 5' end of the mRNA (Sonenberg, Morgan, Merrick, & Shatkin, 1978). Incorporation of eIF4E into the eIF4F complex is an important modulatory point for translational regulation, and is the step specifically controlled by mTORC1 (Brunn et al., 1997; Marcotrigiano, Gingras, Sonenberg, & Burley, 1999). mTORC1 phosphorylates the 4E-binding proteins (4EBPs), rendering them inactive for binding to eIF4E (Brunn et al., 1997). Because 4EBPs compete with eIF4G for binding to eIF4E (Haghighat, Mader, Pause, & Sonenberg, 1995; Marcotrigiano et al., 1999), mTORC1 activity promotes assembly of eIF4F. eIF4G is a large protein that binds eIF4E, mRNA, polyA-binding protein (PABP), eIF3, and eIF4A (Prevot, Darlix, & Ohlmann, 2003). eIF4G performs several functions: stabilizing the RNA-eIF4E interaction by binding RNA directly, bridging the 5' and 3' termini of the mRNA through interactions with eIF4E and PABP, and recruiting the 43S PIC to the 5' end of the mRNA via the interaction with eIF3 (Prevot et al., 2003). Furthermore, eIF4G localizes eIF4A to the mRNA to unwind 5'UTR secondary structure and facilitate scanning of the 40S ribosomal subunit (Prevot et al., 2003). Thus, eIF4F coordinates many interactions required for translation initiation.

eIF4E is the least abundant member of eIF4F (Duncan, Milburn, & Hershey, 1987; Hiremath, Webb, & Rhoads, 1985) and, when mTORC1 is inhibited, becomes even more limited due to sequestration from eIF4F by the 4EBPs (Brunn et al., 1997;

Marcotrigiano et al., 1999; Pause et al., 1994). eIF4E is overexpressed in several types of cancer (Bhat et al., 2015), experimental overexpression of eIF4E in cells leads to oncogenic transformation *in vitro* (Lazaris-Karatzas, Montine, & Sonenberg, 1990), and overexpression in mice leads to increased tumor incidence (Ruggero et al., 2004). Intriguingly, eIF4E overexpression affects translation of some mRNAs more than others (Mamane et al., 2007), suggesting that translation of this subset of messages is particularly dependent on eIF4E. One class of mRNAs whose translation is enhanced by eIF4E overexpression harbors a pyrimidine stretch at the 5' terminus (5' TOPs) (Mamane et al., 2007). mRNAs harboring this 5' TOP motif are also the mRNAs whose translation is the most strongly inhibited when eIF4F is disrupted by acute inhibition of mTOR (Hsieh et al., 2012; Thoreen et al., 2012). The sensitivity of 5' TOP mRNA translation to eIF4F inhibition and eIF4E overexpression suggests that the nucleotides proximal to the 5' cap may play a role in the recruitment of eIF4E to mRNAs.

Biochemical and structural work has primarily detailed the eIF4E-cap interaction (Carberry, Rhoads, & Goss, 1989; Marcotrigiano, Gingras, Sonenberg, & Burley, 1997; Matsuo et al., 1997; Niedzwiecka et al., 2002; Tomoo et al., 2005; Tomoo et al., 2002), and the role of cap-proximal nucleotides in eIF4E recruitment remains unclear. The crystal structure of eIF4E bound to the cap analog ^{me7}GDP reveals a cleft of positive electrostatic potential directly adjacent to ^{me7}GDP, through which the +1 nucleotide could be threaded (Marcotrigiano et al., 1997). In the co-crystal structure of human eIF4E with the cap analog ^{me7}GpppA, contact of eIF4E with adenosine at the +1 position stabilizes an otherwise flexible C-terminal loop of eIF4E, indicating that eIF4E can

contact at least the first juxtacap nucleotide (Tomoo et al., 2002). NMR experiments with yeast eIF4E and ^{me7}GpppA support this finding (Matsuo et al., 1997). However, in a co-crystal structure of murine eIF4E with ^{me7}GpppG, no electron density is observed for the +1 guanosine, consistent with the absence of a direct stabilizing interaction with the +1 nucleotide (Niedzwiecka et al., 2002), and suggesting that an interaction between eIF4E and the +1 nucleotide could depend on nucleotide identity.

How eIF4E binding and translation are influenced by additional nucleotides downstream of the +1 position is even less clear. Subsequent cap-proximal nucleotides can decrease the binding of eIF4E relative to the 5' cap alone, if those nucleotides induce secondary structure in the RNA near the 5' end that physically blocks eIF4E binding (Carberry, Friedland, Rhoads, & Goss, 1992). In another case, cap-proximal secondary structure does not prevent eIF4E binding, but may change the trajectory of the RNA as it exits the cap-binding pocket of eIF4E in a manner inhibitory for translation (O'Leary, Petrov, Chen, & Puglisi, 2013). Conversely, kinetic measurements have shown that the presence of a short oligonucleotide stretch adjacent to the cap can increase the affinity of eIF4E for the RNA (Slepenkov, Darzynkiewicz, & Rhoads, 2006). It appears that the presence of juxtacap nucleotides can influence the binding of eIF4E to mRNA; however, the direction, magnitude, and sequence-dependence of this effect remain unclear.

Here, we systematically identify juxtacap sequences that affect recruitment of eIF4E to a model 5' UTR, and that modulate translation of mRNAs containing the model 5'UTR in a cell-free system. We propose that the juxtacap sequence influences eIF4E binding

and translation indirectly, by influencing secondary structure at the 5' end of the mRNA. Mutations downstream of the juxtacap motif predicted to alter the cap-proximal secondary structure affect translation *in vitro*, in the absence of changes to the juxtacap motif itself. The juxtacap sequence is a previously unappreciated determinant of eIF4E recruitment to mRNAs, and could be an important contributor to establishing translational efficiency *in vivo*.

RESULTS

eIF4E RNA Bind-n-Seq

To systematically identify juxtacap sequences that affect eIF4E binding, we performed a modified RNA Bind-n-Seq (RBNS) protocol (Lambert et al., 2014) using a library of 7-methylguanosine-capped (^{me7}G-capped) RNAs adapted from the murine Rps20 5' UTR (Figure 1A). The T7 RNA polymerase consensus sequence encodes guanosines at positions +1, +2, and +3 from the 5' cap (Dunn & Studier, 1983), which would severely restrict the variability of the juxtacap nucleotides in the library of motifs (Figure 1B). We cloned the Rps20 5' UTR downstream of a modified T7 promoter without guanosines at positions +1, +2, and +3, but were unable to synthesize mRNA from this template (data not shown). Instead, we constructed a template encoding only a single guanosine at position +1 (Figure 1A). We introduced random nucleotides at positions +2 through +6 by PCR of the entire 5' UTR downstream of the modified T7 promoter, and transcribed the randomized 5'UTR *in vitro* (Figure 1A). We capped the 5'UTRs *in vitro*, and performed a series of enzymatic steps to remove any uncapped RNA (Figure 1C). The library contains 1024 motifs of five randomized nucleotides downstream of the ^{me7}G cap and the +1 guanosine, and prepended to the remainder of the murine Rps20 5' UTR.

We incubated the input library with increasing concentrations of recombinant streptavidin binding peptide (SBP)-tagged eIF4E, and isolated eIF4E-bound RNAs through capture of eIF4E with streptavidin-coated magnetic beads. We constructed sequencing libraries from the bound RNAs at each concentration of eIF4E and from the input library, and performed high-throughput sequencing. We normalized the frequency

of each RNA species bound to eIF4E by its frequency in the input library to identify enriched and depleted motifs.

Figure 1: eIF4E Bind-n-Seq

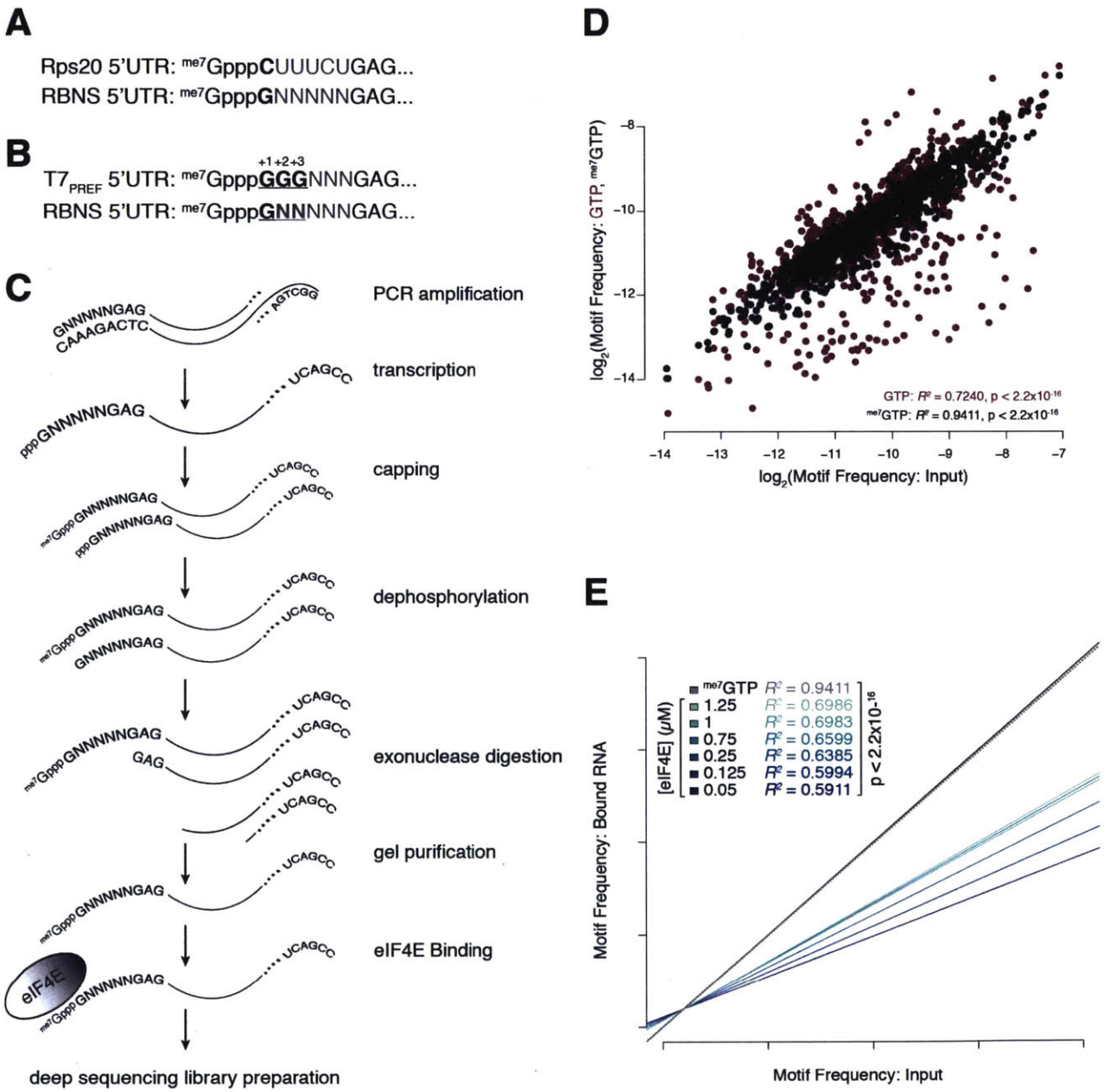


Figure 1: eIF4E Bind-n-Seq

A Diagram of the 5' end of the Rps20 5'UTR and the RNA used for RNA Bind-n-Seq.

B. Diagram of the 5' end of the Rps20 5' UTR transcribed using a template containing a consensus T7 RNA Polymerase promoter versus the RNA Bind-n-Seq template.

C. Overview of RNA Bind-n-Seq library generation. Processing steps to ensure that the RNA Bind-n-Seq library contains only capped RNAs are detailed.

D. eIF4E possesses cap-dependent juxtacap nucleotide preferences. The log_2-transformed frequency of each juxtacap motif in the input library was plotted against either its frequency in the control sample containing 1.25 μM eIF4E and either 10 mM GTP (purple) or ^{me7}GTP (black). Pearson correlation coefficients and p-values are indicated.

E. Juxtacap nucleotide preferences are more apparent when eIF4E is more limiting. Linear regression was performed for the frequency of each motif in the input library versus its frequency in the eIF4E-bound library at each concentration of eIF4E, and the best-fit lines were plotted. Pearson correlation coefficients and p-values are indicated.

Although the interaction of most RNAs with eIF4E requires the 5' cap, some RNA sequences can bind eIF4E in a cap-independent manner (Martin et al., 2011; Mochizuki, Oguro, Ohtsu, Sonenberg, & Nakamura, 2005). To control for cap-independent and background binding, we measured binding at the highest eIF4E concentration (1.25 μM) in the presence of 7-methyl-GTP, (^{me7}GTP), a competitive inhibitor of eIF4E binding to the 5' cap, or GTP, which does not compete with the cap for binding to eIF4E. As expected, ^{me7}GTP addition blocked most RNA binding to eIF4E (data not shown), and the motif frequencies in the ^{me7}GTP sample correlated very highly with input frequencies (R^2 = 0.9411), suggesting primarily cap-dependent binding (Figure 1D). In contrast, we observed less correlation between the frequency of motifs in the eIF4E-bound sample

and the input library for the control sample containing 1.25 μM eIF4E and GTP (R^2 = 0.7420) (Figure 1D), suggesting that eIF4E does have varying affinity for different motifs. In the experimental samples without GTP or ^{me7}GTP, motif representation at the lowest eIF4E concentration (50 nM) was poorly correlated with representation in the input library (R^2 = 0.5911). The correlation increased with increasing eIF4E concentration (Figure 1E), which is expected as eIF4E becomes less limiting and there is less competition between RNAs for binding to eIF4E. Together, these data indicate that eIF4E does indeed bind particular juxtacap nucleotide sequences differently.

Bind-n-Seq Experiment Reveals Juxtacap Sequence Preferences for eIF4E

In order to compare eIF4E binding between motifs present at different initial frequencies in the input library, we calculated the enrichment score. We defined the enrichment score as the frequency of each eIF4E-bound motif, relative to the frequency of the same motif in the input library. We calculated the enrichment score at all concentrations of eIF4E for the 1024 motifs present in the input library. To quantify the total amount of binding across all eIF4E concentrations, we calculated the area under the curve when we plotted eIF4E concentration versus enrichment score (Figure 2A and Figure 2B). The motif with the greatest cumulative binding to eIF4E had a calculated area under the curve that was approximately 78-fold higher than the motif with the least binding to eIF4E, corresponding to a 67-fold higher enrichment score at the lowest eIF4E concentration (Figure 2C). We performed kmeans clustering of the enrichment score at 50 nM eIF4E and examined the juxtacap sequence motifs assigned to each of

the ten clusters (Figure 2D). From these clusters, we classified two groups of motifs: "weak binders," which showed very little eIF4E binding, even at high eIF4E concentrations, and comprise the lowest enrichment kmeans cluster (n=77; median enrichment score for 50 nM eIF4E = 0.308, median AUC = 1.27), and "strong binders," which exhibited high eIF4E binding, even at low eIF4E concentrations, and comprise the highest three enrichment kmeans clusters (n = 46; median enrichment score for 50 nM eIF4E = 3.23; median AUC = 14.34) (Figure 2E). Motifs with weak eIF4E binding were highly enriched for guanosine in the +2 position and depleted for guanosine in the +6 position, with no other major nucleotide preferences (Figure 2E, Figure 2F, and Figure 2H). In contrast, enrichment of +4 guanosine and +5 cytidine, and depletion of +6 adenosine were characteristic of the strong binding motifs (Figure 2E, Figure 2G, and Figure 2H). These data indicate that eIF4E has position-dependent preferences for juxtacap nucleotides.

The Identity of the Juxtacap Nucelotides Affects Translation

Having established that juxtacap sequence identity influences eIF4E binding, we sought to determine whether the juxtacap sequence also influences translation. We individually cloned 5'UTRs of the Rps20 gene containing juxtacap motifs from the clusters of weak and strong eIF4E binding motifs downstream of the modified T7 promoter and upstream of a *Renilla* luciferase-expressing construct harboring an encoded polyA tail (Figure 3A). These constructs contain a 5' UTR that is identical to those in the library used for RNA Bind-n-Seq except for the Kozak sequence, but

Figure 2: eIF4E Bind-n-Seq Reveals Juxtacap Sequence Preferences

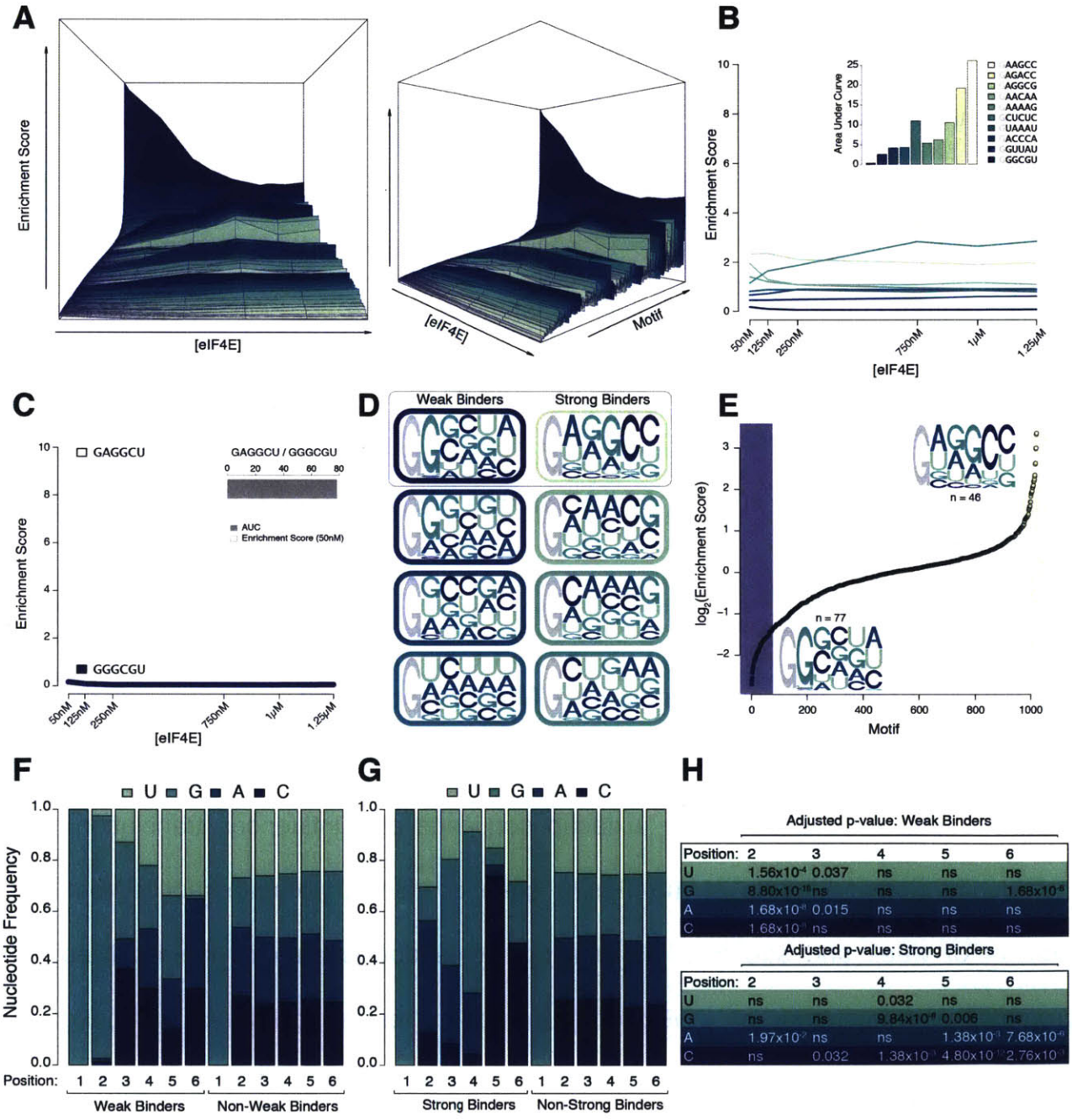


Figure 2: eIF4E Bind-n-Seq reveals juxtacap sequence preferences

A. Overview of the distribution of enrichment scores for each motif across all eIF4E concentrations. Motifs were sorted from least enriched (front) to most enriched (back) in the 50 nM eIF4E sample, and each eIF4E concentration was plotted against the enrichment score at that eIF4E concentration for each motif.

B. Examples of enrichment score distributions and area under the curve for individual motifs from each kmeans group across all eIF4E concentrations. Main: eIF4E concentration is plotted against enrichment score for individual motifs from each kmeans group. Inset: Area under the curve was plotted for each representative motif.

C. The most strongly- and weakly-bound motifs are enriched and depleted, respectively, across all concentrations of eIF4E; area under the curve and enrichment score are similar metrics for eIF4E binding comparisons. Main: eIF4E concentration was plotted against enrichment score for the most strongly bound (ivory) and most weakly bound (purple) motif. Inset: the ratio of most strongly bound to most weakly bound motif was plotted for the area under the curve (gray) and enrichment score at 50 nM eIF4E (white).

D. Distribution of enrichment scores at 50 nM eIF4E and nucleotide frequencies for motifs in each kmeans cluster reveals nucleotide preferences primarily for strongly and weakly bound motifs. The $\log_2$ of the enrichment score at 50 nM eIF4E was plotted for each motif. Juxtacap motifs for all kmeans groups were calculated, with the three highest kmeans groups combined (number of motifs in each group, from most-weakly-bound to most-strongly-bound: n = 77, n = 100, n = 137, n = 202, n = 239, n = 154, n = 69, n = 46). Borders are colored from weakest eIF4E binding (dark) to strongest eIF4E binding (light).

E. Distribution of enrichment scores at 50 nM eIF4E and nucleotide frequencies for strong- and weak-eIF4E-binding motifs reveals nucleotide preferences. The $\log_2$ of the enrichment score at 50 nM eIF4E was plotted for each motif. Highlighted are the lowest kmeans group (n=77) and highest three kmeans groups (n=46). Juxtacap motifs for the lowest and three highest kmeans groups are indicated.

F. Weak-eIF4E-binding motifs are highly enriched for +1 guanosine and depleted for +6 guanosine. The nucleotide frequencies at each position of the juxtacap motif for weak-eIF4E-binding motifs were plotted alongside the nucleotide frequencies for the motifs in the rest of the kmeans enrichment score groups.

G. Strong-eIF4E-binding motifs are highly enriched for +4 guanosine and +5 cytosine, and depleted for +6 adenosine. The nucleotide frequencies at each position of the juxtacap motif for strong-eIF4E-binding motifs and for the motifs in the rest of the kmeans enrichment score groups were plotted.

H. Adjusted p-values for nucleotide frequencies of weak- and strong-eIF4E-binding motifs. P-values were determined by two-sided Fisher's exact test and adjusted by Bonferroni correction, and are summarized for each nucleotide at positions 2 through 6.

encode *Renilla* luciferase so that we could measure the level of translation of each mRNA in a cell-free translation assay. We *in vitro* transcribed and cotranscriptionally capped the RNAs using anti-reverse cap analog, and measured their translation *in vitro* in HeLa cell lysates. Weak eIF4E binders were translated more poorly than strong eIF4E binders (Figure 3B), suggesting that the juxtacap sequence identity can modulate translation.

Figure 3: Juxtacap Nucleotide Identity Affects Translation

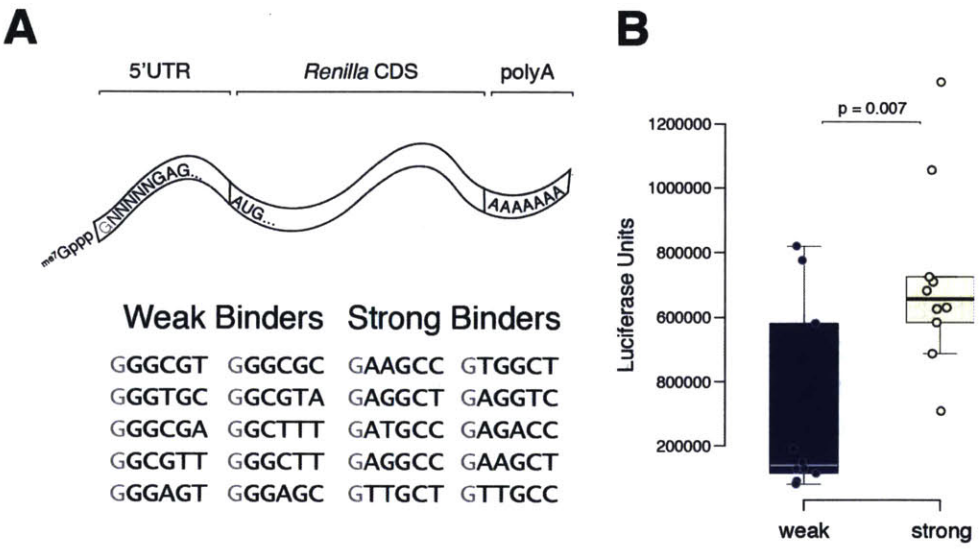


Figure 3: Juxtacap nucleotide identity affects translation

A. Diagram of the *Renilla* luciferase-expressing construct and list of weak- and strong-eIF4E-binding motifs whose translation was measured.

B. Weak- and strong-binding motifs are translated differently in a cell-free translation system. Box-and-whisker plots for weak-eIF4E-binding and strong-eIF4E-binding motifs were generated from the mean luciferase unit values from three independent experiments for each mRNA. The mean luciferase values for each mRNA are indicated as points. Significance was determined by the Wilcoxon rank sum test.

Juxtacap Nucleotide Identity Does not Correlate with Translation in Cells

The observation that eIF4E possesses a range of binding preferences for RNAs containing different juxtacap sequences suggests two alternative hypotheses. The first possibility is that eIF4E directly interacts with nucleotides downstream of the 5' cap, and that the juxtacap nucleotide sequence directly dictates the extent of eIF4E interaction; in this case, the effect of a juxtacap nucleotide sequence would be maintained across all mRNAs containing that juxtacap nucleotide sequence and translated in a cap-dependent fashion, regardless of the 5'UTR context. An alternative possibility is that eIF4E does not directly recognize the nucleotides downstream of the cap, but rather is affected by RNA secondary structures adjacent to the cap, which exist in the context of the longer 5'UTR and remainder of the mRNA. If this were the case, juxtacap motifs would promote specific RNA structures, which would dictate the level of eIF4E binding and subsequent translation. These structures would be independent of the specific sequence of the juxtacap nucleotides, and instead would rely on the greater sequence context of the juxtacap motifs.

To investigate the first possibility, we compared published transcriptional start site (Lizio et al., 2015) and ribosome footprinting (Thoreen et al., 2012) datasets and asked whether our weak- and strong-eIF4E-binding motifs were associated with lowly- and highly-translated mRNAs in cells, respectively. Importantly, we examined only endogenous 5'UTRs that begin with a guanosine, as this feature was invariant in our RNA Bind-n-Seq library. We observed no overall correlation between ribosome occupancy and binding (Figure 4A), and no difference in the ribosome occupancy across kmeans enrichment score groups (data not shown). Furthermore, we saw no difference when we compared the strong-binding motifs represented in the ribosome footprinting dataset (top three kmeans enrichment score clusters; 42/46 motifs mapped to 210 endogenous mRNAs) with the weak-binding motifs represented (lowest kmeans enrichment score group; 53/77 motifs mapped to 2763 endogenous mRNAs) (Figure 4B), nor when we compared both weak- and strong-binding groups to the rest of the mapped motifs (Figure 4B). Additionally, we performed kmeans clustering by ribosome occupancy, and examined the juxtacap nucleotide preferences of highly- and lowly-translated motifs (Figure 4C and Figure 4D). Neither group of motifs showed significant nucleotide preferences (Figure 4D), further indicating that the effect of juxtacap sequence on translation is context-dependent. These analyses suggest that the juxtacap sequence itself is not sufficient to restrict or promote translation in cells.

Juxtacap Nucleotide Sequences Can Modulate Translation by Defining the Cap-Proximal RNA Structure

The lack of correlation between the motifs enriched in eIF4E binding from our Bind-n-Seq and translation of mRNAs in cells suggests that the effect of cap-proximal nucleotide identity on eIF4E binding is mRNA-context-dependent. We sought to test the

Figure 4: Juxtacap Nucleotide Identity Does not Correlate with Translation in Cells

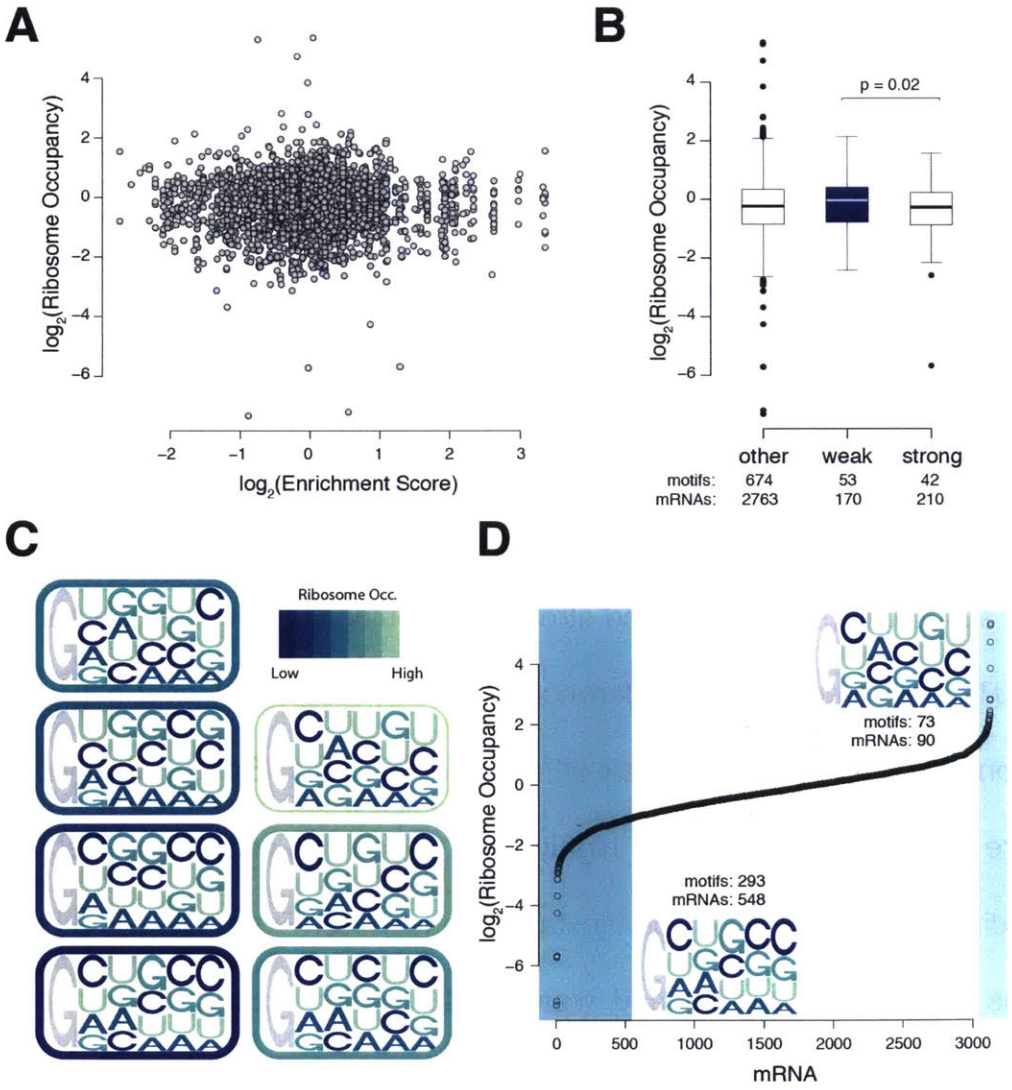


Figure 4: Juxtacap nucleotide identity does not correlate with translation in cells

A. Ribosome occupancy of endogenous mRNAs that begin with a +1 guanosine does not correlate with eIF4E binding. The log_2-transformed Bind-n-Seq enrichment score for each motif at 50 nM eIF4E is plotted against the log_2-transformed ribosome occupancy of endogenous mRNAs.

B. Ribosome occupancy of endogenous mRNAs containing weak- and strong-eIF4E-binding juxtacap motifs is not different from ribosome occupancy of mRNAs containing other juxtacap motifs. Log_2-transformed ribosome occupancy of endogenous mRNAs containing weak-binding, strong-binding, or other motifs is plotted.
C. Motif analysis for motifs in each kmeans cluster reveals no major nucleotide preferences. Kmeans groups were generated based on ribosome occupancy, the four highest kmeans groups were combined, and juxtacap nucleotide frequencies for all kmeans groups were calculated (number of unique motifs/mRNAs in each group, from most-lowly-translated to most-highly-translated: n = 293/548, n = 374/693, n = 363/654, n = 313/558, n = 272/403, n = 121/187, n = 73/90). Borders are colored from lowest ribosome occupancy (dark) to highest ribosome occupancy (light).

D. Highly- and lowly-translated endogenous mRNAs do not exhibit nucleotide preferences. The distribution of log_2-transformed ribosome occupancy was plotted. Highlighted are the lowest kmeans group (n = 293 motifs represented in 548 mRNAs) and four highest kmeans groups (n = 73 motifs represented in 90 mRNAs). The juxtacap motifs for lowest and four highest kmeans groups are indicated.

notion that juxtacap nucleotide identity can define a particular structure that restricts or promotes translation in a constrained 5'UTR context. Secondary structure prediction (Lorenz et al., 2011) for the weak and strong eIF4E-binding groups in the context of the Rps20 5'UTR revealed striking differences in the accessibility of the 5' end of the 5'UTRs containing strong- and weak-eIF4E-binding motifs. Specifically, weak-binding motifs were predicted to have more highly base paired cap-proximal nucleotides, while strong-binding motifs had more open juxtacap sequences (Figure 5A). The base pairing interactions in both the strong- and weak-eIF4E-binding groups were dependent on nucleotide identity downstream of the juxtacap sequence.

In the weak-binding group, the +2 guanosine characteristic of the weak-binding motifs is predicted to be base paired and is associated with concurrent base pairing at the invariant +1 guanosine (Figure 5A). This base pairing could severely restrict eIF4E binding to the cap and subsequent translation. In contrast, the strong-binding motifs were predicted to contain several free juxtacap nucleotides, and the nucleotide with the clearest position-dependent preference, +5 cytidine, is predicted to participate in a base pairing interaction with a downstream guanosine at position +15 of the 5' UTR (Figure 5B).

We mutated this downstream guanosine to an adenosine to disrupt the base pairing interaction and increase the accessibility of the juxtacap sequence (Figure 5B), and measured translation in a cell-free translation system (Figure 5C). Mutation of the downstream guanosine in two strong-eIF4E-binding Rps20 5'UTRs containing motifs predicted to form a base pair between the +5 cytosine and +15 guanosine increased translation, while the compensatory mutation at position +5 in the juxtacap motif decreased translation (Figure 5B and Figure 5C). Modifying the Watson-Crick base pairing between the Rps20 5'UTR and the juxtacap nucleotides can alter translation, suggesting that the juxtacap sequence influences eIF4E binding and translation by defining the accessibility of the 5' end of the mRNA.

Figure 5: Juxtacap Nucleotide Sequence Modulates Translation by Defining Cap-proximal Nucleotide Structure

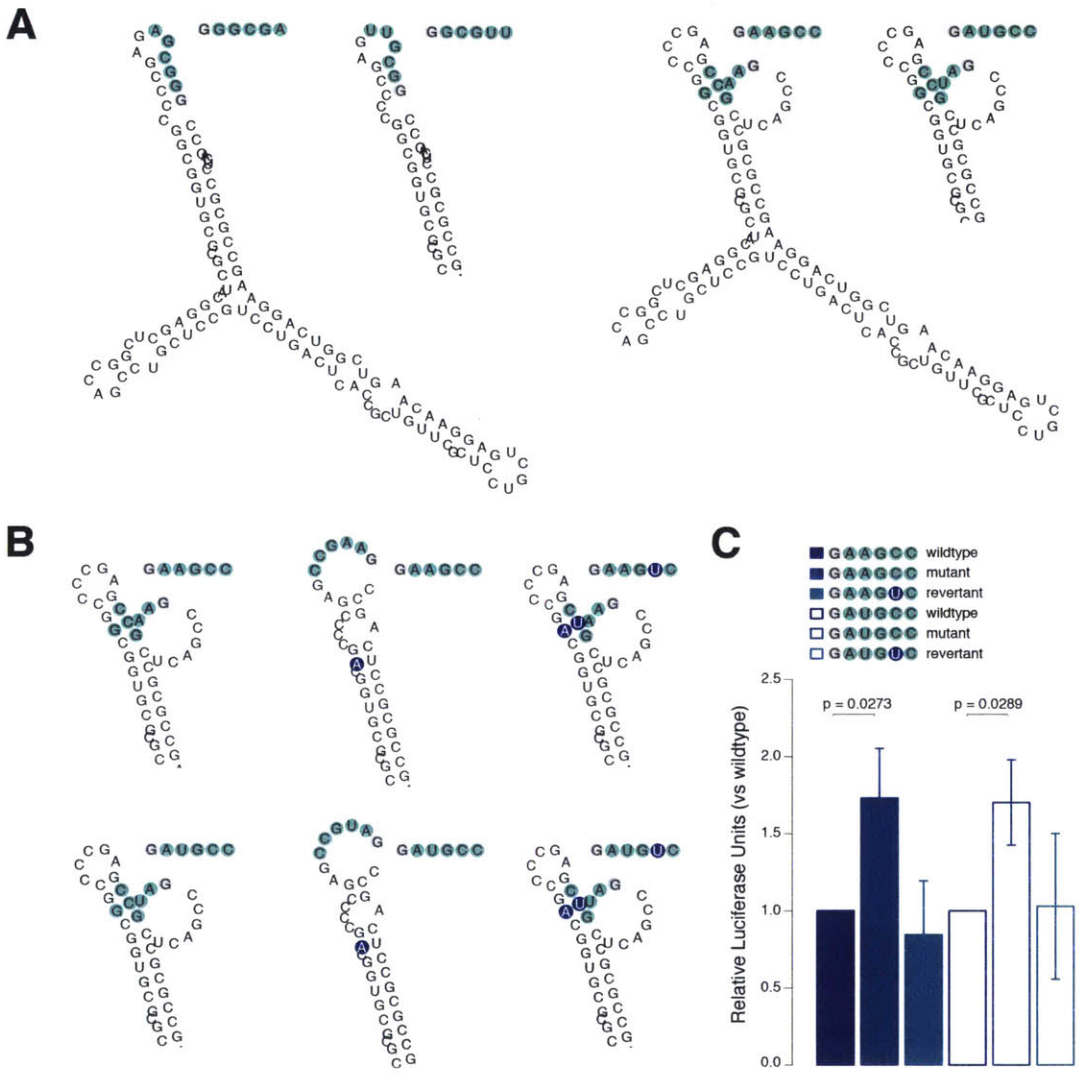


Figure 5: Juxtacap nucleotide sequence modulates translation by defining cap-proximal nucleotide structure

A. The predicted minimum free energy (MFE) structures of 5'UTRs containing weak-eIF4E-binding motifs have less accessible 5' ends than those containing strong-eIF4E-binding motifs. MFE structures for 5'UTRs with two weak-binding motifs and two strong-binding motifs are shown. The portion of the structure not shown for the second 5'UTR in each group is identical to that region in the first 5'UTR. The juxtacap motif nucleotides are highlighted.

B. Mutations downstream of the juxtacap motif are predicted to make the 5' end of 5'UTRs containing strong-eIF4E-binding motifs more accessible. The 5' end of the predicted MFE structures for 5'UTRs with wildtype, mutant, and revertant juxtacap motifs for two strong-eIF4E-binding motifs are shown; the portion of the structures not shown for each 5'UTR are identical to that region in the wildtype 5'UTR shown in Figure 5A. The juxtacap motif nucleotides are highlighted, with mutated nucleotides indicated in purple.

C. Mutations downstream of the juxtacap motif predicted to make the 5' end of 5'UTRs containing strong-eIF4E-binding motifs more accessible also increase translation. The mutant and revertant luciferase unit values for each replicate were normalized to the corresponding wildtype values, and mean ± s.e.m. of the normalized value for three independent experiments were plotted. Significance was determined by a one-tailed Student's t-test with equal variances on non-normalized values (variances were determined to be equal using the F test). P-values are indicated.

DISCUSSION

Initiation is normally the rate-limiting step of translation, and is tightly regulated via multiple mechanisms (Sonenberg & Hinnebusch, 2009). Molecularly defining the preferences of initiation factors for particular mRNAs is important not only for predicting how efficiently a given mRNA will be translated, but also for understanding how different regulatory inputs restrict or promote the translation of specific messages. There is an abundance of structural and biophysical evidence describing how eIF4E binds analogs of the 5' cap (Carberry et al., 1989; Marcotrigiano et al., 1997; Matsuo et al., 1997; Niedzwiecka et al., 2002; Tomoo et al., 2005; Tomoo et al., 2002). However, there was previously very little understanding of how juxtacap nucleotides affect eIF4E binding and subsequent translation.

Studies that thoroughly characterize the 5' ends of mRNA transcripts have uncovered that, for many genes, transcription does not begin at a single well-defined position, but can often occur at multiple discrete sites, or as a distribution around a particular site (Carninci et al., 2006; Kawaji et al., 2006). In some cases, these alternative transcriptional start sites (TSSs) substantially affect the 5'UTR of the mRNA, which can have considerable consequences for translation (Rojas-Duran & Gilbert, 2012). However, it has remained unknown whether the difference of a few nucleotides adjacent to the 5' cap can influence translation. Our findings suggest that the juxtacap sequence can influence the accessibility of the 5' end of an mRNA, and that small changes in this sequence can modulate translation. Our model implies that transcripts encoded by genes possessing even a narrow distribution of TSSs may be translated quite differently

from one another. In the case of a few genes with multiple discrete TSSs, selection of a particular TSS is regulated by transcription factor expression (Carninci et al., 2006; Ohmiya et al., 2014). Although it is unlikely that TSS selection for genes possessing a distribution of TSSs around a particular site would rely on different transcription factors, it is an attractive notion that these distributed TSSs could also be regulated. Regardless of their regulation, we predict that even small differences in TSSs can influence eIF4E binding and translation.

In cells, eIF4E is limiting (Duncan et al., 1987; Hiremath et al., 1985), and therefore is bound to either eIF4G or a 4EBP. Binding of an eIF4G fragment that associates with eIF4E but does not contain the RNA binding domains of eIF4G can alter the structure of eIF4E and increase its affinity for the 5' cap (Gross et al., 2003; Ptushkina et al., 1998). Although so far this RNA-binding-independent conformational coupling has only been directly observed in yeast (Gross et al., 2003; Ptushkina et al., 1998; Tomoo et al., 2005), it raises the possibility that the juxtacap sequence preferences of eIF4E when bound to eIF4G or a 4EBP could be different than for eIF4E alone. We also predict that some differences in eIF4E binding dictated by the juxtacap sequence are masked *in vivo*, due to stabilization of the eIF4E-mRNA interaction by direct binding of eIF4G to the RNA (Haghighat & Sonenberg, 1997; Yanagiya et al., 2009). Furthermore, juxtacap motifs that reduce eIF4E binding primarily through increasing the off-rate of eIF4E would be particularly sensitive to the presence of eIF4G, because eIF4G could act to keep eIF4E tethered to the mRNA after initial binding. The weakly- and strongly-bound motifs we identified by RNA Bind-n-Seq were translated differentially, suggesting that the

presence of eIF4G does not fully conceal all eIF4E-dependent juxtacap binding differences. However, the possibility remains that particular juxtacap sequences could influence eIF4E binding differently in the presence of eIF4G.

The presence of the nucleotide adjacent to the cap, corresponding to the first transcribed nucleotide, can directly contact eIF4E (Tomoo et al., 2002), and reduces the affinity of the 5' cap for eIF4E in biochemical assays (Niedzwiecka et al., 2002; Zuberek et al., 2003). In mammals, the ribose of the +1 nucleotide is typically methylated at the 2' position, and often the +2 nucleotide is methylated as well (Furuichi et al., 1975; Wei, Gershowitz, & Moss, 1975). 2'-*O*-methylation was first identified in tRNA and in functional regions of rRNA (Decatur & Fournier, 2002; Starr & Sells, 1969), and this modification is thought to alter RNA structure, RNA-RNA interactions, and RNA-protein interactions (Motorin & Helm, 2011). Although the functions of this modification have yet to be elucidated for most mRNAs, it may play a role in distinguishing self from non-self RNA (Daffis et al., 2010; Zust et al., 2011). While knockdown of the enzyme responsible for 2'-*O*-methylation at the +1 position did not affect global translation in HeLa cells (Belanger, Stepinski, Darzynkiewicz, & Pelletier, 2010), it was shown for specific mRNAs that 2'-*O*-methylation of the first two nucleotides can affect ribosome binding and translational efficiency (Kuge, Brownlee, Gershon, & Richter, 1998; Kuge & Richter, 1995; Muthukrishnan, Morgan, Banerjee, & Shatkin, 1976; Muthukrishnan, Moss, Cooper, & Maxwell, 1978). It is probable that 2'-*O*-methylation alters translational efficiency at least in part via effects on eIF4E binding, given that at least the modified +1 nucleotide can contact eIF4E directly. It is also possible that modification of these

nucleotides could alter binding of eIF4E differently in the context of different juxtacap sequences, which would allow mRNAs to display even greater variation in eIF4E recruitment and translation than by differences in the juxtacap sequence alone.

The eIF4E used in our study was prepared from *E. coli*, and hence was unphosphorylated. In mammals, eIF4E is phosphorylated by two kinases that bind eIF4G, Mnk1 and Mnk2 (Pyronnet et al., 1999; Waskiewicz, Flynn, Proud, & Cooper, 1997). In the presence of ample nutrients and growth factors, mTORC1 is active and phosphorylates the 4EBPs (Brunn et al., 1997; Marcotrigiano et al., 1999). This event liberates eIF4E and allows it to bind eIF4G, and subsequently become phosphorylated by Mnk1/2 at serine 209. The phosphorylation site is located within the flexible C-terminal loop of eIF4E that is stabilized by the presence of a +1 adenosine in the cocrystal structure (Tomoo et al., 2002). It is positioned directly adjacent to the cap-binding pocket of eIF4E, a prime location for influencing juxtacap nucleotide recognition. Although the precise molecular function of this phosphorylation is not thoroughly understood (Minich, Balasta, Goss, & Rhoads, 1994; Scheper & Proud, 2002; Scheper et al., 2002; Slepenkov et al., 2006; Zuberek et al., 2003), this phosphorylation site is clearly important for translation of a subset of mRNAs that promote transformation and tumorigenesis (Furic et al., 2010). These mRNAs are distinct from the mRNAs whose translation is promoted by wildtype eIF4E overexpression (Mamane et al., 2007), and it is likely that phosphorylated eIF4E and unphosphorylated eIF4E possess different juxtacap nucleotide preferences.

Although the effect of juxtacap sequence on eIF4E binding and translation was primarily mRNA context-dependent in our study, we do not rule out the notion that eIF4E can directly recognize the juxtacap nucleotides. Importantly, our experimental constraints do not allow us to assess the effect of differences in the first juxtacap nucleotide, while structural and biophysical evidence suggests nucleotides at this position are of particular interest (Marcotrigiano et al., 1997; Niedzwiecka et al., 2002; Tomoo et al., 2002; Zuberek et al., 2003). Nucleotides in the +1 position can form different contacts with eIF4E and alter the affinity of eIF4E for the 5' cap, depending on the nucleotide identity (Tomoo et al., 2002; Zuberek et al., 2003), and we predict that differences in the +1 nucleotide will alter the juxtacap sequence preferences of eIF4E. To fully assess the role of the juxtacap sequence in eIF4E binding and translation, it will be important to develop methods to readily synthesize capped mRNAs encoding different +1 nucleotides; for example, by improving existing chemical synthesis methods (Thillier et al., 2012) or by identifying an RNA polymerase that can produce mRNAs with various +1 nucleotides and is adaptable to *in vitro* synthesis. A third possibility would be to identify enzymes that can phosphorylate RNA 5' ends to produce 5'-triphosphorylated RNA (Spencer, Loring, Hurwitz, & Monroy, 1978), which could be used as a substrate for existing *in vitro* capping systems.

In addition to its importance for cap-dependent translation, eIF4E is associated with several other processes in the cell. For instance, eIF4E is involved in the export of particular mRNAs from the nucleus (Osborne & Borden, 2015). eIF4E is also a component of stress granules and P bodies (Kedersha & Anderson, 2007), which serve

as mRNA storage compartments under conditions of stress and as sites of mRNA degradation, respectively. Cap-proximal nucleotide preferences could potentially contribute to mRNA selection for any of these other processes involving eIF4E.

The idea that mRNA affinity for the translational machinery can modulate translation is not new, although there are few specific examples of this phenomenon. Over forty years ago, Lodish proposed the model that translation of mRNAs that initiate protein synthesis at lower rates will be preferentially inhibited when initiation is globally reduced, which was true when comparing translation of alpha and beta globin (Lodish, 1974). Almost a decade later, a difference in the affinity for the general translation initiation factor eIF2 was experimentally shown to mediate selective translation of a particular viral mRNA over globin mRNA in a cell-free system (Rosen, Di Segni, & Kaempfer, 1982). Decades of work on mRNAs containing internal ribosome entry sites (IRESs) has shown that they rely on direct recruitment of general translation initiation factors or even the ribosome for their translation (Deforges, Locker, & Sargueil, 2015; Jackson, 2013). Here, we describe the direction, magnitude, and sequence-dependence of the effect of the juxtacap nucleotides on eIF4E binding to a model 5'UTR. We show that juxtacap sequence identity can affect translation, likely by contextually defining cap-proximal end accessibility. Our work increases the understanding of how mRNAs are chosen for translation, and raises the possibility that initiation factor preferences are a more widespread mechanism for dictating translational efficiency of an mRNA than previously appreciated.

MATERIALS AND METHODS

Reagents

Enzymes were purchased from New England Biolabs, except where noted. Primers were obtained from Integrated DNA Technologies. Antibiotics and chemicals were purchased from Sigma, except where noted. TBE-Urea gels, TBE gels, and sample loading dyes were purchased from ThermoFisher. DMEM and inactivated fetal bovine serum (IFS) were purchased from US Biologicals.

eIF4E cloning, expression, and purification

Murine eIF4E was amplified by PCR from a plasmid encoding eIF4E (Addgene #38239; mouse and rat eIF4E sequences are identical), cloned into the pET302/NT-His bacterial expression vector (ThermoFisher) downstream of a 6xHis tag using XhoI and BamHI restriction sites, and clones were sequence verified using the T7 Promoter Primer.

meIF4E_Nterm_F: AATTCGCTCGAGatggcgactgtcgaaccg

heIF4E_Nterm_R: GCAGCCGGATCCttaaacaacaaacctatt

T7 Promoter Primer: TAATACGACTCACTATAGGG

A streptavidin-binding peptide (SBP) tag was introduced in place of the 6xHis tag using a geneblock (IDT) encoding SBP with a 5' NdeI site and a 3' XhoI site.

SBP_geneblock_NdeI_XhoI:

GATATACATATGGACGAAAAAACGACCGGGTGGCGTGGCGGCCATGTCGTGGAGG
GCCTGGCAGGCGAGCTGGAACAACTGCGCGCACGTCTGGAACACCATCCTCAGGG
ACAGCGCGAACCAGTGAATTCGCTCGAGatggcg

Sequence-verified clones were transformed into Rosetta (DE3) pLysS cells (EMD Millipore), inoculated into 50ml LB supplemented with 34 µg/ml chloramphenicol and 100 µg/ml ampicillin, and grown overnight at 37C. Cultures were diluted 1:100 in LB with chloramphenicol and ampicillin and grown at 37C until the OD was between 0.5 and 0.7. Cultures were shifted to 16C for 20 minutes and eIF4E expression was induced for 18 hours at 16C by addition of 0.5 mM IPTG (FisherScientific). All subsequent purification steps were performed at 4C. Cells were lysed in Resuspension Buffer (20 mM HEPES pH 7.4, 100 mM KCl, 2 mM EDTA, 1 mM DTT) three times in a French press, and CHAPS was immediately added to 0.3%. Lysates were sonicated twice and were cleared sequentially by low speed centrifugation and ultracentrifugation.

^{me7}GTP beads (preparation described below) were washed once with Wash Buffer (20 mM HEPES pH 7.4, 100 mM KCl, 0.3% CHAPS, 1 mM DTT) and added to the cleared lysate. Binding was allowed to proceed for 90 minutes at 4C with gentle mixing. Beads were washed 3 times for 5 minutes each with 20ml Wash Buffer and SBP-eIF4E was eluted with three consecutive washes for 15 minutes each with 1ml Elution Buffer (20 mM HEPES pH 7.4, 500 mM KCl, 0.3% CHAPS, 1 mM DTT). Eluates were pooled and desalted using a PD-Miditrap G-25 size exclusion column (GE Healthcare) into Wash Buffer. Protein was concentrated in an Amicon Ultra centrifugal filter (EMD Millipore)

with a molecular weight cutoff of 10,000kDa. Protein was centrifuged at 17,000*xg* for 10 minutes at 4C, and the soluble material was stored at 4C and used within 3 days. Protein was determined to be pure by Coomassie staining. Protein was centrifuged again for 10 minutes at 17,000*xg* at 4C immediately prior to use, and only the soluble material was used.

Preparation of ^{me7}GTP beads

^{me7}GTP beads were prepared as described previously (Edery, Altmann, & Sonenberg, 1988), with slight modifications. Briefly, one molar equivalent of sodium meta-periodate (FisherScientific) in 0.1 M NaOAc pH 6 was added to ^{me7}GTP in water and incubated for 90 minutes at 4C. Adipic acid dihydrazide agarose beads (Sigma) were washed once with water and with 0.1 M NaOAc pH 6. Beads were resuspended in 0.1 M NaOAc pH 6 and transferred to the tube containing the oxidized ^{me7}GTP. The mixture was incubated with rotation for 90 minutes at 4C, and sodium cyanoborohydride was added. The beads were incubated overnight at 4C with mixing, and washed several times with 1 M NaCl and Storage Buffer (50 mM HEPES pH 7.4, 100 mM KCl, 0.2 mM EDTA). Beads were stored at 4C in 10ml storage buffer with 0.02% sodium azide.

Rps20 5'UTR cloning

The murine Rps20 5'UTR was cloned from mouse embryonic fibroblast cDNA into pIS1-Eef2-5UTR-renilla (Addgene plasmid #38235) by PCR amplification with the following Rps20 5'UTR-specific primers containing a 5' EcoRI restriction site and

modified T7 RNA Polymerase promoter sequence (containing no guanosines at positions +1, +2, and +3), and a 3' NheI restriction site.

mRps20_dbTOP_0G_F:

ttggcgaattcTAATACGACTCACTATAcTTTCTGAGCCCCGGCGG

mRps20_R: tggtggctagcGGCGCGGCTTCCTGACCG

In vitro transcription, RNA processing, and purification

The DNA template for the randomized Bind-n-Seq library was generated by PCR amplification of the Rps20 5'UTR from pIS1-mRps20_dbTOP_0G using an Rps20 5'UTR-specific forward primer containing a 5' EcoRI restriction site, a modified T7 RNA Polymerase promoter containing only a +1 guanosine, randomized nucleotides at positions +2 through +6, and an Rps20 5'UTR-specific reverse primer containing the endogenous murine Rps20 Kozak sequence.

mRps20_db_G1N9_5U_F:

ttggcgaattcTAATACGACTCACTATAGNNNNNNNNNcccggcggtgcgcg

mRps20_Kozak_R: GGCTGAggcgcggcttcctgaccg

An aliquot of the PCR product was run on an agarose gel to confirm the amplification of a single product prior to purification of the rest of the PCR product using the

Quiaquick gel extraction kit (Qiagen). PCR products were eluted in DEPC-treated water, and all subsequent processing steps were carried out using RNAse-free reagents.

The Bind-n-Seq RNA library was transcribed from the DNA template at 37C for two hours in a reaction containing 1X NEB T7 RNA Polymerase buffer, 100 mM DTT, 0.6U SuperaseIn (ThermoFisher), 2 mM ATP, 2 mM UTP, 2 mM CTP, 2 mM GTP, 1 µg randomized DNA template, and 2.5U T7 RNA Polymerase per 100µl reaction volume. DNAse was added directly to the reaction at a concentration of 10U/100µl and the reactions were incubated for an additional 30 minutes at 37C. The transcription reactions were extracted with acid phenol/chloroform (FisherScientific), and samples were incubated 5 minutes at room temperature with periodic vortexing. The aqueous layer was extracted with chloroform and the RNA was precipitated three times with NH_4OAc and isopropanol.

The RNA was denatured at 75C for 10 minutes and put on ice, then capped *in vitro* for two hours at 37C using the ScriptCap m^7G Capping System (Cellscript) according to manufacturer's instructions. Capping reactions were extracted with acid phenol/chloroform and precipitated three times with NH_4OAc and isopropanol.

RNA was treated sequentially with 5' Polyphosphatase (Epicentre) and Terminator 5' Phosphate-Dependent Exonuclease (Epicentre) to remove uncapped RNA. 5' Polyphosphatase reactions contained 1X Polyphosphatase Buffer and 0.6U 5' Polyphosphatase per 1 µg RNA (approximately 20U 5' Polyphosphatase per 1nmol RNA). RNA was denatured for 5 minutes at 75C and placed on ice just prior to addition to the reaction. Reactions were incubated at 37C for 2 hours, and RNA was precipitated

with NaOAc and isopropanol. 5' Polyphosphatase-treated RNA was denatured for 5 minutes at 75C and digested with Terminator Exonuclease in a reaction containing 1X Terminator Reaction Buffer A and 0.3U Terminator Exonuclease per 1 µg RNA for 2 hours at 30C. EDTA was added to 5mM to stop the reaction, and RNA was precipitated with NaOAc and isopropanol. RNA was resuspended in DEPC water, TBE-Urea sample buffer was added, and samples were denatured 10 minutes at 75C and placed on ice. Samples were loaded onto a prerun 6% TBE-Urea gel and run for 50 minutes at 200V. Gels were stained 5 minutes with SYBR gold (FisherScientific) at 1:12,500 in TBE, and gel slices were crushed by centrifugation at 4C at 17,000*xg* for 3 minutes through a 0.65ml microfuge tube pierced with an 18G needle. RNA was eluted in 300 mM NaOAc, 1 mM EDTA overnight at 4C with agitation. Eluate was filtered through a SpinX 0.22 µM filter (Sigma), and RNA was precipitated with isopropanol. RNA was resuspended in DEPC water, and the RNA concentration was measured by absorbance at 260 nm.

RNA Bind-n-Seq

RNA-Bind-n-Seq was adapted from the method of Lambert *et al.* (Lambert et al., 2014). Each 250µl binding reaction contained 1X Binding Buffer (20 mM HEPES pH 7.4, 150 mM KCl, 2.5 mM $MgCl_2$, 0.3% CHAPS, 0.5 mM DTT), SBP-eIF4E (at a concentration of 50 nM, 125 nM, 250 nM, 750 nM, 1 µM, or 1.25 µM), randomized RNA library at a fixed concentration of 9 µM, 15 ng/µl polyI RNA, and 0.4U/µl SuperaseIn. Control reactions also contained either 100 µM GTP or 100 µM ^{me7}GTP. SBP-eIF4E was equilibrated in the reaction mix without RNA library for 30 minutes at 23C, after

which the RNA library was added and the binding reaction allowed to proceed for one hour at 23C with constant mixing at 750rpm in an Eppendorf Thermomixer. During the binding reaction, 100µl streptavidin magnetic beads per reaction were washed in a batch 3 times with 1ml 1X Binding Buffer. Beads were resuspended in 1X Binding Buffer and aliquotted to one tube per reaction. The remaining buffer was removed just prior to addition of each binding reaction; the binding reaction and bead mixture was incubated for an additional one hour at 23C with constant mixing at 750rpm. Reactions were washed once with 1ml 1X Wash Buffer (20 mM HEPES pH 7.4, 150 mM KCl, 0.5 mM EDTA, 0.3% CHAPS) and resuspended in 250µl Elution Buffer (20 mM HEPES pH 7.4, 1 mM EDTA, 1% SDS). Samples were heated for 10 minutes at 70C, and eluate was transferred to a new tube. Samples were extracted with acid phenol/chloroform at 65C for 5 minutes with periodic vortexing. Samples were put on ice for 5 minutes prior to recovery of the aqueous layer. The aqueous layer was extracted with acid phenol/chloroform a second time for 5 minutes at room temperature with periodic vortexing. The aqueous layer was then extracted with one volume chloroform for 30 seconds with constant vortexing, and the RNA was precipitated with NaOAc and isopropanol and resuspended in 20µl 10mM Tris pH 8 per sample.

Sequencing Library Preparation

Sequencing library preparation was performed similarly to Ingolia *et al.* (Ingolia, Brar, Rouskin, McGeachy, & Weissman, 2013). cDNA synthesis was performed with Superscript III Reverse Transcriptase (ThermoFisher). The reverse transcription primer

contained a 5' phosphate to allow enzymatic circularization of the resulting single-stranded DNA, and also a stretch of 6 random nucleotides, to ensure complexity in base composition in the first several cycles during sequencing. The RT primer also contained the reverse complement of the 5' Illumina adapter, an abasic site to allow relinearization of the circularized single-stranded DNA to increase PCR efficiency if necessary, and the reverse complement of the 3' Illumina adapter. Finally, the RT primer contained the reverse complement of the 3' end of the Rps20 5'UTR, which is shared by all RNAs in the library.

dbRps20_RT primer:

/5Phos/NNNNNNGATCGTCGGACTGTAGAACTCTGAAC/iSp18/CACTCA/iSp18/ccttgg cacccgagaattccaGGCTGAggcgcggcttcctg

For the binding reaction samples, 10µl RNA was used in a 20µl cDNA synthesis reaction; for the input library, 0.5 pmol RNA in 10µl 10 mM Tris pH 8 was used. The RNA was mixed with 1µl 50 µM RT primer and 1µl 10 mM dNTP mix. Reactions were incubated for 5 minutes at 65C and placed on ice. To each chilled reaction, a mix containing 0.5µl DEPC water, 4µl 5X First Strand Synthesis Buffer, 1µl 0.1 M DTT, 0.5µl SuperaseIn, and 1µl Superscript III Reverse Transcriptase was added. Reactions were incubated for 45 minutes at 48C. RNA was hydrolyzed by addition of 2.2µl 1 M NaOH and incubation for 20 minutes at 98C, followed by addition of 2.2µl 1 M HCl. DNA was recovered by precipitation with NaCl and isopropanol and resuspended in 10µl 10 mM Tris pH 8. DNA was loaded onto a pre-run 6% TBE-Urea gel and run for 25 minutes at

200V. Gels were stained with SYBR Gold, and gel slices were incubated in 300 mM NaCl with 1 mM EDTA and 10 mM Tris pH 8 overnight at 4C. Eluate was filtered through a SpinX 0.22 µM filter and precipitated with isopropanol.

Pellets were resuspended in 4.5µl Circularization mix without CircLigase (1X CircLigase Buffer, 50 µM ATP, 2.5 mM $MnCl_2$, and 50U CircLigase; Epicentre). CircLigase was added, and reactions were incubated for one hour at 60C. CircLigase was heat inactivated by incubation for 10 minutes at 80C. DNA was precipitated with NaCl and isopropanol, and resuspended in 12µl 10 mM Tris pH 8.

The randomized sequences and a portion of the Rps20 5'UTR were PCR amplified using a forward primer containing the 5' Illumina adapter sequence, and one of several Illumina Tru-Seq small RNA reverse primers containing the Illumina 3' adapter sequence and a unique barcode.

randomized_PCR_F:
AATGATACGGCGACCACCGAGATCTACAC**GTTCAGAGTTCTACAGTCCGA**

Illumina Index Primer 1 (1.25µM eIF4E + GTP):
CAAGCAGAAGACGGCATACGAGAT**CGTGAT**GTGACTGGAGTT**CCTTGGCACCCGA GAATTCCA**

Illumina Index Primer 2 (1.25µM eIF4E + ^{me7}GTP):
CAAGCAGAAGACGGCATACGAGAT**ACATCG**GTGACTGGAGTT**CCTTGGCACCCGA GAATTCCA**

Illumina Index Primer 3 (50nM eIF4E):
CAAGCAGAAGACGGCATACGAGAT**GCCTAA**GTGACTGGAGTT**CCTTGGCACCCGA GAATTCCA**

Illumina Index Primer 4 (125nM eIF4E):
CAAGCAGAAGACGGCATACGAGAT**TGGTCA**GTGACTGGAGTT**CCTTGGCACCCGA GAATTCCA**

Illumina Index Primer 5 (250nM eIF4E):
CAAGCAGAAGACGGCATACGAGAT**CACTGT**GTGACTGGAGTT**CCTTGGCACCCGA**
GAATTCCA

Illumina Index Primer 7 (750nM eIF4E):
CAAGCAGAAGACGGCATACGAGAT**GATCTG**GTGACTGGAGTT**CCTTGGCACCCGA**
GAATTCCA

Illumina Index Primer 9 (1μM eIF4E):
CAAGCAGAAGACGGCATACGAGAT**CTGATC**GTGACTGGAGTT**CCTTGGCACCCGA**
GAATTCCA

Illumina Index Primer 10 (1.25μM eIF4E):
CAAGCAGAAGACGGCATACGAGAT**AAGCTA**GTGACTGGAGTT**CCTTGGCACCCGA**
GAATTCCA

Illumina Index Primer 11 (input library):
CAAGCAGAAGACGGCATACGAGAT**GTAGCC**GTGACTGGAGTT**CCTTGGCACCCGA**
GAATTCCA

For each sample, five reactions of 16.7μl each containing 1X Phusion HF Buffer, 200 μM of each dNTP, 500 nM each of forward and reverse primer, 1μl circularized ssDNA, and 0.334U Phusion High-Fidelity DNA Polymerase were heated for 30 seconds at 98C and amplified for 6, 8, 10, 12, or 14 cycles with the following program:

1. 98C 10 seconds
2. 60C 10 seconds
3. 72C 10 seconds

Reactions were run on a pre-run 8% TBE gel for one hour at 180V. DNA was extracted from bands where the amplification had not reached saturation and was extracted from the gel slices and precipitated as described above. Libraries were resuspended in 5μl water per slice and were pooled and sequenced on an Illumina HiSeq Sequencer according to standard procedures.

Data Analysis

Motifs were extracted from raw sequencing data using a custom Python script to identify reads containing the sequence "GAGCCCCG," which is shared by all 5'UTRs in the library and is directly downstream of the random 5-mer. Reads that did not contain the constant sequence or that did not contain an entire 5-mer preceding the constant sequence were discarded. The five-nucleotide motifs directly preceding the constant sequence were extracted and their frequencies calculated for each binding reaction. All subsequent data manipulation and statistical analysis was performed in R.

Cell-Free Translation Assay

The cell-free translation assay was adapted from Rakotondrafara *et al.* (Rakotondrafara & Hentze, 2011). Briefly, $10x10^6$ Hela cells were seeded in 15cm dishes in 20ml DMEM supplemented with 10% IFS and penicillin/streptomycin. The following day, cells were washed with PBS, trypsinized, and resuspended in ice cold DMEM without serum. Cells were pelleted for 3 min at 318xg at 4C. Cells were washed once with 10ml ice cold PBS and pelleted again. Residual PBS was removed, and cells were resuspended in one pellet volume Hypotonic Lysis Buffer (16 mM HEPES pH 7.4, 10 mM KOAc, 0.5mM MgOAc, 5 mM DTT) with one freshly added EDTA-free protease inhibitor tablet (Roche) per 10ml buffer. Cells were incubated on ice for 10 minutes and homogenized by passage through a 27G1/2 needle 15 times at 4C. Lysate was cleared

by centrifugation for 1 minute at 13,300*xg* at 4C, and soluble material was normalized to 10mg/ml, aliquotted, and stored at -80C.

mRNAs encoding the Rps20 5'UTR with various juxtacap motifs were generated by PCR amplification of pIS1-Rps20_dbTOP_0G using unique forward primers to introduce different juxtacap motifs, together with the mRps20_R primer.

mRps20_GGCGT: ttggcgaattcTAATACGACTCACTATAg**GGCGT**GAGCCCCGGCGG
mRps20_GGTGC: ttggcgaattcTAATACGACTCACTATAg**GGTGC**GAGCCCCGGCGG
mRps20_GGCGA: ttggcgaattcTAATACGACTCACTATAg**GGCGA**GAGCCCCGGCGG
mRps20_GCGTT: ttggcgaattcTAATACGACTCACTATAg**GCGTT**GAGCCCCGGCGG
mRps20_GGCGC: ttggcgaattcTAATACGACTCACTATAg**GGCGC**GAGCCCCGGCGG
mRps20_GCGTA: ttggcgaattcTAATACGACTCACTATAg**GCGTA**GAGCCCCGGCGG
mRps20_GCTTT: ttggcgaattcTAATACGACTCACTATAg**GCTTT**GAGCCCCGGCGG
mRps20_GGCTT: ttggcgaattcTAATACGACTCACTATAg**GGCTT**GAGCCCCGGCGG
mRps20_GGAGT: ttggcgaattcTAATACGACTCACTATAg**GGAGT**GAGCCCCGGCGG
mRps20_GGAGC: ttggcgaattcTAATACGACTCACTATAg**GGAGC**GAGCCCCGGCGG

mRps20_AAGCC: ttggcgaattcTAATACGACTCACTATAg**AAGCC**GAGCCCCGGCGG
mRps20_AGGCT: ttggcgaattcTAATACGACTCACTATAg**AGGCT**GAGCCCCGGCGG
mRps20_ATGCC: ttggcgaattcTAATACGACTCACTATAg**ATGCC**GAGCCCCGGCGG
mRps20_AGGCC: ttggcgaattcTAATACGACTCACTATAg**AGGCC**GAGCCCCGGCGG
mRps20_TGGCT: ttggcgaattcTAATACGACTCACTATAg**TGGCT**GAGCCCCGGCGG
mRps20_AGGTC: ttggcgaattcTAATACGACTCACTATAg**AGGTC**GAGCCCCGGCGG
mRps20_AGACC: ttggcgaattcTAATACGACTCACTATAg**AGACC**GAGCCCCGGCGG
mRps20_AAGCT: ttggcgaattcTAATACGACTCACTATAg**AAGCT**GAGCCCCGGCGG
mRps20_TTGCT: ttggcgaattcTAATACGACTCACTATAg**TTGCT**GAGCCCCGGCGG
mRps20_TTGCC: ttggcgaattcTAATACGACTCACTATAg**TTGCC**GAGCCCCGGCGG

The PCR products were digested with NheI and EcoRI and ligated into the pIS1-Rps20_dbTOP_0G vector upstream of the *Renilla* luciferase gene and an encoded polyA tail. Sequence-verified clones were linearized by BamHI digest and transcribed *in vitro* as described above for the Bind-n-Seq library, except that mRNAs were co-transcriptionally capped using anti-reverse cap analog (ARCA; New England Biolabs),

instead of post-transcriptionally capped. After extraction with acid phenol/chloroform and precipitation with NaOAc and isopropanol, mRNAs were gel separated on a pre-run 6% TBE-Urea gel for 2.5h at 200V. mRNAs were extracted from the gel slices as described above, and resuspended in 10 mM Tris pH8. mRNAs were stored at -80C, and working stocks were stored at -20C.

10 ng of each mRNA was translated in a 10µl reaction containing 40 mM KOAc, 2 mM $MgCl_2$, 0.8 mM ATP, 0.1 mM GTP, and 1X translation buffer (16 mM HEPES pH 7.4, 20 mM creatine phosphate, 0.1 µg/µl creatine kinase, 0.1 mM spermidine (freshly diluted from 1 mM stock), and amino acids at RPMI concentrations) for 30 minutes at 37C. Translation was detected using the *Renilla* Luciferase Assay System (Promega). 10µl 1X *Renilla* Luciferase Lysis Buffer was added on ice to stop the reactions. 10µl of each reaction was added to 50µl of a *Renilla* Luciferase Assay Buffer and *Renilla* Luciferase Assay Substrate mixture immediately prior to reading for 10 seconds in a luminometer. Luciferase readings for experimental samples were background-subtracted using readings from samples containing 10 mM Tris pH 8 instead of RNA.

Statistical Analyses and Data Visualization

The following R packages were used for data visualization and statistical analyses (R Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/):

seqLogo (Bembom,O. (2016) seqLogo: Sequence logos for DNA sequence alignments. R package version 1.34.0.)

colorspace (Ihaka,R., Murrell,P., Hornik,K., Fisher,J.C. and Zeileis,A. (2015). colorspace: Color Space Manipulation. R package version 1.2-6. http://CRAN.R-project.org/package=colorspace)

plot3D (Soetaert,K. (2014) plot3D: Plotting multi-dimensional data. R package version 1.0-2. http://CRAN.R-project.org/package=plot3D)

scales (Wickham,H. (2015) scales: Scale Functions for Visualization. R package version 0.3.0. http://CRAN.R-project.org/package=scales)

flux (Jurasinski,G., Koebsch,F., Guenther,A. and Beetz,S. (2014) flux: Flux rate calculation from dynamic closed chamber measurements. R package version 0.3-0. http://CRAN.R-project.org/package=flux)

Transcriptional Start Site and Ribosome Footprinting Analysis

The FANTOM5 project used cap analysis of gene expression (CAGE) to identify robust transcriptional start sites (TSSs) from numerous mouse and human samples (Consortium et al., 2014). TSS data from NIH3T3 cells (Lizio et al., 2015) was processed to identify dominant TSSs using the CAGEr package (Haberle, Forrest, Hayashizaki, Carninci, & Lenhard, 2015) (with default parameters, except fitRange was set from 10 to 1000 and slope was set to 1.14) and subsetted by TSSs beginning with a guanosine using custom R scripts. TSSs beginning with a guanosine were mapped to overlapping genes using the BEDTools suite (Quinlan & Hall, 2010). Ribosome occupancy of these genes was extracted from ribosome footprinting data from p53-/- mouse embryonic fibroblasts (Thoreen et al., 2012). eIF4E binding was plotted against ribosome occupancy for each juxtacap motif represented in a TSS that overlaps with a gene present in the ribosome footprinting dataset.

RNA Secondary Structure Prediction

RNA secondary structures of model 5'UTRs were predicted using RNAfold (Lorenz et al., 2011), and base pairing probabilities were extracted for nucleotides at the 5' end of the RNA. A nucleotide at a particular position was deemed inaccessible if it had a base pairing probability greater than 0.75.

ACKNOWLEDGEMENTS

We would like to thank Larry Schweitzer for critical reading of the manuscript and Carson Thoreen for generous sharing of custom analysis scripts, as well as invaluable input from L.S. and C.T. throughout all stages of this work. We would like to thank Ellen Edenberg for critical reading of the manuscript, Bingbing Yuan for CAGEr analysis, and the entire Sabatini Lab for helpful feedback. We would also like to thank the Whitehead Institute Genome Technology Center for sequencing. This work was supported by grants from the NIH (R01 CA129105, R01 CA103866, R37 AI047389) to D.M.S. D.M.S. is also an investigator of the Howard Hughes Medical Institute.

REFERENCES

Belanger, F., Stepinski, J., Darzynkiewicz, E., & Pelletier, J. (2010). Characterization of hMTr1, a human Cap1 2'-O-ribose methyltransferase. *J Biol Chem, 285*(43), 33037-33044. doi: 10.1074/jbc.M110.155283

Bhat, M., Robichaud, N., Hulea, L., Sonenberg, N., Pelletier, J., & Topisirovic, I. (2015). Targeting the translation machinery in cancer. *Nat Rev Drug Discov, 14*(4), 261-278. doi: 10.1038/nrd4505

Brunn, G. J., Hudson, C. C., Sekulic, A., Williams, J. M., Hosoi, H., Houghton, P. J., . . . Abraham, R. T. (1997). Phosphorylation of the translational repressor PHAS-I by the mammalian target of rapamycin. *Science, 277*(5322), 99-101.

Carberry, S. E., Friedland, D. E., Rhoads, R. E., & Goss, D. J. (1992). Binding of protein synthesis initiation factor 4E to oligoribonucleotides: effects of cap accessibility and secondary structure. *Biochemistry, 31*(5), 1427-1432.

Carberry, S. E., Rhoads, R. E., & Goss, D. J. (1989). A spectroscopic study of the binding of m7GTP and m7GpppG to human protein synthesis initiation factor 4E. *Biochemistry, 28*(20), 8078-8083.

Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., . . . Hayashizaki, Y. (2006). Genome-wide analysis of mammalian promoter architecture and evolution. *Nat Genet, 38*(6), 626-635. doi: 10.1038/ng1789

Consortium, F., the, R. P., Clst, Forrest, A. R., Kawaji, H., Rehli, M., . . . Hayashizaki, Y. (2014). A promoter-level mammalian expression atlas. *Nature, 507*(7493), 462-470. doi: 10.1038/nature13182

Daffis, S., Szretter, K. J., Schriewer, J., Li, J., Youn, S., Errett, J., . . . Diamond, M. S. (2010). 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature, 468*(7322), 452-456. doi: 10.1038/nature09489

Decatur, W. A., & Fournier, M. J. (2002). rRNA modifications and ribosome function. *Trends Biochem Sci, 27*(7), 344-351.

Deforges, J., Locker, N., & Sargueil, B. (2015). mRNAs that specifically interact with eukaryotic ribosomal subunits. *Biochimie, 114*, 48-57. doi: 10.1016/j.biochi.2014.12.008

Duncan, R., Milburn, S. C., & Hershey, J. W. (1987). Regulated phosphorylation and low abundance of HeLa cell initiation factor eIF-4F suggest a role in translational control. Heat shock effects on eIF-4F. *J Biol Chem, 262*(1), 380-388.

Dunn, J. J., & Studier, F. W. (1983). Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J Mol Biol, 166*(4), 477-535.

Edery, I., Altmann, M., & Sonenberg, N. (1988). High-level synthesis in Escherichia coli of functional cap-binding eukaryotic initiation factor eIF-4E and affinity purification using a simplified cap-analog resin. *Gene, 74*(2), 517-525.

Fraser, C. S. (2015). Quantitative studies of mRNA recruitment to the eukaryotic ribosome. *Biochimie, 114*, 58-71. doi: 10.1016/j.biochi.2015.02.017

Furic, L., Rong, L., Larsson, O., Koumakpayi, I. H., Yoshida, K., Brueschke, A., . . . Sonenberg, N. (2010). eIF4E phosphorylation promotes tumorigenesis and is associated with prostate cancer progression. *Proc Natl Acad Sci U S A, 107*(32), 14134-14139. doi: 10.1073/pnas.1005320107

Furuichi, Y., Morgan, M., Shatkin, A. J., Jelinek, W., Salditt-Georgieff, M., & Darnell, J. E. (1975). Methylated, blocked 5 termini in HeLa cell mRNA. *Proc Natl Acad Sci U S A, 72*(5), 1904-1908.

Golini, F., Thach, S. S., Birge, C. H., Safer, B., Merrick, W. C., & Thach, R. E. (1976). Competition between cellular and viral mRNAs in vitro is regulated by a messenger discriminatory initiation factor. *Proc Natl Acad Sci U S A, 73*(9), 3040-3044.

Gross, J. D., Moerke, N. J., von der Haar, T., Lugovskoy, A. A., Sachs, A. B., McCarthy, J. E., & Wagner, G. (2003). Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell, 115*(6), 739-750.

Haberle, V., Forrest, A. R., Hayashizaki, Y., Carninci, P., & Lenhard, B. (2015). CAGEr: precise TSS data retrieval and high-resolution promoterome mining for integrative analyses. *Nucleic Acids Res, 43*(8), e51. doi: 10.1093/nar/gkv054

Haghighat, A., Mader, S., Pause, A., & Sonenberg, N. (1995). Repression of cap-dependent translation by 4E-binding protein 1: competition with p220 for binding to eukaryotic initiation factor-4E. *EMBO J, 14*(22), 5701-5709.

Haghighat, A., & Sonenberg, N. (1997). eIF4G dramatically enhances the binding of eIF4E to the mRNA 5'-cap structure. *J Biol Chem, 272*(35), 21677-21680.

Hiremath, L. S., Webb, N. R., & Rhoads, R. E. (1985). Immunological detection of the messenger RNA cap-binding protein. *J Biol Chem, 260*(13), 7843-7849.

Hsieh, A. C., Liu, Y., Edlind, M. P., Ingolia, N. T., Janes, M. R., Sher, A., . . . Ruggero, D. (2012). The translational landscape of mTOR signalling steers cancer initiation and metastasis. *Nature, 485*(7396), 55-61. doi: 10.1038/nature10912

Ingolia, N. T., Brar, G. A., Rouskin, S., McGeachy, A. M., & Weissman, J. S. (2013). Genome-wide annotation and quantitation of translation by ribosome profiling. *Curr Protoc Mol Biol, Chapter 4*, Unit 4 18. doi: 10.1002/0471142727.mb0418s103

Jackson, R. J. (2013). The current status of vertebrate cellular mRNA IRESs. *Cold Spring Harb Perspect Biol, 5*(2). doi: 10.1101/cshperspect.a011569

Jackson, R. J., Hellen, C. U., & Pestova, T. V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol, 11*(2), 113-127. doi: 10.1038/nrm2838

Kawaji, H., Frith, M. C., Katayama, S., Sandelin, A., Kai, C., Kawai, J., . . . Hayashizaki, Y. (2006). Dynamic usage of transcription start sites within core promoters. *Genome Biol, 7*(12), R118. doi: 10.1186/gb-2006-7-12-r118

Kedersha, N., & Anderson, P. (2007). Mammalian stress granules and processing bodies. *Methods Enzymol, 431*, 61-81. doi: 10.1016/S0076-6879(07)31005-7

Kuge, H., Brownlee, G. G., Gershon, P. D., & Richter, J. D. (1998). Cap ribose methylation of c-mos mRNA stimulates translation and oocyte maturation in Xenopus laevis. *Nucleic Acids Res, 26*(13), 3208-3214.

Kuge, H., & Richter, J. D. (1995). Cytoplasmic 3' poly(A) addition induces 5' cap ribose methylation: implications for translational control of maternal mRNA. *EMBO J, 14*(24), 6301-6310.

Lambert, N., Robertson, A., Jangi, M., McGeary, S., Sharp, P. A., & Burge, C. B. (2014). RNA Bind-n-Seq: quantitative assessment of the sequence and structural binding specificity of RNA binding proteins. *Mol Cell, 54*(5), 887-900. doi: 10.1016/j.molcel.2014.04.016

Lamphear, B. J., Kirchweger, R., Skern, T., & Rhoads, R. E. (1995). Mapping of functional domains in eukaryotic protein synthesis initiation factor 4G (eIF4G) with picornaviral proteases. Implications for cap-dependent and cap-independent translational initiation. *J Biol Chem, 270*(37), 21975-21983.

Lazaris-Karatzas, A., Montine, K. S., & Sonenberg, N. (1990). Malignant transformation by a eukaryotic initiation factor subunit that binds to mRNA 5' cap. *Nature, 345*(6275), 544-547. doi: 10.1038/345544a0

Lizio, M., Harshbarger, J., Shimoji, H., Severin, J., Kasukawa, T., Sahin, S., . . . consortium, F. (2015). Gateways to the FANTOM5 promoter level mammalian expression atlas. *Genome Biol, 16*, 22. doi: 10.1186/s13059-014-0560-6

Lodish, H. F. (1974). Model for the regulation of mRNA translation applied to haemoglobin synthesis. *Nature, 251*(5474), 385-388.

Lorenz, R., Bernhart, S. H., Honer Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P. F., & Hofacker, I. L. (2011). ViennaRNA Package 2.0. *Algorithms Mol Biol, 6*, 26. doi: 10.1186/1748-7188-6-26

Mamane, Y., Petroulakis, E., Martineau, Y., Sato, T. A., Larsson, O., Rajasekhar, V. K., & Sonenberg, N. (2007). Epigenetic activation of a subset of mRNAs by eIF4E explains its effects on cell proliferation. *PLoS One, 2*(2), e242. doi: 10.1371/journal.pone.0000242

Marcotrigiano, J., Gingras, A. C., Sonenberg, N., & Burley, S. K. (1997). Cocrystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Cell, 89*(6), 951-961.

Marcotrigiano, J., Gingras, A. C., Sonenberg, N., & Burley, S. K. (1999). Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. *Mol Cell, 3*(6), 707-716.

Martin, F., Barends, S., Jaeger, S., Schaeffer, L., Prongidi-Fix, L., & Eriani, G. (2011). Cap-assisted internal initiation of translation of histone H4. *Mol Cell, 41*(2), 197-209. doi: 10.1016/j.molcel.2010.12.019

Matsuo, H., Li, H., McGuire, A. M., Fletcher, C. M., Gingras, A. C., Sonenberg, N., & Wagner, G. (1997). Structure of translation factor eIF4E bound to m7GDP and interaction with 4E-binding protein. *Nat Struct Biol, 4*(9), 717-724.

Merrick, W. C. (2015). eIF4F: a retrospective. *J Biol Chem, 290*(40), 24091-24099. doi: 10.1074/jbc.R115.675280

Minich, W. B., Balasta, M. L., Goss, D. J., & Rhoads, R. E. (1994). Chromatographic resolution of in vivo phosphorylated and nonphosphorylated eukaryotic translation initiation factor eIF-4E: increased cap affinity of the phosphorylated form. *Proc Natl Acad Sci U S A, 91*(16), 7668-7672.

Mochizuki, K., Oguro, A., Ohtsu, T., Sonenberg, N., & Nakamura, Y. (2005). High affinity RNA for mammalian initiation factor 4E interferes with mRNA-cap binding and inhibits translation. *RNA, 11*(1), 77-89. doi: 10.1261/rna.7108205

Motorin, Y., & Helm, M. (2011). RNA nucleotide methylation. *Wiley Interdiscip Rev RNA, 2*(5), 611-631. doi: 10.1002/wrna.79

Muthukrishnan, S., Morgan, M., Banerjee, A. K., & Shatkin, A. J. (1976). Influence of 5'-terminal m7G and 2'--O-methylated residues on messenger ribonucleic acid binding to ribosomes. *Biochemistry, 15*(26), 5761-5768.

Muthukrishnan, S., Moss, B., Cooper, J. A., & Maxwell, E. S. (1978). Influence of 5'-terminal cap structure on the initiation of translation of vaccinia virus mRNA. *J Biol Chem, 253*(5), 1710-1715.

Niedzwiecka, A., Marcotrigiano, J., Stepinski, J., Jankowska-Anyszka, M., Wyslouch-Cieszynska, A., Dadlez, M., . . . Stolarski, R. (2002). Biophysical studies of eIF4E cap-

binding protein: recognition of mRNA 5' cap structure and synthetic fragments of eIF4G and 4E-BP1 proteins. *J Mol Biol, 319*(3), 615-635. doi: 10.1016/S0022-2836(02)00328-5

O'Leary, S. E., Petrov, A., Chen, J., & Puglisi, J. D. (2013). Dynamic recognition of the mRNA cap by Saccharomyces cerevisiae eIF4E. *Structure, 21*(12), 2197-2207. doi: 10.1016/j.str.2013.09.016

Ohmiya, H., Vitezic, M., Frith, M. C., Itoh, M., Carninci, P., Forrest, A. R., . . . Consortium, F. (2014). RECLU: a pipeline to discover reproducible transcriptional start sites and their alternative regulation using capped analysis of gene expression (CAGE). *BMC Genomics, 15*, 269. doi: 10.1186/1471-2164-15-269

Osborne, M. J., & Borden, K. L. (2015). The eukaryotic translation initiation factor eIF4E in the nucleus: taking the road less traveled. *Immunol Rev, 263*(1), 210-223. doi: 10.1111/imr.12240

Pause, A., Belsham, G. J., Gingras, A. C., Donze, O., Lin, T. A., Lawrence, J. C., Jr., & Sonenberg, N. (1994). Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function. *Nature, 371*(6500), 762-767. doi: 10.1038/371762a0

Prevot, D., Darlix, J. L., & Ohlmann, T. (2003). Conducting the initiation of protein synthesis: the role of eIF4G. *Biol Cell, 95*(3-4), 141-156.

Ptushkina, M., von der Haar, T., Vasilescu, S., Frank, R., Birkenhager, R., & McCarthy, J. E. (1998). Cooperative modulation by eIF4G of eIF4E-binding to the mRNA 5' cap in yeast involves a site partially shared by p20. *EMBO J, 17*(16), 4798-4808. doi: 10.1093/emboj/17.16.4798

Pyronnet, S., Imataka, H., Gingras, A. C., Fukunaga, R., Hunter, T., & Sonenberg, N. (1999). Human eukaryotic translation initiation factor 4G (eIF4G) recruits mnk1 to phosphorylate eIF4E. *EMBO J, 18*(1), 270-279. doi: 10.1093/emboj/18.1.270

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics, 26*(6), 841-842. doi: 10.1093/bioinformatics/btq033

Rakotondrafara, A. M., & Hentze, M. W. (2011). An efficient factor-depleted mammalian in vitro translation system. *Nat Protoc, 6*(5), 563-571. doi: 10.1038/nprot.2011.314

Rojas-Duran, M. F., & Gilbert, W. V. (2012). Alternative transcription start site selection leads to large differences in translation activity in yeast. *RNA, 18*(12), 2299-2305. doi: 10.1261/rna.035865.112

Rosen, H., Di Segni, G., & Kaempfer, R. (1982). Translational control by messenger RNA competition for eukaryotic initiation factor 2. *J Biol Chem, 257*(2), 946-952.

Ruggero, D., Montanaro, L., Ma, L., Xu, W., Londei, P., Cordon-Cardo, C., & Pandolfi, P. P. (2004). The translation factor eIF-4E promotes tumor formation and cooperates with c-Myc in lymphomagenesis. *Nat Med, 10*(5), 484-486. doi: 10.1038/nm1042

Scheper, G. C., & Proud, C. G. (2002). Does phosphorylation of the cap-binding protein eIF4E play a role in translation initiation? *Eur J Biochem, 269*(22), 5350-5359.

Scheper, G. C., van Kollenburg, B., Hu, J., Luo, Y., Goss, D. J., & Proud, C. G. (2002). Phosphorylation of eukaryotic initiation factor 4E markedly reduces its affinity for capped mRNA. *J Biol Chem, 277*(5), 3303-3309. doi: 10.1074/jbc.M103607200

Slepenkov, S. V., Darzynkiewicz, E., & Rhoads, R. E. (2006). Stopped-flow kinetic analysis of eIF4E and phosphorylated eIF4E binding to cap analogs and capped oligoribonucleotides: evidence for a one-step binding mechanism. *J Biol Chem, 281*(21), 14927-14938. doi: 10.1074/jbc.M601653200

Sonenberg, N., & Hinnebusch, A. G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell, 136*(4), 731-745. doi: 10.1016/j.cell.2009.01.042

Sonenberg, N., Morgan, M. A., Merrick, W. C., & Shatkin, A. J. (1978). A polypeptide in eukaryotic initiation factors that crosslinks specifically to the 5'-terminal cap in mRNA. *Proc Natl Acad Sci U S A, 75*(10), 4843-4847.

Spencer, E., Loring, D., Hurwitz, J., & Monroy, G. (1978). Enzymatic conversion of 5'-phosphate-terminated RNA to 5'-di- and triphosphate-terminated RNA. *Proc Natl Acad Sci U S A, 75*(10), 4793-4797.

Starr, J. L., & Sells, B. H. (1969). Methylated ribonucleic acids. *Physiol Rev, 49*(3), 623-669.

Thillier, Y., Decroly, E., Morvan, F., Canard, B., Vasseur, J. J., & Debart, F. (2012). Synthesis of 5' cap-0 and cap-1 RNAs using solid-phase chemistry coupled with enzymatic methylation by human (guanine-N(7))-methyl transferase. *RNA, 18*(4), 856-868. doi: 10.1261/rna.030932.111

Thoreen, C. C. (2013). Many roads from mTOR to eIF4F. *Biochem Soc Trans, 41*(4), 913-916. doi: 10.1042/BST20130082

Thoreen, C. C., Chantranupong, L., Keys, H. R., Wang, T., Gray, N. S., & Sabatini, D. M. (2012). A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature, 485*(7396), 109-113. doi: 10.1038/nature11083

Tomoo, K., Matsushita, Y., Fujisaki, H., Abiko, F., Shen, X., Taniguchi, T., . . . Ishida, T. (2005). Structural basis for mRNA Cap-Binding regulation of eukaryotic initiation factor 4E by 4E-binding protein, studied by spectroscopic, X-ray crystal structural, and

molecular dynamics simulation methods. *Biochim Biophys Acta, 1753*(2), 191-208. doi: 10.1016/j.bbapap.2005.07.023

Tomoo, K., Shen, X., Okabe, K., Nozoe, Y., Fukuhara, S., Morino, S., . . . Miura, K. (2002). Crystal structures of 7-methylguanosine 5'-triphosphate (m(7)GTP)- and P(1)-7-methylguanosine-P(3)-adenosine-5',5'-triphosphate (m(7)GpppA)-bound human full-length eukaryotic initiation factor 4E: biological importance of the C-terminal flexible region. *Biochem J, 362*(Pt 3), 539-544.

Walden, W. E., Godefroy-Colburn, T., & Thach, R. E. (1981). The role of mRNA competition in regulating translation. I. Demonstration of competition in vivo. *J Biol Chem, 256*(22), 11739-11746.

Waskiewicz, A. J., Flynn, A., Proud, C. G., & Cooper, J. A. (1997). Mitogen-activated protein kinases activate the serine/threonine kinases Mnk1 and Mnk2. *EMBO J, 16*(8), 1909-1920. doi: 10.1093/emboj/16.8.1909

Wei, C. M., Gershowitz, A., & Moss, B. (1975). Methylated nucleotides block 5' terminus of HeLa cell messenger RNA. *Cell, 4*(4), 379-386.

Yanagiya, A., Svitkin, Y. V., Shibata, S., Mikami, S., Imataka, H., & Sonenberg, N. (2009). Requirement of RNA binding of mammalian eukaryotic translation initiation factor 4GI (eIF4GI) for efficient interaction of eIF4E with the mRNA cap. *Mol Cell Biol, 29*(6), 1661-1669. doi: 10.1128/MCB.01187-08

Zuberek, J., Wyslouch-Cieszynska, A., Niedzwiecka, A., Dadlez, M., Stepinski, J., Augustyniak, W., . . . Darzynkiewicz, E. (2003). Phosphorylation of eIF4E attenuates its interaction with mRNA 5' cap analogs by electrostatic repulsion: intein-mediated protein ligation strategy to obtain phosphorylated protein. *RNA, 9*(1), 52-61.

Zust, R., Cervantes-Barragan, L., Habjan, M., Maier, R., Neuman, B. W., Ziebuhr, J., . . . Thiel, V. (2011). Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat Immunol, 12*(2), 137-143. doi: 10.1038/ni.1979

Chapter 5

Future Directions and Conclusions

Heather R. Keys[1-2]

[1]Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA

[2]Department of Biology, MIT, Cambridge, MA 02139, USA

SUMMARY

Here we have explored the regulation of translation initiation, with the overarching theme of how mTORC1 controls this process, and with a particular focus on how the initiation machinery regulated by mTORC1 influences translation initiation and its regulation. We initially used the global and systematic approach of ribosome footprinting to probe the translational response to a cellular stress that regulates the mTORC1 pathway, amino acid deprivation. We identified a set of novel translationally upregulated mRNAs, and identified putative novel features in a well-characterized translationally upregulated mRNA, although the precise roles of these mRNAs and features in the adaptation of the cell to a lack of amino acids and their mechanism of regulation remain largely unexplored.

We subsequently focused on identifying the mechanism by which acute inhibition of the mTORC1 pathway regulates an important subset of mRNAs, those containing a 5'TOP motif. We showed that the 4EBPs were required for translational repression of 5'TOP mRNAs, and that this class of mRNAs was especially lost from the eIF4F component eIF4E after inhibition of mTOR, relative to mRNAs not harboring a 5'TOP motif. Although it appears that eIF4E does not have a lower affinity for true 5'TOP mRNAs, and thus eIF4E affinity for mRNAs does not explain the dependence of 5'TOP mRNA translation on mTORC1 activity, this work uncovered that the cap-proximal sequence can influence eIF4E binding.

Finally, we systematically assessed how the cap-proximal sequence of mRNA dictates eIF4E binding and translation, using RNA Bind-n-Seq. Our evidence supports a model

in which the juxtacap sequence dictates the cap-proximal RNA secondary structure in an mRNA-context-dependent manner, although we cannot rule out the possibility of direct effects of nucleotide identity on eIF4E binding. Our work has contributed to the understanding of how translation is controlled, but raises many more questions for future study.

FUTURE DIRECTIONS

Translational Response to Amino Acid Deprivation

uORFs in general are thought to inhibit translation of the downstream coding ORF (Wethmar, 2014). However, the mechanism allowing the upregulation of a downstream coding ORF under stress conditions has only been described for a few mRNAs, such as ATF4 (Le & Maizel, 1997; Lee, Cevallos, & Jan, 2009; Lu, Harding, & Ron, 2004; Vattem & Wek, 2004; Zhou et al., 2008). We did not observe an enrichment of mRNAs containing uORFs in our set of mRNAs that were translationally upregulated following amino acid deprivation, suggesting that alternative mechanisms may be utilized in response to this particular stress. IRESs allow the bypass of different steps in the initiation process via noncanonical recruitment of initiation factors, and could be another mechanism for increased translation of mRNAs during amino acid deprivation (Jackson, 2013; Licursi, Komatsu, Pongnopparat, & Hirasawa, 2012). It is often difficult to identify cellular IRESs, but with new sequencing approaches affording us a greater ability to define mRNA structures in cells, it may become easier to systematically search for

common structural features in these mRNAs (Rouskin, Zubradt, Washietl, Kellis, & Weissman, 2014).

It is fascinating that both the mTORC1 and GCN2 pathways exist to regulate translation initiation at two distinct points in response to the identical stress. The mTORC1 pathway is also inhibited downstream of ER stress, which is sensed by another eIF2α kinase, PERK (Deldicque, Bertrand, Patton, Francaux, & Baar, 2011; Harding, Zhang, & Ron, 1999; Nakajima et al., 2011). Thus, it seems that coordinate regulation at multiple points in the initiation process is important in some contexts. It will be important to define the translational programs dependent on mTORC1 and GCN2 downstream of amino acid deficiency, in order to understand why the cell requires multiple pathways to respond to the same stress. Genetic tools already exist to dissect the relative contributions of each pathway to translation (Efeyan et al., 2013; Jiang et al., 2003). This approach could additionally help define mechanisms of translational control downstream of each of these important regulatory pathways.

Regulation of 5'TOP mRNAs

Although our data suggested that eIF4E affinity for mRNAs does not determine their translational sensitivity to mTOR inhibition, it is possible that the dynamics of the eIF4E-eIF4G or eIF4E-4EBP interactions dictate this sensitivity. The interaction of eIF4F with the mRNA could be different for 5'TOP mRNAs, such that the eIF4E-eIF4G interaction is more easily disrupted by 4EBPs, leading to their more dramatic translational repression. Single molecule fluorescence imaging would be a particularly useful tool for

investigating these interactions (Serebrov & Moore, 2016). In fact, this has technique has been performed using components of the yeast eIF4F complex, which uncovered an unexpected influence of PABP on eIF4E binding dynamics (O'Leary, Petrov, Chen, & Puglisi, 2013). A single molecule system could be easily adapted to test how different eIF4F components interact with 5'TOP and non-TOP RNAs. Elucidating how eIF4F dynamically interacts with mRNAs of different nucleotide sequence would be useful in general, as there is very little information on this aspect of translation initiation. Single molecule fluorescence imaging would also be useful for testing whether other candidate proteins affect eIF4F dynamics differently on 5'TOP mRNAs compared with non-TOP mRNAs.

Effects of Juxtacap Nucleotide Sequence on eIF4E Binding and Translation

Due to technical limitations of our *in vitro* transcription system, we were unable to test juxtacap nucleotide sequences with +1 nucleotides other than guanosine. The +1 position is likely particularly important for eIF4E binding in the context of an mRNA, as it could greatly influence 5' end accessibility of an mRNA, and because eIF4E has known preferences for different +1 nucleotides (Niedzwiecka et al., 2002; Tomoo et al., 2002; Zuberek et al., 2003). The +1 and +2 nucleotides are also typically modified by 2'-*O*-methylation in mammalian cells, which could directly modulate eIF4E binding, as well as translation (Furuichi et al., 1975; Wei, Gershowitz, & Moss, 1975). It will be important to develop systems to test the effects of +1 nucleotide identity and of methylation at these

positions, in order to gain a more complete understanding of how the juxtacap nucleotide sequence modulates eIF4E binding and translation.

It is increasingly appreciated that the transcriptional start site (TSS) for many mRNA transcripts does not occur at a single position, but that transcription often occurs as a distribution around a single position, at multiple distinct sites, or is even scattered throughout a broader region (Carninci et al., 2006; Kawaji et al., 2006; Zhao, Valen, Parker, & Sandelin, 2011) (Figure 1). This observation is particularly important, in light of our finding that the juxtacap nucleotide sequence can dictate differences in eIF4E binding and translation. Differences of even a few nucleotides at the very 5' end of the mRNA could have sizeable translational consequences, particularly for mRNAs that utilize a few TSSs equally. Whether there is a regulatory mechanism for selection of different TSSs within a distribution remains to be elucidated, but this could provide another layer of control to the process of protein synthesis.

Figure 1: Examples of Transcriptional Start Site Distributions

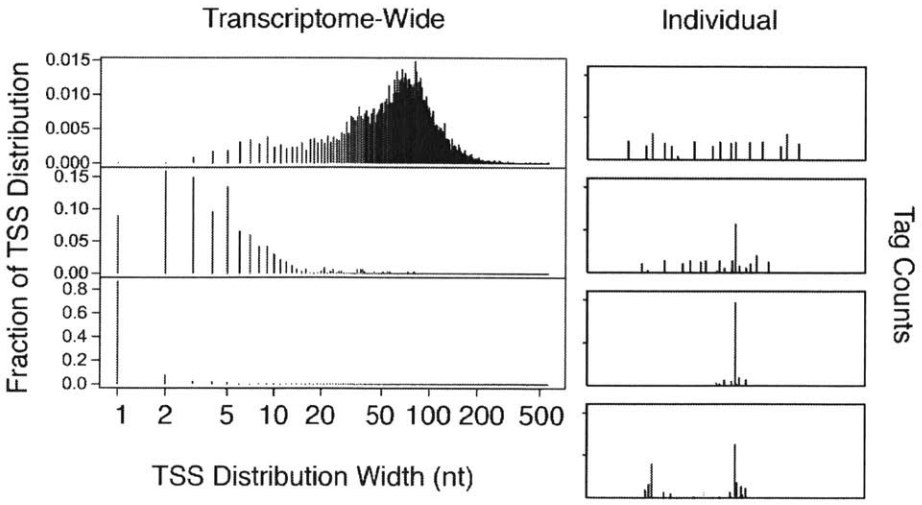


Figure 1: Examples of transcriptional start site distributions

For each of three different classes of TSS distributions, the width, in nucleotides, of each individual distribution is plotted against the fraction of distributions of that width within that class. Examples of individual distributions for four mRNAs are shown on the right. Adapted from (Kawaji et al., 2006; Zhao et al., 2011).

Recently, a method has been developed to combine the capture of mRNA 5' ends with polysome profiling (TATL-seq), which allows the simultaneous identification of TSSs and their degree of association with ribosomes using high throughput sequencing (Arribere & Gilbert, 2013). Use of this technique in mammalian cells would allow us to assess which mRNAs have different TSSs that confer differential translation. Other techniques have also been developed recently to assess mRNA structure in cells (Rouskin et al., 2014). Pairing the TSS information for individual mRNAs and their translation with their corresponding structures would allow us to uncover the relationship between the juxtacap sequence and translation in a transcriptome-wide fashion. Furthermore, these approaches could be combined with eIF4E Bind-n-Seq on endogenous cellular mRNAs, yielding great insights into how mRNA structure affects initiation factor recruitment and translation. These approaches could also be utilized for assessing how the binding of other individual initiation factors or initiation factor complexes affects translation.

CONCLUSION

It is a testament to the complexity of mRNA translation that we continue to uncover important new paradigms for how this process functions and is regulated, despite the fact that it has been studied with great interest for many decades. Several new techniques have allowed recent transcriptome-wide assessment of ribosome occupancy, mRNA modifications, and RNA structure (Carlile et al., 2014; Ingolia, Ghaemmaghami, Newman, & Weissman, 2009; Linder et al., 2015; Rouskin et al., 2014; Schwartz et al., 2014). Despite the exciting advances made with these techniques, many avenues of investigation remain unexplored. Combining these systematic approaches will allow us to describe the process of translation in more exquisite detail than ever before.

Understanding how the mTORC1 pathway regulates translation is crucial to our understanding of how this pathway functions in normal cell growth and disease. Although many attempts have been made to define the most proximal regulator of 5'TOP mRNA translation, this remains an active area of investigation. It is likely that several factors can contribute to the regulation of this important class of mRNAs under various conditions, in different cellular contexts, and with different kinetics. However, the most proximal regulator downstream of mTORC1 will likely be universally essential for the regulation of 5'TOP mRNAs.

Characterization of the affinity of translation initiation factors for different mRNAs is an important step in defining how translation is regulated. Determining how the affinities of difference initiation factors for mRNA individually, cooperatively, and synergistically

affect translation will provide different avenues for targeting the translation of specific mRNAs. In combination with clearly defining the translational programs relevant in disease, this approach could contribute to development of a new set of therapies.

This is an exciting time for research in the field of translation. We now have the tools to explore how the many different mRNA features contribute to translational regulation transcriptome-wide in cells. There is still much insight to be gained from defining translational control in these cellular systems, especially in combination with directed studies to elucidate the molecular underpinnings of this control. However, it is only a matter of time before we will be able to routinely assess the translational landscape in whole organisms and in disease models. Uncovering the principles of translational control will be an active area of study for many years to come.

REFERENCES

Arribere, J. A., & Gilbert, W. V. (2013). Roles for transcript leaders in translation and mRNA decay revealed by transcript leader sequencing. *Genome Res, 23*(6), 977-987. doi: 10.1101/gr.150342.112

Carlile, T. M., Rojas-Duran, M. F., Zinshteyn, B., Shin, H., Bartoli, K. M., & Gilbert, W. V. (2014). Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature, 515*(7525), 143-146. doi: 10.1038/nature13802

Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., . . . Hayashizaki, Y. (2006). Genome-wide analysis of mammalian promoter architecture and evolution. *Nat Genet, 38*(6), 626-635. doi: 10.1038/ng1789

Deldicque, L., Bertrand, L., Patton, A., Francaux, M., & Baar, K. (2011). ER stress induces anabolic resistance in muscle cells through PKB-induced blockade of mTORC1. *PLoS One, 6*(6), e20993. doi: 10.1371/journal.pone.0020993

Efeyan, A., Zoncu, R., Chang, S., Gumper, I., Snitkin, H., Wolfson, R. L., . . . Sabatini, D. M. (2013). Regulation of mTORC1 by the Rag GTPases is necessary for neonatal autophagy and survival. *Nature, 493*(7434), 679-683. doi: 10.1038/nature11745

Furuichi, Y., Morgan, M., Shatkin, A. J., Jelinek, W., Salditt-Georgieff, M., & Darnell, J. E. (1975). Methylated, blocked 5 termini in HeLa cell mRNA. *Proc Natl Acad Sci U S A, 72*(5), 1904-1908.

Harding, H. P., Zhang, Y., & Ron, D. (1999). Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. *Nature, 397*(6716), 271-274. doi: 10.1038/16729

Ingolia, N. T., Ghaemmaghami, S., Newman, J. R., & Weissman, J. S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science, 324*(5924), 218-223. doi: 10.1126/science.1168978

Jackson, R. J. (2013). The current status of vertebrate cellular mRNA IRESs. *Cold Spring Harb Perspect Biol, 5*(2). doi: 10.1101/cshperspect.a011569

Jiang, H. Y., Wek, S. A., McGrath, B. C., Scheuner, D., Kaufman, R. J., Cavener, D. R., & Wek, R. C. (2003). Phosphorylation of the alpha subunit of eukaryotic initiation factor 2 is required for activation of NF-kappaB in response to diverse cellular stresses. *Mol Cell Biol, 23*(16), 5651-5663.

Kawaji, H., Frith, M. C., Katayama, S., Sandelin, A., Kai, C., Kawai, J., . . . Hayashizaki, Y. (2006). Dynamic usage of transcription start sites within core promoters. *Genome Biol, 7*(12), R118. doi: 10.1186/gb-2006-7-12-r118

Le, S. Y., & Maizel, J. V., Jr. (1997). A common RNA structural motif involved in the internal initiation of translation of cellular mRNAs. *Nucleic Acids Res, 25*(2), 362-369.

Lee, Y. Y., Cevallos, R. C., & Jan, E. (2009). An upstream open reading frame regulates translation of GADD34 during cellular stresses that induce eIF2alpha phosphorylation. *J Biol Chem, 284*(11), 6661-6673. doi: 10.1074/jbc.M806735200

Licursi, M., Komatsu, Y., Pongnopparat, T., & Hirasawa, K. (2012). Promotion of viral internal ribosomal entry site-mediated translation under amino acid starvation. *J Gen Virol, 93*(Pt 5), 951-962. doi: 10.1099/vir.0.040386-0

Linder, B., Grozhik, A. V., Olarerin-George, A. O., Meydan, C., Mason, C. E., & Jaffrey, S. R. (2015). Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat Methods, 12*(8), 767-772. doi: 10.1038/nmeth.3453

Lu, P. D., Harding, H. P., & Ron, D. (2004). Translation reinitiation at alternative open reading frames regulates gene expression in an integrated stress response. *J Cell Biol, 167*(1), 27-33. doi: 10.1083/jcb.200408003

Nakajima, S., Hiramatsu, N., Hayakawa, K., Saito, Y., Kato, H., Huang, T., . . . Kitamura, M. (2011). Selective abrogation of BiP/GRP78 blunts activation of NF-kappaB through the ATF6 branch of the UPR: involvement of C/EBPbeta and mTOR-dependent dephosphorylation of Akt. *Mol Cell Biol, 31*(8), 1710-1718. doi: 10.1128/MCB.00939-10

Niedzwiecka, A., Marcotrigiano, J., Stepinski, J., Jankowska-Anyszka, M., Wyslouch-Cieszynska, A., Dadlez, M., . . . Stolarski, R. (2002). Biophysical studies of eIF4E cap-binding protein: recognition of mRNA 5' cap structure and synthetic fragments of eIF4G and 4E-BP1 proteins. *J Mol Biol, 319*(3), 615-635. doi: 10.1016/S0022-2836(02)00328-5

O'Leary, S. E., Petrov, A., Chen, J., & Puglisi, J. D. (2013). Dynamic recognition of the mRNA cap by Saccharomyces cerevisiae eIF4E. *Structure, 21*(12), 2197-2207. doi: 10.1016/j.str.2013.09.016

Rouskin, S., Zubradt, M., Washietl, S., Kellis, M., & Weissman, J. S. (2014). Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature, 505*(7485), 701-705. doi: 10.1038/nature12894

Schwartz, S., Bernstein, D. A., Mumbach, M. R., Jovanovic, M., Herbst, R. H., Leon-Ricardo, B. X., . . . Regev, A. (2014). Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell, 159*(1), 148-162. doi: 10.1016/j.cell.2014.08.028

Serebrov, V., & Moore, M. J. (2016). Single Molecule Approaches in RNA-Protein Interactions. *Adv Exp Med Biol, 907*, 89-106. doi: 10.1007/978-3-319-29073-7_4

Tomoo, K., Shen, X., Okabe, K., Nozoe, Y., Fukuhara, S., Morino, S., . . . Miura, K. (2002). Crystal structures of 7-methylguanosine 5'-triphosphate (m(7)GTP)- and P(1)-7-methylguanosine-P(3)-adenosine-5',5'-triphosphate (m(7)GpppA)-bound human full-length eukaryotic initiation factor 4E: biological importance of the C-terminal flexible region. *Biochem J, 362*(Pt 3), 539-544.

Vattem, K. M., & Wek, R. C. (2004). Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proc Natl Acad Sci U S A, 101*(31), 11269-11274. doi: 10.1073/pnas.0400541101

Wei, C. M., Gershowitz, A., & Moss, B. (1975). Methylated nucleotides block 5' terminus of HeLa cell messenger RNA. *Cell, 4*(4), 379-386.

Wethmar, K. (2014). The regulatory potential of upstream open reading frames in eukaryotic gene expression. *Wiley Interdiscip Rev RNA, 5*(6), 765-778. doi: 10.1002/wrna.1245

Zhao, X., Valen, E., Parker, B. J., & Sandelin, A. (2011). Systematic clustering of transcription start site landscapes. *PLoS One, 6*(8), e23409. doi: 10.1371/journal.pone.0023409

Zhou, D., Palam, L. R., Jiang, L., Narasimhan, J., Staschke, K. A., & Wek, R. C. (2008). Phosphorylation of eIF2 directs ATF5 translational control in response to diverse stress conditions. *J Biol Chem, 283*(11), 7064-7073. doi: 10.1074/jbc.M708530200

Zuberek, J., Wyslouch-Cieszynska, A., Niedzwiecka, A., Dadlez, M., Stepinski, J., Augustyniak, W., . . . Darzynkiewicz, E. (2003). Phosphorylation of eIF4E attenuates its interaction with mRNA 5' cap analogs by electrostatic repulsion: intein-mediated protein ligation strategy to obtain phosphorylated protein. *RNA, 9*(1), 52-61.