JUN mRNA Translation Regulation is Mediated by 5' Untranslated Region (5' UTR) Features and Multiple Translation Initiation Factors

By

Angélica M. González-Sánchez

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Comparative Biochemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jamie H. D. Cate, Chair
Professor Nicholas T. Ingolia
Professor Britt Glaunsinger

Fall 2023

ABSTRACT

JUN mRNA Translation Regulation is Mediated by 5' Untranslated Region (5' UTR) Features and Multiple Translation Initiation Factors

by

Angélica M. González-Sánchez

Doctor of Philosophy in Comparative Biochemistry

University of California, Berkeley

Professor Jamie H. D. Cate, Chair

mRNA translation regulation by eukaryotic initiation factors (eIFs) is crucial for cell survival. In humans, eIF3 stimulates translation of the *JUN* mRNA which encodes the transcription factor JUN, an oncogenic transcription factor that is involved in cell cycle progression, apoptosis, and cell proliferation. Previous studies revealed that eIF3 activates translation of the *JUN* mRNA by interacting with a stem loop in the 5′ untranslated region (5′ UTR) and with the 5′ -7-methylguanosine cap structure. In addition to its interaction site with eIF3, the *JUN* 5′ UTR has a longer than average length, a high degree of secondary structure, high GC content, and an upstream start codon (uAUG). This motivated us to explore the complexity of *JUN* mRNA translation regulation in human cells.

Chapter 2 describes our findings on the contributions of multiple 5' UTR and start codon features in *JUN* translation regulation. We find that *JUN* translation is regulated in a sequence and structure-dependent manner in regions adjacent to the eIF3-interacting site in the *JUN* 5′ UTR. Furthermore, we identify contributions of an additional initiation factor, eIF4A, in *JUN* regulation. We show that enhancing the interaction of eIF4A with *JUN* by using the compound Rocaglamide A (RocA) represses *JUN* translation. We also find that both the upstream AUG (uAUG) and the main AUG (mAUG) contribute to *JUN* translation and that they are conserved throughout vertebrates. Work presented in this chapter demonstrates additional layers of regulation for *JUN* translation.

Chapter 3 describes isolation of a translation initiation multifactor complex (MFC) from *in vitro* translation reactions of a *JUN* 5'UTR reporter mRNA in human cell extracts. The yeast MFC is composed of eIF1, eIF2, eIF3, eIF5 and the initiator methionyl-tRNA (Met-tRNA$_i$) and has been well-characterized. However, knowledge about the human MFC is limited and isolation of the endogenous complex from human cells hasn't been achieved. By using the *JUN* mRNA as a platform for MFC binding, we present the first instance of isolation of an mRNA-bound human MFC. We also present in-depth protocol optimization and propose next steps for complex validation. Work presented in this chapter provides strong evidence for the formation of a human MFC with the presence of novel initiation factors such as eIF4A and eIF4G. Together these findings demonstrate the complexity of *JUN* translation regulation and establish *JUN*'s potential as a model transcript for understanding multiple interacting modes of translation regulation.

DEDICATION PAGE

To my family, especially my mother Ana and my father Alfredo
for their unconditional love and support,

and to my chosen family, especially Luis, Valerie, Mark and Ralphie
for the same thing.

You are the ones that keep my flame burning.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Thank you is truly not enough to express the immense gratitude I have for this journey and for everyone that has been a part of it. All of this work could not have been possible without the support of a plethora of people and resources. I have you all in my heart and hold profound appreciation for all of you.

To my advisor Jamie, thank you for all your incredible support, especially through all the challenging times. Your brilliant ideas and endless enthusiasm were truly the fuel that kept this project going. Thank you for all the time you invested in me and for sharing your love of science every day. I am especially grateful for all your understanding and encouragement throughout this rollercoaster of a journey. You inspire me to always stay curious and perseverant, and for that I thank you.

To all my fellow Cate lab members, for sharing the love of translation and always being willing to lend a helping hand. Thank you to Mia and to Luisa for all your help with experiments when I first arrived in the lab. Thank you to Nadege for being an amazing bench-mate and an inspiring scientist. Thank you to Fred for sharing your enthusiasm for science and for all your help with reagents and all things lab-related. Thank you to Wenfei for your guidance in the beginning of my eIF3 journey, for your patience, and for all the scientific wisdom. Thank you to Kolya for your help with the *in vitro* translation system. Thank you to Amos for being an amazing bay-mate and peer, for always listening, and for offering thoughtful feedback and ideas. You are truly an incredible scientist and I hope you continue making your mark in the field. Thank you to Pooja, Erica and Cynthia, for sharing the love of eIF3 and for always being willing to help with experiments and ideas. Thank you to all the current Cate lab members for making the lab such a great place to do science. I believe in all of you and your capacity to continue doing amazing science.

To the wonderful friends the Cate lab gave me, for your joy and kindness. Firstly to Das, my forever bay-mate. My life at Berkeley and in the lab would truly not have been the same without you. Thank you for always encouraging me, believing in me, listening, and for being a truly genuine friend. Thank you also for keeping me company in all of those late nights of long experiments, for showing me so many techniques, and for helping me troubleshoot so much. You are one of the most wonderful people I've ever met and I am truly lucky to have been able to build a friendship with you. Secondly, to Santi. I am so grateful that you joined the lab when you did! The lab truly lit up whenever you were around because of your positive attitude and incredible work-ethic, which are genuinely inspiring. Thank you for always being there with a listening ear and with words of encouragement, and for always being open to have coffee or lunch and talk about anything. You, just like Das, were an incredible escape from stressful times without even having to leave the lab. You are an amazing scientist and an even more incredible person and I am grateful to be able to call you a friend. Also, to Claudia. Thank you for being an incredibly supportive, encouraging, and caring friend. I hold all your advice, wisdom, and compassion close to my heart. And lastly, to Lorena. Thank you for all the conversations in the kitchen and for motivating me to be an outstanding scientist like you.

To my amazing students, for inspiring me in more ways than one. Eimy, Sona, Czara and Gaby, thank you for all your hard work and commitment to the science. Your enthusiasm and fresh perspectives truly inspired me and the project would not be the

same without each one of you. You are all incredible scientists and have the potential to achieve anything you put your mind into. I wish you all the best in your future careers.

To the valuable resources that Berkeley has provided me, for supporting me as a scientist. Thank you to the Comparative Biochemistry Graduate Program, for their guidance and assistance. To my committee members Professors Glaunsinger and Ingolia, thank you for your valuable scientific feedback and support. To the incredible collaborators I've had along the years, especially Aaron and Michael, thank you for your valuable scientific insights and incredible teamwork. To both Stanley Hall and IGI staff, thank you for making these wonderful places to do science. To QB3, especially Rosanne Lurie, thank you for all the extraordinarily helpful resources you provide for us students.

To all the student organizations, for helping me feel part of something bigger. Thank you to Boricuas at Berkeley for being an incredible community and a piece of home away from home, especially during tough times. To Beyond Academia, thank you for the incredible teamwork and outstanding commitment. You truly helped me find my sense of purpose, pushed me to discover my potential, and inspired me to pursue my future career. To BASIS and Be a Scientist, thank you for creating a space to give back to the community and embrace our inner child. You helped me remember the reason why I love science and reconnect with the excitement of learning. Special thanks to my BASIS team: Maya, Giselle and Jenny. Thank you for all the early mornings, all your hard work, and for sharing your passion for teaching and learning.

To the friends that grad school gave me, for being incredible role models. Jenny, thank you for always believing in me and helping me see all my potential. You truly made me braver and more perseverant and, even though you are now so far away, you continue to inspire me to do more and to be a better person. Likewise, Nisha and Laura, thank you for being incredibly inspiring women and for all the chats and encouragement.

To the friends that life has given me, for becoming my family and my rock. Luis, thank you for understanding me even when I don't understand myself, for always being there to listen without judgement, and for always sharing your joy even from the distance. I am lucky to have you in my life for over a decade now and I hope we stay this close until we are golden girls. Valerie, you are without doubt one of the best things that I have gained from my life in academia. Thank you for always listening and knowing exactly what to say, for sharing all the woes and joys of grad school with me, and for always encouraging me to keep going. You were truly an oasis in this journey. I'm forever grateful that our coffee dates have turned into a truly genuine friendship and hope to have you as a friend for the rest of our lives.

To Mark, thank you for being my partner, my support, and my safe place. Thank you for reminding me that I am capable and that I am loved every single day. Thank you for this adventure of a life that we are building together and for being such an incredible dad to Ralphie. Thank you for bringing joy to each and every one of our days. You make our lives better. I'm so excited for what's to come and so grateful for our little family.

Last but not least, a massive thank you to my family, for their continued love and support. Mami y papi, gracias por siempre apoyarme y motivarme a dar lo mejor de mi. Gracias por creer en mi y confiar en mis capacidades. Sobre todo, gracias por su amor incondicional y por su entendimiento especialmente en los tiempos difíciles. Verdaderamente, no estaría donde estoy sin ustedes.

Even though thank you is not enough, thank you to each and every one of you.

Chapter 1: General introduction

Portions of this chapter were adapted from the following publication:
González-Sánchez, A.M., Castellanos-Silva, E.A., Díaz-Figueroa, G., Cate, J.H.D. 2023. *JUN* mRNA Translation Regulation is Mediated by Multiple 5' UTR and Start Codon Features. bioRxiv. https://doi.org/10.1101/2023.11.17.567602

1.1 Translation initiation is a complex and crucial step in eukaryotic translation

Protein translation is one of the most energetically expensive cellular processes and is highly regulated, especially during translation initiation (Buttgereit and Brand 1995; Sonenberg and Hinnebusch 2009; Topisirovic and Sonenberg 2011; Hershey et al. 2012; Leibovitch and Topisirovic 2018). Translation initiation is a complex process which regulates expression of eukaryotic genes and employs over a dozen eukaryotic translation initiation factors (eIFs) (Sachs and Varani 2000; Jackson et al. 2010; Hinnebusch 2011; Aitken and Lorsch 2012). These include eIF1, eIF1A, eIF3, eIF5, eIF2 and the eIF4F complex, which is composed of eIF4E, eIF4A and eIF4G (Jackson et al. 2010; Aitken and Lorsch 2012). During eukaryotic translation initiation, a ternary complex made up of initiator methionyl-tRNA (Met-tRNA$_i$), eIF2, and GTP is formed (Olsen et al. 2003; Hinnebusch 2014). The 43S pre-initiation complex (PIC) then comes together by recruitment of the ternary complex, the 40S ribosomal subunit, and eukaryotic initiation factors 1, 1A, 3 and 5 (Pestova et al. 1998; Asano et al. 2001; Algire et al. 2002; Majumdar et al. 2003; Kolupaeva et al. 2005). After adopting an open conformation, the 43S PIC joins eIF4F at the mRNA 5′ -7-methylguanosine cap structure in order to recruit the mRNA to form the 48S initiation complex (Hinnebusch 2014). This newly formed 48S initiation complex is then capable of scanning the mRNA through its 5′ untranslated region (5' UTR) until it locates a start codon (Pestova and Kolupaeva 2002). Once the start codon is recognized, several initiation factors are released in order for the ribosome to begin elongation (Aitken and Lorsch 2012; Hinnebusch 2014).

1.2 Eukaryotic initiation factors regulate specialized translation

During initiation, the roles of several eIFs have been linked to translation regulation of subsets of mRNAs. For example, experiments performed in human cells revealed that eIF3 regulates the translation of specific mRNAs by direct interactions (Lee et al. 2015, 2016; De Silva et al. 2021). These eIF3-mRNA interactions are important for homeostasis but also play essential roles upon nutrient deprivation and drive the integrated stress response, among other functions (Xu et al. 2012; Lee et al. 2015, 2016; Pulos-Holmes et al. 2019; Tacca et al. 2019; Cate 2017; Gomes-Duarte et al. 2017; Lamper et al. 2020; Lin et al. 2020; Wolf et al. 2020; De Silva et al. 2021; Mukhopadhyay et al. 2023; Mestre-Fos et al. 2023). eIF4A, an RNA helicase, has also been associated with translation regulation of a subset of mRNAs in human cells, more specifically by unwinding 5′ UTRs that are highly structured and polypurine rich and many of which are related to cell-cycle progression and apoptosis (Svitkin et al. 2001; Rubio et al. 2014; Iwasaki et al. 2016, 2019). Moreover, eIF1 and eIF5 play important roles in the selection of translational start sites, depending not only on the translational context of the start codon (AUG), but also

on the abundance of these initiation factors and specific cellular conditions (Hann et al. 1992; Fletcher et al. 1999; Sonenberg and Dever 2003; Loughran et al. 2012; Ivanov et al. 2008, 2010, 2022).

1.3 Eukaryotic initiation factor 3 (eIF3) regulates translation of specific mRNAs

Translation initiation factor eIF3 is a crucial player in protein expression regulation through its roles in bridging the 43S PIC and eIF4F complexes (Figure 1.1), and also by performing specialized regulatory roles (Kolupaeva et al. 2005; Hinnebusch 2006; Valášek et al. 2003). eIF3 specifically binds to and regulates translation of a subset of mRNAs, many of which are involved in cell cycle regulation, cell growth, differentiation, and other crucial cellular functions. The interaction between eIF3 and mRNAs was shown to be mediated by RNA structural elements in the 5′ UTR of specific mRNAs in human embryonic kidney (HEK293T) cells and to cause translational activation or repression of these mRNAs (Lee et al. 2015). eIF3 has also been shown to have cell-specific regulatory roles in T cells, with eIF3 interactions throughout the entire length of the transcript for specific mRNAs, such as the ones encoding the T cell receptor alpha and beta subunits (TCRA and TCRB, respectively), mediating a translational burst essential for T cell activation (De Silva et al. 2021). In yeast, eIF3 has also been linked to mRNA recruitment and scanning as a mediator of mRNA-PIC interactions (Jivotovskaya et al. 2006; Chiu et al. 2010; Mitchell et al. 2010). Furthermore, in zebrafish eIF3 subunit H (EIF3H) was shown to regulate translation of mRNAs encoding the eye lens protein crystallin during embryogenesis (Choudhuri et al. 2013). These examples demonstrate that eIF3 plays a variety of mRNA-specific regulatory roles.

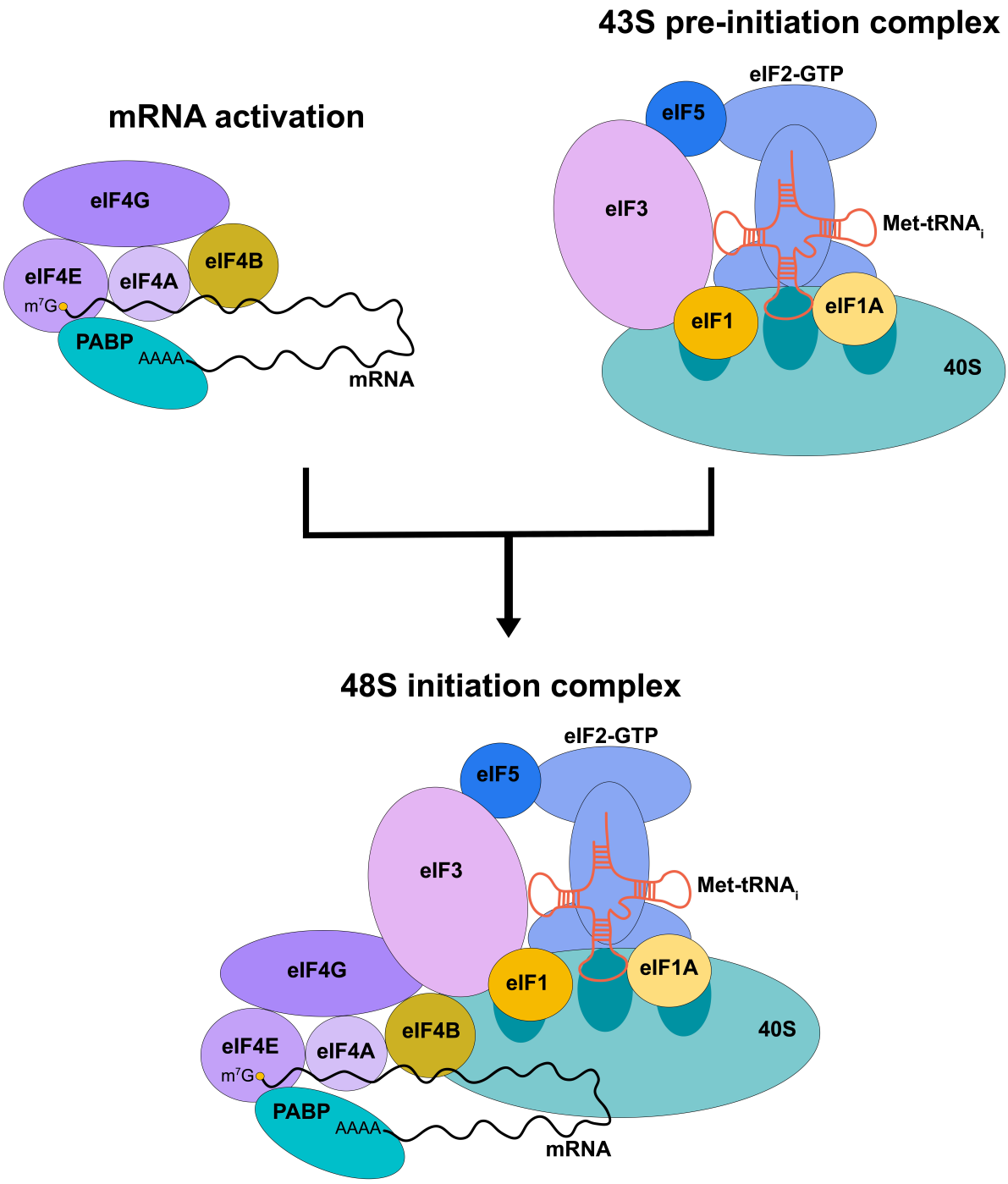


Figure 1.1. eIF3 acts as a scaffold in the formation of the 48S initiation complex. Schematic depicting the first steps of eukaryotic translation initiation. eIF: eukaryotic initiation factor, m^{7}G: 5′-7-methylguanosine cap structure, PABP: poly(A)-binding protein, Met-tRNA$_i$: initiator methionyl-tRNA.

1.4 *JUN* translation regulation is not fully understood

One of the reported eIF3-target mRNAs in human cells, *JUN*, encodes the transcription factor JUN, also known as c-Jun, which regulates gene expression in response to different stimuli (Wisdom et al. 1999; Meng and Xia 2011). As a component of the activator protein-1 (AP-1) complex, JUN regulates transcription of a large number of genes and acts mainly as a transcriptional activator (Smeal et al. 1991). JUN is therefore highly involved in various cellular processes including cell proliferation, apoptosis, tumorigenesis, and it was the first oncogenic transcription factor discovered (Bohmann et al. 1987; Ryder et al. 1988; Meng and Xia 2011). Regulation of JUN expression is particularly important because its downregulation can lead to cell cycle defects and its upregulation can lead to accelerated cell proliferation, which occurs in some cancers (Johnson et al. 1996; Gee et al. 2000; Briggs et al. 2002; Vasilevskaya and O'Dwyer 2003; Nateri et al. 2005; Hui et al. 2007; Blau et al. 2012; Chen and Bourguignon 2014). Therefore, it is not surprising for JUN expression regulation to be complex and to occur at both the transcriptional and translational levels. At the transcriptional level, *JUN* mRNA expression is regulated by its own protein product, which binds a high-affinity AP-1 binding site in the *JUN* promoter region and in turn induces its transcription (Nakamura et al. 1991; Angel et al. 1988; Lamph et al. 1988). *JUN* expression regulation at the translational level is mediated by its mRNA interaction with eIF3. Binding of eIF3 subunits EIF3A, EIF3B, EIF3D, and EIF3G to a stem loop in the *JUN* 5′ UTR results in activation of translation (Lee et al. 2015). Moreover, eIF3 subunit D (EIF3D) acts as a 5' cap-binding protein on the *JUN* mRNA, mediated by a cis-acting RNA element located in the 153 nucleotides immediately downstream of the *JUN* 5′-7-methylguanosine cap structure (Lee et al. 2016). This RNA element is also thought to block recruitment of the eIF4F complex (Lee et al. 2016). *JUN* expression regulation at the translational level has also been shown to be affected by m^6A methylation by METTL3 in its 3′ UTR and by contributions of an RNA structural element which activates its translation in glioblastoma (Blau et al. 2012; Suphakhong et al. 2022).

JUN possesses a longer than average 977-nucleotide 5′ UTR that is highly GC rich. Due to its length and complexity, *JUN*'s 5′ UTR might present additional layers of translational regulation of its mRNA through novel structural and/or sequence elements. Previously reported involvement of several initiation factors, including eIF3 and eIF4A, in the recruitment of mRNAs with long and structurally complex 5′ UTRs further supports a 5′ UTR-mediated mechanism for *JUN* translation regulation and suggests that additional factors may be involved in *JUN* regulation (Parsyan et al. 2011; Stanciu et al. 2022). For example, most recently *JUN* was shown to be sensitive to RocA, an anti-cancer drug that clamps eIF4A onto specific polypurine sequences - mainly GAA(G/A) - in the 5′ UTRs of a subset of mRNAs (Iwasaki et al. 2016, 2019). However, the implications of this interaction on *JUN* translation have not been previously evaluated. *JUN* also possesses two potential translational start sites, an upstream start codon (uAUG) located 4 codons upstream of the main start codon (mAUG). However, translational start site selection for the *JUN* mRNA has not been previously explored. The work presented in Chapter 2 focuses on investigating *JUN* translation regulation in human cells by exploring different regions of the *JUN* 5′ UTR (Figure 1.2) and how mRNA features and the interaction of initiation factors in these regions contribute to *JUN* translation.

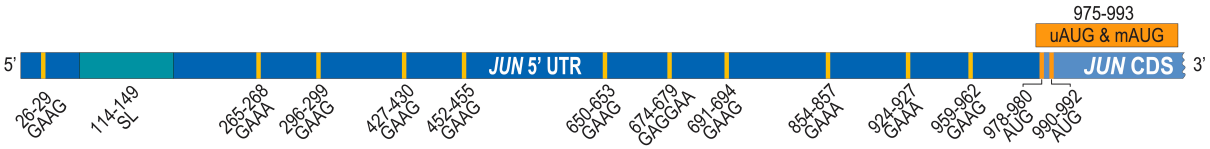


Figure 1.2. *JUN* possesses a complex 5' UTR and two start codons
Schematic showing the *JUN* 5' UTR features investigated in this study and their nucleotide position. SL: eIF3-interacting stem loop, GAA(G/A): polypurine sequence onto which RocA clamps eIF4A, uAUG: upstream start codon, mAUG: main start codon, CDS: coding sequence.

1.5 Isolation of a human multifactor complex (MFC) is yet to be achieved

A crucial step in eukaryotic translation initiation is the formation of a ternary complex made up of initiator methionyl-tRNA (Met-tRNA$_i$), eIF2, and GTP (Olsen et al. 2003; Hinnebusch 2014). This ternary complex delivers the Met-tRNA$_i$ to the 40S subunit of the ribosome, which it accomplishes with the aid of eukaryotic initiation factors 1, 1A, 3 and 5 (Asano et al. 2001; Algire et al. 2002; Majumdar et al. 2003; Kolupaeva et al. 2005; Pestova et al. 1998). The resulting complex, known as the 43S preinitiation complex (PIC), then recruits an activated mRNA for scanning. During scanning, the Met-tRNA$_i$ is transferred to the P-site of the ribosome through a reaction promoted by eIF5 in which the eIF2-GTP complex is hydrolyzed, releasing eIF2-GDP and the Met-tRNA$_i$. It is widely accepted that this collaboration between the ternary complex and other initiation factors occurs mainly upon formation of the PIC. Interestingly, studies have shown that it is possible for a multifactor complex (MFC) composed of eIF3, eIF5 and eIF1 in addition to the ternary complex to form prior to recruitment of the 40S ribosomal subunit (Asano et al. 2000; Valášek et al. 2003; Sokabe et al. 2012).

The presence of the translation initiation MFC composed of eIF1, eIF2, eIF3, eIF5 and the initiator methionyl-tRNA (Met-tRNA$_i$) has been observed in yeast (Figure 1.3) (Asano et al. 2000; Valášek et al. 2003). The yeast MFC was identified both *in vitro* and *in vivo* by sucrose density gradients and by affinity purification using tagged components (Asano et al. 2000; Valášek et al. 2002). Insights on the existence of this complex in yeast were initially suggested by interactions of different eIF3 subunits including TIF32 (EIF3A), NIP1 (EIF3C) and PRT1 (EIF3B) with SUI1 (eIF1), eIF5, and eIF2 in yeast (Phan et al. 1998; Asano et al. 1998, 1999, 2000; Valášek et al. 2002, 2003). Other studies established that eIF1, eIF2, eIF3 and eIF5 form stable interactions with each other in yeast, with eIF5 being the main mediator of MFC formation by bridging the interaction between eIF3 and eIF2 (Asano et al. 2000; Singh et al. 2004; Yamamoto et al. 2005). The yeast MFC has also been reconstituted from purified components and its structure, both on its own and bound to the 40S ribosomal subunit, has been elucidated at low resolution using cryo-EM (Gilbert et al. 2007).

Meanwhile, MFC characterization in other eukaryotes has mainly been achieved *in vitro*. For example, MFC formation has been reported *in vitro* from *Arabidopsis thaliana* and wheat components (Dennis et al. 2009). Here, it was demonstrated that a plant MFC is formed *in vitro* and that phosphorylation by the protein kinase CK2 stabilizes the interactions between MFC components both from purified components and in plant cell extracts. Disrupting phosphorylation of eIF5 resulted in a decrease in binding of MFC components as well as in lower levels of translation, which may point to the importance of the MFC interactions for the integrity of translation initiation in plants (Dennis et al. 2009). First insights on a mammalian MFC were provided by experiments showing that eIF3 is able to stimulate Met-tRNA$_i$ binding to the 40S ribosomal subunit using purified components *in vitro* (Benne and Hershey 1978). This was validated *in vivo* using yeast strains carrying eIF3 mutations (Feinberg et al. 1982; Naranda et al. 1994; Danaie et al. 1995; Phan et al. 1998). Additional studies using purified human components *in vitro* demonstrated interactions between MFC components, including eIF3-eIF1 and eIF5-eIF2 (Fletcher et al. 1999; Bieniossek et al. 2006). Moreover, transient transfection of polyhistidine-tagged eIF3 into human cells and subsequent purification revealed binding of endogenous eIF5 to the expressed eIF3 *in vivo* (Bandyopadhyay and Maitra 1999). All of these findings demonstrate the conservation of interactions between MFC components both in yeast and humans. These observations posed the question of whether an MFC forms endogenously in human cells as well. A later study provided evidence for the presence of a ribosome-free MFC in human and rabbit cell extracts (Sokabe et al. 2012). This study was also able to reconstitute the human MFC *in vitro* using purified human MFC components at physiological concentrations. Moreover, this study characterized the interactions between different combinations of purified human MFC components and proposed the eIF2-eIF3 interaction as the main mediator of MFC formation in humans (Sokabe et al. 2012). As a whole, although MFC interactions have been reconstituted *in vitro* using purified human components, a fully endogenous human MFC is yet to be isolated.

A concrete function for the MFC is also yet to be established. However, the general consensus is that the MFC acts as a translation initiation intermediate. In yeast, the consistent stable association of eIF1, eIF3, eIF5 and the ternary complex as part of the MFC supports an integral role for the MFC in translation initiation and links the functions of these individual initiation factors as a whole entity (Valášek et al. 2002, 2003). It has been proposed that formation of the MFC stimulates binding of the ternary complex to the 40S ribosomal subunit in yeast, both *in vitro* and *in vivo*, and that this is mainly mediated by interactions between eIF2 and eIF3 (Asano et al. 2001; Valášek et al. 2002). Moreover, mutations in the eIF5 motif that bridges the eIF3-eIF2 interaction in the yeast MFC significantly disrupted translation initiation *in vivo* (Asano et al. 2000). These studies also revealed the presence of Met-tRNA$_i$ as part of the yeast MFC, specifically bound to eIF2 (Asano et al. 2000). All of this suggests that the MFC plays an integral role in mediating transfer of the Met-tRNA$_i$ to the 40S, therefore promoting engagement of ribosomes for translation initiation in yeast. However, *in vitro* studies have shown that in mammalian systems the presence of the MFC does not accelerate binding of the ternary complex to the 40S subunit (Sokabe et al. 2012). In addition, even though the MFC may be playing a role in the release of eIF2-GDP from the ribosome after recognition of the start codon in yeast, this doesn't seem to be the case from MFC reconstituted from mammalian

components (Jennings and Pavitt 2010; Sokabe et al. 2012). It has also been shown that the Met-tRNA$_i$ can be delivered to the 40S ribosomal subunit by the human MFC just as efficiently as by the ternary complex on its own, which would suggest a redundant role for these complexes in mammals (Sokabe et al. 2012). Given this, a more feasible hypothesis is that the MFC may be acting as a reservoir for initiation factors, facilitating the initial steps of translation by making its main players more accessible (Aitken and Lorsch 2012). Moreover, structural studies from *in vitro* reconstituted yeast MFC suggest that binding of the MFC to the 40S subunit may be causing a conformational change that facilitates mRNA loading onto the 40S subunit (Gilbert et al. 2007). This finding is yet to be explored in a mammalian system. It is also unclear whether the MFC is formed by the same components for all mRNAs in the cell, or whether it plays a role in specialized translation by combining different initiation factors depending on the mRNA. This could be true for example in cases where translation occurs independent of the eIF4F complex (Kwan and Thompson 2019). Because of all this, further mechanistic studies are needed in order to understand the formation and function of the MFC, especially in mammalian cells. In Chapter 3, we explore isolation of the human MFC using the *JUN* mRNA as a platform for complex formation.

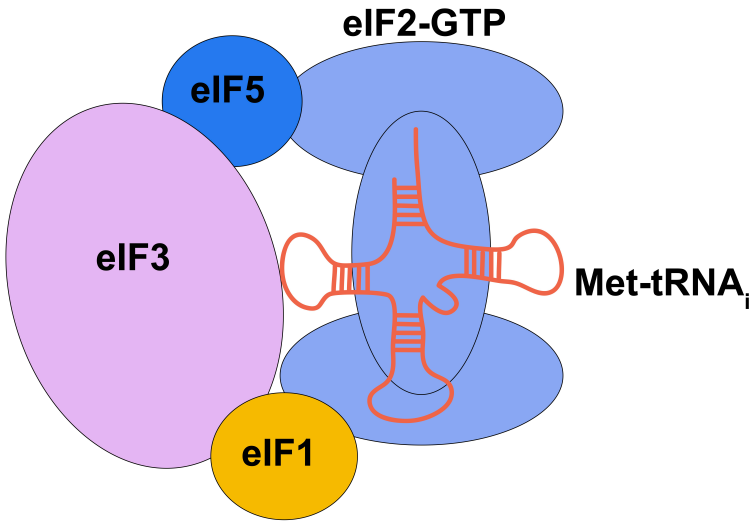


Figure 1.3. Composition of the previously reported yeast multifactor complex (MFC). Schematic depicting components of the yeast MFC. eIF: eukaryotic initiation factor, Met-tRNAi: initiator methionyl-tRNA.

Chapter 2: *JUN* mRNA Translation Regulation is Mediated by Multiple 5' UTR and Start Codon Features

Portions of this chapter were adapted from the following publication:
González-Sánchez, A.M., Castellanos-Silva, E.A., Díaz-Figueroa, G., Cate, J.H.D. 2023. *JUN* mRNA Translation Regulation is Mediated by Multiple 5' UTR and Start Codon Features. bioRxiv. https://doi.org/10.1101/2023.11.17.567602

2.1 Introduction

As presented in Chapter 1, JUN is an oncogenic transcription factor that acts in response to external stimuli, such as cellular stress (Bohmann et al. 1987; Ryder et al. 1988; Wisdom et al. 1999; Meng and Xia 2011). The transcription factor JUN is encoded by the *JUN* mRNA, which is a transcript targeted by eIF3 for translation regulation (Lee et al. 2015). EIF3 subunits A, B, D and G bind the 5' UTR of the *JUN* mRNA and activate *JUN* translation (Lee et al. 2015). EIF3 subunit D also binds the 5′ -7-methylguanosine cap structure of the *JUN* transcript independently of eIF4E, which in turn suggests a non-canonical mode of translation initiation for *JUN* (Lee et al. 2016). Both of these findings point to a mechanism of specialized translation for the *JUN* mRNA, mediated by its 5' UTR.

The 5' UTR of mRNAs has been well-established as a region for translation regulation (Pesole et al. 2001; Mignone et al. 2002; Araujo et al. 2012; Leppek et al. 2018). This is due to the presence of regulatory elements within this region which may adopt secondary structures or contain sequence motifs that mediate their interaction with a variety of regulatory proteins, such as initiation factors. For example, it is well known that eIF3 interacts with RNA structural elements, as previously mentioned, and that it contributes to the recruitment of mRNAs with long 5' UTRs (Lee et al. 2015; Pulos-Holmes et al. 2019; Stanciu et al. 2022). Moreover, several studies have shown the effects of GC content and high levels of RNA secondary structure in the 5' UTR in decreasing translational efficiency (Pelletier and Sonenberg 1985; Babendure et al. 2006). Others have identified sequence motifs in the 5' UTR with which initiation factors can interact. For example, a previous study identified GAA(G/A) as the motif to which the cancer compound Rocaglamide A (RocA) clamps eIF4A, suggesting in turn an interaction of this initiation factor with a specific subset of mRNAs (Iwasaki et al. 2016, 2019). In addition, 5' UTRs can contain more than one start codon, for example an upstream start codon (uAUG) and a main start codon (mAUG). These help regulate translation of a transcript by modulating start codon selection, which is in turn mediated by eIF1 and eIF5 (Maag et al. 2006; Ivanov et al. 2010; Loughran et al. 2012; Ivanov et al. 2022). Regulatory elements within the 5' UTR are therefore capable of modulating translation in a variety of cellular conditions and disease (Holcik and Sonenberg 2005; Barbosa et al. 2013; Xiang et al. 2023).

Interestingly, the *JUN* 5' UTR harbors all of the aforementioned regulatory elements. These include a long sequence (977 nucleotides) with a high GC content, high levels of secondary structure (Lee et al. 2015), 11 instances of the GAA(G/A) RocA-eIF4A binding motif and an uAUG. Given this, in this chapter we present experiments investigating *JUN* translation regulation in human cells by exploring mRNA features in

different regions of the *JUN* 5′ UTR and how the interaction of initiation factors in these regions contributes to *JUN* translation. Firstly, we applied mutagenesis to the *JUN* 5′ UTR near the eIF3 binding site to determine whether other sequence or structural elements in this region contribute to *JUN* translation regulation. We also further investigated the contributions of eIF4A to *JUN* translation both by mRNA mutagenesis and through cellular treatment with RocA. Finally, we explored how the translational context of both *JUN* start codons affect start site selection. We also explored conservation of the *JUN* sequence containing both the uAUG and mAUG with their translational contexts. With this work we aim to reveal additional layers of regulation for the *JUN* mRNA and to provide insights into its potential participation in a non-canonical pathway of translation initiation. As a whole, results shown in this chapter demonstrate that *JUN* translation regulation is a complex process that involves various initiation factors, including eIF3 and eIF4A, and mRNA features such as secondary structures in the 5' UTR.

2.2 Results

2.2.1 *JUN* translation is regulated by 5' UTR sequence and structural elements

Binding of eIF3 to a stem-loop in the 5' UTR of the *JUN* mRNA leads to its translational activation (Lee et al. 2015). Mutations in this stem loop have been shown to disrupt the interaction with eIF3 and to repress *JUN* translation (Lee et al. 2015). However, the effects of other mutations in the *JUN* 5′ UTR remain to be explored. We first tested whether mutations in other regions within and near the *JUN*-eIF3 interacting stem loop (SL) affect *JUN* translation. We generated mRNA reporter constructs containing the full-length *JUN* 5′ UTR and Nanoluciferase (Nluc) coding sequence (CDS) that included mutations in a 208 nucleotide (nt) SL proximal region whose secondary structure was previously determined (Lee et al. 2015) by selective 2'-hydroxyl acylation analyzed by primer extension, also known as SHAPE (Figure 2.1A , SHAPE). All of the mutations disrupt either the secondary structure or the sequence of highly structured regions within the SL proximal region (Figure 2.1B). For each of these constructs transfected into HEK293T cells, together with an mRNA reporter with the Hemoglobin Beta Subunit (*HBB)* 5′ UTR and a Firefly luciferase (Fluc) CDS as an internal control, we assessed translation using luciferase assays.

As expected, deletion of the *JUN*-eIF3 interacting stem loop (Figure 2.1C, mutant ΔSL) significantly represses *JUN* reporter translation when compared to the WT construct. Mutations to SL loop nucleotides C128-U129, previously shown to be unreactive by SHAPE mapping *in vitro* and therefore likely to be involved in RNA-RNA contacts, also significantly affected *JUN* reporter translation, with U129G dramatically increasing translation (Figure 2.1C, mutant A) (Lee et al. 2015). Interestingly, replacing the SL loop with a much smaller and possibly more stable UUCG tetraloop substantially increased *JUN* reporter translation (Figure 2.1C, mutant E) (Antao et al. 1991). However, replacing all of the U's with A's in the loop sequence had little effect on *JUN* translation (Figure 2.1C, mutant D). As a whole, these findings support the importance of the SL loop in *JUN* translation regulation, yet reveal a complexity in its role maintaining and stabilizing the secondary structure of the SL region. Mutations in other structured regions of the *JUN* 5′ UTR near the eIF3 binding site also significantly affected *JUN* translation. For example,

disrupting the stem loop between nucleotides 23 and 33 with point mutations in nucleotides 24 and 30 repressed *JUN* reporter translation (Figure 2.1C, mutant F). By contrast, deleting the bulge loop formed by nucleotides 42 to 47 increased *JUN* reporter translation (Figure 2.1C, mutant G). These findings suggest that these secondary structure features in the *JUN* 5′ UTR outside the originally identified eIF3 binding site play opposing roles in regulating *JUN* translation. However, mutations to two other loop and bulge regions near the SL (nts 160-166 and 184-187) had little or no effect on *JUN* reporter translation (Figure 2.1C, mutants H-J).

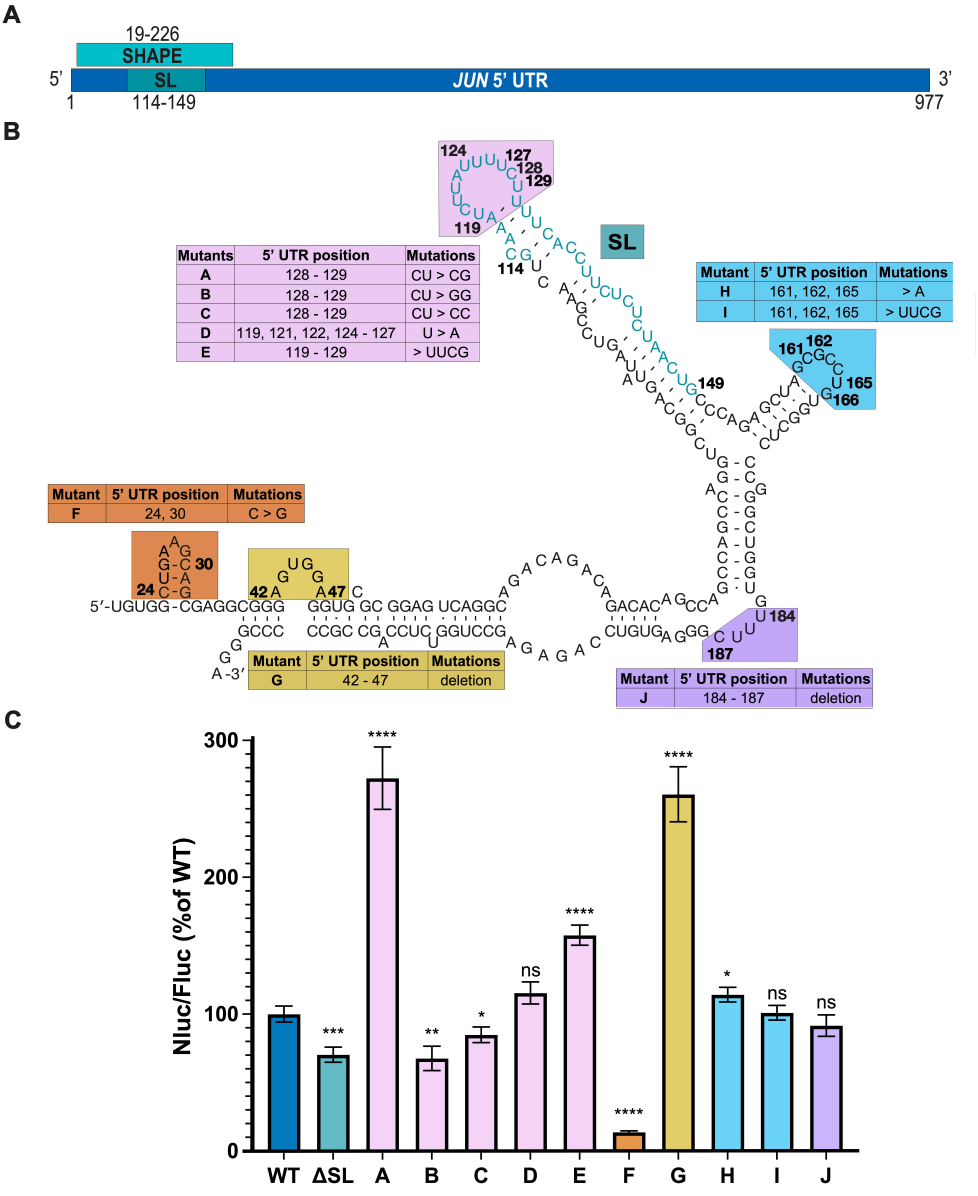


Figure 2.1. *JUN* translation is regulated by 5' UTR sequence and structural elements. (A) Depiction of the full *JUN* 5' UTR. The locations of the 208-nt region studied by SHAPE (SHAPE) and the eIF3-interacting stem loop (SL) are marked, along with the

nucleotides involved in each region. (B) Secondary structure of the 208-nt region in the *JUN* 5' UTR mapped by SHAPE is shown. Nucleotides are numbered according to their position in the 5′ UTR. Mutant *JUN* 5′ UTR mRNA constructs and their corresponding mutations are described in their associated tables. (C) Luminescence measured from HEK293T cells transfected with the *JUN* 5′ UTR reporter mRNAs expressing Nanoluciferase (Nluc). Translation was assessed using a dual-luciferase assay and normalized to a control mRNA harboring an *HBB* 5′ UTR and a Firefly luciferase (Fluc) CDS. Nluc/Fluc ratios were normalized to the WT *JUN* 5′ UTR, set as 100%. Technical triplicates for each biological replicate, and a total of at least three biological replicates were taken for each measurement. P values determined using a one-sample t test versus a hypothetical value of 100 are shown as follows: *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$, ****$p \leq 0.0001$. The mean value of the replicates and standard error of the mean are shown.

2.2.2 *JUN* is highly sensitive to RocA treatment

Rocaglamide A (RocA) is an anti-cancer compound that specifically clamps eIF4A onto polypurine sequences in a subset of mRNAs, in an ATP-independent manner. This clamping of eIF4A blocks 43S scanning, leading to premature, upstream translation initiation and reducing protein expression from transcripts containing RocA–eIF4A target sequences (Iwasaki et al. 2016, 2019). Interestingly, *JUN* is one of the mRNAs identified as highly sensitive to RocA treatment (Iwasaki et al. 2016). However, little is known about how promoting or disrupting the *JUN* interaction with eIF4A affects *JUN* translation. To this end, we first transfected *JUN* 5′ UTR and Nluc mRNA reporter constructs designed above (Figure 2.1B) together with the *HBB* 5′ UTR and Fluc CDS control mRNA, into HEK293T cells and treated these with increasing concentrations of RocA or DMSO (as a negative control). In all the cases we tested, including the WT, ΔSL, and the UUCG tetraloop mutation in the SL loop, treatment with RocA strongly suppressed *JUN* reporter translation (Figure 2.2A). This effect was not observed for the control Nluc reporter mRNA harboring the *HBB* 5′ UTR, which has not been reported as RocA sensitive. The fact that constructs with mutations that affect the eIF3-interacting stem loop in the *JUN* 5′ UTR were still highly sensitive to RocA treatment suggests that the RocA-mediated effects on the *JUN* 5′ UTR are independent of eIF3 regulation. The persistent repressive trend of RocA treatment on *JUN* translation also suggests that eIF4A serves an important role in *JUN* translation regulation.

RocA-sensitive mRNAs are enriched in the polypurine sequence GAA(G/A) (Iwasaki et al. 2016). As shown in Figure 2.2B, *JUN* possesses 11 of these polypurine sequences across the entire length of its 5′ UTR, with none present in the eIF3-interacting stem loop. In order to evaluate the effect of disrupting these sequences in the *JUN* 5′ UTR, we mutated these polypurine (GAA(G/A)) sequences to the mixed purine/pyrimidine sequence CAAC, previously reported to disrupt RocA-mediated eIF4A binding to mRNAs (Iwasaki et al. 2019). Interestingly, the *JUN* reporter mRNAs with these mutations (mutants CAAC or CAAC + ΔSL) remained highly sensitive to RocA (Figure 2.2C). This indicates that there are additional eIF4A target sequences in the *JUN* 5′ UTR that are not necessarily equivalent to the reported predominant GAA(G/A) motif. Moreover, deleting the eIF3 interacting stem loop, together with the GAA(G/A) mutations (mutant CAAC + ΔSL), has no further effect on translation. We observed similar effects with the *JUN* mRNA

reporters *in vitro* using HEK293T cell extracts (Figure 2.2D). Taken together, these results support a model in which eIF4A regulates *JUN* translation in an eIF3 independent manner, pointing to further layers of regulation for *JUN* translation, mediated by additional initiation factors.

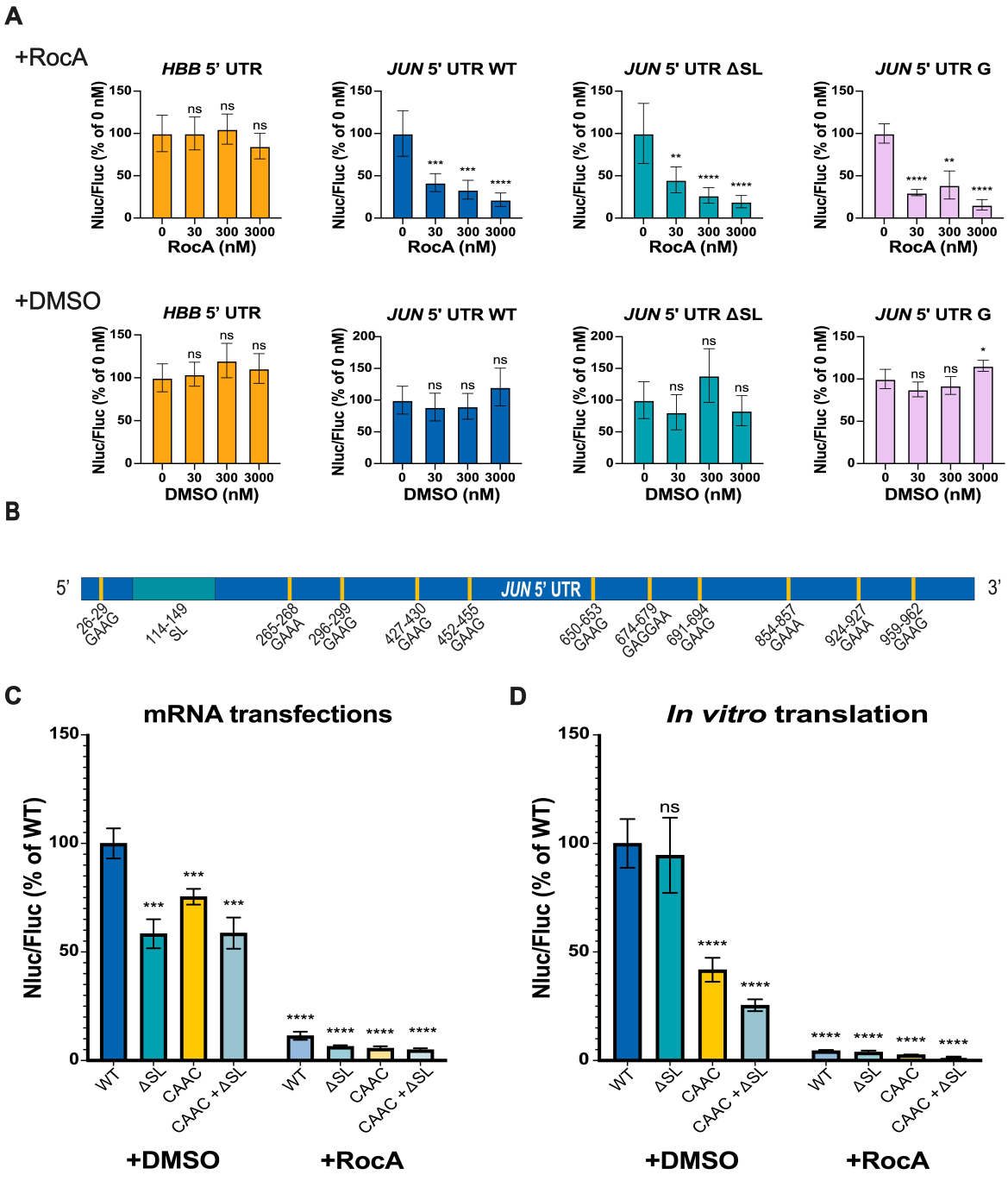


Figure 2.2. *JUN* is highly sensitive to RocA treatment. (A) HEK293T cells co-transfected with *JUN* 5′ UTR and Nluc CDS reporter mRNAs (WT, ΔSL or mutant G, Figure 1) and with an *HBB* 5' UTR and Fluc mRNA as an internal control, were treated

with increasing concentrations of RocA (+RocA) or DMSO control (+DMSO) 3 hours post-transfection, as previously reported (Iwasaki et al. 2016). An mRNA with the *HBB* 5′ UTR and Nluc CDS mRNA was also used as a RocA-insensitive control. Translation was assessed using a dual-luciferase assay as in Figure 2.1. Nluc/Fluc measurements were normalized to the corresponding untreated condition (0 nM RocA) and reported as a percentage of this measurement. (B) The location of polypurine (GAA(G/A)) sequences in the *JUN* 5′ UTR are indicated with yellow lines. Each of these 11 sequences was mutated to CAAC. (C) Luminescence of HEK293T cells transfected with *JUN* 5′ UTR and Nluc CDS reporter mRNAs (WT, ΔSL, CAAC or CAAC + ΔSL), together with the *HBB* 5′ UTR and Fluc CDS mRNA control. Transfected cells were treated with 300 nM RocA (+RocA) or DMSO (+DMSO) 3 hours post-transfection. Translation was assessed using a dual-luciferase assay as in Figure 2.1, and Nluc/Fluc measurements were normalized to the WT *JUN* 5′ UTR and Nluc CDS +DMSO measurements, reported as percentages. (D) Luminescence from *in vitro* translation reactions using the *JUN* 5′ UTR and Nluc CDS reporter mRNAs (WT, ΔSL, CAAC or CAAC + ΔSL). Reactions were treated with 300 nM RocA (+RocA) or DMSO (+DMSO). Luminescence values of each mutant were normalized to the WT *JUN* 5′ UTR and Nluc CDS +DMSO measurements and reported as percentages. In panels A, C, and D, technical triplicates for each biological replicate, and a total of at least three biological replicates were taken for each measurement. P values determined using a one-sample t test versus a hypothetical value of 100 are shown as follows: *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$, ****$p \leq 0.0001$. The mean value of the replicates and standard error of the mean are shown.

2.2.3 Two start codons contribute to *JUN* translation in cells

Start codon selection regulates the translation of many transcripts (Ivanov et al. 2008, 2010, 2017, 2022; Loughran et al. 2012; Kozak 1986). Recently, it was reported that the translational context of start codons on transcripts with an upstream open reading frame (uORF) and a main open reading frame (mORF) affects which of these is preferentially selected for translation, mediated by eukaryotic initiation factor 1 (eIF1) and eukaryotic initiation factor 5 (eIF5) (Ivanov et al. 2022). While eIF1 promotes skipping of weak translational start sites, eIF5 increases initiation at these sites. The relative abundance of these two factors determines which start codon is used. The strongest translational context, also known as the ideal Kozak sequence context, contains a purine at the -3 position, preferably an adenosine (A), and a guanosine (G) at the +4 position, relative to the AUG start codon (Figure 2.3A). A weak translational context results when either of these purines at the -3 and +4 positions is substituted by a pyrimidine. The *JUN* mRNA possesses two AUG start codons, an in-frame upstream AUG (uAUG) four codons before a main AUG (mAUG), with different translational contexts (Figure 2.3A). The *JUN* uAUG possesses a weak translational context, with a uridine (U) at the -3 position and an adenosine (A) at the +4 position. By contrast, the *JUN* mAUG has a strong translational context, with an adenosine (A) at the -3 position and a guanosine (G) at the +4 position. It is not known which of these *JUN* AUGs is preferentially selected for translation and there currently is no evidence of JUN peptides that initiate at the uAUG.

To investigate whether *JUN* translation can initiate at either AUG or whether one is preferentially selected, we designed mRNA reporter constructs containing the *JUN* 5′ UTR and the first 51 nucleotides of the *JUN* CDS (corresponding to 17 amino acids), followed by the full Nluc CDS (Figure 2.3B). The WT version of this construct therefore contains both *JUN* AUG start codons and their intact translational contexts. We then mutated start codons individually or their translational context to test their roles in *JUN* translation. We transfected these mRNA reporters into HEK293T cells, together with the *HBB* 5′ UTR and Fluc CDS control, and monitored translation using luciferase assays. In general, disrupting either AUG or changing their translational context significantly represses *JUN* translation, which in turn suggests that translation can initiate at both AUGs (Figure 2.3C). We found that disrupting either AUG by mutation to AAG repressed *JUN* reporter translation, consistent with both AUGs contributing to JUN translation (Figures 2.3B and 2.3C). The more substantial decrease in *JUN* reporter translation due to the mAUG (*JUN* ΔmAUG, 95% reduction) compared to mutation of the uAUG (*JUN* ΔuAUG, 75% reduction) suggests that the mAUG start codon may be preferred in our experimental conditions.

Changing the translational context of either AUG also repressed *JUN* translation. Interestingly, making the sequence context for the uAUG stronger – either by introducing an A in the -3 position of the upstream AUG (Figure 2.3B) or by also including a G mutation in the +4 position to make it an ideal Kozak sequence – resulted in a 50% decrease in translation (Figure 2.3C). Moreover, using the uAUG in a strong Kozak context while weakening the translational context of the mAUG further represses *JUN* translation, to about 10% of the WT levels (Figure 2.3C, mutant S-uAUG W-mAUG). Taken together, these results strongly support the hypothesis that both AUGs are used for translation, and that the preference for which AUG is selected for initiation depends partly on its translational context.

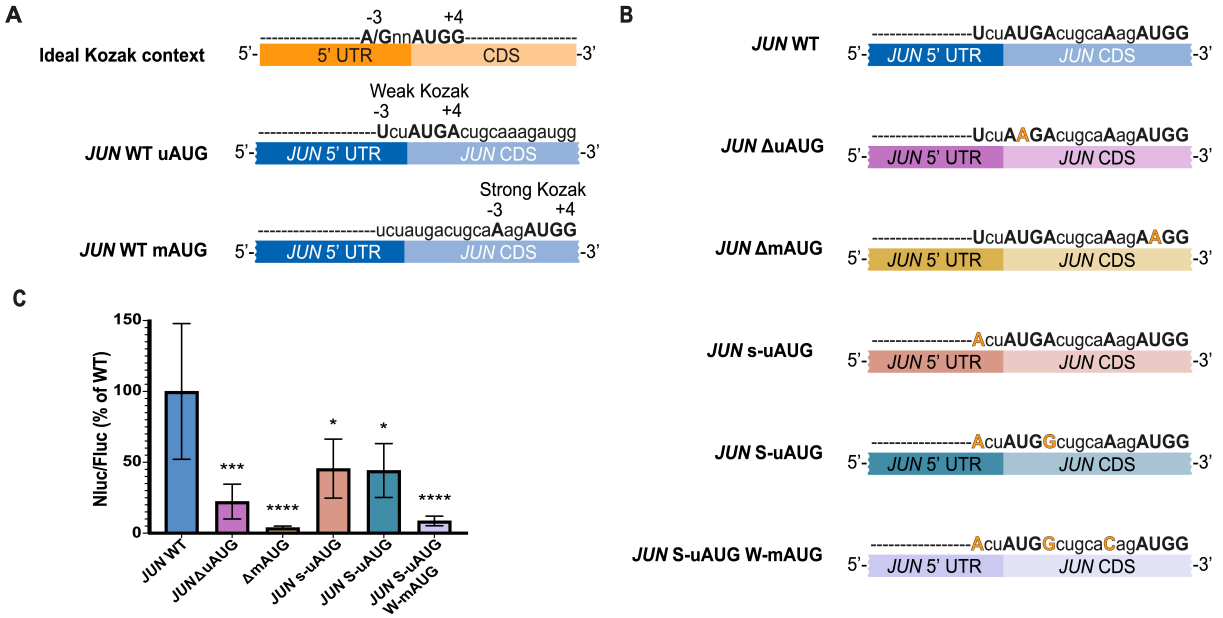


Figure 2.3. Two start codons contribute to *JUN* translation in cells. (A) Diagram depicting the ideal Kozak context for a generic open reading frame (A/GnnAUGG). Below,

diagrams depicting each of the *JUN* start codons (AUG) and their translational contexts. (B) Diagram depicting *JUN* mRNA reporter constructs, with their corresponding mutations in each of the *JUN* start codons and their translational contexts. The constructs contained the full *JUN* 5′ UTR sequence along with the first 51 nucleotides of the *JUN* CDS, upstream of the full Nluc CDS. (C) Luminescence from HEK293T cells transfected with *JUN* 5′ UTR and 51nt *JUN* CDS and Nluc CDS reporter mRNAs (WT, ∆uAUG, ∆mAUG, s-uAUG, S-uAUG or S-uAUG W-mAUG), together with an *HBB* 5′ UTR and Fluc CDS control, assessed using a dual-luciferase assay as in Figure 2.1. Nluc/Fluc measurements of each mutant were normalized to the WT *JUN* 5′ UTR and 51nt *JUN* CDS and Nluc CDS measurements and reported as percentages. Technical triplicates for each biological replicate, and a total of at least three biological replicates were taken for each measurement. P values determined using a one-sample t test versus a hypothetical value of 100 are shown as follows: *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$, ****$p \leq 0.0001$. The mean value of the replicates and standard error of the mean are shown.

2.2.4 *JUN* uAUG and mAUG are conserved in vertebrates

To further investigate whether both *JUN* AUGs contribute to its translation, we examined sequence conservation of the *JUN* 5′ UTR and early CDS region that contains both AUG start codons and their translational context. We searched the 19-nucleotide region spanning the Kozak contexts of both AUGs in 100 species using the Genome Data Viewer (NLM-NCBI) and Ensembl for sequence confirmation (Table A.1). Remarkably, sequences in this region are conserved both at the nucleotide and at the amino acid level in the species examined (Figure 2.4A and 2.4B). Conservation of both *JUN* AUGs is present in all vertebrates, whereas only the mAUG is present in the invertebrates we investigated (Figure 2.4C). This conservation of both of *JUN*'s AUGs and their translational context suggests an ancient mechanism for *JUN* translation regulation and highlights the importance of both *JUN* AUGs.

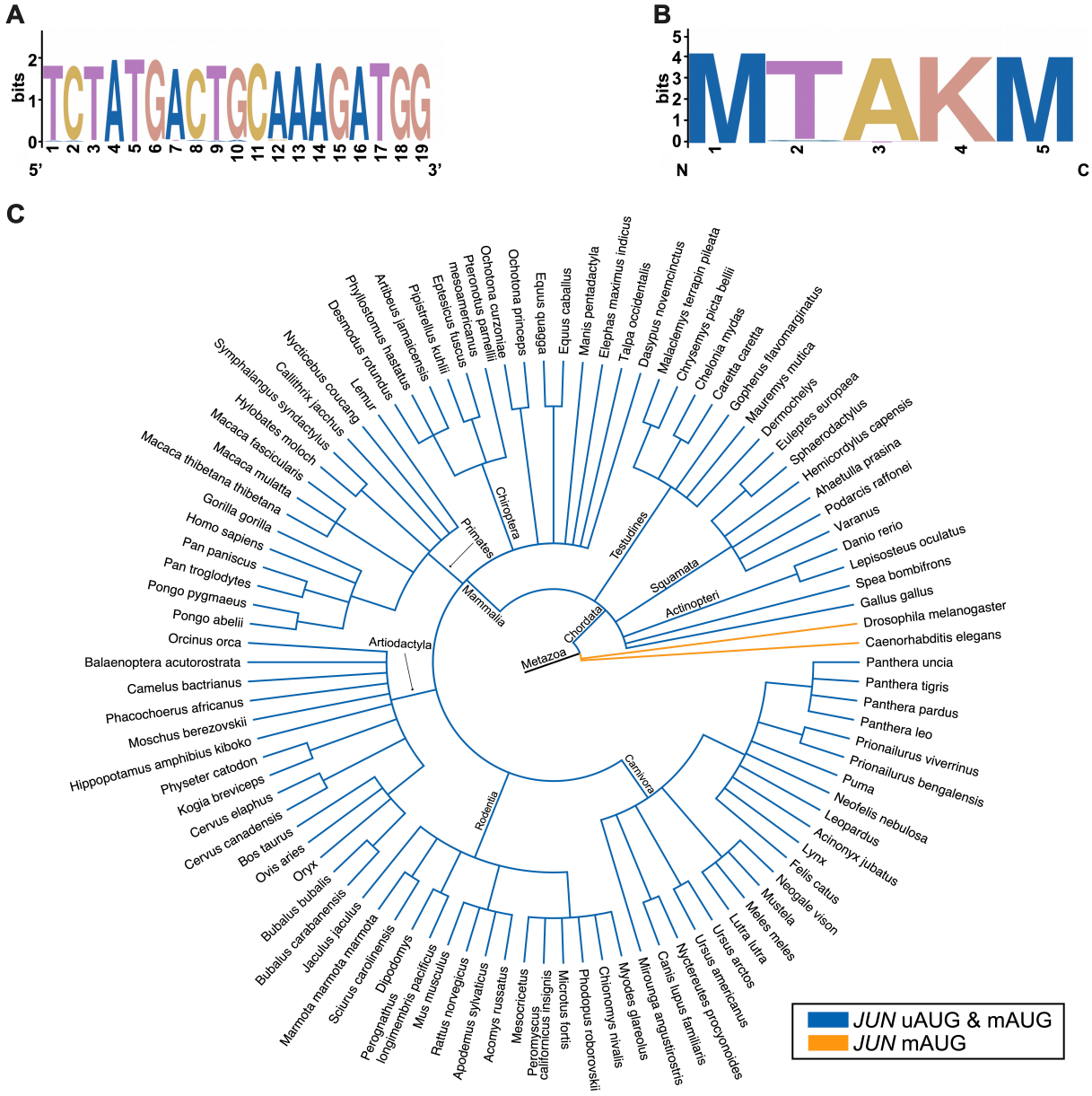


Figure 2.4. Both *JUN* start codons are conserved in vertebrates. (A) Sequence logo depicting the conservation of the 19 nucleotide *JUN* sequence spanning both start codons and their translational context amongst 100 species. (B) Sequence logo depicting the conservation of the 5 amino acid *JUN* sequence containing both start codon methionines amongst 100 species. (C) Phylogenetic tree depicting the conservation of both *JUN* AUGs amongst 100 species. Species with both the *JUN* uAUG and the *JUN* mAUG are depicted with blue branches, while species with only one *JUN* mAUG are depicted in orange.

2.3 Discussion

Since JUN was the first oncogenic transcription factor identified (Bohmann et al. 1987; Ryder et al. 1988) it is notable how little is known mechanistically about how *JUN* expression is controlled at the translational level. In this work we probed the contributions of mRNA features and initiation factors to *JUN* translation regulation in human cells. Our study reveals that *JUN* translation regulation is a complex process that is mediated by mRNA target sequences and structural elements spanning the entire *JUN* 5′ UTR. Moreover, we provide evidence that initiation factors in addition to eIF3 (Lee et al. 2015) contribute to *JUN* translation regulation. We also found that both the uAUG and mAUG contribute to *JUN* translation. Given that the *JUN* 5′ UTR has a length that exceeds the average 218 nt human 5′ UTR (Leppek et al. 2018), a high level of secondary structure (Lee et al. 2015), and a GC rich sequence, our hypothesis is that many features within its 5′ UTR that participate in its regulation are still unknown.

Previous results found that eIF3 can directly bind structures in the 5’ UTR of specific mRNA transcripts to regulate their translation, with *JUN* serving as a prototypical example (Lee et al. 2015). Here we explored the regulatory roles of RNA structural elements within or near the eIF3-interacting stem loop (SL) region of the *JUN* 5’ UTR (Figure 2.1). In addition to the importance of this SL for enhancing *JUN* translation in cells, we found that replacing the SL loop by a highly-stable UUCG tetraloop (Antao et al. 1991) increases *JUN* translation. It is possible that significant local rearrangements may be required for the canonical *JUN* SL loop sequence to bind eIF3 or that the canonical SL loop is highly dynamic, and insertion of the UUCG tetraloop locks this structure in the most favorable conformation for eIF3 binding. Additionally, there may be some sequence specificity in the SL loop, as mutation of two nucleotides in the context of the wild-type loop at positions 128 and 129 also affect *JUN* translation levels (Figure 2.1). Interestingly, the ability of this eIF3-interacting SL structure to promote translation is shown by the fact that it can be inserted in a modular way into the 3′ UTR of reporter mRNAs to promote translation, as shown in activated T cells (De Silva et al. 2021). We also found additional structural elements besides the eIF3-interacting SL that contribute to *JUN* translation. Most notably, these are a stem loop between nucleotides 23 and 33 of the *JUN* 5′ UTR and a bulge loop between nucleotides 42 and 47, which enhance or repress translation, respectively (Figure 2.1). These secondary structure elements may serve potential regulatory roles, similar to the one shown for the eIF3-interacting SL. These results are consistent with previous findings which have correlated long and highly structured 5′ UTRs with complex regulation mediated by eIF3 (Stanciu et al. 2022). However, it remains to be determined whether eIF3 interacts directly with these regions. It is also possible that these secondary structure elements mediate additional regulatory interactions (Leppek et al. 2018). Further evidence will be required to determine whether additional initiation factors interact with the structural elements studied in this SL-proximal region.

We also found evidence for a role for eIF4A in *JUN* translation regulation. We demonstrated that *JUN* is highly sensitive to RocA, consistent with prior transcriptome-wide experiments (Iwasaki et al. 2016) and with *JUN* being a target of eIF4A regulation (Figure 2.2). Interestingly, RocA sensitivity is independent of *JUN* interactions with eIF3, since mutations in the *JUN* eIF3-interacting SL did not affect its sensitivity to RocA (Figure

2.2). RocA was shown to clamp eIF4A onto GAA(G/A) polypurine sequences in a subset of RocA sensitive mRNAs and these mRNAs are in fact rich in these tetramer motifs (Iwasaki et al. 2016). Notably, *JUN* possesses 11 of these GAA(G/A) motifs in its 5' UTR; however, mutating these sequences to CAAC did not overcome *JUN* sensitivity to RocA, suggesting that RocA may clamp eIF4A onto additional polypurine sequences in the *JUN* 5′ UTR different from the predominant motif previously identified (Iwasaki et al. 2016). A potential polypurine sequence present in the *JUN* 5′ UTR and to which RocA might clamp eIF4A is AGAG (Iwasaki et al. 2019). Within eIF3, subunit EIF3D can bind to the *JUN* mRNA 5′-7-methylguanosine cap structure, while an RNA structural element adjacent to the cap blocks recruitment of the eIF4F complex (Lee et al. 2016). However, our results with RocA treatment suggest that at least some of the eIF4F components may contribute to *JUN* mRNA recruitment and scanning. This suggests that there may be a novel mRNA recruitment complex for *JUN*, in which eIF4A is present despite the absence of eIF4E, with EIF3D possibly acting as the cap-binding protein in this context.

Although *JUN* possesses a 5' UTR nearly 1 kb in length, it also has two closely-spaced potential start codons, an upstream start codon (uAUG) 4 codons away from a downstream "main" AUG (mAUG). However, which of these start codons is preferentially selected and whether they both contribute to *JUN* translation is currently unknown. Notably, experimental evidence for usage of the uAUG would be missed in published mass spectrometry experiments due to presence of a lysine at codon -1 relative to the mAUG, which would lead to removal of the leading peptide in commonly-used trypsin digests. Using reporters with the full-length *JUN* 5' UTR and both AUGs, we find that both AUGs likely contribute to *JUN* translation, albeit in a complex way (Figure 2.3). For example, deleting each AUG individually, repressed *JUN* translation significantly, with deletion of the mAUG causing a more severe reduction. However, the contexts of the uAUG and mAUG do not always correlate with translational output. For example, changing the weak context of the uAUG seen in WT *JUN* into a strong context decreased translation by 50% rather than increasing it. In this case, the mAUG is also likely used, as weakening the translational context of the mAUG in the strong uAUG context background further repressed translation to about 10% of WT levels. These results suggest that while both AUGs contribute to *JUN* translation, perhaps the mAUG plays a major role. These results also raise the possibility that translational efficiency of the first 4 codons including the uAUG may be lower than that of the mAUG, which would result in a lower translational output from the uAUG when it is used.

The fact that both AUGs may contribute to *JUN* translation suggests they may be part of a regulatory switch in varying cellular conditions. For example, unwinding of an RNA secondary structure downstream of an uAUG in an immune response promotes translation initiation at the mAUG of specific mRNAs in *Arabidopsis thaliana* (Xiang et al. 2023). Other cellular conditions, such as stress, starvation, or polyamine abundance could influence start codon selection (Hann et al. 1992; Hinnebusch 2005; Ivanov et al. 2008; Starck et al. 2016). Finally, the relative abundance of eIF1 and eIF5 – which regulate the stringency of start codon selection (Loughran et al. 2012; Ivanov et al. 2010, 2022) – could influence which *JUN* start codon is used, and thus the translational output of the *JUN* mRNA. Further experiments will be needed in order to test this hypothesis.

When exploring the evolution of *JUN*'s AUGs we found that both are conserved in vertebrates, which suggests an ancient mechanism of regulation for *JUN* by means of

translational start site selection. Importantly, the translational context is also conserved for most of the examined species (Figures 2.4A and 2.4B, Table A.1), suggesting that the translational context plays a significant role in determining which start codon is selected. Our observations align with previous reports which showed that uAUGs are highly conserved in higher eukaryotes due to their roles in modulating translation initiation under regulatory circumstances (Chew et al. 2016; Zhang et al. 2021). In addition, the evolutionary conservation suggests that more than one JUN polypeptide may be expressed by initiation of translation at both the uAUG and the mAUG. This type of alternative initiation has been shown previously by leaky scanning of uAUGs in a weak translational context, especially of those that are close to their downstream mAUG which allows for backward oscillation of the ribosome (Smith et al. 2005; Matsuda and Dreher 2006). Further studies are needed in order to test whether *JUN* leads to expression of more than one polypeptide, depending on the start codon selected. For example, this would require using a different protease for mass spectrometry besides trypsin to avoid cleavage after the lysine at position -1 relative to the mAUG, to retain N-terminal peptides originating at the uAUG.

The fact that JUN was the first oncogenic transcription factor identified (Bohmann et al. 1987; Ryder et al. 1988) makes it notable that many different mechanisms regulate JUN expression at the translational level. Results summarized in this chapter demonstrate the potential of the *JUN* mRNA as a model transcript for understanding new mechanisms of mRNA translation regulation. Our findings open the doors for further exploration of the regulatory roles of long and highly structured 5′ UTRs and the initiation factors that participate in translation regulation. It also points to possible new roles for *JUN* mRNA translation levels in mediating cellular response to a wide array of physiological conditions.

2.4 Materials and Methods

Reporter plasmids

To generate the *JUN* 5′ UTR and the *HBB* 5′ UTR Nluc reporter plasmids, the *JUN* 5′ UTR (ENST00000371222.4) previously generated by amplification from human cDNA (Lee et al. 2015) and the *HBB* 5′ UTR (ENST00000335295.4) commercially generated (IDT) sequences were each inserted into the pNL1.1 *NanoLuc* luciferase reporter plasmid (Promega, GenBank Accession Number JQ437370) downstream of a T7 promoter using overlap-extension PCR with Q5 High-Fidelity DNA Polymerase (NEB) and InFusion cloning (Takara Bio). For the *JUN* AUG mutants, the first 51 nucleotides of the *JUN* CDS were inserted downstream of the full *JUN* 5′ UTR sequence and upstream of the full Nluc CDS in the pNL1.1 plasmid. For the Fluc reporter plasmid, the *HBB* 5′ UTR Nluc reporter plasmid was amplified and the *NanoLuc* luciferase sequence was replaced by a commercially generated *Firefly* luciferase sequence (IDT) (Giacomelli et al. 2018). Subsequent mutant versions of the *JUN* reporter plasmids were made by amplifying the plasmid using overlap-extension PCR with Q5 High-Fidelity DNA Polymerase (NEB) and primers containing the corresponding mutations, insertions, or deletions, followed by InFusion cloning (Takara). All primers used for amplification can be found in Table 2.1. All sequences were verified by Sanger sequencing.

Table 2.1: Primers used for cloning

Primer ID	**Sequence**
HBB 5' UTR Nluc CDS vector – Forward	GGCCGCGACTCTAGAGTCGG
HBB 5' UTR Nluc CDS vector – Reverse	GGTGGCGGTGTCTGTTTGAGG
Fluc CDS insert - Forward	ACAGACACCGCCACCATGGAAGACGCCAAAAACATAAAG
Fluc CDS insert - Reverse	TCTAGAGTCGCGGCCTTAcacggcgatctttccgccct
JUN 5' UTR Mutant A - Forward	TTATTTTCGTTTCACCTTCTCTCTAACTGCCC
JUN 5' UTR Mutant A - Reverse	GGTGAAACGAAAATAAGATTTGCAGTTCGGAC
JUN 5' UTR Mutant B - Forward	TTATTTTGGTTTCACCTTCTCTCTAACTGCCC
JUN 5' UTR Mutant B - Reverse	GTGAAACCAAAATAAGATTTGCAGTTCGGAC
JUN 5' UTR Mutant C - Forward	TTATTTTCCTTTCACCTTCTCTCTAACTGCCC
JUN 5' UTR Mutant C - Reverse	GTGAAAGGAAAATAAGATTTGCAGTTCGGAC
JUN 5' UTR Mutant D - Forward	GCAAAACAAAAAAACATTTCACCTTCTCTCTAACTGCCC
JUN 5' UTR Mutant D - Reverse	ATGTTTTTTTGTTTTGCAGTTCGGACTATACTGCCG
JUN 5' UTR Mutant E - Forward	TGCAAATTCGTTTCACCTTCTCTCTAACTGCCC
JUN 5' UTR Mutant E - Reverse	GGTGAAACGAATTTGCAGTTCGGACTATACTGCC
JUN 5' UTR Mutant F - Forward	CTGAAGGAGGGAGGCGGGAGTGGAGGTG
JUN 5' UTR Mutant F - Reverse	GCCTCCCTCCTTCAGCCACACTCAGTGCAAC
JUN 5' UTR Mutant G - Forward	GAGGCGGGGGTGCGCGGAGTCAGGCAG
JUN 5' UTR Mutant G - Reverse	GCGCACCCCCGCCTCGCTGCTTCAGC
JUN 5' UTR Mutant H - Forward	AGCTAGAACCAGTGGCTCCCGGGCTG
JUN 5' UTR Mutant H - Reverse	CCACTGGTTCTAGCTCTGGGCAGTTAGAGAGAAGGT
JUN 5' UTR Mutant I - Forward	GAGCTATTCGTGGCTCCCGGGCTGGTG
JUN 5' UTR Mutant I - Reverse	AGCCACGAATAGCTCTGGGCAGTTAGAGAGAAGG

JUN 5' UTR Mutant J - Forward	GGCTGGTGGGGAGTGTCCAGAGAGCCTG
JUN 5' UTR Mutant J - Reverse	CACTCCCCACCAGCCCGGGAGCCAC
JUN 5' UTR Mutant CAAC insert - Forward	ACGACTCACTATAGGGCTCAGAGTTGCACTGAGT GTG
JUN 5' UTR Mutant CAAC insert - Reverse	TCTAGAGTCGCGGCCGATCAAAAACATGGGTGAT CCTCA
JUN 5' UTR Mutant CAAC ΔSL insert - Forward	GGCCGCGACTCTAGAGTCGG
JUN 5' UTR Mutant CAAC ΔSL insert - Reverse	cctatagtgagtcgtattaGGTGGCTTTACC
JUN 5' UTR Nluc CDS vector - Forward	ATGGTCTTCACACTCGAAGATTTCGTTGGGGACT GGCGACAGACAGCCGG
JUN 5' UTR Nluc CDS vector - Reverse	GAGTGTGAAGACCATCGAGGCGTTGAGGGCATC GTCATAGAAGGTCGTTTCCATCTTTGCAGTCATAG AACAGTCCGTCACTTCACG
JUN 5' UTR s-uAUG Nluc CDS vector - Reverse	GAGTGTGAAGACCATCGAGGCGTTGAGGGCATC GTCATAGAAGGTCGTTTCCATCTTTGCAGTCATag Tacagtccgtcacttcacg
JUN 5' UTR ΔuAUG Nluc CDS vector - Reverse	GAGTGTGAAGACCATCGAGGCGTTGAGGGCATC GTCATAGAAGGTCGTTTCCATCTTTGCAGTCTTag Aacagtccgtcacttcacg
JUN 5' UTR ΔmAUG Nluc CDS vector - Reverse	GAGTGTGAAGACCATCGAGGCGTTGAGGGCATC GTCATAGAAGGTCGTTTCCTTCTTTGCAGTCATag Aacagtccgtcacttcacg
JUN 5' UTR S-uAUG Nluc CDS vector - Reverse	GAGTGTGAAGACCATCGAGGCGTTGAGGGCATC GTCATAGAAGGTCGTTTCCATCTTTGCAGCCATag Tacagtccgtcacttcacg
JUN 5' UTR S-uAUG W-mAUG Nluc CDS vector - Reverse	GAGTGTGAAGACCATCGAGGCGTTGAGGGCATC GTCATAGAAGGTCGTTTCCATCTGTGCAGCCATag Tacagtccg

In vitro transcription

All RNA reporters were made by *in vitro* transcription with a standard T7 RNA polymerase protocol using DNA template gel extracted using the Zymoclean Gel DNA Recovery Kit (Zymo), 1x T7 RNA Polymerase buffer (NEB), 5 mM ATP (Thermo Fisher Scientific), 5 mM CTP (Thermo Fisher Scientific), 5 mM GTP (Thermo Fisher Scientific), 5mM UTP (Thermo Fisher Scientific), 5 µg BSA (NEB), 9 mM DTT, 25 mM $MgCl_2$, 200U T7 RNA polymerase (NEB), 50U Murine RNAse inhibitor (NEB) and incubating for 4 hours at 37 °C. The DNA template used for *in vitro* transcription was generated by PCR amplification from the corresponding reporter plasmid using the Q5 High-Fidelity DNA Polymerase (NEB) with a reaction including a forward primer containing the T7 promoter sequence and a 60T reverse primer for polyadenylation. Primers used for each transcript can be found in Table 2.2. After *in vitro* transcription, RNAs were treated with RQ1 DNAse (Promega) following the manufacturer's protocol and precipitated with 7.5 M lithium chloride. RNAs were then capped using Vaccinia D1/D2 (Capping enzyme) (NEB) and 2′ O-methylated using Vaccinia VP39 (2′ O Methyltransferase) (NEB) in a reaction that also included 1X capping buffer (NEB), 10 mM GTP (Thermo Fisher Scientific) and 4 mM SAM (NEB). RNAs were then purified with the RNA Clean and Concentrator-5 Kit (Zymo). In order to verify the integrity of the *in vitro* transcribed mRNAs, 6% polyacrylamide TBE-

Urea denaturing gels were run using 1X TBE (Invitrogen), a ssRNA ladder (NEB) and SYBR safe stain (see representative gel in Figure S2.1).

Table 2.2: Primers used for DNA template preparation

Primer ID	**Sequence**
JUN 5' UTR - Transcript - Forward	taatacgactcactatagggctcagagttgcactgag
Nluc 60T - Transcript - Reverse	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTTTTTTTTTACGCCAGAATGCGT TCGCAC
HBB 5'UTR - Transcript - Forward	taatacgactcactataggACATTTGCTTCTGACACAACTG TG
Fluc 60T - Transcript - Reverse	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTTTTTTTTTAcacggcgatctttccgccct tcttgg
Nluc MS2 60T - Transcript - Reverse	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTTTTTTTTTACATGGGTGATCCT CATGTaaatgatc

HEK293T cells and mRNA transfections

HEK293T cells were maintained in DMEM (Gibco) supplemented with 10% FBS (VWR) and 1% Pen/Strep (Gibco). Cells were grown at 37 °C in 5% carbon dioxide and 100% humidity. Luciferase reporter mRNAs were transfected into these cells using the TransIT-mRNA Transfection Kit (Mirus), with the following protocol modifications. HEK293T cells were seeded into opaque 96-well plates (Thermo Fisher Scientific) about 16 hours prior to transfections. The next day, once the cells reached 80% confluency, transfections were performed by adding the following to each well: 7 µL of pre-warmed OptiMEM media (Invitrogen), 500 ng of 5′ -capped and 3′ -polyadenylated Nluc reporter mRNA, 150 ng of 5′ -capped and 3′ -polyadenylated Fluc reporter mRNA, 2 µL of Boost reagent (Mirus Bio) and 2 µL of TransIT mRNA reagent (Mirus Bio). Transfection reactions were incubated at room temperature for 3 minutes prior to drop-wise addition into each well. Transfected cells were incubated at 37 °C for 8 hours, after which luciferase assays were performed using the NanoGlo Dual-Luciferase Reporter Assay System (Promega) following the manufacturer's protocol. Luminescence was then measured using a Spark multimode microplate reader (TECAN). Nluc/Fluc ratios were normalized to the corresponding control condition, set as 100%. Technical triplicates for each biological replicate, and a total of at least three biological replicates were taken for each measurement. P values were determined using a one-sample t test versus a hypothetical value of 100. The mean value of the replicates and standard error of the mean were plotted.

HEK293T pSB-HygB-GADD34-K3L cells and extract preparation

HEK293T pSB-HygB-GADD34-K3L cells (Aleksashin et al. 2023) were maintained in DMEM media (Gibco) supplemented with 10% Tet-system approved FBS (Gibco) and 1% Pen/Strep (Gibco). Cells were grown at 37 °C in 5% carbon dioxide and 100% humidity. Cells were grown for extract preparation as follows. The day after plating cells

from a frozen stock into a T25 flask (Cell Star), media was exchanged and supplemented with 200 µg/mL Hygromycin B (Invitrogen). The following day, cells were transferred to a T75 flask (Corning) with media supplemented with 200 µg/mL Hygromycin B. Once cells reached 100% confluency, half of the cells were transferred to a T175 flask (Falcon) with media supplemented with 200 µg/mL Hygromycin B. Once cells reached 100% confluency, cells were passaged onto 25 150 mm plates (Corning) at a 1 to 25 ratio. The next day, cells were treated overnight with 20 µg Doxycycline (Takara Bio) per plate.
In vitro translation extracts were made from HEK293T pSB-HygB-GADD34-K3L cells using a previously described protocol (Aleksashin et al. 2023). Cells were placed on ice, scraped and collected by centrifugation at 1000 xg for 5 minutes at 4 °C. Cells were washed once with ice-cold DPBS (Gibco) and collected once again by centrifugation at 1000 xg for 5 minutes at 4°C. After this, cells were homogenized with an equal volume of freshly made ice-cold hypotonic lysis buffer (10 mM HEPES-KOH pH 7.6, 10 mM KOAc, 0.5 mM $Mg(OAc)_2$, 5 mM dithiothreitol). After hypotonic-induced swelling for 45 min on ice, cells were homogenized using a syringe attached to a 26G needle (BD). Extract was then centrifuged at 15000 xg for 1 minute at 4 °C. The resulting supernatant was aliquoted, frozen with liquid nitrogen, and stored at -80 °C.

In vitro translation

In vitro translation reactions were performed using HEK293T pSB-HygB-GADD34-K3L translation-competent cell extract, as previously described (Aleksashin et al. 2023). Translation reactions contained 50% translation-competent cell extract, 52 mM HEPES pH 7.4 (Takara), 35 mM potassium glutamate (Sigma), 1.75 mM $Mg(OAc)_2$ (Invitrogen), 0.55 mM spermidine (Sigma), 1.5% Glycerol (Fisher Scientific), 0.7 mM putrescine (Sigma), 5 mM DTT (Thermo Scientific), 1.25 mM ATP (Thermo Fisher Scientific), 0.12 mM GTP (Thermo Fisher Scientific), 10 mM L-Arg; 6.7 mM each of L-Gln, L-Ile, L-Leu, L-Lys, L-Thr, L-Val; 3.3 mM each of L-Ala, L-Asp, L-Asn, L-Glu, Gly, L-His, L-Phe, L-Pro, L-Ser, L-Tyr; 1.7 mM each of L-Cys, L-Met; 0.8 mM L-Trp, 20 mM creatine phosphate (Roche), 60 µg/mL creatine kinase (Roche), 4.65 µg/mL myokinase (Sigma), 0.48 µg/mL nucleoside-diphosphate kinase (Sigma), 0.3 U/mL inorganic pyrophosphatase (Thermo Fisher Scientific), 100 µg/mL total calf tRNA (Sigma), 0.8 U/µL RiboLock RNase inhibitor (Thermo Scientific), and 1000 ng of the corresponding mRNA. Reactions were then incubated for 60 minutes at 32 °C, and Nanoluciferase activity was monitored using the Nano-Glo Luciferase Assay Kit (Promega) using a Spark multimode microplate reader (TECAN). The average of each biological replicate was normalized to the control condition, set as 100%. Technical triplicates for each biological replicate, and a total of at least three biological replicates were taken for each measurement. P values determined using a one-sample t test versus a hypothetical value of 100 are shown as follows: *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$, ****$p \leq 0.0001$. The mean value of the replicates and standard error of the mean were plotted.

Conservation analysis for *JUN* AUGs

The 19-nucleotide *JUN* 5′ UTR and *JUN* CDS region that spans both AUG start codons and their translational context was searched in 100 species. Species were selected randomly, starting with *Homo sapiens* and increasing the evolutionary distance throughout the vertebrates up to the invertebrates (Table A.1). Species sequences were compiled using the Genome Data Viewer (NLM-NCBI) and Ensembl. Sequence logos for the conserved nucleotide and amino acid sequences were created using WebLogo (https://weblogo.berkeley.edu/) (Crooks et al. 2004; Schneider and Stephens 1990). Taxonomy analysis for the species of interest was performed using the NCBI Taxonomy Browser (Schoch et al. 2020; Sayers et al. 2019). Phylogenetic tree was generated using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

2.5 Supplemental Figures

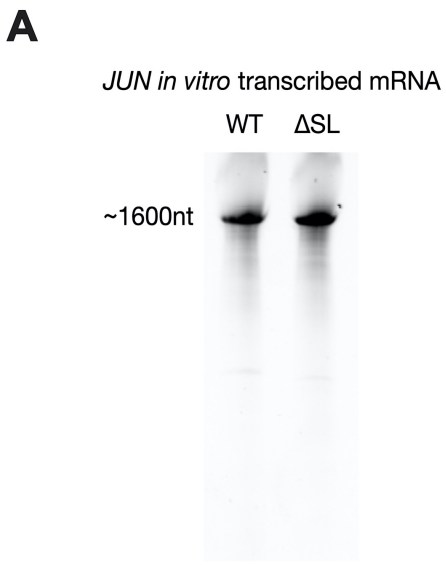


Figure S2.1. Representative mRNA TBE-Urea gel.
6% TBE-Urea gel for *in vitro* transcribed mRNA for the WT or ΔSL *JUN* 5′ UTR and Nluc CDS reporter constructs. nt, nucleotide.

Chapter 3: Isolation of a *JUN* Translation Initiation Multifactor Complex (MFC)

3.1 Introduction

As discussed in Chapter 1, the multifactor complex (MFC) is a pre-43S translation initiation intermediate composed of eIF1, eIF2, eIF3, eIF5 and Met-tRNA$_i$ in yeast (Asano et al. 2000). Although its role is still not well understood, it is thought to mediate transfer of the initiator methionyl-tRNA (Met-tRNA$_i$) onto the 40S ribosomal subunit (Asano et al. 2000, 2001; Valášek et al. 2002, 2003). However, it is also possible that the MFC is playing other roles such as acting as a reservoir to facilitate access to initiation factors (Aitken and Lorsch 2012) or mediating non-canonical modes of translation initiation, as we suggest may occur on the *JUN* mRNA. Though MFC formation has been studied in various eukaryotes including yeast and plants, little is known about its roles and composition in humans. One potential reason for this is that the MFC's intermediate nature renders it a challenging complex to isolate, particularly in the context of a mammalian system. However, as previously reported by our lab, we now know that eIF3 binds to the 5′ untranslated region (5′ UTR) and to the 5′ -7-methylguanosine cap structure of the *JUN* mRNA (Lee et al. 2015, 2016). Given that eIF3 is one of the core components of the MFC, this binding might render the *JUN* mRNA as a useful platform for isolation of this complex. Due to the absence of eIF4E in 48S sucrose gradient fractions from *JUN in vitro* translation reactions (Lee et al. 2016), our hypothesis is that eIF3 recruits the *JUN* mRNA for translation initiation by binding to its 5′ -7-methylguanosine cap structure together with the rest of the components of the human MFC. This may facilitate *JUN* mRNA translation initiation which, as shown in Chapter 2, is a complex and highly regulated process.

Because of this, in the experiments we present in Chapter 3 we aimed to purify a *JUN* MFC from human cells. Our approach combines *in vitro* translation using human cell extract with biochemical techniques used for purification of ribonucleoprotein (RNP) complexes. Isolation of RNPs can be achieved by many methods, including RNA-centric approaches which focus on capturing the RNA of interest and using it as a "bait" for capture of proteins bound to it (Jazurek et al. 2016). Proteins in the RNP can then be identified by a variety of methods including proteomics. This type of approach can be applied *in vitro* and *in vivo* and determining which specific method to use depends on the properties of the RNP of interest. For example, isolation of a complex that forms uniquely under metabolic stress would be more feasible to isolate *in vivo* from cells subjected to the conditions of interest. Notably, isolation of RNP complexes presents unique challenges. For example, it is difficult to isolate and identify complexes formed by low-abundance proteins and that are present in a complex mixture, such as cell extract, because many unspecific interactions may occur in such a mixture. Another potential challenge is the highly dynamic nature of numerous RNP complexes, many of which are transient (Jazurek et al. 2016). Isolation of the MFC complex presents both of these challenges. Because of this, we employed MS2-tagged RNA affinity purification (MS2-TRAP) which is one of the most-well established ribonucleoprotein (RNP) complex purification methods (Yoon et al. 2012; Yoon and Gorospe 2016; Jazurek et al. 2016). This method is based on the incorporation of an RNA element known as the MS2 RNA

stem loop into an RNA of interest, therefore tagging it for isolation. The MS2 stem loop is an RNA sequence from an *Escherichia coli* (*E. coli*) bacteriophage, which is recognized with high affinity and specificity by the MS2 coat protein (MS2) (Bernardi and Spahr 1972; Stripecke and Hentze 1992; Keryer-Bibens et al. 2008). This RNA-binding protein can in turn be fused to a protein such as the maltose-binding protein (MBP), which binds to amylose beads and provides a surface for immobilization of the MBP-MS2 bound RNP. The RNP of interest can then be eluted using maltose, which competes with the MBP-amylose interaction (Das et al. 2000; Zhou et al. 2002). This allows for specific isolation of the MBP-MS2 bound RNP complex as well as for removal of proteins bound non-specifically with the use of salt washes. For our experiments, a *JUN* 5' UTR reporter mRNA was tagged with the MS2 stem loop sequence and isolation of the MFC complex was pursued by addition of MBP-MS2 fusion protein to *in vitro* translation reactions using human cell extracts.

Since the MFC is a translation initiation complex, we also employed sucrose gradients for isolation of this complex. Sucrose gradients involve separation of translation complexes, including ribosomes and translation factors, using ultracentrifugation. They are widely applied for isolation of RNA-protein complexes formed during translation (Mašek et al. 2011). They also pose the advantage of allowing precise monitoring and isolation of complexes that form during specific stages of translation by means of fractionation. In terms of the MFC, the first full isolation of the yeast MFC was achieved by sucrose gradients from whole cell extracts (Asano et al. 2000). This makes these a promising approach for isolation of human MFC.

We explored the composition of the human *JUN* MFC using a combination of MS2 pulldowns, sucrose gradients, mass spectrometry and additional techniques such as real-time quantitative polymerase chain reaction (RT-qPCR) and western blots. As a whole, this chapter shows the establishment of an efficient protocol for human *JUN* MFC isolation, provides insights into the components of this complex including the presence of novel factors, and proposes future steps for validation of this complex.

3.2 Results

3.2.1 Establishing an MS2-TRAP protocol for human *JUN* MFC purification

Binding of eIF3 to the *JUN* mRNA has been previously confirmed both at the *JUN* 5' UTR and at its 5′ -7-methylguanosine cap structure (Lee et al. 2015, 2016). However, whether eIF3 binds *JUN* in conjunction with other initiation factors, for example as part of a translation initiation MFC, is currently unknown. Because of this, we sought to isolate a *JUN* reporter mRNA from *in vitro* translation reactions using human cell extract to capture the translation initiation factors that bind to it during translation. More specifically we used translationally active human cell extract from HEK293T cells, cells in which the eIF3 interaction with *JUN* was previously observed (Lee et al. 2015, 2016). Due to this interaction, we used an RNA pulldown approach in which the *JUN* mRNA acts as a "bait" for the factors that bind to it during translation. We adapted an MS2-tagged RNA affinity purification (MS2-TRAP, also referred to as MS2 pulldown) approach, previously used in our lab for isolation of bacterial ribosomes (Youngman and Green 2005; Ward et al. 2019). This approach uses the well-established MS2-binding RNA stem loop (Bernardi and

Spahr 1972; Stripecke and Hentze 1992; Yoon et al. 2012; Yoon and Gorospe 2016; Jazurek et al. 2016) as an RNA tag and the Maltose Binding Protein (MBP)- MS2 coat protein (MS2) fusion (MBP-MS2). In this context, the MBP-MS2 fusion protein will bind the MS2-tagged RNA and the RNA-protein complex will be isolated by binding of MBP to amylose beads with further elution using maltose.

As shown in Figure 3.1A, for this MS2-TRAP approach, we designed an mRNA reporter construct, made up of the *JUN* 5' UTR, the Nanoluciferase (Nluc) coding sequence (CDS), a short non-coding linker sequence, and the MS2 stem loop. The MS2 stem loop was specifically located at the 3' of this reporter mRNA to avoid interference with binding of factors to the *JUN* 5' UTR, which is where the MFC presumably engages. We prepared translationally active HEK293T cell extract and used it for *in vitro* translation reactions (IVTs) of the *JUN* mRNA reporter construct (Rakotondrafara and Hentze 2011). Initially, translation of this reporter was assessed using luciferase assays and once translation was confirmed, translation reactions were used as the input for the MS2-TRAP. *In vitro* translation of mRNA reporter constructs is highly dependent on the mRNA of interest (Rakotondrafara and Hentze 2011). As described in Chapter 2, the *JUN* 5' UTR is 977 nucleotides in length, is highly structured, has a high GC content, and its translation regulation is complex. All of this affects the rate and efficiency of translation of it both *in vivo* and *in vitro* due to effects in scanning and a greater requirement for the involvement of initiation factors (Pelletier and Sonenberg 1985; Svitkin et al. 2001; Pestova and Kolupaeva 2002; Babendure et al. 2006; Mitchell et al. 2010; Hinnebusch 2011; Vassilenko et al. 2011; Aitken and Lorsch 2012; Leppek et al. 2018). This caused a need for optimization of the *in vitro* translation protocol (Rakotondrafara and Hentze 2011) for efficient translation of the *JUN* reporter construct. As stated in the previously established IVT protocol, concentrations of magnesium and potassium greatly affect translation efficiency *in vitro* (Rakotondrafara and Hentze 2011). We therefore sought to find the optimal concentrations of magnesium and potassium in our *in vitro* translation reactions, by doing independent and contiguous titrations of these. Optimal levels of translation of the *JUN* 5' UTR WT reporter construct, as measured by Nluc luminescence were obtained when using 2 mM of magnesium and 60 mM of potassium per *in vitro* translation reaction (Figure S3.1). Moreover, under these conditions, we were able to replicate the results observed *in vivo* for both the *JUN* 5' UTR WT and the *JUN* 5' UTR mutant ΔSL (Figure 3.1B), in which the eIF3-interacting stem loop is deleted and which was described in Chapter 2. The concentration of RNA per *in vitro* translation reaction was also optimized and established at 1 μg (25 ng/μl of IVT reaction) (Figure S3.2). Cell lysate used for *in vitro* translation was also treated with an optimized concentration of Micrococcal nuclease (MNAse) in order to promote translation of the reporter mRNA (Figure S3.3).

Additional steps in the MS2-TRAP protocol needed optimization for isolation of a human RNA-protein complex. We first optimized the binding of the *JUN* reporter mRNA to the MBP-MS2 fusion protein. For this, we established a step previous to the *in vitro* translation reaction in which the reporter mRNA is incubated with the MBP-MS2 fusion protein at a ratio previously established (1:1000) for 2 hours at 4 °C (Ward et al. 2019). This promotes binding of the MBP-MS2 to the 3' of the *JUN* reporter mRNA, since it's the only protein present in the mixture at that time, and allows the 5' to be available for access of initiation factors during translation. Binding of the *in vitro* translation reaction sample to

the amylose beads was also optimized to 2 hours at 4 °C with rotation. Washes to remove unspecific binding of proteins to the amylose beads were optimized to 3 washes for 2 minutes each at room temperature using IVT buffer, a buffer similar to that included in the *in vitro* translation reaction (see Materials and Methods, section 3.4.6). Elution time was optimized to 45 minutes at 4 °C with rotation after addition of IVT buffer with 10 mM maltose. A schematic of the final optimized protocol is shown in Figure 3.1C. After protocol optimization, pulldown efficiency was evaluated with real time quantitative polymerase chain reaction (RT-qPCR) using primers that amplify the *JUN* reporter mRNA (see Table 3.3). RT-qPCR analysis revealed that 62% of the input mRNA was recovered in the elution, which is indicative of the pulldown efficiency (Figure 3.1D).

Optimization of the MS2-pulldown protocol led to successful isolation of a *JUN* translation initiation multifactor complex. The presence of translation initiation factors in the elution was determined using western blot analysis. As shown in Figure 3.2A, elution from MS2 pulldowns prepared using IVT reactions of the *JUN* 5' UTR WT construct revealed previously reported components of the MFC, such as eIF3 and eIF2, but lacked eIF1 and eIF5 which were previously reported in yeast (Asano et al. 2000). Interestingly, when probing for components of the eIF4F complex, eIF4E was absent while eIF4G and eIF4A1 were present. These findings support previous data in which eIF3 subunit D acts as the cap-binding protein for the *JUN* mRNA (Lee et al. 2016). These findings also suggest that components of the eIF4F complex are participating as part of a complex, together with eIF3 and eIF2, possibly for recruitment of the *JUN* mRNA during translation initiation. This also supports our findings shown in Chapter 2, where we demonstrated *JUN*'s sensitivity to the cancer compound Rocaglamide A (RocA), which clamps eIF4A to its target mRNAs and therefore demonstrates eIF4A's involvement in *JUN* translation regulation. As a whole, our findings up to this point show the initial steps for establishment of an MS2-TRAP protocol for isolation of a *JUN* translation initiation multifactor complex from human cell extract.

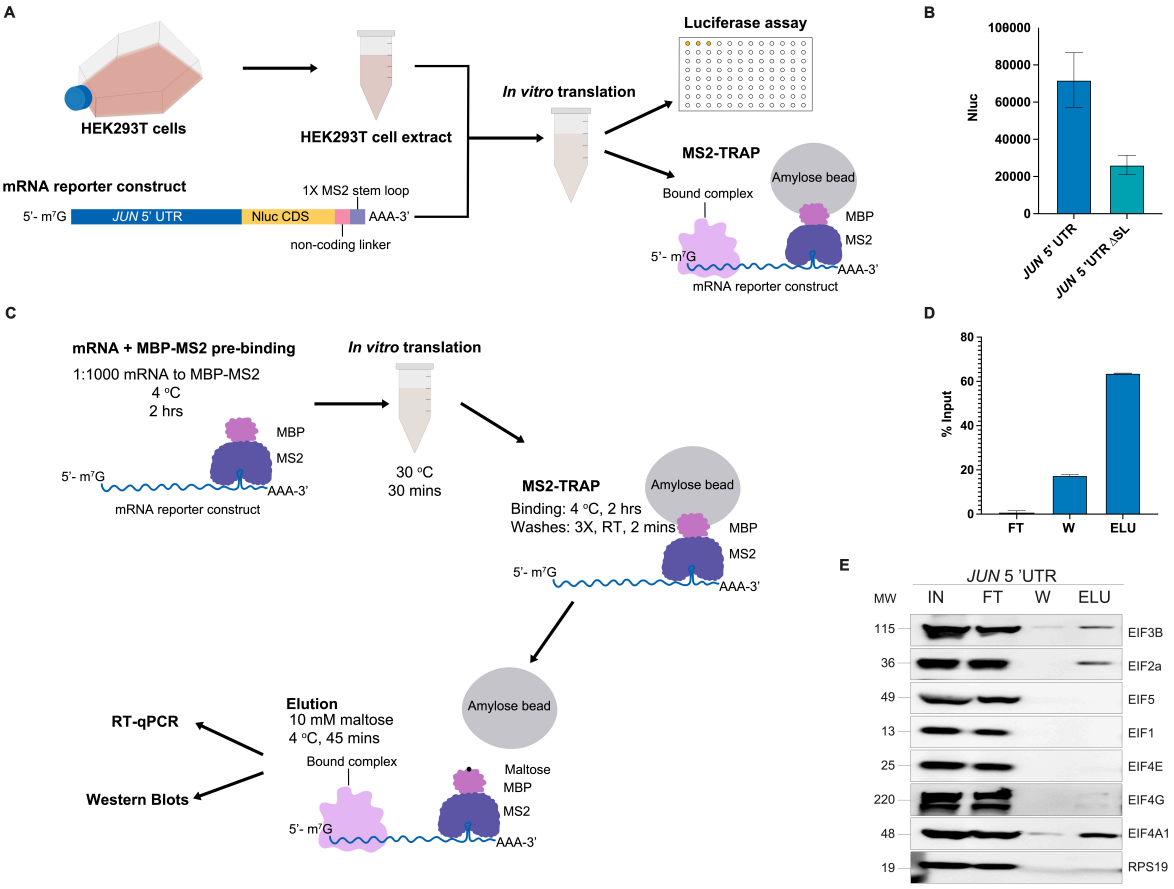


Figure 3.1. MS2-TRAP reveals components of the human *JUN* MFC
(A) Schematic of MS2-TRAP approach. (B) Representative luminescence levels for *in vitro* translation reactions using HEK293T cell extract and the *JUN* 5' UTR WT or ΔSL with the Nluc CDS, non-coding linker and MS2 stem loop construct. Reactions were incubated for 30 minutes at 30 °C, and Nanoluciferase activity was monitored using the Nano-Glo Luciferase Assay Kit (Promega) using a Spark multimode microplate reader (TECAN). Technical triplicates for each biological replicate, and a total of at least three biological replicates were taken for each measurement. The mean value of the replicates and standard error of the mean were plotted. (C) Schematic of optimized MS2-TRAP protocol. (D) Real time quantitative polymerase chain reaction (RT-qPCR) showing pulldown efficiency. FT: flow-through, W: wash, ELU: elution. Technical triplicates for each biological replicate were taken for each measurement. Values were normalized to those of the input sample and reported as a percentage of this (% Input). The mean value of the replicates and standard error of the mean were plotted. (E) Representative western blots from MS2 pulldown samples using the *JUN* 5' UTR WT *in vitro* translation as input. IN: input, FT: flow-through, W: wash, ELU: elution.

3.2.2 Validation of human *JUN* MFC using mass spectrometry

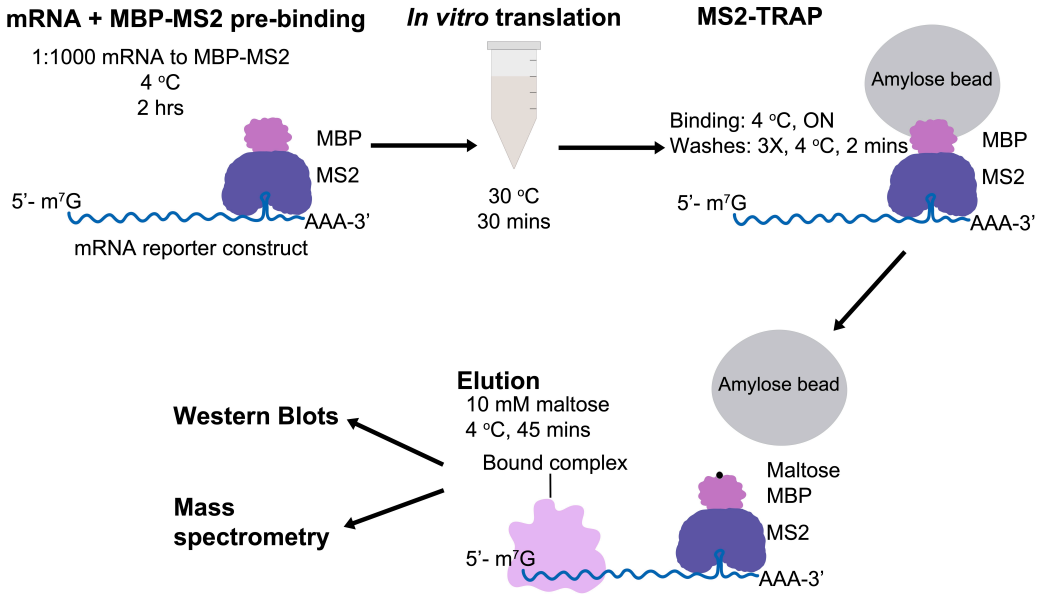
In order to validate components of the human *JUN* MFC, we sought to perform mass spectrometry, more specifically one-dimensional liquid chromatography coupled with mass spectrometry (LC-MS/MS), using elution samples from MS2-TRAP. With the goal of obtaining sufficient complex yield for mass spectrometry, we further optimized the MS2-TRAP protocol as shown in Figure 3.2A. One of the primary goals of the optimization is to increase binding of the MBP-MS2 captured RNA to the amylose beads. For this, we increased the binding of the IVT reaction to the beads from 2 hours to overnight, at 4 °C with rotation. In addition, in order to prevent loss of factors bound to *JUN* during the washes, these were done at 4 °C instead of at room temperature. Furthermore, we used different versions of the *JUN* 5' UTR reporter mRNA in our *in vitro* translation reactions. Firstly, we used the *JUN* 5' UTR WT and the *JUN* 5' UTR mutant ΔSL in order to evaluate if deleting the eIF3-interacting stem loop in *JUN* affects the resulting components of the MFC. Moreover, we used the *JUN* 5' UTR UUCG mutant, described in Chapter 2 (Mutant E, Figure 2.1), in which the loop in *JUN*'s SL is substituted by a highly stable UUCG tetraloop. Since this mutant resulted in a much higher level of translation both *in vivo* (Figure 2.1) and *in vitro* (Figure S3.4), our hypothesis was that using it would result in a higher yield of MFC in the elution. Lastly, we used the *JUN* 5' UTR WT mRNA with the addition of Rocaglamide A (RocA) to the *in vitro* translation reaction. As previously shown, RocA clamps eIF4A onto polypurine sequences on specific mRNAs (Iwasaki et al. 2016, 2019), one of which is *JUN*. As we previously established in Chapter 2, *JUN* is highly sensitive to RocA and possesses 11 of such polypurine sequences (Figure 2.2). As shown in Figure 3.1, eIF4A is one of the *JUN* MFC components that we have been able to isolate using MS2 pulldowns. Because of this, our hypothesis was that treating the *in vitro* translation reaction with RocA would lead to clamping of eIF4A onto the *JUN* mRNA and that, in consequence, any other initiation factors that may be interacting with eIF4A and *JUN* as part of the MFC would be captured more efficiently.

After performing *in vitro* translation reactions and MS2-TRAP using each of these constructs, we sought to analyze the MS2 pulldown samples using western blots. As shown in Figure 3.2B, *JUN* MFC components previously established (Figure 3.1E), such as eIF4G, eIF3, eIF2 and eIF4A1, were present in the elution samples for all of the *JUN* constructs. Also consistent with our previous pulldowns, eIF4E was absent in the elution samples of all constructs. Moreover, we were curious to see if other factors involved in translation initiation of mRNAs with long and highly structured 5' UTRs could be captured in our *JUN* MS2 pulldowns as well. For example, we probed for eukaryotic initiation factor 4B (eIF4B), a component of the eIF4F complex which has been shown to stimulate translation of long mRNAs with structured 5′ UTRs (Shahbazian et al. 2010; Sen et al. 2016). We also probed for Death Associated Protein 5 (DAP5), which has been found to associate with eIF2β and eIF4A1 and to selectively regulate translation of mRNAs with structured 5' UTRs and with upstream open reading frames (uORFs) (Liberman et al. 2015; Weber et al. 2022). Moreover we probed for DDX3, which is a DEAD-box RNA chaperone that facilitates translation initiation of mRNAs with high GC content and highly structured 5' UTRs and which has been thought to bind to the eIF4F complex and/or to eIF3 (Calviello et al. 2021; Mo et al. 2021). Interestingly, all three of these factors were

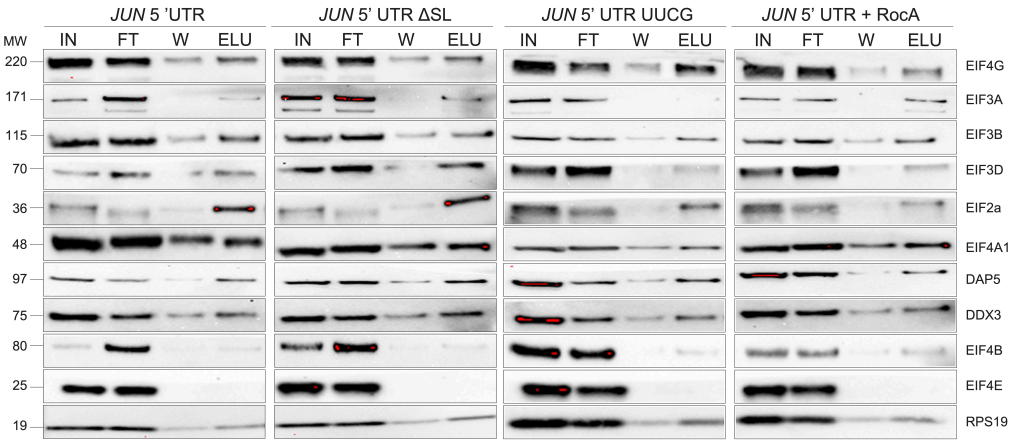
present in the elution samples of all *JUN* 5' UTR variants, albeit eIF4B was found in much lower abundance.

Elution samples for each of the *JUN* 5' UTR variants were then prepared and submitted for LC-MS/MS analysis (as described in Materials and Methods section 3.4.11). Results from this experiment were inconclusive, as the most prominent protein identified in all of the samples was the MBP-MS2 fusion protein, which we added to the *in vitro* translation reactions in order to perform MS2-TRAP. The high excess of MBP-MS2 in our samples prevented identification of a significant amount of any other peptides. Therefore, we were unable to validate components of the human *JUN* MFC using mass spectrometry and further sought to optimize sample preparation and pursue validation with additional techniques.

A



B



C

Sample	Most prominent protein	Sequence counts	Sequence coverage (%)
JUN 5' UTR WT	MBP-MS2	217	81.4
JUN 5' UTR ΔSL	MBP-MS2	230	85.1
JUN 5' UTR UUCG	MBP-MS2	201	86.9
JUN 5' UTR WT + RocA	MBP-MS2	176	85.9

Figure 3.2. Validation of human *JUN* MFC using mass spectrometry
(A) Schematic of human *JUN* MFC sample preparation for mass spectrometry. (B) Representative western blot from MS2 pulldowns using *in vitro* translation reactions of the *JUN* 5' UTR WT and Nluc CDS, *JUN* 5' UTR ΔSL and Nluc CDS, *JUN* 5' UTR UUCG

and Nluc CDS or *JUN* 5' UTR WT and Nluc CDS with Rocaglamide A (RocA) treatment. (C) Results from mass spectrometry (one-dimensional LC-MS/MS) analysis using MS2-pulldown samples from *in vitro* translation reactions shown in (B).

3.2.3 Validation of human *JUN* MFC using sucrose gradients

Sucrose gradients are a well-established technique for the separation of translation complexes (Mašek et al. 2011). In fact, the original MFC was identified as a pre-48S complex from yeast whole cell extracts that were fractionated using sucrose gradients (Asano et al. 2000). Because of this, we sought to use sucrose gradients for isolation of the human *JUN* MFC. Our intent is to validate the results of our MS2-TRAP and to directly compare our results with the MFC reported from yeast. For this, we performed *in vitro* translation reactions (IVTs) of the *JUN* 5' UTR reporter mRNA, layered these onto sucrose density gradients, collected fractions from these, and monitored the presence of initiation factors in these fractions using western blots (Figure 3.3A). For the IVTs, we used a previously established optimized system based on efficient translation using HEK293T pSB-HygB-GADD34-K3L translation-competent cell extract (Aleksashin et al. 2023). As shown in Figure 3.3B, this system renders much higher levels of translation, as measured by luciferase assays, than the previously employed system (labeled as HEK293T) (Rakotondrafara and Hentze 2011). Our hypothesis is that a highly efficient IVT system will yield higher levels of translation initiation complexes, which we can then identify and isolate by sucrose gradient fractionation. Translation of the *JUN* 5' UTR WT, Nluc CDS, non-coding linker, and MS2 stem loop construct was performed at increasing mRNA concentrations and monitored by luciferase assays. This mRNA titration revealed that 4 μg of *JUN* reporter mRNA yielded the highest level of translation, in comparison to lower concentrations (Figure 3.3C). Moreover, a course of incubation times for the IVTs revealed that translation in this system increases over time (Figure 3.3D). However, in order to capture translation initiation factors, we opted for a 30 minute incubation, since this showed to be the inflection point were translation is consistently increasing. This suggests that active engagement of ribosomes in translation is occurring at this timepoint, which in turn suggest the active formation of initiation complexes in order to engage and scan the mRNA. *In vitro* translation reactions of the *JUN* 5' UTR reporter mRNA performed under these conditions were therefore layered onto sucrose gradients and these were then centrifuged and fractionated. The presence of RNA in the fractions was monitored by measurement of absorbance at 254 nm (Abs254) and fractions were collected from the top to the bottom of the gradient. Even though Abs254 does not show an evident pre-48S peak, which would be indicative of a translation initiation multifactor complex (Figure 3.3E), fractions collected showed interesting results in this pre-48S region as shown by western blot analysis (Figure 3.3F). In fact, components of the MFC such as eIF3, eIF4G and eIF4A were observed in the pre-48S region, while eIF2 is absent. This is consistent with our previous results from MS2 pulldown samples, except for the absence of eIF2. It is possible that the sucrose gradient conditions disrupted the eIF2 interaction with the rest of the MFC, which would suggest that this interaction is weaker or more transient than that of eIF3, eIF4G and eIF4A as part of the endogenous human *JUN* MFC. Though further validation is needed, sucrose gradients served to confirm the presence of eIF3, eIF4A, and eIF4G as part of a pre-48S *JUN* MFC.

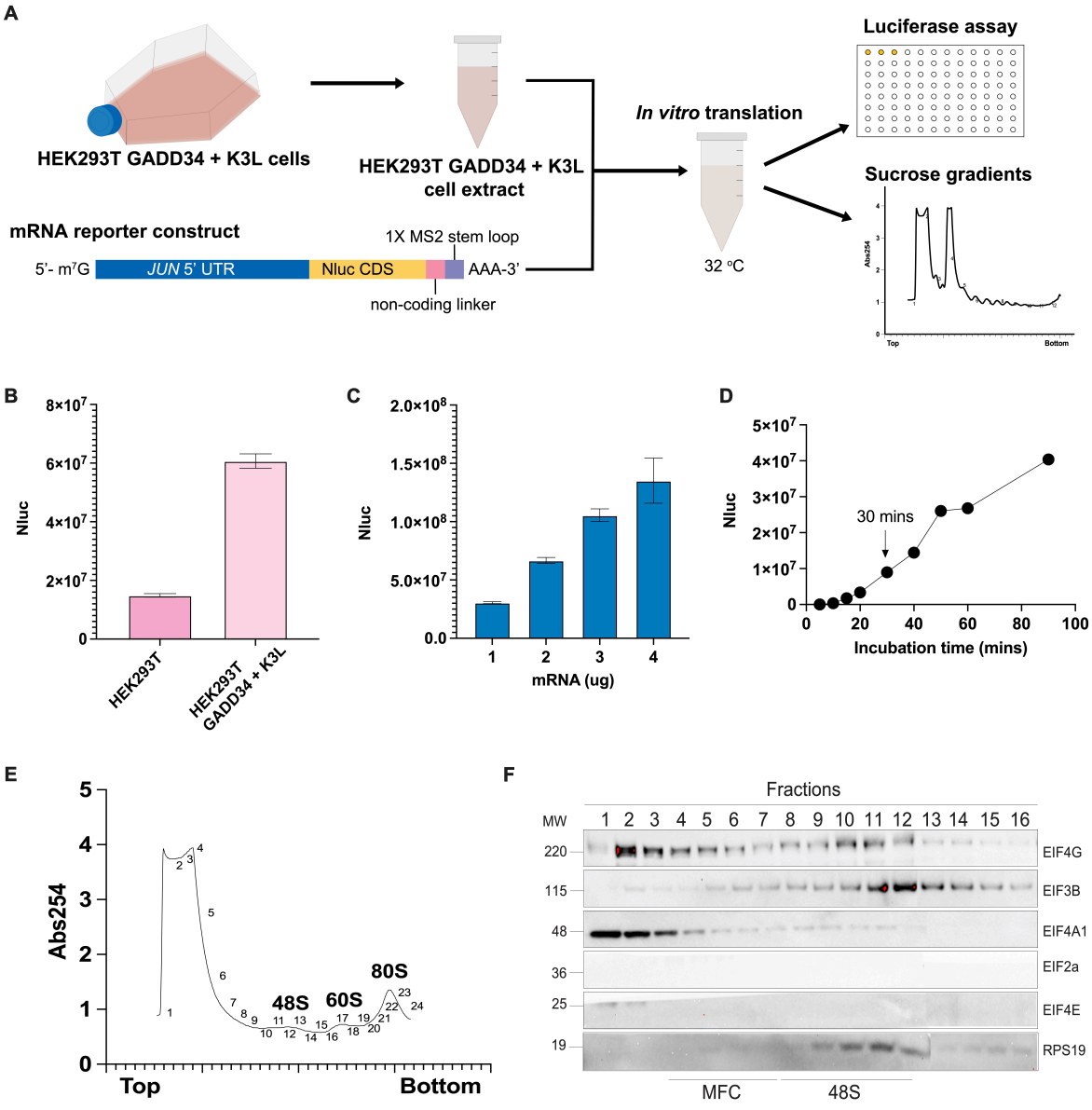


Figure 3.3. Some components of the human *JUN* MFC are validated by sucrose gradients (A) Schematic of sucrose gradient approach from *in vitro* translation reactions using the *JUN* 5' UTR WT and Nluc CDS reporter construct. (B) Optimization of *in vitro* translation conditions using the *JUN* 5' UTR WT and Nluc CDS reporter construct and HEK293T cell extract from either WT or GADD34 + K3L engineered cells. Reactions were incubated for 30 minutes at 32 °C, and Nanoluciferase activity was monitored as described in Figure 3.1. Technical triplicates for each biological replicate, and a total of at least two biological replicates were taken for each measurement. The mean value of the replicates and standard error of the mean were plotted. (C) mRNA titration for optimization of *in vitro* translation conditions using the *JUN* 5' UTR WT and Nluc CDS reporter

construct and HEK293T GADD34 + K3L cell extract. Reactions were incubated for 30 minutes at 32 °C, and Nanoluciferase activity was monitored as described in Figure 3.1. Technical triplicates for each biological replicate, and a total of at least two biological replicates were taken for each measurement. The mean value of the replicates and standard error of the mean were plotted. (D) Time course of incubation of *in vitro* translation reaction using the *JUN* 5' UTR WT and Nluc CDS reporter construct and HEK293T GADD34 + K3L cell extract. Reactions were incubated for 30 minutes at 32 °C, and Nanoluciferase activity was monitored as described in Figure 3.1. Technical triplicates for each biological replicate, and a total of at least two biological replicates were taken for each measurement. The mean value of the replicates and standard error of the mean were plotted. (E) Representative sucrose gradient from WT *JUN* 5'UTR and Nluc CDS *in vitro* translation reactions using optimized system. (F) Representative western blots from sucrose gradient fractions in (E).

3.2.4 Strategies for optimization of human *JUN* MFC sample for validation

Given the slight discrepancies between the human *JUN* MFC isolated from MS2-TRAP and that observed after sucrose gradients, further validation using a more robust method such as mass spectrometry is needed. As shown in section 3.2.2, previous attempts at mass spectrometry for this complex resulted in an overwhelming identification of the MBP-MS2 fusion protein used for MS2-TRAP, which prevented identification of significant amounts of any other protein. This could be both due to an excess of MBP-MS2 in the sample, as well as due to a low yield of the *JUN* MFC. Therefore, we sought to address both of these issues in order to prepare a more optimal *JUN* MFC sample for validation.

It has been previously reported that EIF3D-specialized translation is regulated through a phosphorylation switch which promotes cell survival during chronic glucose deprivation (Lamper et al. 2020). Therefore, upon glucose starvation translation of EIF3D-target mRNAs increases, which aids cell survival under those conditions. *JUN* has been established as an EIF3D-target mRNA, specifically by binding of EIF3D to *JUN*'s 5′ -7-methylguanosine cap structure (Lee et al. 2016). Because of this, we hypothesized that cells which were grown under glucose starvation conditions would yield higher levels of *JUN* translation and would result in a higher yield of the *JUN* MFC. In order to investigate this, we prepared translationally active cell extract from HEK293T cells that were grown in media without glucose and we used this extract for *in vitro* translation reactions of Nluc reporter mRNAs. These included the *JUN* 5' UTR WT, the *JUN* 5' UTR ΔSL, and the Hemoglobin Beta Subunit (*HBB*) 5' UTR as a negative control. As shown in Figure 3.4A, luciferase assays from these *in vitro* translation reactions revealed that cell extract from cells grown under glucose starvation conditions were generally less translationally active for all constructs than that of cells grown under standard conditions. Despite this, we performed MS2 pulldowns using these *in vitro* translation reactions for the *JUN* 5' UTR WT. As shown in Figure 3.4B, western blot analysis from these MS2 pulldown samples revealed a generally low yield of all MFC components in these samples. This suggests that glucose starvation is not an optimal strategy for increasing yields of *JUN* MFC *in vitro*. Given that the number of cells used for preparation of extract was normalized to be equal

in both conditions (+/- Gluc), it is possible that lower levels of translation in the glucose deprived condition are a consequence of the metabolic stress that this condition imposes on the cells. It is also possible that the mechanism that promotes *JUN* translation in cells under glucose starvation requires factors that are not functional *in vitro*.

In order to address the issue of an excess of MBP-MS2 in the sample used for mass spectrometry, we decided to evaluate if a lower ratio of mRNA to MBP-MS2 fusion protein could be used for the MS2 pulldowns. Since our MS2 pulldown protocol involves pre-binding of the MBP-MS2 fusion protein to the mRNA, altering this ratio could affect the levels of translation obtained. Therefore, we first evaluated whether the amount of MBP-MS2 in the *in vitro* translation reactions affects translation of the *JUN* 5' UTR reporter mRNA *in vitro*. As shown in Figure 3.4C, a significant drop in translation levels was observed with the 1:1000 mRNA to MBP-MS2 ratio previously used for our MS2 pulldowns. Because of this, we decided to move forward using the next highest ratio that didn't affect translation significantly, which was 1:100, for subsequent MS2 pulldowns. In addition to this, in order to decrease the amount of MBP-MS2 fusion protein in elution samples from MS2 pulldowns, we decided to optimize the elution step of the MS2 pulldowns. We previously used maltose to act as a competitor for binding to the maltose binding protein (MBP, as part of the MBP-MS2 fusion protein), therefore releasing it from binding to the amylose beads. This approach results in release of the MBP-MS2 protein bound to the MS2 stem loop on the target mRNA into the elution sample.

In order to prevent this release of MBP-MS2 into the elution, we decided to substitute the elution using maltose with an elution using Ribonuclease H (RNAse H). RNAse H is an endoribonuclease that hydrolyzes the phosphodiester bonds of RNA that is hybridized to DNA, which results in cleavage at the site of the RNA-DNA hybrid (Hyjek et al. 2019). We therefore designed DNA oligos that bind the 3' end of the Nluc CDS on the mRNA reporter construct containing the *JUN* 5' UTR with the Nluc CDS, a non-coding linker, and an MS2 stem loop. RNAse H cleavage at the site of these oligos results in a separation of the *JUN* 5' UTR and Nluc CDS portion of the mRNA from the MS2 stem loop portion which is bound to the MBP-MS2. As depicted in Figure 3.4D, in the absence of maltose, the MBP-MS2 fusion protein will remain bound to the amylose beads and will therefore be absent from the elution samples. As shown by Coomassie staining of SDS PAGE gel in Figure 3.4E, elution using RNAse H resulted in a significantly lower amount of MBP-MS2 fusion protein in the elution sample when compared with the elution using maltose. When evaluating the yield of components of the *JUN* MFC using western blots, elution using RNAse H was comparable to that using maltose (Figure 3.4F). However, our control condition using a mock RNAse H elution, which consists of all the RNAse H buffers but is depleted of the RNAse H enzyme, revealed that elution of MFC components was also achieved under these conditions. This points to an unspecific release of initiation factors under these conditions, which may not be specifically bound to the *JUN* mRNA. Therefore, it is difficult to determine whether elution of the *JUN* MFC is being achieved by RNAse H elution. Further optimization is needed in order to achieve precise elution of the human *JUN* MFC, in high yields and depleted of the MBP-MS2 fusion protein, for validation using mass spectrometry.

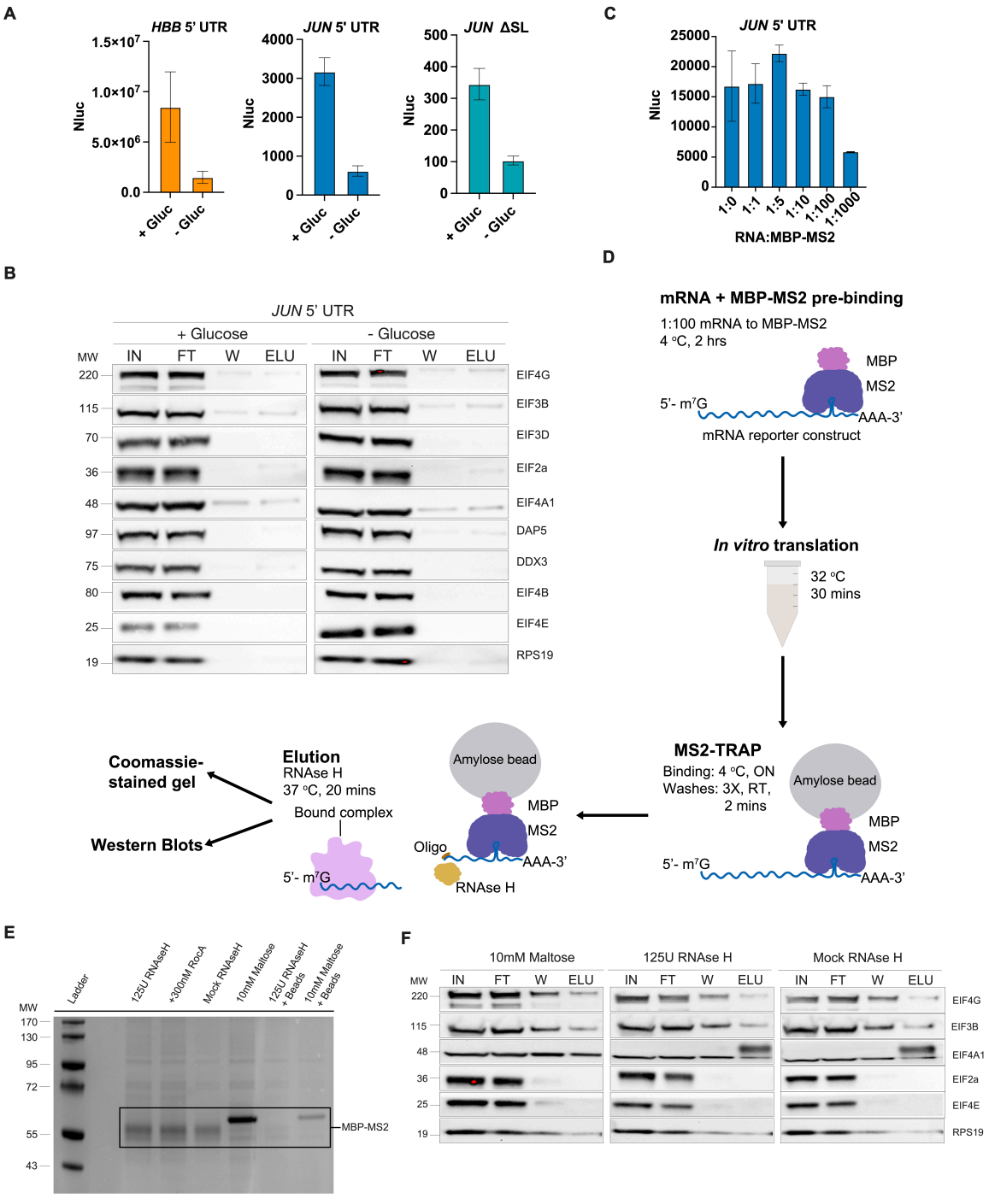


Figure 3.4. Strategies for optimization of human *JUN* MFC sample for validation
(A) *In vitro* translation reaction using Nluc reporter mRNAs (*JUN* 5' UTR WT, *JUN* 5' UTR ΔSL, *HBB* 5' UTR) with cell extract from glucose starved HEK293T cells. +Gluc: HEK293T cells grown under standard conditions, -Gluc: HEK293T cells grown under glucose

starvation conditions, *HBB*: Hemoglobin Beta Subunit. Reactions were incubated for 30 minutes at 32 °C, and Nanoluciferase activity was monitored as described in Figure 3.1. Technical triplicates for each biological replicate, and a total of at least three biological replicates were taken for each measurement. The mean value of the replicates and standard error of the mean were plotted. (B) Representative western blots of samples prepared using MS2 pulldowns from *in vitro* translation reactions using the *JUN* 5' UTR WT and Nluc CDS reporter construct and cell extract from glucose starved HEK293T cells. IN: input, FT: flow-through, W: wash, ELU: elution. (C) Luciferase assay from *in vitro* translation reactions of *JUN* 5'UTR WT and Nluc CDS reporter mRNA which was pre-incubated with varying concentrations of MBP-MS2 protein. Ratio reported is that of mRNA to MBP-MS2 protein. Reactions were incubated for 30 minutes at 32 °C, and Nanoluciferase activity was monitored as described in Figure 3.1. Technical triplicates for each biological replicate were taken for each measurement. The mean value of the replicates and standard error of the mean were plotted. (D) Schematic of sample preparation using MS2 pulldown with RNAse-H elution from *in vitro* translation reaction of *JUN* 5' UTR WT and Nluc CDS using HEK293T GADD34 + K3L cell extract. (E) Representative Coomassie-stained gel from samples prepared using *JUN* 5' UTR WT and Nluc CDS *in vitro* translation MS2 pulldown with RNAse-H elution. (F) Representative western blots from samples prepared using *JUN* 5'UTR WT and Nluc CDS *in vitro* translation MS2 pulldown with RNAse-H elution.

3.3 Discussion

Despite the fact that the multifactor complex (MFC) was identified as being composed of eIF1, eIF2, eIF3, eIF5 and the initiator methionyl-tRNA (Met-tRNA$_i$) in yeast over two decades ago (Asano et al. 2000), not much is known about this complex in higher eukaryotes. Though previous studies have been able to recapitulate the yeast MFC using purified human components and these interactions have been well characterized *in vitro* (Bandyopadhyay and Maitra 1999; Bieniossek et al. 2006; Sokabe et al. 2012), the endogenous human MFC is yet to be elucidated. In this chapter we established a protocol for the isolation of the human MFC. We used the *JUN* mRNA as a platform for MFC binding, since binding of one of the well-established yeast MFC components, eIF3, has been shown on the *JUN* 5' UTR and 5′ -7-methylguanosine cap structure in human cells (Lee et al. 2015, 2016). This mechanism of binding suggests an integral role for eIF3 in recruitment of the *JUN* mRNA for translation, which we hypothesize could be achieved in the context of the MFC. Given this, we applied MS2-tagged RNA affinity purification (MS2-TRAP) to isolate the human MFC bound to the *JUN* mRNA. In this approach, introduction of an MS2 RNA hairpin loop into a *JUN* 5' UTR reporter mRNA allowed isolation of this mRNA by binding of the MBP-MS2 fusion protein, which recognizes the MS2 RNA element, in *in vitro* translation reactions using human cell extract. Results from this RNA pulldown revealed reproducible isolation of a core human MFC composed of eIF3, eIF2, eIF4A and eIF4G (Figure 3.1). This finding represents the first instance of isolation of an endogenous mRNA-bound human MFC and interestingly suggests the presence of novel MFC components, such as eIF4A and eIF4G, in this context. Validation of this MFC was achieved using sucrose density gradients, with minimal discrepancies observed for eIF2 (Figure 3.3). Despite our best efforts, validation

using mass spectrometry was not achieved due to issues with sample yield and homogeneity (Figures 3.2 and 3.4). This demonstrated some of the challenges in isolation of translation complexes and motivated further optimization of these methods, some of which was shown in this chapter as well.

Although a variety of approaches have been developed for the isolation of ribonucleoprotein complexes (RNPs), we employed the MS2-TRAP approach for several reasons. Firstly, it allowed us to introduce minimal changes to the *JUN* 5' UTR reporter mRNA in order to capture it and immobilize it. Secondly, it allows us to capture the *JUN* mRNA precisely in the context of MFC formation during translation, by using *in vitro* translation reactions. It also allows us to prepare cell extracts with cells grown under standard conditions without introducing any significant changes. Moreover, it avoids using agents, such as crosslinkers, which may alter the interaction profile in *JUN*'s 5' UTR. In turn, all of this allows us to isolate the endogenous human MFC under conditions as close to physiological as possible while performing the experiment *in vitro*. Design of the MS2-tagged reporter mRNA and general approach were based on a protocol previously employed by our lab for MS2-tagged ribosome purification from *E. coli* crude ribosome extracts (Ward et al. 2019). Optimization of this method for our mammalian system was firstly achieved by titration of mRNA and of the salts of magnesium and potassium as suggested by a previously established *in vitro* translation protocol (Rakotondrafara and Hentze 2011). Method was further adapted at a smaller scale and binding and washes were adapted to the appropriate buffers, temperature, and incubation times for preservation of the integrity of translation and of the MFC (Asano et al. 2000; Rakotondrafara and Hentze 2011; Aleksashin et al. 2023).

Even though pulldown efficiency for the *JUN* 5' UTR reporter mRNA was achieved, as demonstrated by RT-qPCR (Figure 3.1), and initiation factors involved in the human MFC were captured, as shown by western blots (Figure 3.1), use of the MS2-TRAP approach had some limitations. Firstly, use of this approach at the scale required for our system rendered low yields of the MFC, which made it difficult to use for downstream applications. In addition, the traditional elution method for MS2-TRAP, which is release of the MS2 protein by competition with maltose, results in excess of the MS2 protein in the elution. This can cause issues with detection of the RNP complex of interest, which was in fact the case when we attempted to validate MFC components using mass spectrometry. Our mass spectrometry results showed a predominant detection of the MBP-MS2 fusion protein (Figure 3.2), which presumably hindered detection of lower-abundance proteins. In fact, a major caveat of the MS2-TRAP approach, as well as of other RNA-pulldown approaches, is their limitation in capturing low abundance complexes, as may be the case of the human MFC. To overcome this limitation, it is possible to increase the concentration of the components of the RNP of interest, for example of the *JUN* 5' UTR reporter mRNA, though this may cause issues with unspecific binding. Despite these caveats, the MS2-TRAP approach served as an informative experiment and provided insights into the formation of a *JUN* MFC using *in vitro* translation from human cell extract.

Another approach that we employed for isolation of a human *JUN* MFC was *in vitro* translation combined with sucrose gradient fractionation. This approach was selected because a similar approach was used for identification of the yeast MFC (Asano et al. 2000). In this previous study however, whole cell extract was layered on a sucrose

gradient instead of an *in vitro* translation reaction. The use of an *in vitro* translation reaction was chosen for our study because it allows us to selectively isolate a *JUN* MFC instead of a complex formed by binding to other endogenous transcripts. Western blot analysis from fractions collected after sucrose gradient centrifugation revealed a pre-48S complex composed of eIF3, eIF4A and eIF4G (Figure 3.3). This served in part as validation for the presence of these MFC components in our MS2-TRAP elution samples. However, the absence of eIF2 in comparison with these elution samples suggests that this method may be affecting complex composition either due to the stringency of the centrifugation or due to the potential different rates of association and dissociation of the MFC components. Further studies are needed in order to determine these rates since previous experiments characterized these interactions only *in vitro* using purified human MFC components and not in the context of human cell extract (Sokabe et al. 2012). It is important to note that an optimized *in vitro* translation system established in our lab (Aleksashin et al. 2023) was used for the translation reactions layered onto sucrose gradients. This was necessary because *in vitro* translation reactions with lower efficiency were not able to be detected by sucrose gradient fractionation (data not shown). This observation serves to demonstrate the need for a highly efficient translation system for isolation of the human MFC. As a whole, sucrose gradient experiments served to validate some *JUN* MFC components though more robust validation is still needed.

In addition to providing valuable insights about a human MFC composition, our study demonstrated some of the challenges that arise when isolating a translation initiation complex. Even though the role of the MFC is not fully understood, data from *in vivo* yeast experiments suggests that it is a translation initiation intermediate that somehow mediates recruitment of the Met-tRNA$_i$ to the 40S (Asano et al. 2000, 2001; Valášek et al. 2002, 2003). This suggests that the MFC is a highly dynamic complex, which would make it difficult to capture especially in a heterogeneous mixture such as that of an *in vitro* translation reaction. Since stability of the human MFC *in vivo* has not been established, it is difficult to determine whether some of its components may be dissociating under our experimental conditions. This could very well be the reason for the absence of eIF1, eIF5, and eIF2 in some of our isolated complexes. Moreover, even though the incubation time of the *in vitro* translation reactions was optimized to approximate reaction termination before significant elongation has occurred, it is difficult to determine whether we are actually capturing a translation initiation complex. Given the low abundance of 40S subunits in our MS2 pulldown elution, as evidenced by western blots of RPS19, it is reasonable to assume that we are in fact capturing an initiation complex, but further validation is also needed to confirm this. It is also possible that our *in vitro* approach may be causing unspecific binding of factors, as suggested by western blots performed after RNAse H elution. In general, validation of the human *JUN* MFC without proteomics is complicated. Therefore, further experiments are needed in order to optimize sample yield and purity for mass spectrometry. Further optimization of the MS2 pulldowns, of the binding for example, or of the sucrose gradients is feasible. However, it is worth considering using a cell-based approach which may be more robust and physiologically relevant. For example, we could express both the MS2-tagged *JUN* mRNA and the MBP-MS2 protein in human cells and allow the MFC to form in cells before isolating it with the MS2-TRAP approach (Yoon et al. 2012; Yoon and Gorospe 2016). If pursuing this approach, we could also use crosslinking to stabilize MFC interactions and

ensure capture of the complex. A cell-based approach may also allow to further investigate the possibility that the amount of MFC may differ depending on the cell type and the physiological state, as previously suggested (Sokabe et al. 2012). In addition to MFC isolation and validation, further experiments are needed in order to dissect the role and dynamics of formation of the human MFC. As a whole, our experiments pave the way for in-depth exploration of an mRNA-bound human MFC and provide evidence for the existence of such a complex by using the eIF3-target mRNA *JUN* as a platform.

3.4 Materials and Methods

3.4.1 Reporter plasmids

To generate the *JUN* 5′ UTR with Nanoluciferse (Nluc) CDS, linker and MS2 stem loop plasmids (WT, ΔSL, UUCG), the corresponding construct (described in Chapter 2) was amplified using primers for insertion of the linker and the MS2 stem loop (see Table 3.1 below). Insertion was achieved using overlap-extension PCR with Q5 High-Fidelity DNA Polymerase (NEB) and InFusion cloning (Takara Bio). All sequences were verified by Sanger sequencing.

Table 3.1: Primers used for cloning

Primer ID	**Sequence**
JUN 5' UTR + Nluc CDS + Linker + MS2 - Forward	GGCCGCGACTCTAGAGTCGGGGCG
JUN 5' UTR + Nluc CDS + Linker + MS2 - Reverse	TCTAGAGTCGCGGCCgatcaaaaACATGGGTGATCC TCATGTaaatgatcgttcttggggcacaggaactggTTACGCCA GAATGCGTTCGC

3.4.2 *In vitro* transcription

All RNA reporters were made by *in vitro* transcription with a standard T7 RNA polymerase protocol using DNA template gel extracted using the Zymoclean Gel DNA Recovery Kit (Zymo), 1x T7 RNA Polymerase buffer (NEB), 5 mM ATP (Thermo Fisher Scientific), 5 mM CTP (Thermo Fisher Scientific), 5 mM GTP (Thermo Fisher Scientific), 5mM UTP (Thermo Fisher Scientific), 5 µg BSA (NEB), 9 mM DTT, 25 mM $MgCl_2$, 200U T7 RNA polymerase (NEB), 50U Murine RNAse inhibitor (NEB) and incubating for 4 hours at 37 °C. The DNA template used for *in vitro* transcription was generated by PCR amplification from the corresponding reporter plasmid using the Q5 High-Fidelity DNA Polymerase (NEB) with a reaction including a forward primer containing the T7 promoter sequence and a 60T reverse primer for polyadenylation. Primers used for each transcript can be found in Table 3.2 below. After *in vitro* transcription, RNAs were treated with RQ1 DNAse (Promega) following the manufacturer's protocol and precipitated with 7.5 M lithium chloride. RNAs were then capped using Vaccinia D1/D2 (Capping enzyme) (NEB) and 2′ O-methylated using Vaccinia VP39 (2′ O Methyltransferase) (NEB) in a reaction that also included 1X capping buffer (NEB), 10 mM GTP (Thermo Fisher Scientific) and 4 mM SAM (NEB). RNAs were then purified with the RNA Clean and Concentrator-5 Kit (Zymo). In order to verify the integrity of the *in vitro* transcribed mRNAs, 6% polyacrylamide TBE-Urea denaturing gels were run using 1X TBE (Invitrogen), a ssRNA ladder (NEB) and SYBR safe stain.

Table 3.2: Primers used for DNA template preparation

Primer ID	Sequence
JUN 5' UTR - Transcript - Forward	taatacgactcactatagggctcagagttgcactgag
Nluc MS2 60T - Transcript - Reverse	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTTTTTTTTTTTACATGGGTGATCCT CATGTaaatgatc

3.4.3 HEK293T cells and extract preparation

HEK293T cells were maintained in DMEM (Gibco) supplemented with 10% FBS (VWR) and 1% Pen/Strep (Gibco). Cells were grown at 37 °C in 5% carbon dioxide and 100% humidity. For cells grown under glucose starvation conditions, 16 hours after previous passage, media was exchanged to DMEM without D-glucose and Sodium Pyruvate (Gibco), and supplemented with 10% FBS (VWR). Cells were left to recover for 36 to 48 hours under glucose starvation conditions, until they reached a minimum of 90% confluency and then they were used for cell extract preparation.
In vitro translation extracts were made from HEK293T cells using a previously described protocol (Rakotondrafara and Hentze 2011). Cells were scraped and collected by centrifugation for 2 minutes at 376 x g at 4°C. Cells were washed once with cold PBS (137 mM NaCl, 2.7 mM KCl, 100 mM Na_2HPO_4, 2 mM KH_2PO_4) then homogenized with an equal volume of freshly made cold hypotonic lysis buffer (10 mM HEPES-KOH pH 7.6, 10 mM KOAc, 0.5 mM Mg(OAc)2, 5 mM dithiothreitol (DTT), and 1 Complete EDTA-free Proteinase Inhibitor Cocktail tablet (Roche) per 10 ml of buffer). After hypotonic-induced swelling for 45 minutes on ice, cells were homogenized using a syringe attached to a 27G needle until 95% of cells burst as determined by trypan blue staining. Lysate was then centrifuged at 14,000 x g for 1 minute at 4°C. The resulting supernatant was moved to a new tube, avoiding the top lipid layer. Lysate aliquots were quickly frozen with liquid nitrogen and stored at 80°C.

3.4.4 HEK293T pSB-HygB-GADD34-K3L cells and extract preparation

HEK293T pSB-HygB-GADD34-K3L cells (Aleksashin et al. 2023) were maintained in DMEM media (Gibco) supplemented with 10% Tet-system approved FBS (Gibco) and 1% Pen/Strep (Gibco). Cells were grown at 37 °C in 5% carbon dioxide and 100% humidity. Cells were grown for extract preparation as follows. The day after plating cells from a frozen stock into a T25 flask (Cell Star), media was exchanged and supplemented with 200 µg/mL Hygromycin B (Invitrogen). The following day, cells were transferred to a T75 flask (Corning) with media supplemented with 200 µg/mL Hygromycin B. Once cells reached 100% confluency, half of the cells were transferred to a T175 flask (Falcon) with media supplemented with 200 µg/mL Hygromycin B. Once cells reached 100% confluency, cells were passaged onto 25 150 mm plates (Corning) at a 1 to 25 ratio. The next day, cells were treated overnight with 20 µg Doxycycline (Takara Bio) per plate.
In vitro translation extracts were made from HEK293T pSB-HygB-GADD34-K3L cells using a previously described protocol (Aleksashin et al. 2023). Cells were placed on ice,

scraped and collected by centrifugation at 1000 xg for 5 minutes at 4 °C. Cells were washed once with ice-cold DPBS (Gibco) and collected once again by centrifugation at 1000 xg for 5 minutes at 4°C. After this, cells were homogenized with an equal volume of freshly made ice-cold hypotonic lysis buffer (10 mM HEPES-KOH pH 7.6, 10 mM KOAc, 0.5 mM $Mg(OAc)_2$, 5 mM dithiothreitol). After hypotonic-induced swelling for 45 minutes on ice, cells were homogenized using a syringe attached to a 26G needle (BD). Extract was then centrifuged at 15000 xg for 1 minute at 4 °C. The resulting supernatant was aliquoted, frozen with liquid nitrogen, and stored at -80 °C.

3.4.5 *In vitro* translation

In vitro translation reactions prepared for luciferase assays were performed using HEK293T translation-competent cell lysate, as previously described, with modifications (Rakotondrafara and Hentze 2011). Translation reactions contained 50% HEK293T translation-competent cell extract, 2 mM ATP, 0.42 mM GTP, 7 mM tris(2-carboxyethyl) phosphine, 28 mM HEPES pH 7.5, 2 mM creatine phosphate (Roche), 0.01 µg/µl creatine kinase (Roche), 2 mM $Mg(OAc)_2$, 60 mM KOAc, 10 µM amino acids (Promega), 0.21 mM spermidine, 0.6 mM putrescine and 0.8 U/µl murine RNase inhibitor (NEB). Translation reactions were incubated at 30°C for 30 min and Nanoluciferase activity was monitored using the Nano-Glo Luciferase Assay Kit (Promega) using a Spark multimode microplate reader (TECAN). Technical triplicate measurements were taken for each biological replicate.
Optimized *in vitro* translation reactions were performed using HEK293T pSB-HygB-GADD34-K3L translation-competent cell extract, as previously described (Aleksashin et al. 2023). Translation reactions contained 50% translation-competent cell extract, 52 mM HEPES pH 7.4 (Takara), 35 mM potassium glutamate (Sigma), 1.75 mM $Mg(OAc)_2$ (Invitrogen), 0.55 mM spermidine (Sigma), 1.5% Glycerol (Fisher Scientific), 0.7 mM putrescine (Sigma), 5 mM DTT (Thermo Scientific), 1.25 mM ATP (Thermo Fisher Scientific), 0.12 mM GTP (Thermo Fisher Scientific), 10 mM L-Arg; 6.7 mM each of L-Gln, L-Ile, L-Leu, L-Lys, L-Thr, L-Val; 3.3 mM each of L-Ala, L-Asp, L-Asn, L-Glu, Gly, L-His, L-Phe, L-Pro, L-Ser, L-Tyr; 1.7 mM each of L-Cys, L-Met; 0.8 mM L-Trp, 20 mM creatine phosphate (Roche), 60 µg/mL creatine kinase (Roche), 4.65 µg/mL myokinase (Sigma), 0.48 µg/mL nucleoside-diphosphate kinase (Sigma), 0.3 U/mL inorganic pyrophosphatase (Thermo Fisher Scientific), 100 µg/mL total calf tRNA (Sigma), 0.8 U/µL RiboLock RNase inhibitor (Thermo Scientific), and 1000 ng of the corresponding mRNA. Reactions were then incubated for 60 minutes at 32 °C and Nanoluciferase activity was monitored using the Nano-Glo Luciferase Assay Kit (Promega) using a Spark multimode microplate reader (TECAN). Technical triplicate measurements were taken for each biological replicate.

3.4.6 MS2-TRAP

MBP-MS2 fusion protein was expressed and purified as previously described (Ward et al. 2019). Optimized MS2-TRAP protocol begins with pre-incubation of the MBP-MS2 fusion

protein with the corresponding mRNA reporter construct containing the MS2 stem loop (ACATGAGGATCACCCATGT) in a 1:100 ratio of mRNA to MBP-MS2 for 2 hours at 4 °C. The amount of mRNA used for optimized MS2 pulldowns was determined by titration into IVT reactions and was 25 ng/μl. MS2 pulldowns were performed using IVT reactions at different scales including 140 μl, 280 μl, 400 μl and 700 μl. HEK293T cell extract was prepared for *in vitro* translation reactions by treatment with 0.015 U/μl Micrococcal Nuclease (MNase) (NEB) and 0.75 mM $CaCl_2$ for 15 minutes at room temperature. MNase reactions were stopped by addition of 3 mM EGTA and incubation at 4 °C until addition into the *in vitro* translation reaction. *In vitro* translation reactions prepared for MS2 pulldowns contained 50% Micrococcal Nuclease-treated HEK293T translation-competent cell extract, 28 mM HEPES pH 7.5, 2 mM $Mg(OAc)_2$, 60 mM KOAc, 22 mM tris(2-carboxyethyl) phosphine, 0.2 mM spermidine, 0.6 mM putrescine, 25 ng/μl mRNA pre-bound to 250 ng/μl MBP-MS2, 0.8 U/μl murine RNase inhibitor (NEB), 2 mM ATP, 2 mM GTP, 2 mM DL-Methionine (Sigma). Translation reactions were mixed thoroughly by pipetting and incubated at 30°C for 30 minutes. Amylose resin (NEB) was prepared for pulldowns by washing with MS2-150 buffer (20 mM HEPES pH 7.5, 150 mM KCl, 1 mM EDTA, 2 mM 2-mercaptoethanol). The amount of amylose resin used was determined by the volume of the *in vitro* translation sample and it should be a 3:1 ratio of packed volume of beads to volume of IVT. In order to prepare the *in vitro* translation reaction for binding to the beads, the volume of the reaction was made equal to that of the beads by addition of IVT buffer (4 mM HEPES pH 7.5, 30 mM KOAc, 0.5 mM $Mg(OAc)_2$, 0.02 mM spermidine, 0.12 mM putrescine, 0.2 mM tris(2-carboxyethyl) phosphine or TCEP). *In vitro* translation reaction, IVT buffer and amylose beads were combined (as Input) and incubated on a rotation platform overnight (approximately 16 hours) at 4 °C. The next morning, binding reaction was centrifuged at 2000 xg for 2 minutes at 4 °C and supernatant (Flow Through) was saved. Amylose beads were washed 3 times with double the volume of IVT buffer (relative to the volume of beads) with an addition of 0.1% Triton X with 2 minutes incubation at room temperature. Centrifugation at 2000 xg for 2 minutes at 4 °C was performed between washes and supernatant was saved (Wash 1, 2 or 3). Elution was achieved by addition of half the volume (relative to the volume of beads) of IVT buffer with 10 mM maltose and incubation for 45 minutes at 4 °C on a rotation platform. Elution sample was obtained as the supernatant after centrifugation at 2000 xg for 2 minutes at 4 °C.

For MS2 pulldowns using optimized IVT system, pulldowns were set up as described above, except that IVT reactions were assembled using HEK293T pSB-HygB-GADD34-K3L translation-competent cell extract as described in section 3.4.5. Extract for these reactions was not treated with MNAse.

3.4.7 RNA isolation and RT-qPCR

Total RNA was isolated from MS2 pulldown samples (I = Input, FT = Flow Through, W = Wash, ELU = Elution) using TRIzol Reagent (Invitrogen) following the manufacturer's protocol. Samples were equalized to the same volume (for example 400 μl) using IVT buffer and TRIzol was added at a 1:1 ratio. RT-qPCR analysis was performed using the Power SYBR Green RNA-to-Ct 1-Step kit (Applied Biosystems) according to the

manufacturer's instructions, and the Bio-Rad CFX96 Touch Real-Time PCR qPCR system (Bio-Rad). *JUN* reporter mRNA was quantified using specific primers (shown in Table 3.3) at a concentration of 100 nM each and using an equal volume of RNA per reaction in a 20 μl reaction. Quantification was done in three biological replicates, with each biological replicate having three technical replicates.

Table 3.3: Primers used for RT-qPCR

Primer ID	**Sequence**
JUN 5' UTR - RT-qPCR - Forward	gctcagagttgcactgagtgtg
JUN 5' UTR - RT-qPCR - Reverse	agaacagtccgtcacttcacg

3.4.8 Western Blot analysis

SDS PAGE gels were run using MS2 pulldown samples (I = Input, FT = Flow Through, W = Wash, ELU = Elution). Samples were loaded with NuPAGE loading dye (Invitrogen) onto NuPAGE 4-12% gels (Invitrogen) using 1X NuPAGE SDS Running Buffer (Invitrogen) at 110 V for approximately 2 hours. Proteins were then transferred from the gels to polyvinylidene difluoride (PVDF) membranes (Bio-Rad) using a wet-transfer apparatus (Bio-Rad) with transfer buffer (25 mM tris base, 190 mM glycine, 20% methanol, pH 8.3) at 80 V for 2 hours and 20 minutes at 4 °C. Membranes were then cut and blocked using 5% milk in TBST (20 mM tris base, 150 mM NaCl, 0.1% Tween 20, pH 7.6). Membranes were subsequently incubated with the appropriate dilutions of primary antibodies in 5% milk in TBST, overnight at 4 °C. Antibodies used are listed in Table 3.4 below. The next morning, membranes were washed with 4 washes of TBST, the first one for 15 minutes, and the subsequent 3 washes for 5 minutes each. Membranes were then incubated with their corresponding secondary antibody at the corresponding dilution in 5% milk in TBST for 1 hour at room temperature. Membranes were then washed 3 times with TBST for 5 minutes each. SuperSignal West Pico Chemiluminescent Substrate (Thermo Fisher Scientific) developing reagents were added to the membranes and these were imaged using an iBright FL1500 Imaging System (Thermo Fisher Scientific).

Table 3.4: Antibodies used for western blots

Antibody name	**Company**	**Catalog number**	**Dilution**
anti-eIF3B/EIF3S9	Bethyl	A301-761A	1:1000
anti-eIF2α (FL-315)	Santa Cruz Biotechnology	SC-11-386	1:1000
anti-eIF5	Bethyl	A301-771A	1:1000
anti-eIF1 (D7G3L)	Cell Signaling	12496S	1:1000
anti-eIF4E	BD	610269	1:1000
anti-eIF4G (A-10)	Santa Cruz Biotechnology	sc-13315	1:1000
anti-eIF4A1	Cell Signaling	2490	1:1000
anti-RPS19	Bethyl	A304-002A	1:2000

anti-eIF3A	Santa Cruz Biotechnology	sc-365789	1:1000
anti-eIF3D	Bethyl	A301-758A	1:1000
anti-DAP5	Santa Cruz Biotechnology	sc137011	1:1000
anti-DDX3	Bethyl	A300-474A	1:1000
anti-eIF4B	Bethyl	A301-767A	1:2000
ECL Anti-Rabbit IgG, Horseradish Peroxidase linked whole antibody	GE Healthcare	NA934VS	1:10000
Goat Anti-Mouse IgG-HRP	Santa Cruz Biotechnology	sc-2055	1:10000

3.4.9 Sucrose gradients

In vitro translation reactions were performed using the optimized *in vitro* translation system (Aleksashin et al. 2023), described in section 3.4.5, with the following modifications. Reactions were 200 μl in volume, contained 4 μg (100 ng/μl) *JUN* reporter mRNA and were incubated for 30 minutes at 32 °C. After incubation, reactions were layered onto a 13 ml 5%–30% sucrose gradient, made with gradient buffer consisting of: 5% sucrose (w/v) or 30% sucrose (w/v), 4 mM HEPES pH 7.5, 30 mM KOAc, 0.5 mM Mg $(OAc)_2$, 0.22 mM Spermidine, 0.72 mM Putrescine and 1.2 mM TCEP. The gradient was centrifuged at 38,000 rpm for 5 hours at 4 °C in a SW-41 rotor. The gradient was then fractionated using the Brandel gradient fractionator and ISCO UA-6 UV detector with a sensitivity of 0.1 and 500 μl fractions were collected through the entirety of the gradient. From each of the fractions, 10 μl was used for western blot analysis, which was performed as described in section 3.4.8.

3.4.10 MS2-TRAP using RNAse-H elution

MS2-TRAP of the *JUN* reporter mRNAs was performed as described in section 3.4.6, with the following modifications. *In vitro* translation reactions were performed using the optimized *in vitro* translation system described in section 3.4.5, but with a 30 minute incubation at 32 °C. Elution from amylose beads was achieved using 125 U of RNAse H (NEB), which was added to the amylose beads post-washes, together with 1X RNAse H reaction buffer (NEB), 3 mM $MgCl_2$, 10 mM DTT, 400 U SUPERase•In RNase Inhibitor (Thermo Fisher Scientific) and 1 μM of each of the oligos designed to bind the *JUN* reporter mRNA (see Table 3.4). Beads were incubated with the described RNAse H solution for 20 minutes at 37 °C on a rotation platform. Reaction was then centrifuged at 2000 xg for 2 minutes at 4 °C. Supernatant was collected and used for western blot analysis (as described in section 3.4.8) and for Coomassie stained gel analysis. For Coomassie stained gel analysis, samples were loaded with NuPAGE loading dye (Invitrogen) onto NuPAGE 4-12% gels (Invitrogen) using 1X NuPAGE SDS Running Buffer (Invitrogen) at 110 V for approximately 2 hours. Gels were then stained with SimplyBlue SafeStain (Thermo Fisher Scientific) following the manufacturer's protocol

and subsequently imaged using an iBright FL1500 Imaging System (Thermo Fisher Scientific).

Table 3.5: Oligos used for RNAse-H elution

Primer ID	**Sequence**
RNAse H - Elution - oligo 1	TTACGCCAGAATGCG
RNAse H - Elution - oligo 2	TTCGCACAGCCGCCA
RNAse H - Elution - oligo 3	GCCGGTCACTCCGTT
RNAse H - Elution - oligo 4	ATGGTTACTCGGAAC

3.4.11 Mass spectrometry sample preparation

MS2-TRAP was performed as described in section 3.4.6 using the *JUN* reporter mRNAs (WT, ΔSL, UUCG, or WT with 0.3 μM RocA treatment). Elution samples from MS2 pulldowns were used both for western blot analysis (described in section 3.4.8) to confirm the presence of *JUN* MFC components, and for SDS PAGE gel loading for mass spectrometry sample preparation. RNA in the elution samples was quantified using the NanoDrop One/One (Thermo Fisher Scientific) to confirm the presence of sufficient amount of complex for the mass spectrometry analysis (between 10-200 ng of sample in 27 μl is recommended). Samples were then concentrated to ~50 μl using a 30K 0.5 ml concentrator (Millipore Sigma) and IVT buffer (4 mM HEPES pH 7.5, 30 mM KOAc, 0.5 mM Mg $(OAc)_2$, 0.02 mM spermidine, 0.12 mM putrescine, 0.2 mM TCEP) with 10 mM maltose was exchanged to IVT buffer without maltose. Concentrated elution samples (27 μl of each) were loaded with NuPAGE loading dye (Invitrogen) onto NuPAGE 4-12% gels (Invitrogen) using 1X NuPAGE SDS Running Buffer (Invitrogen) at 100 V until samples were around 1-2 centimeters into the resolving gel section. Lanes for each sample were excised from the gel and submitted for identification using one-dimensional LC-MS/MS at the Vincent J. Coates Proteomics/Mass Spectrometry Laboratory at the University of California Berkeley.

3.5 Supplemental Figures

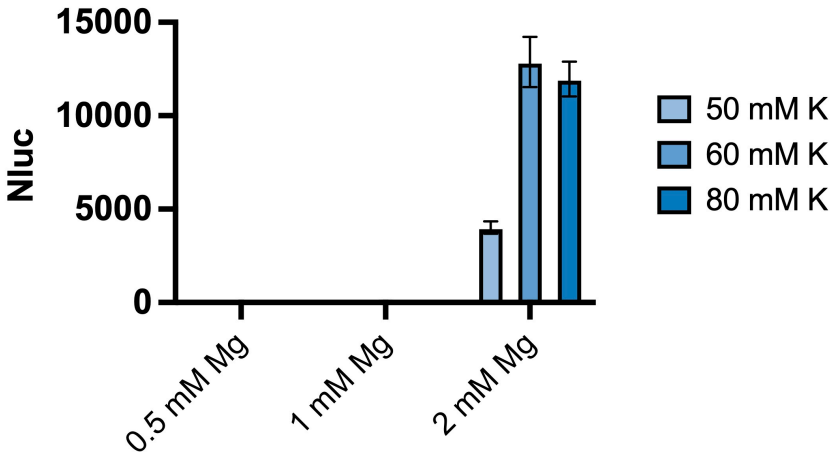


Figure S3.1. *In vitro* translation reactions with potassium and magnesium titration
Luminescence measured from *in vitro* translation reactions using HEK293T cell lysate and 240 ng of the *JUN* 5' UTR and Nluc CDS reporter mRNA, with increasing concentrations of potassium (K) or magnesium (Mg). Reactions were incubated for 30 minutes at 30 °C, and Nanoluciferase activity was monitored as described in Figure 3.1. Technical triplicates for each biological replicate, and a total of at least three biological replicates were taken for each measurement.

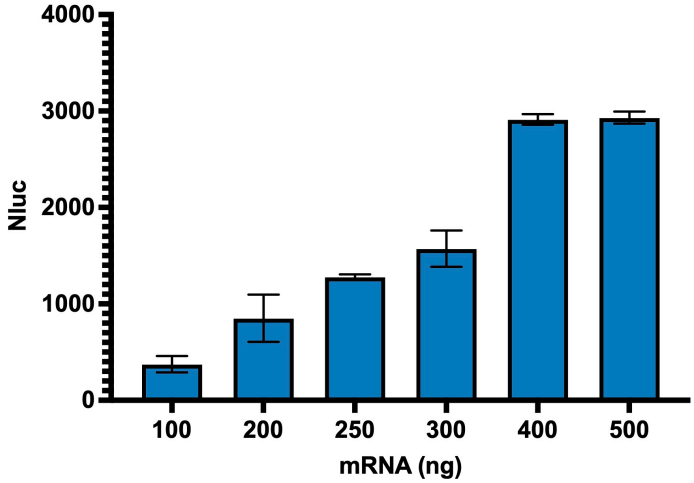


Figure S3.2. *In vitro* translation reactions with mRNA titration
Luminescence measured from *in vitro* translation reactions using HEK293T cell lysate and increasing concentrations of the *JUN* 5' UTR and Nluc CDS reporter mRNA. Reactions were incubated for 30 minutes at 30 °C, and Nanoluciferase activity was monitored as described in Figure 3.1. Technical triplicates were taken for each measurement.

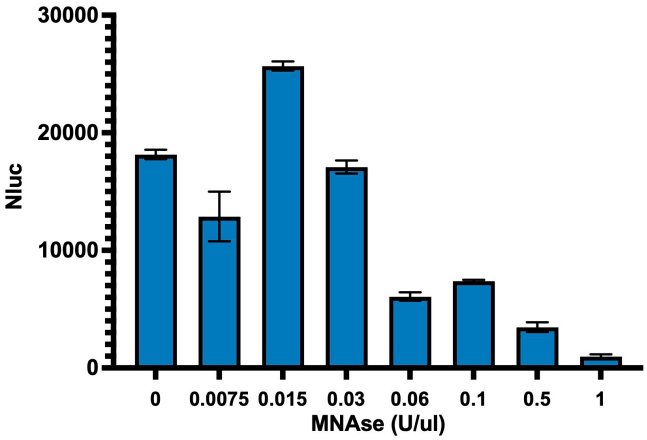


Figure S3.3. *In vitro* translation reactions with MNAse titration
Luminescence measured from *in vitro* translation reactions using HEK293T cell lysate and the *JUN* 5' UTR and Nluc CDS reporter mRNA, with increasing concentrations of micrococcal nuclease (MNAse). Reactions were incubated for 30 minutes at 30 °C, and Nanoluciferase activity was monitored as described in Figure 3.1. Technical triplicates for each biological replicate, and a total of at least two biological replicates were taken for each measurement.

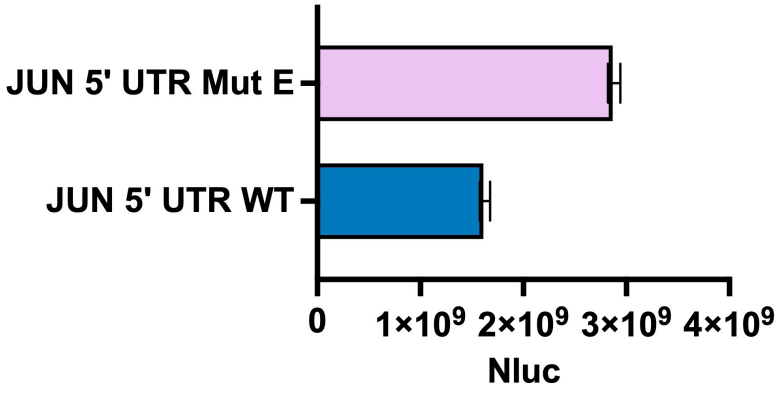


Figure S3.4. *In vitro* translation reactions with mutant E
Luminescence measured from *in vitro* translation reactions using HEK293T GADD34 + K3L cell lysate and either the *JUN* 5' UTR WT and Nluc CDS or the *JUN* 5' UTR mutant E and Nluc CDS reporter mRNAs. Mut E: insertion of UUCG tetraloop in place of loop in SL region. Reactions were incubated for 30 minutes at 32 °C, and Nanoluciferase activity was monitored as described in Figure 3.1. Technical triplicates were taken for each measurement.

Chapter 4: Conclusions and Future Directions

4.1 *JUN* mRNA: a model transcript for understanding complex translation regulation

A plethora of studies have tied JUN expression to disease, particularly cancer (Gee et al. 2000; Wulf et al. 2001; Briggs et al. 2002; Vasilevskaya and O'Dwyer 2003; Nateri et al. 2005; Hui et al. 2007; Blau et al. 2012; Chen and Bourguignon 2014; Suphakhong et al. 2022). Though much work has been done to understand *JUN* transcription regulation and how it relates to disease (Angel et al. 1988; Nakamura et al. 1991), our current understanding of its translation regulation has predominantly resulted from work on eIF3 (Lee et al. 2015, 2016; Lamper et al. 2020). However, the complexity of the *JUN* 5' UTR suggests that many more factors are involved in its translation regulation. Work presented in this dissertation aimed at expanding our understanding of the *JUN* mRNA in the context of translation and at uncovering the different layers of regulation mediated by its 5' UTR. To this end, we have presented experiments both *in vitro* and in human cells exploring the different contributors in this regulation, including RNA structural elements and initiation factors. We have found that additional structured regions near the eIF3-binding stem loop in the *JUN* 5' UTR contribute to *JUN* translation regulation. Interestingly, we have also uncovered a potential role for the initiation factor eIF4A in *JUN* translation regulation, both due to the sensitivity of the *JUN* transcript to the eIF4A-targeting compound Rocaglamide A (RocA) and due to the presence of eIF4A in a *JUN*-bound translation initiation multifactor complex (MFC). Moreover, we have revealed the contribution of two start codons in *JUN* translation and the conservation of this region amongst vertebrates. These findings point to an important evolutionary role of this portion of the *JUN* mRNA in translation regulation of this transcript. This also suggests a potential regulatory role for eIF1 and eIF5 in *JUN* start codon selection (Hann et al. 1992; Fletcher et al. 1999; Sonenberg and Dever 2003; Ivanov et al. 2008, 2010, 2022). We have also established a protocol for *in vitro* isolation of a *JUN* translation initiation MFC from human cell extract. Our findings in this aspect not only suggest novel components for the human MFC, but also establish the *JUN* mRNA as a useful platform for isolation of this complex. As a whole, this work demonstrates that the complexity of *JUN* translation regulation is much larger than previously appreciated and shows the importance of expanding our understanding on the translation regulation of transcripts with such complexity.

Along the course of this work, we found additional interesting facts regarding the *JUN* mRNA. For instance, we observed certain discrepancies when comparing the *JUN* cDNA sequence that our lab amplified from HEK293T cells with annotated sequences. We therefore evaluated these discrepancies in order to determine whether these were reported variants such as single nucleotide polymorphisms (SNPs). After an extensive search in a variety of databases (Table A.2), all but one of these discrepancies were found to be reported variants. These could represent cell-specific *JUN* 5' UTR sequence variations, which may or may not be relevant for *JUN* regulation in this cell line. This is yet to be investigated. This observation demonstrates the need for proper and consistent annotation of sequence variants reported in the literature, especially for mRNAs.

Our work also raised additional questions about translation regulation of the *JUN* mRNA in human cells. For example, it is still unknown whether initiation factors, including eIF3, bind to the *JUN* 5' UTR structured elements that we found to be relevant for *JUN* translation. Given that our mutagenesis analysis only focused on a short region of the *JUN* 5' UTR, there are potentially many more sequence and structural elements in the entirety of this region that are important for regulation. This is especially plausible given the length and high GC content of this region. Studies focused on determining the secondary structure of the entirety of the *JUN* 5' UTR would be beneficial in order to target structured regions in future mutagenesis analyses. In regards to *JUN*'s sensitivity to RocA further studies are needed in order to dissect all of the motifs to which RocA is clamping eIF4A in *JUN*. Mechanistic studies would also be useful for understanding the precise role of eIF4A in *JUN* translation regulation. Similarly, though we pioneered evidence for the use of both *JUN* start codons for translation initiation, we have yet to dissect the full mechanism for start codon selection in *JUN*. Additional studies in cells would be helpful for exploring use of both start codons in the endogenous transcript. Biochemical assays are also needed to determine levels of eIF1 and eIF5 in our cellular conditions, to then be able to correlate these with our findings on start codon usage. In addition, proteomics analysis would be valuable to determine whether a JUN peptide can in fact initiate at the upstream start codon. Though there may be technical challenges in studying the *JUN* transcript in cells, expanding our investigation to this context would allow us to correlate our findings with the physiological roles of JUN. This would be impactful for the understanding of JUN expression in a disease background.

4.2 One step closer to the isolation of a human multifactor complex (MFC)

Isolation of a human *JUN* MFC was a particularly exciting aspect of this work. By adapting an MS2-TRAP approach for binding of a *JUN* 5' UTR reporter mRNA from an *in vitro* translation reaction, we were able to isolate a human *JUN* translation initiation complex composed of eIF3, eIF2, eIF4A and eIF4G. Sucrose gradients from *JUN in vitro* translation reactions validated the presence of eIF3, eIF4A and eIF4G in this complex. Though confirmation of the isolated MFC components using proteomics was unsuccessful, our results provide strong evidence for the existence of a human MFC and pave the way for future validation of this complex. These findings also support the idea of a substantial cooperation of translation initiation factors in *JUN* translation regulation. In turn, this points to a potential mode of specialized translation for this transcript.

Work presented here also demonstrates the many challenges of isolating a translation initiation complex. For example, as expected for a translation intermediate, the human MFC seems to be a transient and low abundance complex. Therefore, capturing a substantial amount of this complex at the precise timeframe of formation would require robust enhancement and stabilization of this complex, which was difficult to achieve under our experimental conditions. Given this, it would be beneficial to explore alternative methods for human MFC isolation. For example, it is feasible to consider the use of a cell-based approach, for example MS2-TRAP, expressing the MS2 pulldown components directly in cells (Yoon et al. 2012; Yoon and Gorospe 2016). This would allow for formation of the MFC under physiological conditions, although it is unknown whether it will allow capture of such a transient and low abundance complex. One advantage of this method

is that it also allows for stabilization of the complex, for example with the use of UV crosslinking. Validation of the isolated human *JUN* MFC with proteomics is also absolutely crucial, especially given the intricate nature of the *JUN* 5' UTR which may interact with numerous initiation factors both specifically and non-specifically. Optimization of the MS2-TRAP protocol either *in vivo* or *in vitro* is also reasonable, for example by additional fine-tuning of transcript concentration to promote binding with eIF3 (Lee et al. 2015) or by increasing the amount of MS2 stem loops at the 3' of the transcript (Yoon et al. 2012; Yoon and Gorospe 2016). Once successful purity and high yield of the human *JUN* MFC is achieved, structural studies would be very informative for dissecting the interactions between the MFC components and the *JUN* 5' UTR. Initial cryo-EM trials were also attempted during the course of this work (data not shown), but they were unsuccessful due to the low MFC yield and heterogeneity of our sample. Other strategies that may help to tackle the issue of sample heterogeneity are additional steps of affinity purification, such as size exclusion chromatography, though these would in turn require substantial sample yield.

Despite the challenges with our approach, our results motivated many additional questions regarding the human MFC that are worth exploring. For example, the presence of novel initiation factors in the isolated *JUN* MFC, such as eIF4A and eIF4G, makes us wonder about the roles of these factors in *JUN* translation regulation. A possible initial approach for exploring this is the use of factor-depleted *in vitro* translation reactions of *JUN* 5' UTR reporter mRNAs (Gallie 2007). Moreover, mechanistic studies are still needed in order to characterize the rate and mechanism of formation for the human MFC. Isolation of the human MFC from different mammalian cell types and under different cellular conditions would also be beneficial in order to explore whether MFC composition varies in those contexts.

4.3 Final thoughts

By investigating translation regulation of the eIF3-target mRNA *JUN* we unveiled new contributors for regulation of this transcript during translation initiation. This has in turn inspired a number of ideas on the exploration of novel modes of specialized translation. As a transcript with a unique mode of regulation itself, as suggested both by our findings and by previous studies (Lee et al. 2015, 2016; Lamper et al. 2020), *JUN* supports the notion that such a pathway is mediated by complex sequence and structural features on specific regions of specific mRNAs. Our work is therefore a great example about how understanding translation initiation will help us understand expression of transcripts that are relevant under specific cellular conditions and disease. Although questions about *JUN* mRNA regulation still remain, such as what is the role of the 3' UTR in this process, our work has provided important insights into translational regulation of this important oncogenic factor and about translation initiation in general.

References

Aitken CE, Lorsch JR. 2012. A mechanistic overview of translation initiation in eukaryotes. *Nat Struct Mol Biol* **19**: 568–576.

Aleksashin N, Chang ST-L, Cate J. 2023. A highly efficient human cell-free translation system. *RNA* rna.079825.123.

Algire MA, Maag D, Savio P, Acker MG, Tarun SZ, Sachs AB, Asano K, Nielsen KH, Olsen DS, Phan L, et al. 2002. Development and characterization of a reconstituted yeast translation initiation system. *RNA* **8**: 382–397.

Angel P, Hattori K, Smeal T, Karin M. 1988. The jun proto-oncogene is positively autoregulated by its product, Jun/AP-1. *Cell* **55**: 875–885.

Antao VP, Lai SY, Tinoco I Jr. 1991. A thermodynamic study of unusually stable RNA and DNA hairpins. *Nucleic Acids Res* **19**: 5901–5905.

Araujo PR, Yoon K, Ko D, Smith AD, Qiao M, Suresh U, Burns SC, Penalva LOF. 2012. Before It Gets Started: Regulating Translation at the 5′ UTR. *Int J Genomics* **2012**: e475731.

Asano K, Clayton J, Shalev A, Hinnebusch AG. 2000. A multifactor complex of eukaryotic initiation factors, eIF1, eIF2, eIF3, eIF5, and initiator tRNA(Met) is an important translation initiation intermediate in vivo. *Genes Dev* **14**: 2534–2546.

Asano K, Krishnamoorthy T, Phan L, Pavitt GD, Hinnebusch AG. 1999. Conserved bipartite motifs in yeast eIF5 and eIF2Bepsilon, GTPase-activating and GDP-GTP exchange factors in translation initiation, mediate binding to their common substrate eIF2. *EMBO J* **18**: 1673–1688.

Asano K, Phan L, Anderson J, Hinnebusch AG. 1998. Complex formation by all five homologues of mammalian translation initiation factor 3 subunits from yeast Saccharomyces cerevisiae. *J Biol Chem* **273**: 18573–18585.

Asano K, Shalev A, Phan L, Nielsen K, Clayton J, Valášek L, Donahue TF, Hinnebusch AG. 2001. Multiple roles for the C-terminal domain of eIF5 in translation initiation complex assembly and GTPase activation. *EMBO J* **20**: 2326–2337.

Babendure JR, Babendure JL, Ding J-H, Tsien RY. 2006. Control of mammalian translation by mRNA structure near caps. *RNA* **12**: 851–861.

Bandyopadhyay A, Maitra U. 1999. Cloning and characterization of the p42 subunit of mammalian translation initiation factor 3 (eIF3): Demonstration that eIF3 interacts with eIF5 in mammalian cells. *Nucleic Acids Res* **27**: 1331–1337.

Barbosa C, Peixeiro I, Romão L. 2013. Gene Expression Regulation by Upstream Open Reading Frames and Human Disease. *PLOS Genet* **9**: e1003529.

Benne R, Hershey JW. 1978. The mechanism of action of protein synthesis initiation factors from rabbit reticulocytes. *J Biol Chem* **253**: 3078–3087.

Bernardi A, Spahr P-F. 1972. Nucleotide Sequence at the Binding Site for Coat Protein on RNA of Bacteriophage R17. *Proc Natl Acad Sci U S A* **69**: 3033–3037.

Bieniossek C, Schütz P, Bumann M, Limacher A, Uson I, Baumann U. 2006. The crystal structure of the carboxy-terminal domain of human translation initiation factor eIF5. *J Mol Biol* **360**: 457–465.

Blau L, Knirsh R, Ben-Dror I, Oren S, Kuphal S, Hau P, Proescholdt M, Bosserhoff A-K, Vardimon L. 2012. Aberrant expression of c-Jun in glioblastoma by internal ribosome entry site (IRES)-mediated translational activation. *Proc Natl Acad Sci* **109**: E2875–E2884.

Bohmann D, Bos TJ, Admon A, Nishimura T, Vogt PK, Tjian R. 1987. Human proto-oncogene c-jun encodes a DNA binding protein with structural and functional properties of transcription factor AP-1. *Science* **238**: 1386–1392.

Briggs J, Chamboredon S, Castellazzi M, Kerry JA, Bos TJ. 2002. Transcriptional upregulation of SPARC, in response to c-Jun overexpression, contributes to increased motility and invasion of MCF7 breast cancer cells. *Oncogene* **21**: 7077–7091.

Buttgereit F, Brand MD. 1995. A hierarchy of ATP-consuming processes in mammalian cells. *Biochem J* **312**: 163–167.

Calviello L, Venkataramanan S, Rogowski KJ, Wyler E, Wilkins K, Tejura M, Thai B, Krol J, Filipowicz W, Landthaler M, et al. 2021. DDX3 depletion represses translation of mRNAs with complex 5′ UTRs. *Nucleic Acids Res* **49**: 5336–5350.

Cate JHD. 2017. Human eIF3: from ‘blobology’ to biological insight. *Philos Trans R Soc B Biol Sci* **372**: 20160176.

Chen L, Bourguignon LYW. 2014. Hyaluronan-CD44 interaction promotes c-Jun signaling and miRNA21 expression leading to Bcl-2 expression and chemoresistance in breast cancer cells. *Mol Cancer* **13**: 52.

Chew G-L, Pauli A, Schier AF. 2016. Conservation of uORF repressiveness and sequence features in mouse, human and zebrafish. *Nat Commun* **7**: 11663.

Chiu W-L, Wagner S, Herrmannová A, Burela L, Zhang F, Saini AK, Valásek L, Hinnebusch AG. 2010. The C-terminal region of eukaryotic translation initiation factor 3a (eIF3a) promotes mRNA recruitment, scanning, and, together with eIF3j and the eIF3b RNA recognition motif, selection of AUG start codons. *Mol Cell Biol* **30**: 4415–4434.

Choudhuri A, Maitra U, Evans T. 2013. Translation initiation factor eIF3h targets specific transcripts to polysomes during embryogenesis. *Proc Natl Acad Sci* **110**: 9818–9823.

Crooks GE, Hon G, Chandonia J-M, Brenner SE. 2004. WebLogo: A Sequence Logo Generator. *Genome Res* **14**: 1188–1190.

Danaie P, Wittmer B, Altmann M, Trachsel H. 1995. Isolation of a Protein Complex Containing Translation Initiation Factor Prt1 from Saccharomyces cerevisiae. *J Biol Chem* **270**: 4288–4292.

Das R, Zhou Z, Reed R. 2000. Functional Association of U2 snRNP with the ATP-Independent Spliceosomal Complex E. *Mol Cell* **5**: 779–787.

De Silva D, Ferguson L, Chin GH, Smith BE, Apathy RA, Roth TL, Blaeschke F, Kudla M, Marson A, Ingolia NT, et al. 2021. Robust T cell activation requires an eIF3-driven burst in T cell receptor translation eds. B. Malissen and T. Taniguchi. *eLife* **10**: e74272.

Dennis MD, Person MD, Browning KS. 2009. Phosphorylation of Plant Translation Initiation Factors by CK2 Enhances the in Vitro Interaction of Multifactor Complex Components *. *J Biol Chem* **284**: 20615–20628.

Feinberg B, McLaughlin CS, Moldave K. 1982. Analysis of temperature-sensitive mutant ts 187 of Saccharomyces cerevisiae altered in a component required for the initiation of protein synthesis. *J Biol Chem* **257**: 10846–10851.

Fletcher CM, Pestova TV, Hellen CUT, Wagner G. 1999. Structure and interactions of the translation initiation factor eIF1. *EMBO J* **18**: 2631–2637.

Gallie DR. 2007. Use of In Vitro Translation Extract Depleted in Specific Initiation Factors for the Investigation of Translational Regulation. In *Methods in Enzymology*, Vol. 429 of, pp. 35–51.

Gee JMW, Barroso AF, Ellis IO, Robertson JFR, Nicholson RI. 2000. Biological and clinical associations of c-jun activation in human breast cancer. *Int J Cancer* **89**: 177–186.

Giacomelli AO, Yang X, Lintner RE, McFarland JM, Duby M, Kim J, Howard TP, Takeda DY, Ly SH, Kim E, et al. 2018. Mutational processes shape the landscape of TP53 mutations in human cancer. *Nat Genet* **50**: 1381–1387.

Gilbert RJC, Gordiyenko Y, von der Haar T, Sonnen AF-P, Hofmann G, Nardelli M, Stuart DI, McCarthy JEG. 2007. Reconfiguration of yeast 40S ribosomal subunit domains by the translation initiation multifactor complex. *Proc Natl Acad Sci* **104**: 5788–5793.

Gomes-Duarte A, Lacerda R, Menezes J, Romão L. 2017. eIF3: a factor for human health and disease. *RNA Biol* **15**: 26–34.

Hann SR, Sloan-Brown K, Spotts GD. 1992. Translational activation of the non-AUG-initiated c-myc 1 protein at high cell densities due to methionine deprivation. *Genes Dev* **6**: 1229–1240.

Hershey JWB, Sonenberg N, Mathews MB. 2012. Principles of translational control: an overview. *Cold Spring Harb Perspect Biol* **4**: a011528.

Hinnebusch AG. 2006. eIF3: a versatile scaffold for translation initiation complexes. *Trends Biochem Sci* **31**: 553–562.

Hinnebusch AG. 2011. Molecular Mechanism of Scanning and Start Codon Selection in Eukaryotes. *Microbiol Mol Biol Rev* **75**: 434–467.

Hinnebusch AG. 2014. The Scanning Mechanism of Eukaryotic Translation Initiation. *Annu Rev Biochem* **83**: 779–812.

Hinnebusch AG. 2005. Translational regulation of GCN4 and the general amino acid control of yeast. *Annu Rev Microbiol* **59**: 407–450.

Holcik M, Sonenberg N. 2005. Translational control in stress and apoptosis. *Nat Rev Mol Cell Biol* **6**: 318–327.

Hui L, Bakiri L, Mairhorfer A, Schweifer N, Haslinger C, Kenner L, Komnenovic V, Scheuch H, Beug H, Wagner EF. 2007. p38α suppresses normal and cancer cell proliferation by antagonizing the JNK–c-Jun pathway. *Nat Genet* **39**: 741–749.

Hyjek M, Figiel M, Nowotny M. 2019. RNases H: Structure and mechanism. *DNA Repair* **84**: 102672.

Ivanov IP, Loughran G, Atkins JF. 2008. uORFs with unusual translational start codons autoregulate expression of eukaryotic ornithine decarboxylase homologs. *Proc Natl Acad Sci* **105**: 10079–10084.

Ivanov IP, Loughran G, Sachs MS, Atkins JF. 2010. Initiation context modulates autoregulation of eukaryotic translation initiation factor 1 (eIF1). *Proc Natl Acad Sci* **107**: 18056–18060.

Ivanov IP, Saba JA, Fan C-M, Wang J, Firth AE, Cao C, Green R, Dever TE. 2022. Evolutionarily conserved inhibitory uORFs sensitize Hox mRNA translation to start codon selection stringency. *Proc Natl Acad Sci* **119**: e2117226119.

Ivanov IP, Wei J, Caster SZ, Smith KM, Michel AM, Zhang Y, Firth AE, Freitag M, Dunlap JC, Bell-Pedersen D, et al. 2017. Translation Initiation from Conserved

Non-AUG Codons Provides Additional Layers of Regulation and Coding Capacity. *mBio* **8**: 10.

Iwasaki S, Floor SN, Ingolia NT. 2016. Rocaglates convert DEAD-box protein eIF4A into a sequence-selective translational repressor. *Nature* **534**: 558–561.

Iwasaki S, Iwasaki W, Takahashi M, Sakamoto A, Watanabe C, Shichino Y, Floor SN, Fujiwara K, Mito M, Dodo K, et al. 2019. The Translation Inhibitor Rocaglamide Targets a Bimolecular Cavity between eIF4A and Polypurine RNA. *Mol Cell* **73**: 738-748.e9.

Jackson RJ, Hellen CUT, Pestova TV. 2010. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* **11**: 113–127.

Jazurek M, Ciesiolka A, Starega-Roslan J, Bilinska K, Krzyzosiak WJ. 2016. Identifying proteins that bind to specific RNAs - focus on simple repeat expansion diseases. *Nucleic Acids Res* **44**: 9050–9070.

Jennings MD, Pavitt GD. 2010. eIF5 has GDI activity necessary for translational control by eIF2 phosphorylation. *Nature* **465**: 378–381.

Jivotovskaya AV, Valášek L, Hinnebusch AG, Nielsen KH. 2006. Eukaryotic Translation Initiation Factor 3 (eIF3) and eIF2 Can Promote mRNA Binding to 40S Subunits Independently of eIF4G in Yeast. *Mol Cell Biol* **26**: 1355–1372.

Johnson R, Spiegelman B, Hanahan D, Wisdom R. 1996. Cellular transformation and malignancy induced by ras require c-jun. *Mol Cell Biol* **16**: 4504–4511.

Keryer-Bibens C, Barreau C, Osborne HB. 2008. Tethering of proteins to RNAs by bacteriophage proteins. *Biol Cell* **100**: 125–138.

Kolupaeva VG, Unbehaun A, Lomakin IB, Hellen CUT, Pestova TV. 2005. Binding of eukaryotic initiation factor 3 to ribosomal 40S subunits and its role in ribosomal dissociation and anti-association. *RNA* **11**: 470–486.

Kozak M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**: 283–292.

Kwan T, Thompson SR. 2019. Noncanonical Translation Initiation in Eukaryotes. *Cold Spring Harb Perspect Biol* **11**: a032672.

Lamper AM, Fleming RH, Ladd KM, Lee ASY. 2020. A phosphorylation-regulated eIF3d translation switch mediates cellular adaptation to metabolic stress. *Science* **370**: 853–856.

Lamph WW, Wamsley P, Sassone-Corsi P, Verma IM. 1988. Induction of proto-oncogene JUN/AP-1 by serum and TPA. *Nature* **334**: 629–631.

Lee ASY, Kranzusch PJ, Cate JHD. 2015. eIF3 targets cell-proliferation messenger RNAs for translational activation or repression. *Nature* **522**: 111–114.

Lee ASY, Kranzusch PJ, Doudna JA, Cate JHD. 2016. eIF3d is an mRNA cap-binding protein that is required for specialized translation initiation. *Nature* **536**: 96–99.

Leibovitch M, Topisirovic I. 2018. Dysregulation of mRNA translation and energy metabolism in cancer. *Adv Biol Regul* **67**: 30–39.

Leppek K, Das R, Barna M. 2018. Functional 5′ UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat Rev Mol Cell Biol* **19**: 158–174.

Liberman N, Gandin V, Svitkin YV, David M, Virgili G, Jaramillo M, Holcik M, Nagar B, Kimchi A, Sonenberg N. 2015. DAP5 associates with eIF2β and eIF4AI to promote Internal Ribosome Entry Site driven translation. *Nucleic Acids Res* **43**: 3764–3775.

Lin Y, Li F, Huang L, Polte C, Duan H, Fang J, Sun L, Xing X, Tian G, Cheng Y, et al. 2020. eIF3 Associates with 80S Ribosomes to Promote Translation Elongation, Mitochondrial Homeostasis, and Muscle Health. *Mol Cell* **79**: 575-587.

Loughran G, Sachs MS, Atkins JF, Ivanov IP. 2012. Stringency of start codon selection modulates autoregulation of translation initiation factor eIF5. *Nucleic Acids Res* **40**: 2898–2906.

Maag D, Algire MA, Lorsch JR. 2006. Communication between Eukaryotic Translation Initiation Factors 5 and 1A within the Ribosomal Pre-initiation Complex Plays a Role in Start Site Selection. *J Mol Biol* **356**: 724–737.

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**: 6580–6587.

Mašek T, Valášek L, Pospíšek M. 2011. Polysome Analysis and RNA Purification from Sucrose Gradients. In *RNA: Methods and Protocols* (ed. H. Nielsen), *Methods in Molecular Biology*, pp. 293–309.

Matsuda D, Dreher TW. 2006. Close spacing of AUG initiation codons confers dicistronic character on a eukaryotic mRNA. *RNA* **12**: 1338–1349.

Meng Q, Xia Y. 2011. c-Jun, at the crossroad of the signaling network. *Protein Cell* **2**: 889–898.

Mestre-Fos S, Ferguson L, Trinidad MI, Ingolia N, Cate JHD. 2023. eIF3 engages with 3′-UTR termini of highly translated mRNAs in neural progenitor cells. https://doi.org/10.1101/2023.11.11.566681

Mignone F, Gissi C, Liuni S, Pesole G. 2002. Untranslated regions of mRNAs. *Genome Biol* **3**: reviews0004.1.

Mitchell SF, Walker SE, Algire MA, Park E-H, Hinnebusch AG, Lorsch JR. 2010. The 5′-7-Methylguanosine Cap on Eukaryotic mRNAs Serves Both to Stimulate Canonical Translation Initiation and to Block an Alternative Pathway. *Mol Cell* **39**: 950–962.

Mo J, Liang H, Su C, Li P, Chen J, Zhang B. 2021. DDX3X: structure, physiologic functions and cancer. *Mol Cancer* **20**: 38.

Mukhopadhyay S, Amodeo ME, Lee ASY. 2023. eIF3d controls the persistent integrated stress response. *Mol Cell* **83**: 3303-3313.

Nakamura T, Datta R, Kharbanda S, Kufe D. 1991. Regulation of jun and fos gene expression in human monocytes by the macrophage colony-stimulating factor. *Cell Growth Differ Mol Biol J Am Assoc Cancer Res* **2**: 267–272.

Naranda T, MacMillan SE, Hershey JW. 1994. Purified yeast translational initiation factor eIF-3 is an RNA-binding protein complex that contains the PRT1 protein. *J Biol Chem* **269**: 32286–32292.

Nateri AS, Spencer-Dene B, Behrens A. 2005. Interaction of phosphorylated c-Jun with TCF4 regulates intestinal cancer development. *Nature* **437**: 281–285.

Olsen DS, Savner EM, Mathew A, Zhang F, Krishnamoorthy T, Phan L, Hinnebusch AG. 2003. Domains of eIF1A that mediate binding to eIF2, eIF3 and eIF5B and promote ternary complex recruitment in vivo. *EMBO J* **22**: 193–204.

Parsyan A, Svitkin Y, Shahbazian D, Gkogkas C, Lasko P, Merrick WC, Sonenberg N. 2011. mRNA helicases: the tacticians of translational control. *Nat Rev Mol Cell Biol* **12**: 235–245.

Pelletier J, Sonenberg N. 1985. Insertion mutagenesis to increase secondary structure within the 5′ noncoding region of a eukaryotic mRNA reduces translational efficiency. *Cell* **40**: 515–526.

Pesole G, Mignone F, Gissi C, Grillo G, Licciulli F, Liuni S. 2001. Structural and functional features of eukaryotic mRNA untranslated regions. *Gene* **276**: 73–81.

Pestova TV, Borukhov SI, Hellen CUT. 1998. Eukaryotic ribosomes require initiation factors 1 and 1A to locate initiation codons. *Nature* **394**: 854–859.

Pestova TV, Kolupaeva VG. 2002. The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes Dev* **16**: 2906–2922.

Phan L, Zhang X, Asano K, Anderson J, Vornlocher HP, Greenberg JR, Qin J, Hinnebusch AG. 1998. Identification of a translation initiation factor 3 (eIF3) core complex, conserved in yeast and mammals, that interacts with eIF5. *Mol Cell Biol* **18**: 4935–4946.

Pulos-Holmes MC, Srole DN, Juarez MG, Lee AS-Y, McSwiggen DT, Ingolia NT, Cate JH. 2019. Repression of ferritin light chain translation by human eIF3 eds. A.G. Hinnebusch, J.L. Manley, and L.S. Valasek. *eLife* **8**: e48193.

Rakotondrafara AM, Hentze MW. 2011. An efficient factor-depleted mammalian in vitro translation system. *Nat Protoc* **6**: 563–571.

Rubio CA, Weisburd B, Holderfield M, Arias C, Fang E, DeRisi JL, Fanidi A. 2014. Transcriptome-wide characterization of the eIF4A signature highlights plasticity in translation regulation. *Genome Biol* **15**: 476.

Ryder K, Lau LF, Nathans D. 1988. A gene activated by growth factors is related to the oncogene v-jun. *Proc Natl Acad Sci U S A* **85**: 1487–1491.

Sachs AB, Varani G. 2000. Eukaryotic translation initiation: there are (at least) two sides to every story. *Nat Struct Biol* **7**: 356–361.

Sayers EW, Cavanaugh M, Clark K, Ostell J, Pruitt KD, Karsch-Mizrachi I. 2019. GenBank. *Nucleic Acids Res* **47**: D94–D99.

Schneider TD, Stephens RM. 1990. Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res* **18**: 6097–6100.

Schoch CL, Ciufo S, Domrachev M, Hotton CL, Kannan S, Khovanskaya R, Leipe D, Mcveigh R, O'Neill K, Robbertse B, et al. 2020. NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database* **2020**: baaa062.

Sen ND, Zhou F, Harris MS, Ingolia NT, Hinnebusch AG. 2016. eIF4B stimulates translation of long mRNAs with structured 5′ UTRs and low closed-loop potential but weak dependence on eIF4G. *Proc Natl Acad Sci U S A* **113**: 10464–10472.

Shahbazian D, Parsyan A, Petroulakis E, Hershey J, Sonenberg N. 2010. eIF4B controls survival and proliferation and is regulated by proto-oncogenic signaling pathways. *Cell Cycle* **9**: 4106–4109.

Singh CR, Yamamoto Y, Asano K. 2004. Physical Association of Eukaryotic Initiation Factor (eIF) 5 Carboxyl-terminal Domain with the Lysine-rich eIF2β Segment Strongly Enhances Its Binding to eIF3 *. *J Biol Chem* **279**: 49644–49655.

Smeal T, Binetruy B, Mercola DA, Birrer M, Karin M. 1991. Oncogenic and transcriptional cooperation with Ha-Ras requires phosphorylation of c-Jun on serines 63 and 73. *Nature* **354**: 494–496.

Smith E, Meyerrose TE, Kohler T, Namdar-Attar M, Bab N, Lahat O, Noh T, Li J, Karaman MW, Hacia JG, et al. 2005. Leaky ribosomal scanning in mammalian genomes: significance of histone H4 alternative translation in vivo. *Nucleic Acids Res* **33**: 1298–1308.

Sokabe M, Fraser CS, Hershey JWB. 2012. The human translation initiation multi-factor complex promotes methionyl-tRNAi binding to the 40S ribosomal subunit. *Nucleic Acids Res* **40**: 905–913.

Sonenberg N, Dever TE. 2003. Eukaryotic translation initiation factors and regulators. *Curr Opin Struct Biol* **13**: 56–63.

Sonenberg N, Hinnebusch AG. 2009. Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. *Cell* **136**: 731–745.

Stanciu A, Luo J, Funes L, Galbokke Hewage S, Kulkarni SD, Aitken CE. 2022. eIF3 and Its mRNA-Entry-Channel Arm Contribute to the Recruitment of mRNAs With Long 5′-Untranslated Regions. *Front Mol Biosci* **8**.

Starck SR, Tsai JC, Chen K, Shodiya M, Wang L, Yahiro K, Martins-Green M, Shastri N, Walter P. 2016. Translation from the 5′ untranslated region shapes the integrated stress response. *Science* **351**: aad3867.

Stripecke R, Hentze MW. 1992. Bacteriophage and spliceosomal proteins function as position-dependent cis/trans repressors of mRNA translation in vitro. *Nucleic Acids Res* **20**: 5555–5564.

Suphakhong K, Terashima M, Wanna-udom S, Takatsuka R, Ishimura A, Takino T, Suzuki T. 2022. m6A RNA methylation regulates the transcription factors JUN and JUNB in TGF-β-induced epithelial–mesenchymal transition of lung cancer cells. *J Biol Chem* **298**: 102554.

Svitkin YV, Pause A, Haghighat A, Pyronnet S, Witherell G, Belsham GJ, 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**: 382–394.

Tacca LMA de, Pulos-Holmes MC, Floor SN, Cate JHD. 2019. PTBP1 mRNA isoforms and regulation of their translation. *RNA* **25**: 1324–1336.

Topisirovic I, Sonenberg N. 2011. mRNA translation and energy metabolism in cancer: the role of the MAPK and mTORC1 pathways. *Cold Spring Harb Symp Quant Biol* **76**: 355–367.

Valášek L, Mathew AA, Shin B-S, Nielsen KH, Szamecz B, Hinnebusch AG. 2003. The yeast eIF3 subunits TIF32/a, NIP1/c, and eIF5 make critical connections with the 40S ribosome in vivo. *Genes Dev* **17**: 786–799.

Valášek L, Nielsen KH, Hinnebusch AG. 2002. Direct eIF2–eIF3 contact in the multifactor complex is important for translation initiation in vivo. *EMBO J* **21**: 5886–5898.

Vasilevskaya I, O'Dwyer PJ. 2003. Role of Jun and Jun kinase in resistance of cancer cells to therapy. *Drug Resist Updat* **6**: 147–156.

Vassilenko KS, Alekhina OM, Dmitriev SE, Shatsky IN, Spirin AS. 2011. Unidirectional constant rate motion of the ribosomal scanning particle during eukaryotic translation initiation. *Nucleic Acids Res* **39**: 5555–5567.

Ward FR, Watson ZL, Ad O, Schepartz A, Cate JHD. 2019. Defects in the Assembly of Ribosomes Selected for β-Amino Acid Incorporation. *Biochemistry* **58**: 4494–4504.

Weber R, Kleemann L, Hirschberg I, Chung M-Y, Valkov E, Igreja C. 2022. DAP5 enables main ORF translation on mRNAs with structured and uORF-containing 5′ leaders. *Nat Commun* **13**: 7510.

Wisdom R, Johnson RS, Moore C. 1999. c-Jun regulates cell cycle progression and apoptosis by distinct mechanisms. *EMBO J* **18**: 188–197.

Wolf DA, Lin Y, Duan H, Cheng Y. 2020. eIF-Three to Tango: emerging functions of translation initiation factor eIF3 in protein synthesis and disease. *J Mol Cell Biol* **12**: 403–409.

Wulf GM, Ryo A, Wulf GG, Lee SW, Niu T, Petkova V, Lu KP. 2001. Pin1 is overexpressed in breast cancer and cooperates with Ras signaling in increasing the transcriptional activity of c-Jun towards cyclin D1. *EMBO J* **20**: 3459–3472.

Xiang Y, Huang W, Tan L, Chen T, He Y, Irving PS, Weeks KM, Zhang QC, Dong X. 2023. Pervasive downstream RNA hairpins dynamically dictate start-codon selection. *Nature* **621**: 423–430.

Xu T-R, Lu R-F, Romano D, Pitt A, Houslay MD, Milligan G, Kolch W. 2012. Eukaryotic Translation Initiation Factor 3, Subunit a, Regulates the Extracellular Signal-Regulated Kinase Pathway. *Mol Cell Biol* **32**: 88–95.

Yamamoto Y, Singh CR, Marintchev A, Hall NS, Hannig EM, Wagner G, Asano K. 2005. The eukaryotic initiation factor (eIF) 5 HEAT domain mediates multifactor assembly and scanning with distinct interfaces to eIF1, eIF2, eIF3, and eIF4G. *Proc Natl Acad Sci* **102**: 16164–16169.

Yoon J-H, Gorospe M. 2016. Identification of mRNA-interacting factors by MS2-TRAP (MS2-tagged RNA affinity purification). *Methods Mol Biol Clifton NJ* **1421**: 15–22.

Yoon J-H, Srikantan S, Gorospe M. 2012. MS2-TRAP (MS2-tagged RNA affinity purification): Tagging RNA to identify associated miRNAs. *Methods San Diego Calif* **58**: 81–87.

Youngman EM, Green R. 2005. Affinity purification of in vivo-assembled ribosomes for in vitro biochemical analysis. *Methods* **36**: 305–312.

Zhang H, Wang Y, Wu X, Tang X, Wu C, Lu J. 2021. Determinants of genome-wide distribution and evolution of uORFs in eukaryotes. *Nat Commun* **12**: 1076.

Zhou Z, Licklider LJ, Gygi SP, Reed R. 2002. Comprehensive proteomic analysis of the human spliceosome. *Nature* **419**: 182–185.

Appendix

Appendix Table 1. Conservation analysis for *JUN* uAUG and mAUG.

Compilation of species investigated for *JUN* AUGs conservation analysis, including the nucleotide and amino acid sequences of the 19-nucleotide *JUN* 5′ UTR and *JUN* CDS region that spans both AUG start codons and their translational context for each species, the corresponding Reference Sequence (RefSeq) accession numbers for each sequence, and the percent similarity of each sequence to the human *JUN* sequence.

Organism (common name)	**Organism (scientific name)**	**RefSeq accession number (Nucleotide)**	**RefSeq accession number (Protein)**
Human	*Homo sapiens*	NM_002228.4	NP_002219.1
Chimpanzee	*Pan troglodytes*	XM_513442.6	XP_513442.2
Pygmy chimpanzee	*Pan paniscus*	XM_003824222.6	XP_003824270.1
Western lowland gorilla	*Gorilla gorilla*	XM_004025880.3	XP_004025929.1
Sumatran orangutan	*Pongo abelii*	XM_002810763.5	XP_002810809.2
Bornean orangutan	*Pongo pygmaeus*	XM_054488470.1	XP_054344445.1
Silvery gibbon	*Hylobates moloch*	XM_032136524.2	XP_031992415.1
Pere David's macaque	*Macaca thibetana thibetana*	XM_050804328.1	XP_050660285.1
Rhesus monkey	*Macaca mulatta*	NM_001265850.2	NP_001252779.1
Crab-eating macaque	*Macaca fascicularis*	XM_005543232.3	XP_005543289.1
Ring-tailed lemur	*Lemur catta*	XM_045547914.1	XP_045403870.1
Slow loris	*Nycticebus coucang*	XM_053576050.1	XP_053432025.1
Siamang	*Symphalangus syndactylus*	XM_055234226.1	XP_055090201.1
White-tufted-ear marmoset	*Callithrix jacchus*	XM_002750880.6	XP_002750926.1
European snow vole	*Chionomys nivalis*	XM_057785267.1	XP_057641250.1
Reed vole	*Microtus fortis*	XM_050136224.1	XP_049992181.1
Bank vole	*Myodes glareolus*	XM_048453087.1	XP_048309044.1
Creeping vole	*Microtus oregoni*	XM_041651928.1	XP_041507862.1
European marmot	*Marmota marmota marmota*	XM_048815299.1	XP_048671256.1
Gray squirrel	*Sciurus carolinensis*	XM_047562473.1	XP_047418429.1
Golden hamster	*Mesocricetus auratus*	XM_013116518.3	XP_012971972.1
Desert hamster	*Phodopus roborovskii*	XM_051188187.1	XP_051044144.1
Golden spiny mouse	*Acomys russatus*	XM_051164333.1	XP_051020290.1
California mouse	*Peromyscus californicus insignis*	XM_052715886.1	XP_052571846.1
House mouse	*Mus musculus*	NM_010591.2	NP_034721.1
European woodmouse	*Apodemus sylvaticus*	XM_052176774.1	XP_052032734.1
Little pocket mouse	*Perognathus longimembris pacificus*	XM_048351427.1	XP_048207384.1
North American deer mouse	*Peromyscus maniculatus b airdii*	XM_006974021.3	XP_006974083.1
Banner-tailed kangaroo rat	*Dipodomys spectabilis*	XM_042681035.1	XP_042536969.1
Norway rat	*Rattus norvegicus*	NM_021835.3	NP_068607.1
Lesser Egyptian jerboa	*Jaculus jaculus*	XM_045151203.1	XP_045007138.1
Iberian mole	*Talpa occidentalis*	XM_037498772.2	XP_037354669.1
Bactrian camel	*Camelus bactrianus*	XM_010958378.2	XP_010956680.1
Red deer	*Cervus elaphus*	XM_043877187.1	XP_043733122.1

Elk	*Cervus canadensis*	XM_043461159.1	XP_043317094.1
Sheep	*Ovis aries*	XM_004002020.5	XP_004002069.2
Cattle	*Bos taurus*	NM_001077827.1	NP_001071295.1
Scimitar-horned oryx	*Oryx dammah*	XM_040232750.	XP_040088684.1
Water buffalo	*Bubalus bubalis*	XM_006048272.4	XP_006048334.1
Carabao	*Bubalus carabanensis*	XM_055585704.1	XP_055441679.1
Chinese forest musk deer	*Moschus berezovskii*	XM_055401793.1	XP_055257768.1
Black-lipped pika	*Ochotona curzoniae*	XM_040978479.1	XP_040834413.1
American pika	*Ochotona princeps*	XM_004599929.3	XP_004599986.1
Eurasian river otter	*Lutra lutra*	XM_047726657.1	XP_047582613.1
European polecat	*Mustela putorius furo*	XM_004751637.3	XP_004751694.1
American mink	*Neogale vison*	XM_044238339.1	XP_044094274.1
Eurasian badger	*Meles meles*	XM_046026981.1	XP_045882937.1
Chinese pangolin	*Manis pentadactyla*	XM_036911237.2	XP_036767132.2
American black bear	*Ursus americanus*	XM_045802220.1	XP_045658176.1
Brown bear	*Ursus arctos*	XM_026497983.4	XP_026353768.1
Polar bear	*Ursus maritimus*	XM_040627242.1	XP_040483176.1
Jamaican fruit-eating bat	*Artibeus jamaicensis*	XM_037148528.2	XP_037004423.1
Big brown bat	*Eptesicus fuscus*	XM_008139371.3	XP_008137593.1
Common vampire bat	*Desmodus rotundus*	XM_053921007.1	XP_053776982.1
Greater spear-nosed bat	*Phyllostomus hastatus*	XM_045858290.1	XP_045714246.1
Northern elephant seal	*Mirounga angustirostris*	XM_045897588.2	XP_045753544.1
Parnell's mustached bat	*Pteronotus parnellii mesoamericanus*	XM_054573889.1	XP_054429864.1
Kuhl's pipistrelle	*Pipistrellus kuhlii*	XM_036444790.2	XP_036300683.1
Dog	*Canis lupus familiaris*	XM_005620245.4	XP_038393322.1
Racoon dog	*Nyctereutes procyonoides*	XM_055306966.1	XP_055162941.1
Asiatic elephant	*Elephas maximus indicus*	XM_049879393.1	XP_049735350.1
Common warthog	*Phacochoerus africanus*	XM_047788888.1	XP_047644844.1
Killer whale	*Orcinus orca*	XM_004273791.4	XP_004273839.1
Sperm whale	*Physeter catodon*	XM_024119922.3	XP_023975690.1
Pygmy sperm whale	*Kogia breviceps*	XM_059054023.1	XP_058910006.1
Minke whale	*Balaenoptera acutorostrata*	XM_007164749.2	XP_007164811.2
Hippopotamus	*Hippopotamus amphibius kiboko*	XM_057712398.1	XP_057568381.1
Lion	*Panthera leo*	XM_042952350.1	XP_042808284.1
Leopard cat	*Prionailurus bengalensis*	XM_043575115.1	XP_043431050.1
Fishing cat	*Prionailurus viverrinus*	XM_047872112.1	XP_047728068.1
Bobcat	*Lynx rufus*	XM_047085545.1	XP_046941501.1
Leopard	*Panthera pardus*	XM_019449108.2	XP_019304653.1
Snow leopard	*Panthera uncia*	XM_049617076.1	XP_049473033.1
Clouded leopard	*Neofelis nebulosa*	XM_058717233.1	XP_058573216.1
Geoffroy's cat	*Leopardus geoffroyi*	XM_045477588.1	XP_045333544.1
Domestic cat	*Felis catus*	XM_011284967.4	XP_011283269.3
Tiger	*Panthera tigris*	XM_042997384.1	XP_042853318.1
Cheetah	*Acinonyx jubatus*	XM_027059048.2	XP_026914849.1
Jaguarundi	*Puma yagouaroundi*	XM_040470409.1	XP_040326343.1
Plains zebra	*Equus quagga*	XM_046662656.1	XP_046518612.1
Horse	*Equus caballus*	XM_023628887.1	XP_023484655.1
Nine-banded armadillo	*Dasypus novemcinctus*	XM_004470671.4	XP_004470728.1
Tarantolino	*Euleptes europaea*	XM_056845674.1	XP_056701652.1
Aeolian wall lizard	*Podarcis raffonei*	XM_053392286.1	XP_053248261.1

Chicken	*Gallus gallus*	NM_001031289.2	NP_001026460.2
Plains spadefoot toad	*Spea bombifrons*	XM_053469550.1	XP_053325525.1
Oriental whip snake	*Ahaetulla prasina*	XM_058174348.1	XP_058030331.1
Diamondback terrapin	*Malaclemys terrapin pileata*	XM_054036757.1	XP_053892732.1
Leatherback sea turtle	*Dermochelys coriacea*	XM_038414824.2	XP_038270752.1
Green sea turtle	*Chelonia mydas*	XM_043520950.1	XP_043376885.1
Loggerhead turtle	*Caretta caretta*	XM_048861394.1	XP_048717351.1
Yellowpond turtle	*Mauremys mutica*	XM_045026087.1	XP_044882022.1
Painted turtle	*Chrysemys picta bellii*	XM_005284797.3	XP_005284854.1
Mexican gopher tortoise	*Gopherus flavomarginatus*	XM_050962067.1	XP_050818024.1
Komodo dragon	*Varanus komodoensis*	XM_044430327.1	XP_044286262.1
Townsend's dwarf sphaero	*Sphaerodactylus townsendi*	XM_048497630.1	XP_048353587.1
Graceful crag lizard	*Hemicordylus capensis*	XM_053249389.1	XP_053105364.1
Zebrafish	*Danio rerio*	NM_199987.1	NP_956281.1
Spotted gar	*Lepisosteus oculatus*	XM_00635001.2	XP_006635064.1
Fruit fly	*Drosophila melanogaster*	NT_033778.4	ALC41668.1
Nematode	*Caenorhabditis elegans*	NC_003280.10	NP_001022366.1

Appendix Table 2. Annotation revision for the *JUN* 5' UTR sequence.
Compilation of variants found in the cDNA of the *JUN* 5' UTR and their annotation on databases including IGV, ENSEMBL, gnomAD, BLAST, ENSEMBL variants, and NLM-NCBI.

Position in 5'UTR (5'->3')	Position in 5'UTR (3'->5')	Chromosome position (3'->5')	Annotated sequence (IGV)	Annotated sequence (ENSEMBL)	Mutation observed in cDNA	Type of mutation	Reported in gnomAD	Reported in BLAST	Reported in ENSEMBL variants	Reported in NLM-NCBI
293	-685	Chr1: 58, 783, 755	G	G	A	substitution	NO	NO	NO	NO
373	-605	Chr1: 58, 783, 675	C	C	G	substitution	YES	NO	YES	YES
689	-289	Chr1: 58, 783, 359	A	A	C	substitution	YES	NO	YES	YES
798-800	(-180) – (-178)	Chr1: downstream of 58, 783, 248	none	GAG	none	deletion	NO	YES	NO	NO
Between 800-801	(-178) – (-177)	Chr1: between 58, 783, 247 - 58, 783, 248	none	none	CC	insertion	NO	YES	PARTIAL	NO
907	-71	Chr1: 58, 783, 141	C	C	G	substitution	YES	NO	YES	YES