

DISSERTATION

STRANGE TRANSLATION - INVESTIGATING IRES-MEDIATED AND CODON NON-  
OPTIMAL TRANSLATION DYNAMICS AT THE SINGLE MRNA LEVEL IN LIVING  
CELLS

Submitted by

Amanda Lynn Koch

Department of Biochemistry and Molecular Biology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2021

Doctoral Committee:

Advisor: Timothy Stasevich

Brian Munsky

Steven Markus

Olve Peersen

Jeffrey Wilusz

Copyright by Amanda L. Koch 2021

All Rights Reserved

## ABSTRACT

### STRANGE TRANSLATION - INVESTIGATING IRES-MEDIATED AND CODON NON-OPTIMAL TRANSLATION DYNAMICS AT THE SINGLE MRNA LEVEL IN LIVING CELLS

With the advent of Nascent Chain Tracking (NCT), a technique used to visualize single-molecule events of translation in living cells, answering detailed questions about how, when, and where translation is occurring in living cells is possible. Since its publishing debut in 2016, NCT has provided a wealth of information about translation initiation and elongation dynamics, subcellular localization, translation site structure, and reaction to stress for both canonical and non-canonical translation in living cells. Here, we slightly modify the NCT assay to quantify translation dynamics when a ribosome is recruited to an mRNA in a non-canonical fashion and when a ribosome encounters codon non-optimal stretches on a transcript.

The first step of translation requires a primed ribosome to be recruited to a readied mRNA. Canonically, this recruitment takes place on the 5' cap of an mRNA and is termed cap-dependent initiation. However, some eukaryotic messages and many viral RNAs use an internal ribosome entry site (IRES) to recruit ribosomes and initiate translation in a cap-independent manner. Specifically, viruses use IRES elements to hijack host ribosomes to translate viral proteins and properly propagate in host cells. While well-studied in bulk, the dynamics of IRES-mediated translation remain unexplored at the

single-molecule level. Here, we developed a bicistronic biosensor encoding distinct repeat epitopes in two open reading frames (ORFs), one translated from the 5'-cap, the other from the Encephalomyocarditis Virus IRES. When combined with a pair of complementary probes that bind the epitopes co-translationally, the biosensor lights up in different colors depending on which ORF is translated. Using the sensor together with single-molecule tracking and computational modeling, we measured the kinetics of cap-dependent versus IRES-mediated translation in living human cells. We show that bursts of IRES translation are shorter and rarer than bursts of cap translation, although the situation reverses upon stress. Collectively our data support a model for translational regulation primarily driven by transitions between translationally active and inactive RNA states.

Once the ribosome has been recruited to the mRNA and a start codon located, the ribosome will begin decoding the mRNA in nucleotide triplets or codons to ultimately create a protein. In some cases, the ribosome encounters a codon that it cannot decode efficiently. The relationship between codons and ribosome efficiency is termed codon optimality. It has been shown that codon non-optimal mRNA are less stable in cells. However, little is known about the effects of codon non-optimality on translation kinetics and overall translation regulation. In an ongoing collaboration with the Rissland group, we use bulk assays and NCT to address unanswered questions about how codon non-optimality leads to translation regulation along with mRNA instability. Thus far, we have evidence to support that translation repression is occurring in codon non-optimal conditions through inhibition of ribosome initiation and slower elongation. Further investigations of exact translation repression mechanisms are ongoing.

## ACKNOWLEDGEMENTS

I would like to first and foremost thank Tim Stasevich for your mentorship and support over the last five years. You have guided me into becoming a young scientist while giving me the freedom and support to explore my passions outside of science.

Thank you to all the Stasevich lab members. Your feedback and general conversations have made me a much better critical thinker and experimentalist.

I would like to thank our collaborators, the Munsky lab and the Rissland lab. It has been great getting to know and work with you all. You have provided me with such great insight into computational work and the biology of translational repression mechanisms.

To my student advisory committee, thank you for your time, confidence, and unwavering support. I am lucky to have such wonderfully insightful people in my corner keeping me on track.

I would like to thank Tatsuya Morisaki for your mentorship and friendship. You have been an incredible mentor not only in science but in life. You reassured me with such kindness every time I thought I wanted to quit graduate school and gave me advice at some of the hardest times of my life. You push me to be a better person. Thank you.

I am eternally grateful for the friendship of Charlotte, Dylan, Katie B., Saxton, and Rob TP Williams. Each of you has an extremely special place in my heart. You all are honest, supportive, and caring. It is an understatement to say that I wouldn't be here if it

weren't for each of you. I hope that we continue our friendships well past the glue of graduate school.

Thank you to my forever best friends, Taytay, Philly, El, Hayl, and Kels. You all are the most intelligent, most driven, kindest, and strongest women that I have ever met. I strive to be more like each of you every day. Thank you for your friendship, support, and love through college and our 20s.

I would like to thank my family. Mom and Dad, thank you for your constant love, honesty, and support. These were tough years. You were always there to listen to me and offer much-needed guidance and wisdom. You are amazing parents. I am extremely lucky to have you. Adam, it is rare that a best friend is also your sibling. I didn't choose you as a brother, but I choose you as my best friend. Thank you for being there for me always. You helped get me through graduate school. I am excited to repay the favor.

Finally, thank you to the entire BMB department. Thank you for accepting me into this program and guiding me through it. I am a proud CSU Ram and an even prouder CSU BMBer.

## DEDICATION

*To all those who have ever been waitlisted*

*and*

*To all those who have ever thought about quitting*

# Chapter 1

## Introduction

### 1.1 A historical view of translation and its discovery

In 1958, Francis Crick first proposed the concept of the central dogma, the process in which DNA blueprints are converted to functional protein products needed by the cell [1]. This process is broken up into two highly regulated steps, transcription, where DNA blueprints are encoded into RNA instructions, and translation where RNA instructions are decoded into final protein products [2]. Since 1958, researchers have tirelessly worked to understand how each of these steps is regulated and importantly how disease-causing misregulation occurs.

Though Crick realized the central dogma in its entirety, he could not have done so without the monumental discovery by George Palade in 1955 [3]. With the newly developed electron microscope and methodology for separating cell parts, George Palade discovered organelles in the cell where protein formation takes place [4]. These organelles are complex molecular factories called ribosomes that conduct the process of protein synthesis or translation.

The final missing piece to this puzzle was the middleman or messenger RNA – what the ribosome reads to make the desired protein. Interestingly, the discovery of mRNA came after that of the ribosome and Crick's hypothesized central dogma [5]. Up until this point, it was hypothesized that at least two RNA molecules were needed during protein synthesis – the rRNA in the ribosome and the tRNA that transferred the amino



acids to the ribosome for protein synthesis [6]. How the tRNA and ribosomes got together was still unknown. It was hypothesized that a short-lived intermediate template existed that brings all the protein synthesis components together, but no one knew for sure [7–9].

During an after-hour gathering of the minds, Crick and several other colleagues discussed ongoing experiments [10]. Two separate groups at this informal meeting had been studying the roles of RNA in cells. Through radiolabeling experiments, one group became convinced that RNA was the precursor to DNA, but the results were confounding [11]. The other group showed through bacteria induction that an immediate chemical signal existed allowing for host cells to synthesize needed proteins rapidly but had no idea what that signal could be [12,13]. Both groups realized simultaneously that the RNA molecule studied by one group was indeed the mysterious chemical messenger of the other. Shortly thereafter, in 1961, several groups described the nature of messenger RNA and the role it plays as the middle man in the central dogma [14–16].

Since Palade's discovery of ribosomes and the establishment of messenger RNA molecules, much wisdom has been gained about how translation is conducted within cells and how it can be regulated. Translation can be broken down into three stages: ribosome recruitment and initiation, elongation, and termination. Each step has its own set of required factors that involves coordinated binding, unbinding, and regulation.

## 1.2 Introduction to translation initiation

Typically, mRNA consist of the following sections: a 5' cap, 5' untranslated region (UTR), open reading frame (ORF), 3' UTR, and poly(A) tail. First, mRNA are capped by a methylated guanosine unusually linked to the next nucleotide in a 5' to 5' triphosphate

fashion. Adjacent nucleotides can also be methylated in a 2'-O-methylated fashion leading to different flavors of the 5' cap including the cap 0, cap 1, and cap 2 structures [17]. This is the site for factor binding and mRNA protection [18]. In between the cap and the start codon, is the 5' UTR. Varying in length between different transcripts, this is typically the region where ribosomal scanning occurs [19]. Through complex structure formation, the 5' UTR can also regulate translation [20]. After the start codon is the coding region of the mRNA or the open reading frame (ORF). This is the section of the mRNA that is translated into protein products [21]. Following a stop codon is the 3' UTR which often contains protein and small RNA binding sites that regulate translation [22,23]. Finally, the mRNAs are usually tailed with a stretch of adenosines. The poly(A) tail is known to aid in nuclear export and mRNA stability [24].

For translation initiation to occur, both the small subunit of the ribosome (40S) and the mRNA must be primed by the binding of numerous translation initiation factors [25,26]. In eukaryotes, a complex orchestra of at least twelve different initiation factors is required for translation to initiate [27]. First, the mRNA is readied by the binding of initiation factors in the eIF4 family [28]. Though other factors bind to the mRNA, the factors listed here have the most established function. eIF4E which binds the 5' cap is interestingly phylogenetically conserved across eukaryotes from yeast to human highlighting the importance of this factor in successful translation initiation in multiple species [29]. Because of its importance, eIF4E is also a main target in translation regulation especially during stress which will be discussed in detail in Chapter 2 [18]. eIF4G binds to eIF4E and acts as a scaffold between the ribosome and the mRNA. eIF4G also has a binding site for the polyA binding protein (PABP) implying its role in looping mRNAs cap to tail

[30]. Because the 5' untranslated regions (UTR) of an mRNA can be highly structured, a helicase, eIF4A, is necessary to unwind a path for the ribosome from the cap to the start codon [31].

With the mRNA prepared, next, the small subunit of the ribosome must likewise be primed by the assembly of initiation factors. The two key initiation factors responsible for this are eIF2 and eIF3 [32]. In charge of bringing the initiator tRNA or Met-tRNA to the ribosome and recognizing the start codon of an mRNA, eIF2 is essential for cap-dependent translation initiation [33]. When the start codon is recognized, eIF2 hydrolyzes GTP to GDP which leads to the recruitment of the 60S ribosomal subunit and the beginning of translation elongation. Because GDP-bound eIF2 cannot bind Met-tRNA, GDP must be exchanged for GTP such that other rounds of translation can be initiated. A subunit of eIF2, eIF2B is required for the recycling of inactive eIF2-GDP to active eIF2-GTP [34]. The largest initiation factor, eIF3 serves multiple functions in translation initiation. It participates in recruiting the eIF2-Met-tRNA complex to the 40S ribosome. Importantly, eIF3 facilitates the recruitment of the mRNA to the ribosome [35].

Once these initiation factors are in the proper place, the ribosome can be recruited to the mRNA but initiation is not done yet. Now the ribosome must move along the 5' UTR until proper base pairing occurs between Met-tRNA and an AUG start codon [19]. Obstacles such as complicated structures in the 5' UTR can dramatically reduce start site recognition and therefore translational efficiencies. Through all obstacles, important initiation factors eIF1 and eIF1A are responsible for keeping the ribosome in a necessary open conformation until a proper start codon is reached [36]. This search for an AUG in the 5' to 3' direction along the 5'UTR is termed ribosomal scanning [37,38].

As the ribosome scans, it will recognize a start codon depending on the surrounding sequences, a phenomenon thoroughly investigated by Marilyn Kozak. Impressively, from 1979 – 1987, Kozak published six single first-author papers describing the influence of AUG flanking sequences on start codon recognition in mammalian cells [39–44]. She found through careful comparisons of hundreds of mRNAs and systematic mutagenesis that CCAGCC**AUGG** was the consensus sequence for translation initiation with the bolded AUG as the start codon [40]. Once the proper base pairing between the start codon and the initiator tRNA occurs, the 60S subunit is recruited and the process of elongation can begin [33].

While most eukaryotic mRNAs are translated in the canonical cap-dependent manner, as described above, some eukaryotic messages and many viral RNAs use an internal ribosome entry site (IRES) to recruit ribosomes and initiate translation in a cap-independent manner [45]. This concept will be discussed in greater detail in Chapter 2.

### 1.3 Introduction to translation elongation

The final product after the initiation cycle is complete is an 80S ribosome with a start codon paired to an initiator tRNA in the P site of the ribosome [46]. From here, the ribosome will decode the mRNA in nucleotide triplicates or codons until a stop codon is reached. With the help of the ternary complex composed of elongation factor eEF1A, some power (GTP), and an aminoacyl-tRNA, the ribosome can begin elongation [47,48]. When proper base pairing between the tRNA-anti-codon and the codon on the mRNA happens, peptide bond formation occurs between amino acids in the P and A sites of the ribosome [49]. Simultaneously, a ribosomal conformational change occurs that triggers

the hydrolysis of GTP ultimately transitioning the ribosome and releasing the now GDP-bound ternary complex from the ribosome in a process called translocation [50,51]. The ribosome, now positioned with the next codon in the A site, is ready for aminoacyl-tRNA pairing [52]. This cycle continues – ribosomes decode mRNA in a 5' – 3' direction synthesizing proteins from the amino terminus (N) to the carboxy terminus (C) – until the ribosome reaches a stop codon. When the ribosome reaches the stop codon, termination occurs, ejecting the ribosome and fully synthesized protein from the mRNA.

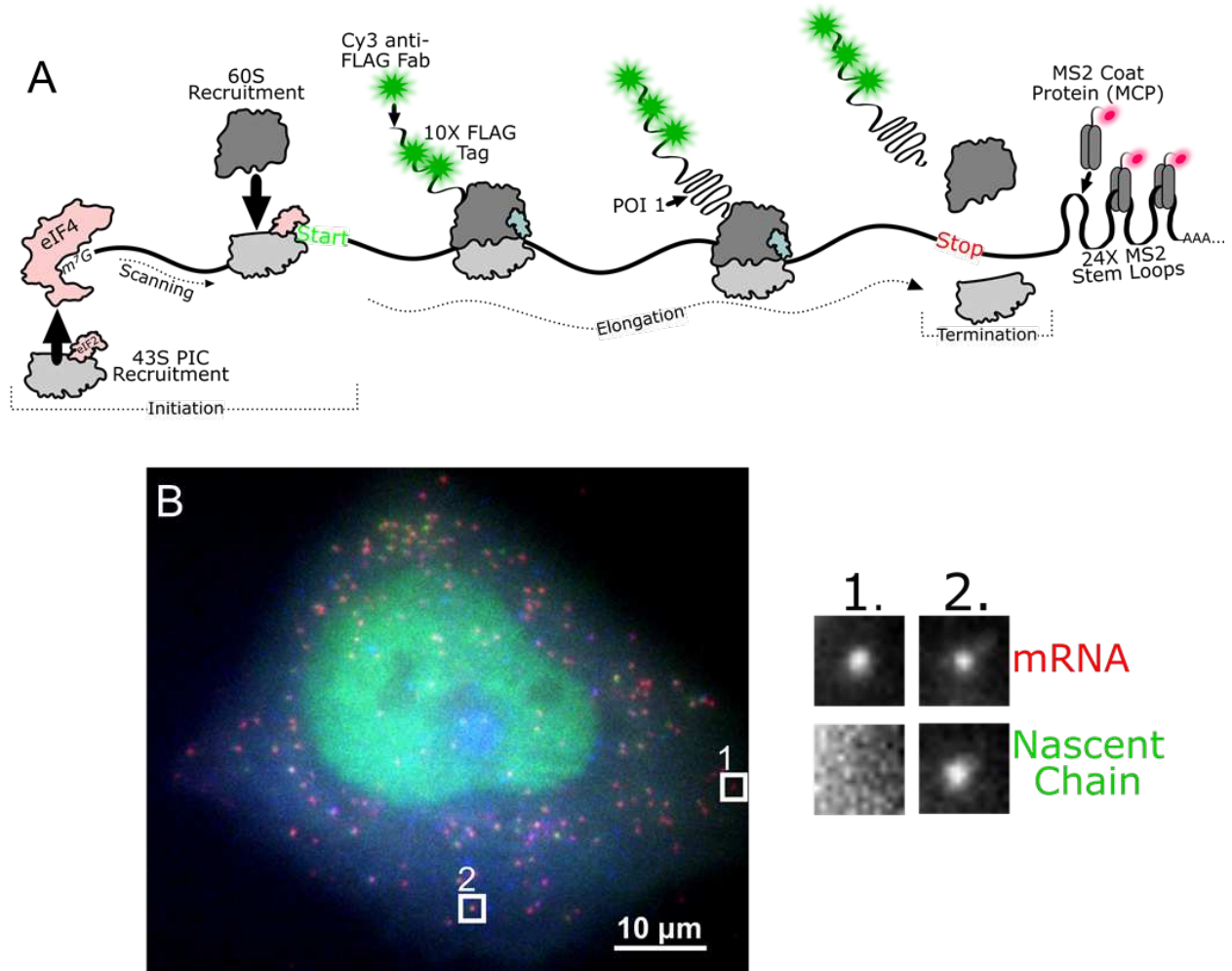
It has been well established that a major source of translational regulation occurs at the initiation stage [27]. However, more evidence is emerging that regulation occurs at the elongation stage as well. Through post-translational modifications of elongation factors, elongation rates can be modulated ultimately changing the amount of protein product output [53]. Even more recently, other regulation pathways have been described that activate when ribosomes collide while elongating [54,55]. This mode of regulation will be discussed in detail in Chapters 3 and 4.

## 1.4 A way to visualize translation in real-time - NCT

Traditionally, translation output and dynamics have been investigated through population studies such as tagging a protein of interest with a fluorescent protein, using radiolabeling, or employing bioluminescence [56]. A newer method, called ribosome profiling has been developed to see how translating ribosomes are distributed across the transcriptome [57,58]. These studies are great at deducing the relative abundance of proteins in different cellular conditions and on average where ribosomes are on transcripts. However, because they rely on population averaging, cell to cell and RNA to

RNA heterogeneity information is lost. To study translation heterogeneity within single living cells, an assay with higher spatio-temporal resolution is necessary.

Luckily, in 2016, five papers were published describing technology to visualize single-molecule events of translation in living cells [59–63]. This technology allows for the quantification of the heterogeneity of translation dynamics from one mRNA to another [64]. To image and quantify single mRNA translation dynamics, fluorescent intrabodies (such as Fab [65], scFv [66,67], or nanobodies[68] ) are required. These intrabodies bind and label repeated epitopes inserted on the N-terminus of a protein of interest. As the protein of interest is translated, the repeated epitopes emerge from the ribosome and are bound within seconds by the fluorescent intrabodies [64,69]. This strategy amplifies fluorescence within translation sites at two levels: first, multiple fluorescent intrabodies can bind the repeated epitopes within a single nascent peptide chain at the same time; second, multiple ribosomes can translate the mRNA in polysomes to produce multiple nascent peptide chains within a single translation site. These two levels of amplification produce bright fluorescent puncta that can be detected with single-molecule precision above background using a sensitive fluorescence microscope (Figure 1.1B). As ribosomes initiate, elongate, and terminate translation, the fluorescence intensity within individual translation sites fluctuates up or down, yielding insights into translation dynamics.



**Figure 1.1: Steps of translation and nascent chain tracking assay.**

(A) Schematic briefly illustrating the steps of translation – initiation, elongation, and termination. Schematic also displays the components needed for the nascent chain tracking (NCT) assay – 10x FLAG Tag at the N-terminus of a protein of interest (POI), Cy3 labeled anti-FLAG Fab, 24x MS2 stem loops in the 3' UTR, and Halo tagged MS2 coat protein (MCP). (B) NCT assay representative U2OS cell. Crops indicate 1. non-translating mRNA and 2. translating mRNA.

## 1.5 New applications for NCT

With this technology in hand, it is now possible to answer detailed questions about the stages of translation and the regulation occurring at each stage (Figure 1.1). For the

rest of this chapter, I will describe some of the recent biological questions that have been answered using the NCT assay and new advances in probe technology.

*The following subsections of chapter 1 are based on the following work (Cialek et al., 2020). I was the co-first author on this work along with C.A.C. Listed here are the author contributions: A.L.K., C.A.C., and T.J.S. wrote the manuscript. G.G. designed and created the figures. A.L.K., C.A.C. and T.J.S. edited the manuscript and figures.*

### 1.5.1 mRNA Structure During Translation

During canonical translation, interactions between initiation factors bound to the 5' cap and poly-A tail have led researchers to believe mRNA exist in a stable closed-loop conformation [70]. Recent single-molecule experiments are beginning to question this model. Specifically, Adivarahan et al. and Khong et al. used smFISH ([71,72]; reviewed in [73]) to precisely measure distances from the 5' to 3' end of single mRNA in untreated and translation inhibited cells. These studies revealed mRNA ends are further apart in translating versus non-translating mRNA species [74,75]. These data suggest either mRNA is not in a closed loop during translation, or a closed-loop conformation is a transient event [76].

### 1.5.2 Ribosome Recruitment Mechanisms

Different mechanisms have evolved to recruit ribosomes to mRNA and facilitate translation. The canonical mechanism employed by most eukaryotic transcripts relies on the Cap Binding Complex (CBC) to recruit the first round of ribosomes after nuclear export [77]. To investigate this at the single-molecule level, Hoek et al. used NCT to image translation in wildtype cells and cells with eIF4E inactivated by ha4EBP1 over-expression



(leaving only CBC for ribosome recruitment) [78]. According to this work, CBC can recruit multiple ribosomes in bursts during early rounds of translation before being replaced by eIF4E. This work highlights how NCT can be combined with inactivation/knock-down assays to dissect the contribution of a specific factor to translational gene regulation.

### 1.5.3 Open Reading Frame Selection

Once the ribosome is recruited to an mRNA, it must scan the 5' untranslated region (UTR) to find a start codon and thereby choose an open reading frame (ORF) [19]. However, some transcripts have multiple ORFs and/or multiple AUGs, so the ribosome must decide where to initiate translation. NCT is now being used to investigate the stochasticity and heterogeneity of this choice. In one study, NCT was used to examine the regulation of the stress-response gene *ATF4*, the transcript of which contains two upstream ORFs [62]. According to this work, the third ORF, which codes for the ATF4 protein, was rarely chosen by ribosomes under normal conditions but was dramatically upregulated in a short-lived burst upon stress. NCT revealed these bursts were initiated in a matter of seconds and lasted for minutes at a time.

More recently, Boersma et al. used the MASH tag to investigate ribosomal start-site selection [68]. In their MASH reporter, nascent chains were labeled in different colors depending on the start site chosen. This revealed a surprising degree of heterogeneity amongst mRNA, with multiple start sites used intermittently and to varying degrees. Different 5' UTRs resulted in different start site preferences, directly demonstrating a role for 5' UTRs in dictating which start site is chosen. In addition, computational fits suggested ribosomes often reinitiate after upstream ORF translation.

#### 1.5.4 Subcellular Localization and Translation

To regulate genes in a spatiotemporal manner, mRNAs are localized to specific cellular compartments for translation. Multiple groups have now used NCT to study translation dynamics in neurons ([61,79–81]; reviewed in [82]). In particular, Wu et al. showed that translation is not necessarily repressed during active mRNA transport in neurons [61]. More recently, Cioni & Lin et al. showed that endosome-associated mRNA are actively translated [80]. Further, blocking the maturation of endosomes with a drug mutating an important endosomal protein inhibited translation without disrupting the mRNA association with endosomes.

Translation can also be targeted to specific subcellular locations. mRNAs encoding membrane and secreted proteins are translated mainly in the endoplasmic reticulum (ER), whereas mRNAs encoding cytosolic proteins are translated mainly in the cytosol. To further study this, Voigt et al. combined ER staining with NCT to track mRNA encoding cytosolic proteins [83]. Surprisingly, they found that a subset of these mRNA were localized to the ER during translation. Furthermore, they showed that mRNA localized to the ER were translated by more ribosomes on average, directly demonstrating subcellular localization can alter translation efficiency.

#### 1.5.5 Heterogeneity in Ribosome Elongation and Frameshifting

The regulation of ribosomal elongation rates is an important form of translation control. For example, ribosomes could pause or stall at specific nucleotide sequences to modulate the folding of nascent chains (Reviewed in [55]). Pausing can also lead to

ribosomal traffic jams that are known to trigger quality control and the unfolded protein response [84]. NCT is now being used to investigate ribosome elongation dynamics in the context of specific mRNA sequences. First, Yan et al. demonstrated that ribosomes pause for extended times at the XBP1 pause site (which is known to induce ribosome pausing [85]), after which they exit together in bursts [63]. More recently, Lyon et al. used the multi-frame tag to visualize frameshifting dynamics at the HIV-1 frameshift sequence [86]. This sequence contains a ribosomal pause site and a slippery sequence that causes ribosomes to occasionally slip from the 0 frame to the -1 frame. NCT revealed frameshifting occurs in bursts and is associated with long pauses at the frameshift sequence that induce ribosomal traffic jams.

## Chapter 2

# Quantifying the dynamics of IRES and cap translation with single-molecule resolution in live cells

## 2.1 Introduction to viral and cellular Internal Ribosome Entry Sites (IRES)

Canonically, ribosomes are recruited to the 5' cap of an mRNA as described in Chapter 1. However, situations exist where the ribosome can be recruited to an mRNA in a non-canonical fashion [45,87]. A specific example of cap-independent translation initiation is called Internal Ribosome Entry Sites (IRES) where the ribosome is recruited to an area on the transcript that is not the 5' cap [88,89].

In 1988, the first IRES elements were discovered in the poliovirus (PV) and encephalomyocarditis virus (EMCV) by laboratories of Nahum Sonenberg and Eckard Wimmer, respectively [90,91]. Both investigators reasoned that since these viral mRNAs are naturally un-capped ribosome recruitment must be occurring in a novel cap-independent manner [92]. Indeed, it was shown by both groups *in vitro* and *in vivo* that the uncapped 5'UTR regions of both PV and EMCV successfully initiated translation.

Since their initial discovery, many other viral IRES elements in viruses such as human hepatitis C (HCV), classic swine fever virus, and cricket paralysis virus (CrPV) were discovered and classified [93,94]. Through sophisticated biochemical assays, it was

found that IRES elements from various viruses adopt different complex three-dimensional structures and require subsets of translation initiation factors to recruit the ribosome [94]. It has since been deduced that some viruses have evolved IRES sequences to hijack host ribosomes during infection when cap-dependent initiation is compromised [95,96]. Though IRES elements carry out the same function in cells – recruit the ribosome independent of the cap – no primary sequence similarity or universal structure has been found between different viral or cellular IRES elements. Regardless, viral IRES elements can be categorized into several distinct classes, each containing unique structural RNA motifs that attract different subsets of host translation factors [88,97,98]. Along with subsets of initiation factors, other endogenous proteins called IRES trans-acting factors or ITAFs have been shown to bind and modulate IRES activity [99].

Not surprisingly, a similar hunt occurred for IRES elements in human cellular mRNA. Though precise mechanisms of recruitment are poorly understood, IRES elements have been detected in mRNAs involved in development, differentiation, cell cycle progression, cell growth, apoptosis, and stress all situations when cap-dependent initiation is repressed [100–102]. Likewise, viruses exploit cap-dependent translation repression during cellular stress to free up host ribosomes. The result is a large pool of host initiation factors and ribosomal subunits that are free to bind and initiate at IRES sequences [103–106].

## 2.2 Biosensor design and rationale

With slight modifications of the NCT assay, described in Chapter 1.5, we designed a biosensor capable of fairly comparing and quantifying IRES and cap translation initiation dynamics in living U2OS cells [107] (Figure 2.1A-B).

Most previous experimental analyses of IRES-mediated translation use bicistronic transcripts encoding an IRES between two reporter proteins, and quantify IRES activity by the ratio of upstream and downstream reporter expression [87,108]. Although bicistronic reporters contain IRES sequences out of context, their inherent one-to-one cap:IRES stoichiometry ensures that both cap and IRES experience nearly identical subcellular environments, making it possible to fairly compare cap-dependent and IRES-mediated translation. Another advantage is they allow precise dissection of the IRES element itself, independent of other compounding factors.

So far, bicistronic reporters have been beneficial for deducing the relative IRES activity in cells hours or days after transfection [87]. However, these studies have lacked the spatiotemporal resolution needed to visualize the sites of IRES translation and quantify translation initiation and elongation kinetics in real-time. This has made it difficult to assess the heterogeneity of IRES-mediated translation among individual RNA or within specific subcellular environments. Methods to study IRES-mediated translation with higher spatiotemporal resolution are needed to precisely understand how IRES-mediated translation differs from cap-dependent translation.

To fairly compare IRES and cap translation, we constructed a bicistronic NCT biosensor that is bound by different probes depending on the manner of translation initiation (Figure 2.1A). Encoded under cap-dependent translation is a lysine demethylase

KDM5B, N-terminally fused to a spaghetti monster tag (SM) encoding 10×FLAG epitopes. The FLAG SM tag is bound by fluorescently conjugated fragments of anti-FLAG antibodies (Fab), allowing visualization of cap translation soon after the first FLAG epitope emerges from the ribosome [59,109]. Encoded under IRES-mediated translation is a Kinesin-like protein Kif18b, N-terminally tagged with 24×SunTag epitopes that are quickly bound by single-chain variable fragments (scFv) fused to GFP as the SunTag epitope emerges from the ribosome [63]. In addition, the biosensor contains 24×MS2 stem-loops in the 3' untranslated region (UTR) that are bound by Halo-tagged MS2 coat proteins (MCP) (Figure 2.1A) [109,110].

We built this construct initially without an IRES element. This construct which we term the no IRES construct can be used not only as a starting point for IRES sequence insertion but also as a ribosomal run-through control. To insert different IRES sequences with ease, the no IRES construct was designed with a restriction enzyme site between the end of KDM5B and the beginning of the SunTag. This design allowed for IRES sequences to be inserted between the two tags via Isothermal Assembly. Though this chapter describes results from a single IRES sequence, other IRES sequences have been successfully inserted.

As a first application, we inserted the IRES element from EMCV into our no IRES biosensor. EMCV is a small single-stranded RNA virus that causes many mammalian diseases [111,112]. The EMCV IRES sequence is 553 nucleotides in length and contains a methionine start codon for the preferred open reading frame (ORF) (Figure 2.1A) [112]. Previous assays have shown the EMCV IRES recruits ribosomes without the need for a 5' cap or many canonical translation initiation factors [91,113,114]. However, little is

known about when, where, and how EMCV IRES-mediated translation occurs in living cells at the single-molecule level.

Because our biosensor fluoresces in three colors, multiple components need to be expressed within cells at appropriate levels. We use a loading technique called bead loading to achieve this with high efficiency and minimal effort [65,115]. With bead loading, various probes -- including the MS2 coat protein used to label mRNA with an MS2 tag, Fab to label FLAG-tagged epitopes, and anti-SunTag GFP-scFv to label SunTag epitopes -- can all be purified and loaded together in a single, simple step [59,107]. We find this approach to be very convenient, enabling the fine-tuning of probe levels, rapid testing, and combinatorial experimentation.

Along with optimizing probe levels, how probes are expressed in cells was also optimized. For example, initially, the anti-SunTag GFP-scFv probe used to label the SunTag epitopes were expressed on a plasmid [63]. Loading this plasmid led to anti-SunTag GFP-scFv concentration variability cell-to-cell. Some cells that were loaded with many copies of anti-SunTag GFP-scFv plasmid expressed the protein at levels that were too high to see translation signals above background. We noticed when anti-SunTag GFP-scFv concentrations are too high in our cells, probe aggregation occurred further adding to the difficulty of detecting and tracking translation sites. Because loading purified MCP and Fab into our cells yielded beautiful results, we set out, with the help of Dr. Mam Scherman, to purify and load anti-SunTag GFP-scFv. Purified protein in hand, we found the optimal loading concentration that resulted in detectable and trackable translation spots.

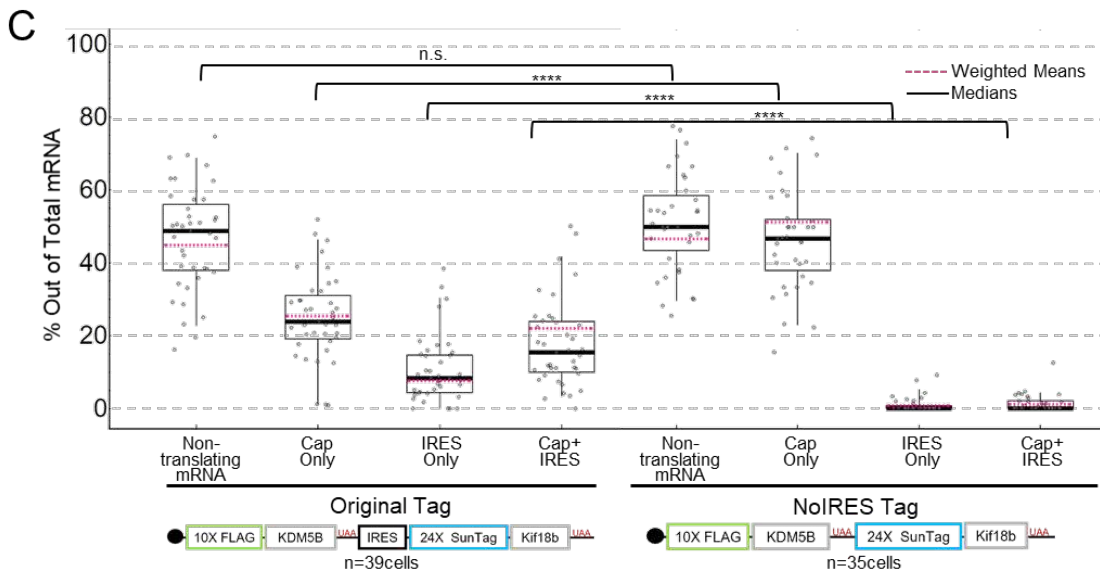
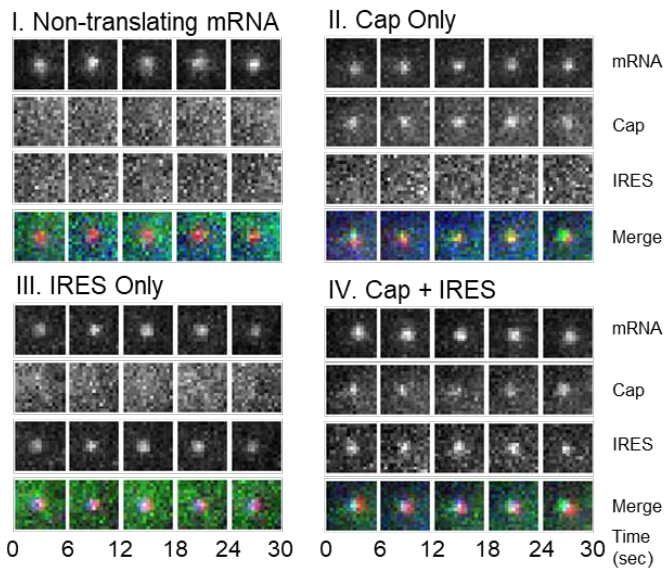
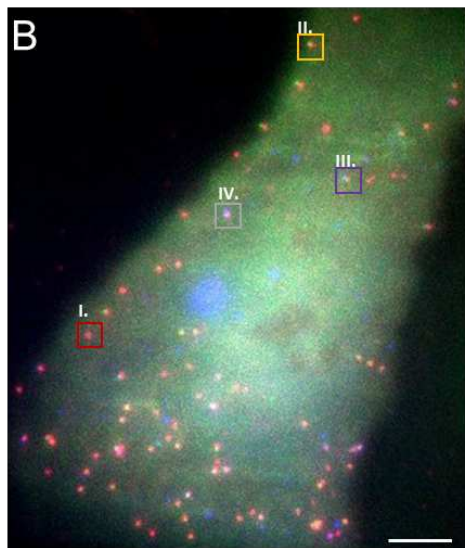
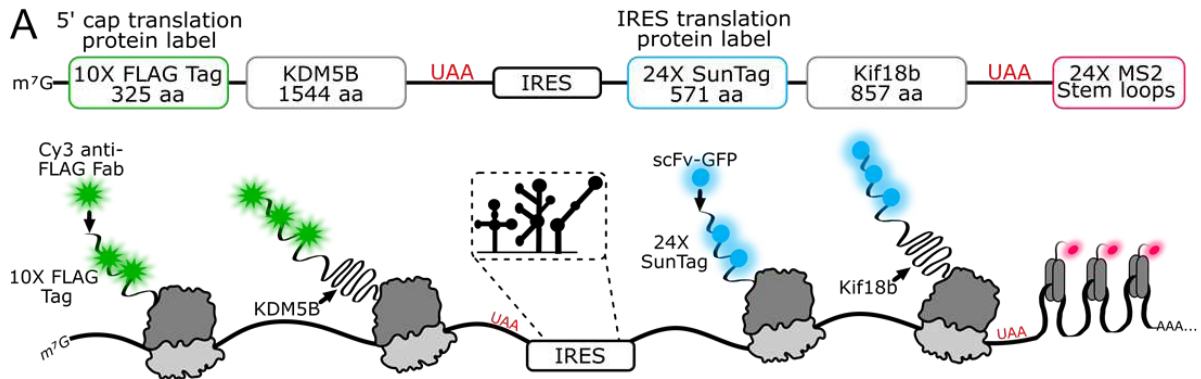


To visualize these aspects of EMCV IRES translation, we bead loaded plasmid DNA encoding our biosensor, along with purified anti-FLAG Cy3-Fab, anti-SunTag GFP-scFv, and HaloTag-MCP into U2OS cells 3-6 hours before imaging. With this combination of probes, translation sites exhibit protein labeled by Fab or scFv co-moving with mRNA labeled by MCP. In addition to non-translating mRNA (Figure 2.1B panel I), we identified translation sites labeled by just Fab (Figure 2.1B panel II), just scFv (Figure 2.1B panel III), and both Fab and scFv (Figure 2.1B panel IV), indicating Cap-only translation, IRES-only translation, and Cap+IRES translation, respectively.

We performed two control studies to confirm that spots were active translation sites. First, to rule out fluorescence bleed-through from the protein channels to the mRNA channel, we repeated experiments without labeling mRNA. Regardless of the intensity of the translation signal, no bleed-through was observed in the mRNA channel (Figure 2.2A-B). All other bleed-through possibilities were ruled out by direct observations of distinct populations of non-translating mRNA, IRES-only translation sites, and Cap-only translation sites. Second, to show that the translation sites were active, we treated cells with 50 ug/mL puromycin, an elongation inhibitor that releases nascent chains from ribosomes, and we confirmed a rapid disappearance of all Fab or scFv translation signals within translation sites (Figure 2.2C) [116].

To better quantify the heterogeneity of translation, we took 2.5-minute movies (25 frames x 13 z-planes per volume x 3 colors = 975 images per movie) of 39 cells from eight experiments. Movies were acquired using HILO illumination to maximize signal-to-noise [117]. In total, we observed 3748 mRNA, of which 1784 were being translated. 24% of mRNA were translated in a Cap-only manner, 8% in an IRES-only manner, and 15%

in a cap- and IRES- manner simultaneously (Cap+IRES) (Figure 2.1C, left). As a control, we imaged the no IRES construct, and we observed an almost complete reduction in anti-SunTag GFP-scFv signals (Figure 2.1C, right), ruling out ribosomal run-through from cap to IRES. All else equal, these data demonstrate that the IRES element alone can capture and initiate host ribosomes, but not as efficiently as the cap in a bicistronic context. Given the large fraction of transcripts we observed being translated in both a cap- and IRES-manner, these data also demonstrate that a single bicistronic transcript can be translated simultaneously in two open reading frames.



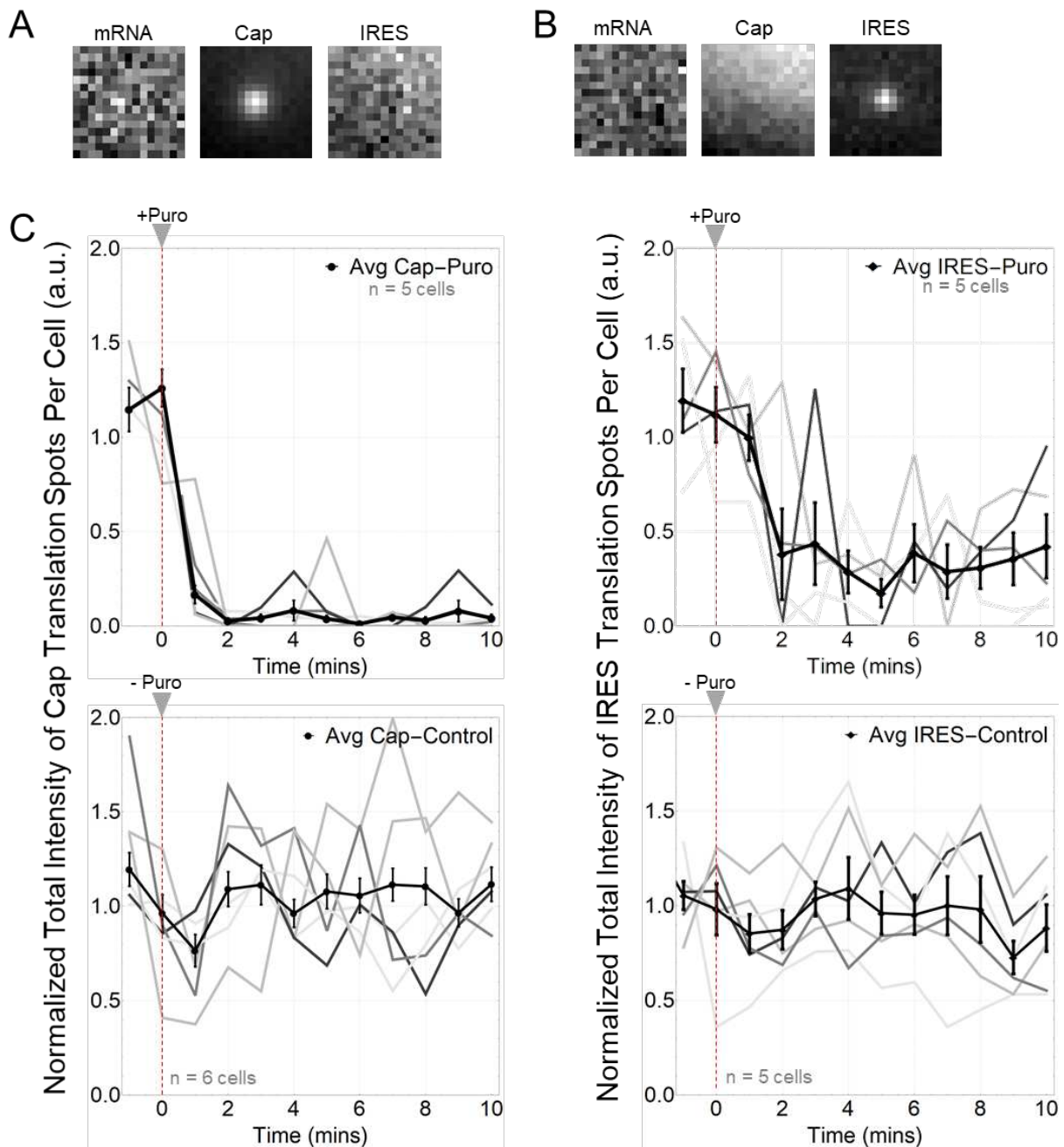
**Figure 2.1 A multicolor biosensor to compare Cap and IRES translation at the single-molecule level in living cells.**

(A) Overview of the construct. (B) Schematic of the system. Cap-dependent protein reporter (green) is labeled by anti-FLAG Fab conjugated to Cy3 that binds the 10× FLAG peptide epitopes in the N-terminus. IRES-mediated protein reporter (blue) is labeled by a GCN4 scFv fused to a GFP that binds the 24× SunTag peptide epitopes. RNA (red) is marked by MCP-Halo labeled with JF646 that binds to repeated MS2 stem loops in the 3' UTR. (C) Left. Representative cell imaged 6 hours after plasmid and probe loading. Different colored boxes within the cell illustrate different types of translation spots seen within a single cell. Right. Examples of co-moving spots. I – non-translating mRNA (red). II – single mRNA translating Cap Only (yellow). III – single mRNA translating IRES Only (purple). IV – single mRNA translating in a Cap and IRES manner (gray). (D) Quantification of species percentages out of total mRNA for both the Original Tag and the noIRES control. Each point represents the percent of that species in a cell. The p-values are based on a two-tailed Mann-Whitney test: \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ . The thick black line indicates the median and the dashed red line represents the weighted (by mRNA/cell) mean, the box indicates the 25-75% range, and the whiskers indicate the 5-95% range.

## 2.3 Sub-cellular localization and size of translation sites in cells

We wondered why there were fewer EMCV IRES translation spots compared to cap. One possibility is the IRES sequence itself requires a specific environment. For example, EMCV IRES-mediated translation is known to occur around replication complexes that localize to specific regions of cells [111]. While this compartmentalization is likely mediated by many factors, the IRES sequence itself could play an active role in the process or require proper localization for translation. To test this, we quantified the propensity for IRES-mediated translation sites to compartmentalize in three ways. We first measured (1) the tendency of IRES-mediated translation sites to cluster and (2) their average distance to the nuclear periphery. This revealed the EMCV IRES has no preference for clustering (Figure 2.3A) and only a minor preference for translation in the perinuclear region (Figure 2.3B). We next tracked individual translation sites to quantify (3) their mobilities. From the 3771 mRNA we tracked, we calculated their cumulative distributions of diffusion coefficients and average mean squared displacements (Figure

2.3C-D). In both analyses, IRES-only translation sites were statistically indistinguishable from cap-only translation sites. According to all metrics, IRES-mediated translation sites localize and move within cells similarly to cap-dependent translation sites. Collectively, these data suggest the EMCV IRES does not require a specialized environment for translation.



**Figure 2.2: Using a translation inhibitor, puromycin, to test for action translation.**

(A-B) Five frame average of a Cap Only and IRES Only translation spot. mRNA marker dye, JF646, was not added to these cells. (C) Top graph shows normalized total intensity over time curves for Cap-dependent and IRES-mediated translation spots, respectively, after addition of puromycin. Gray lines indicate individual cell examples. Thick dark line indicates the average total intensity of all cells. Red dashed line indicates time at which puromycin was added. n=5 cells for both cap and IRES translation. Bottom graph shows normalized total intensity Cap-dependent without the addition of puromycin. n=6 cells for cap translation and n=5 cells for IRES translation.

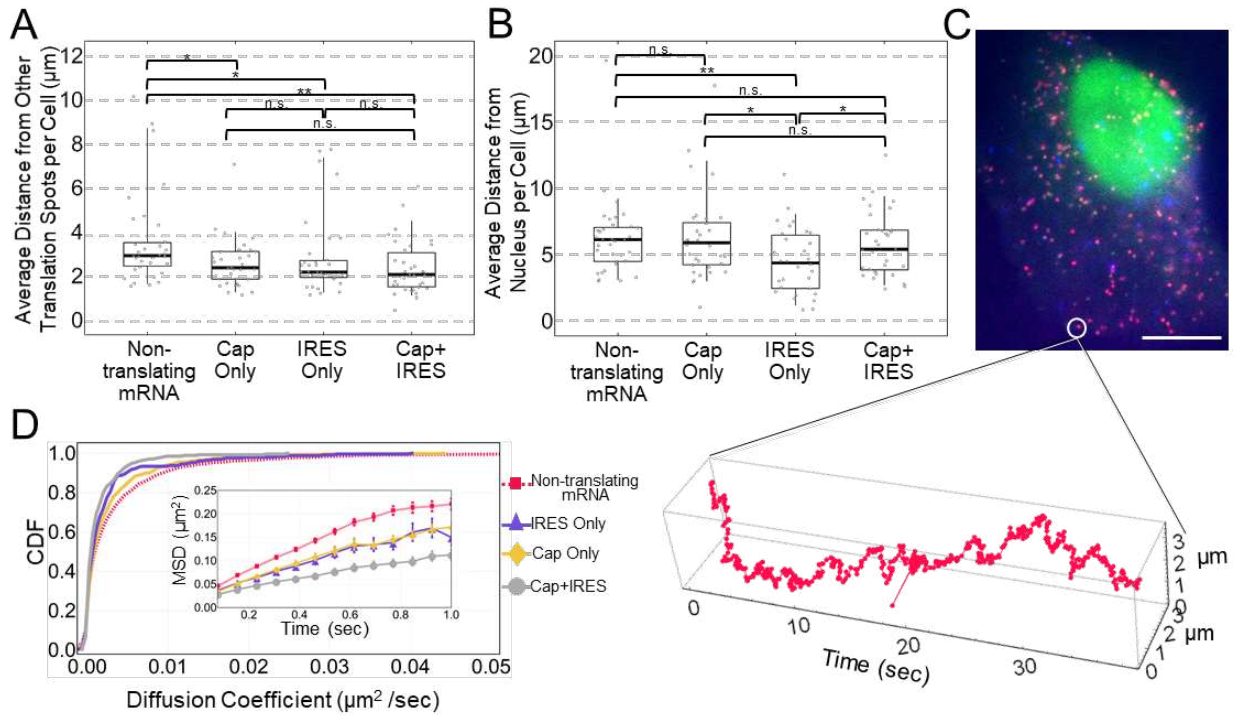
Since we could not distinguish IRES-mediated translation sites based on their subcellular localization or mobility, we wondered if other features could explain the lower probability of observing IRES-mediated translation sites. In particular, we wondered if there were differences in the size of single IRES translation sites compared to single cap translation sites. Recently, the Zenklusen [74] and Parker [75] laboratories used smFISH to show that cap translating ribosomes tend to stretch out translation sites, i.e., heavily translated mRNA cover a greater volume within the cell, in contrast to models of mRNA looping [70]. Because IRES-mediated translation sites do not require looping factors, they could have a different ribosomal organization than cap translation sites [113]. To test this, we took advantage of the long ORFs in our biosensor. The 1D distance from the center of the cap ORF and IRES ORF to the center of the 3'UTR marked by the MS2 signal is 8.5 kb and 3 kb, respectively (Figure 2.1A). Assuming cap and IRES ribosomes initiate stochastically, their average positions provide an approximation for the center of each ORF. This allowed us to measure the distance between the centers of the IRES ORF, cap ORF and 3'UTR, all within single translation sites (Figure 2.4A-B, Figure 2.5).

We began with Cap-only and IRES-only translation sites, to see if they differ in size. In contrast to this, the median distance between cap-dependent ribosomes and the

3'UTR in Cap-only sites was 159 nm, a value statistically indistinguishable from the 149 nm median distance we measured between IRES-mediated ribosomes and the 3'UTR in IRES-only translation sites (Figure 2.4C). This similarity suggests that when the IRES ORF is not being translated (as in Cap-only sites), it is compact. Furthermore, by ranking translation sites by their total intensity (i.e., total ribosomal content or degree of translation), we found that as the brightness of the Cap-only or IRES-only signals increased, the distance between those signals and the MS2 signal marking the 3' UTR also increased (Figure 2.4C). These data corroborate in living cells what the Zenklusen [74] and Parker [75] labs observed in fixed cells; namely, that translating ribosomes tend to stretch out mRNA. However, the similarity in the size of Cap-only and IRES-only translation sites suggests these features alone cannot account for the relatively low probability of IRES-mediated translation.

For completeness, we examined the size of Cap+IRES translation sites (Figure 2.4D). In these sites, the median distance between cap-dependent ribosomes and the 3'UTR was 146 nm, similar to the 159 nm distance we measured in Cap-only translation sites (p-value = 0.42). In contrast, the median distance between IRES-mediated ribosomes and the 3'UTR was just 101 nm, significantly less than the 149 nm distance we measured in IRES-only translation sites (p-value = 6E-6). These data indicate the upstream cap-dependent ribosomes restrict the freedom of downstream IRES-mediated ribosomes, causing them to spread out less. Here it makes sense that IRES-mediated ribosomes tend to be closer to the 3'UTR than cap-dependent ribosomes because the IRES itself is considerably 3' of the cap. Finally, we again confirmed that as the brightness

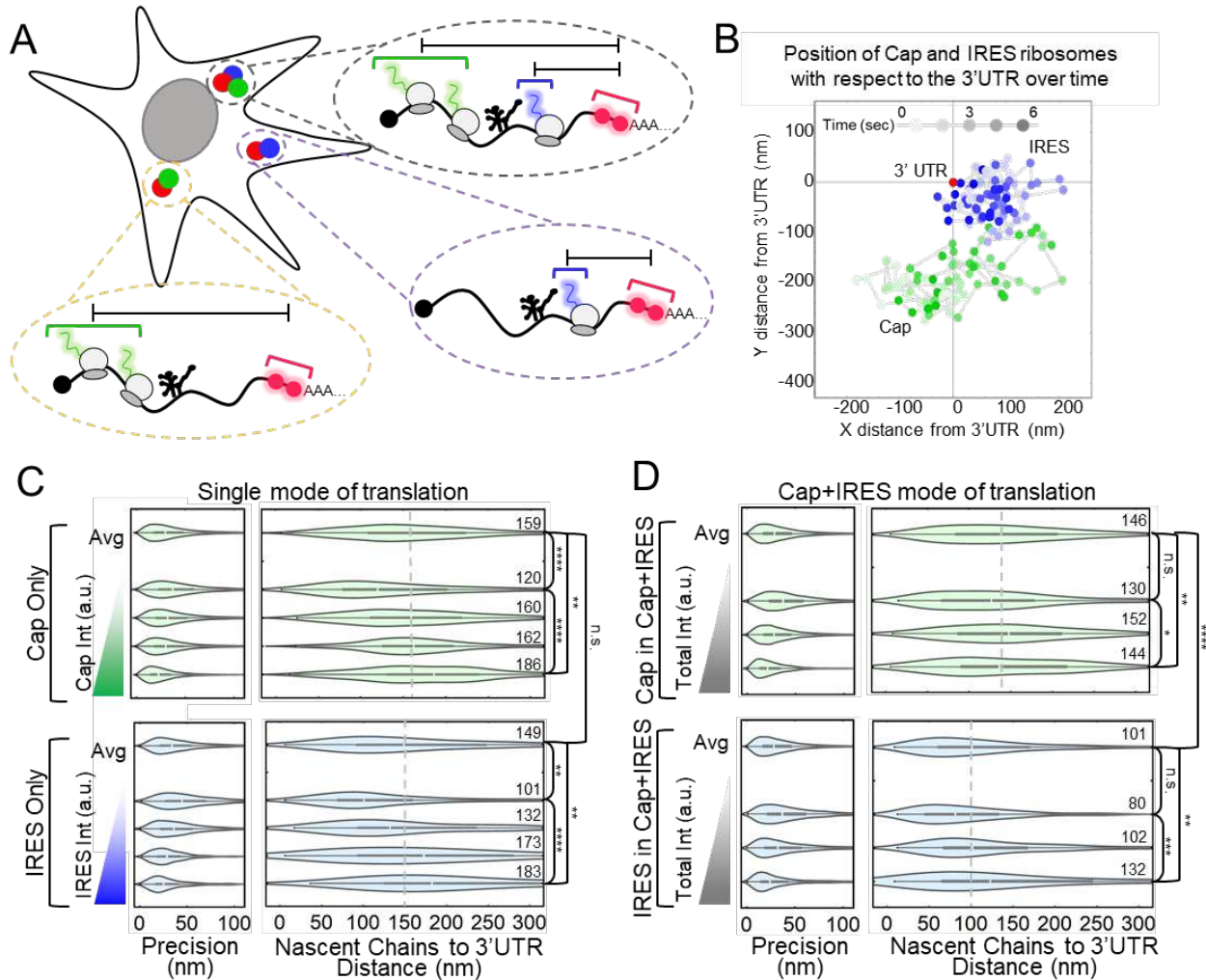
of translation signals increased, the distances between these signals and the MS2 signal marking the 3' UTR also increased (Figure 2.4D).



**Figure 2.3: Quantifying the biophysical properties of Cap and IRES translation sites.**

(A) Quantification of translating and non-translating mRNA distances in  $\mu\text{m}$  to translation spots within single cells. Each point represents the average distance per cell. (B) Quantification of distance in  $\mu\text{m}$  from the nucleus of translating and non-translating mRNA. Each point represents the average distance from the nucleus per cell. (C) Representative cell imaged with fast imaging conditions. An example mRNA is highlighted with white circle and a track through time of that mRNA is graphed below. (D) Cumulative distribution function plot of mRNA Only (red), IRES Only (purple), Cap Only (yellow), and Cap+IRES (gray) species based on their diffusion coefficients ( $\mu\text{m}^2/\text{sec}$ ). Inset shows the Mean Square Displacements of the different species over time in seconds.  $n=3771$  total tracked mRNA,  $n=11$  cells.

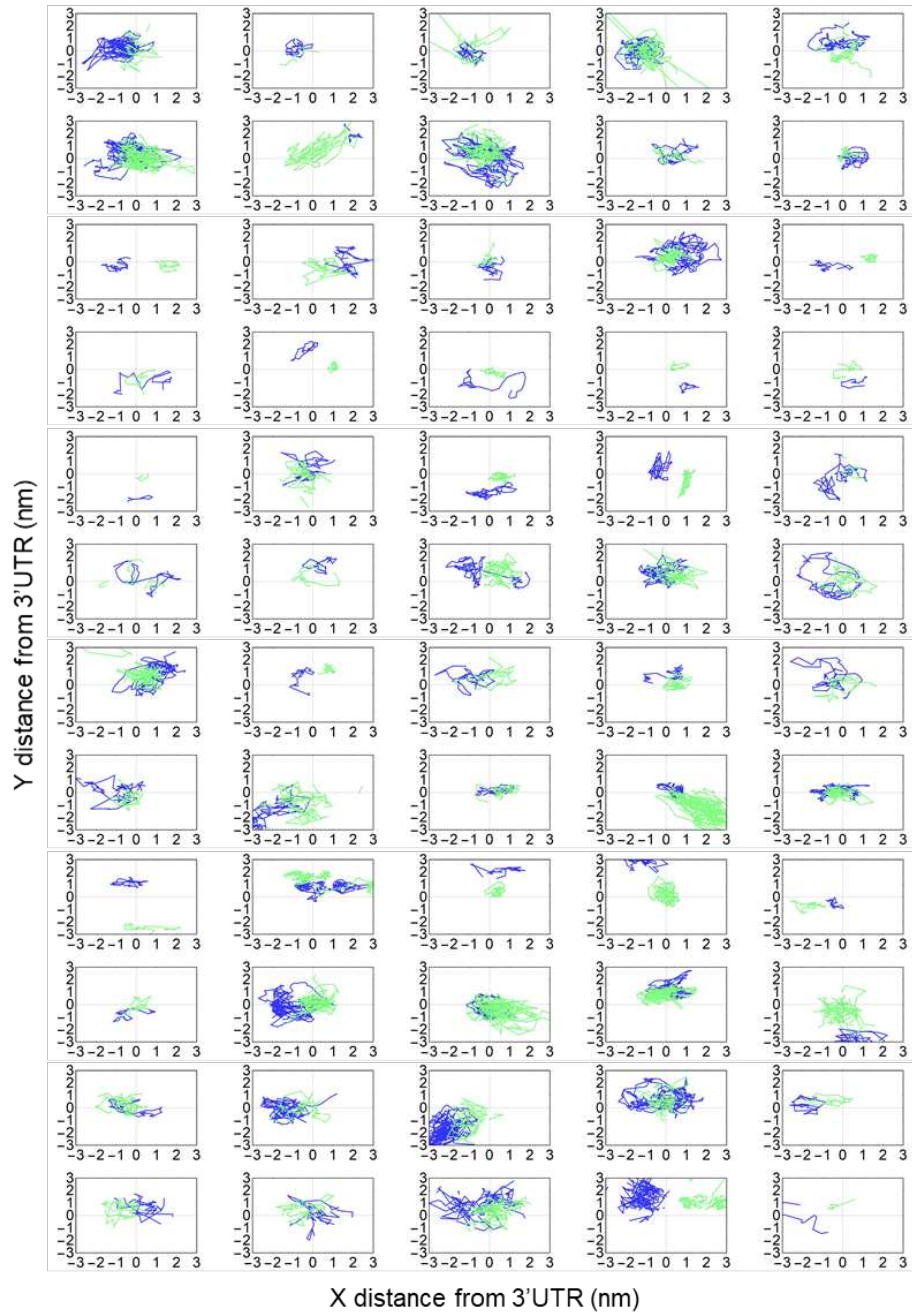
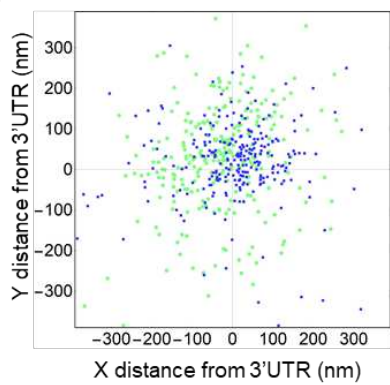




**Figure 2.4: Measuring distances from the 3'UTR to Cap and IRES nascent chains**

(A) Schematic showing how the measurements from the 3'UTR to the Cap and IRES nascent chains were conducted.

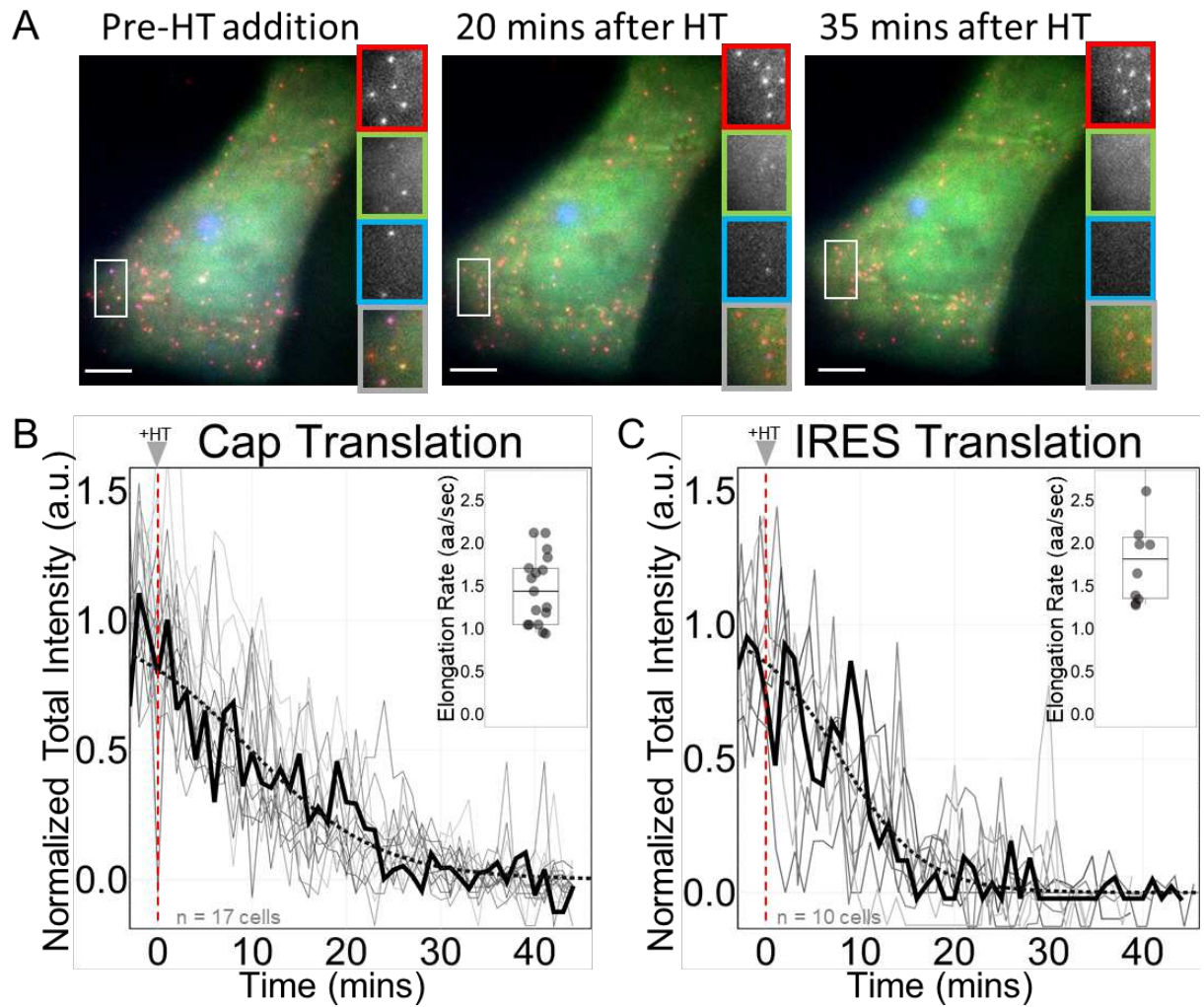
(B) Graph showing IRES and Cap nascent chain positions relative to 3'UTR through time of a representative Cap+IRES translation spot. X and Y distances displayed in nm and time represented as gradient of either blue or green. (C) Cap in Cap+IRES spots distances from the Cap nascent chains to the 3'UTR averaged across all intensities and sorted by total intensity. IRES in Cap+IRES spots distances from the IRES nascent chains to the 3'UTR averaged across all intensities and sorted by total intensity. Brightest (bottom) and dimmest (top). (D) Cap in Cap Only spots distances from the 3'UTR averaged across cap intensities and sorted by increasing intensity. IRES in IRES Only spots distances from the 3'UTR averaged across cap intensities and sorted by increasing intensity. Light gray line marks the median averaged distances for Cap and IRES. For the box and whisker plot, the white line indicates the medians, the box indicates the 25%-75% range, and the whiskers indicate the 5%-95% range.

**A****B**

**Figure 2.5: Measuring distances between Cap and IRES nascent chains in Cap+IRES translation spots**  
(A) Measured distances of Cap and IRES nascent chains to 3'UTR through time in representative single Cap+IRES translation tracks.  
(B) Median distances of Cap and IRES nascent chains to 3'UTR of each Cap+IRES track. Distances are measured in nanometers (nm). 3' UTR coordinates were fixed at (0,0) for all analysis.

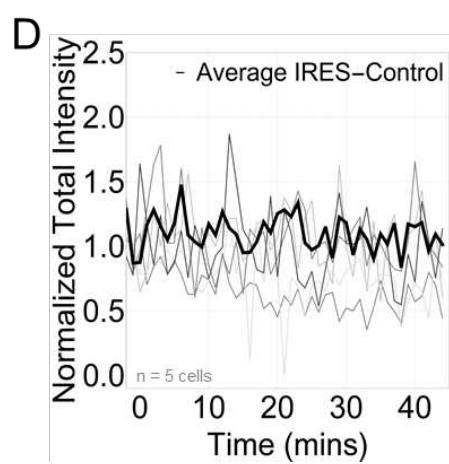
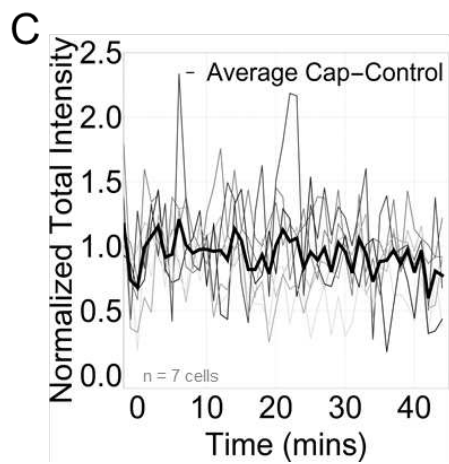
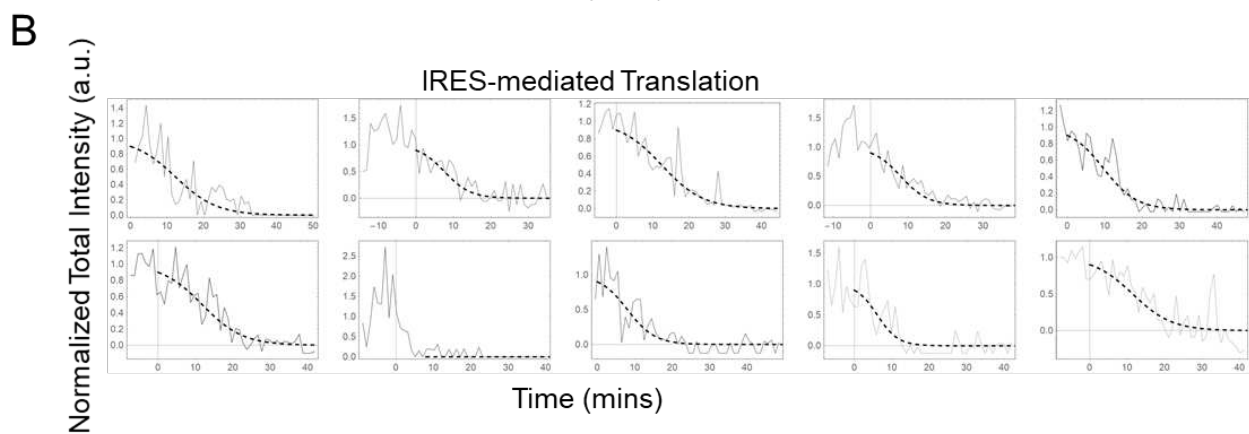
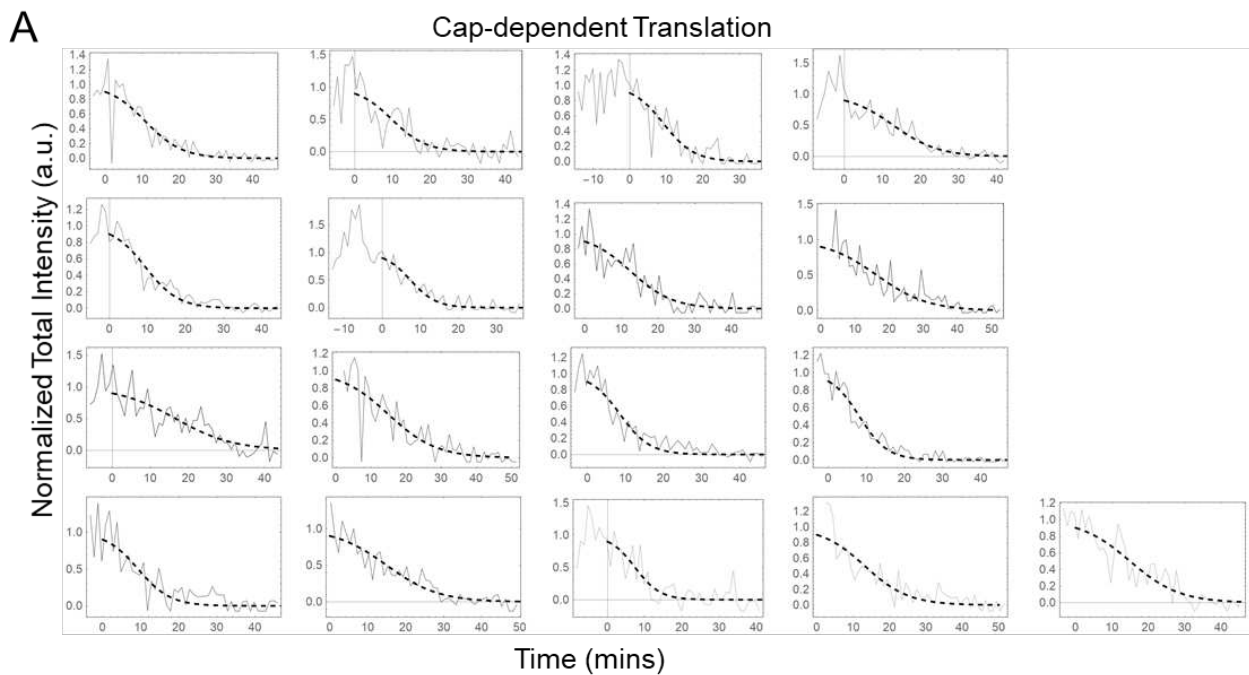
## 2.4 Quantifying the elongation, ribosome recruitment, and initiation dynamics of cap and IRES translation

Despite our inability to detect any clear distinguishing features that predict IRES-mediated translation, the simple observation that it is rare compared to cap-dependent translation (Figure 2.1C) suggested either the IRES is less efficient than the cap at recruiting and initiating ribosomes or the IRES-mediated ORF is translated more quickly than that of the cap (which would lead to fewer ribosomes along the ORF at any given time). To rule out the latter possibility, we estimated cap-dependent and IRES-mediated elongation rates by measuring the time it takes ribosomes to run-off the biosensor after translation initiation is blocked by harringtonine (HT) [118]. Consistent with ribosomal run-off, harringtonine led to a steady decay in the nascent chain signals within translation sites (Figure 2.6 and Figure 2.7A-B). This decay was not due to photobleaching because it was not observed in untreated control cells (Figure 2.7C-D). Fitting the linear portion of the intensity decay curve gave elongation rates of  $1.44 \pm 0.40$  aa/sec for cap-dependent translation and  $1.81 \pm 2.39$  aa/sec for IRES-mediated translation, corresponding to run-off times of 45 min and 43 min, respectively. The similarity of these rates and run-off times demonstrates elongation is not responsible for the lower number of IRES-mediated translation sites compared to cap-dependent translation sites.



**Figure 2.6: Calculating the elongation rates of Cap-dependent and IRES-mediated translation.**

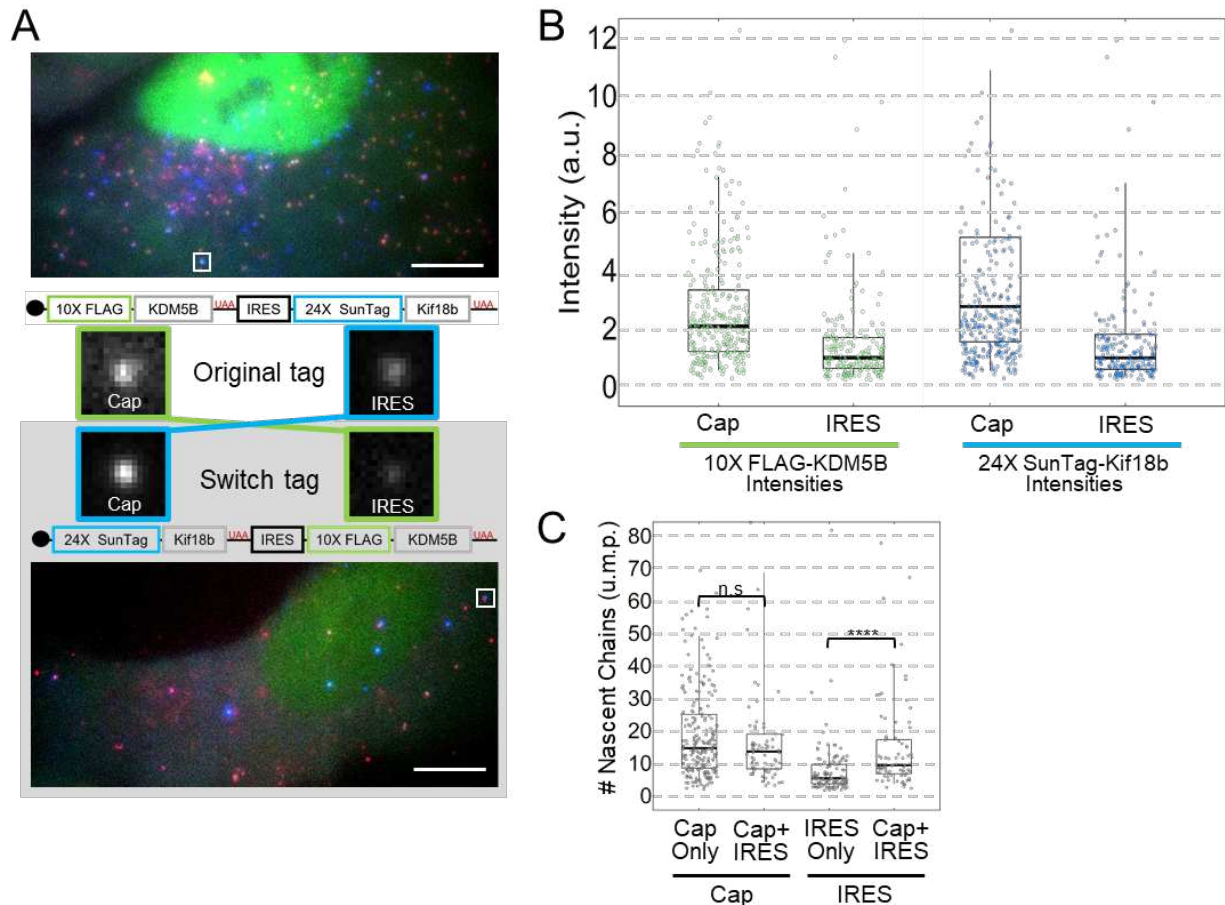
(A) Cells before HT addition, 20 minutes after HT addition, and 35 minutes after HT addition with crops of mRNA channel (red), Cap channel (green), IRES channel (blue), and merge (black). (B-C) Normalized total intensity decay curves of Cap-dependent and IRES-mediated translation, respectively. Each gray line represents a single cell treated with HT. The black line is a representative cell. The dotted black line is the Tanx fit of the representative cell. The inset is the calculated Cap elongation rates of each cell. All cells were imaged for 45 minutes with a 1-minute interval between each capture. Intensity values are in arbitrary units (a.u)



**Figure 2.7: Ribosomal run-off curves from single cells after addition of harringtonine.**

(A) Harringtonine-induced ribosomal run-off curves from single cells. Each curve shows the decay in nascent chain signal intensity from all Cap-dependent and (B) IRES-mediated translation sites within a single cell post-harringtonine. Run-off curves were phenomenologically fit to a Tanh function to align curves in time for averaging in Figure 4. The slope of fitted curves at a normalized intensity value of 0.5 was used to estimate the elongation rate. (C) Cap-dependent (n=7 cells) and (D) IRES-mediated (n=5 cells) translation controls in which no drugs were added. Each gray line shows the total nascent chain signal intensity from all translation sites in an individual cell. The thick black line is a representative cell. Intensity in arbitrary units (a.u.). All cells were imaged for 45 minutes with a 1-minute interval between each capture.

Having demonstrated elongation does not distinguish cap and IRES translation sites, we turned our attention to ribosome recruitment and translation initiation. Assuming one of these factors limits IRES-mediated translation, we would predict fewer IRES-mediated ribosomes than cap-dependent ribosomes in single translation sites. To test this prediction, we needed to fairly compare the intensities of nascent chain signals within single translation sites. A direct comparison was not possible because the cap-dependent and IRES-mediated nascent chains differ in sequence, have a different number of tags, and are labeled by complementary fluorophores and probes that have different binding kinetics and different quantum efficiencies. To enable a fairer comparison, we developed a Switch Tag in which the reporters were swapped (Figure 2.8A). This allowed us to compare the exact same reporter under the control of both the cap (in the Original Tag, for example) and the IRES (in the Switch Tag). In this way, we could ensure any differences in the intensity of translation sites would reflect differences in ribosome recruitment and/or initiation dynamics rather than differences in fluorophore/probe detection kinetics and/or codon biases within epitope tags.



**Figure 2.8: The cap recruits and initiates 2-3 times more ribosomes than the IRES.**

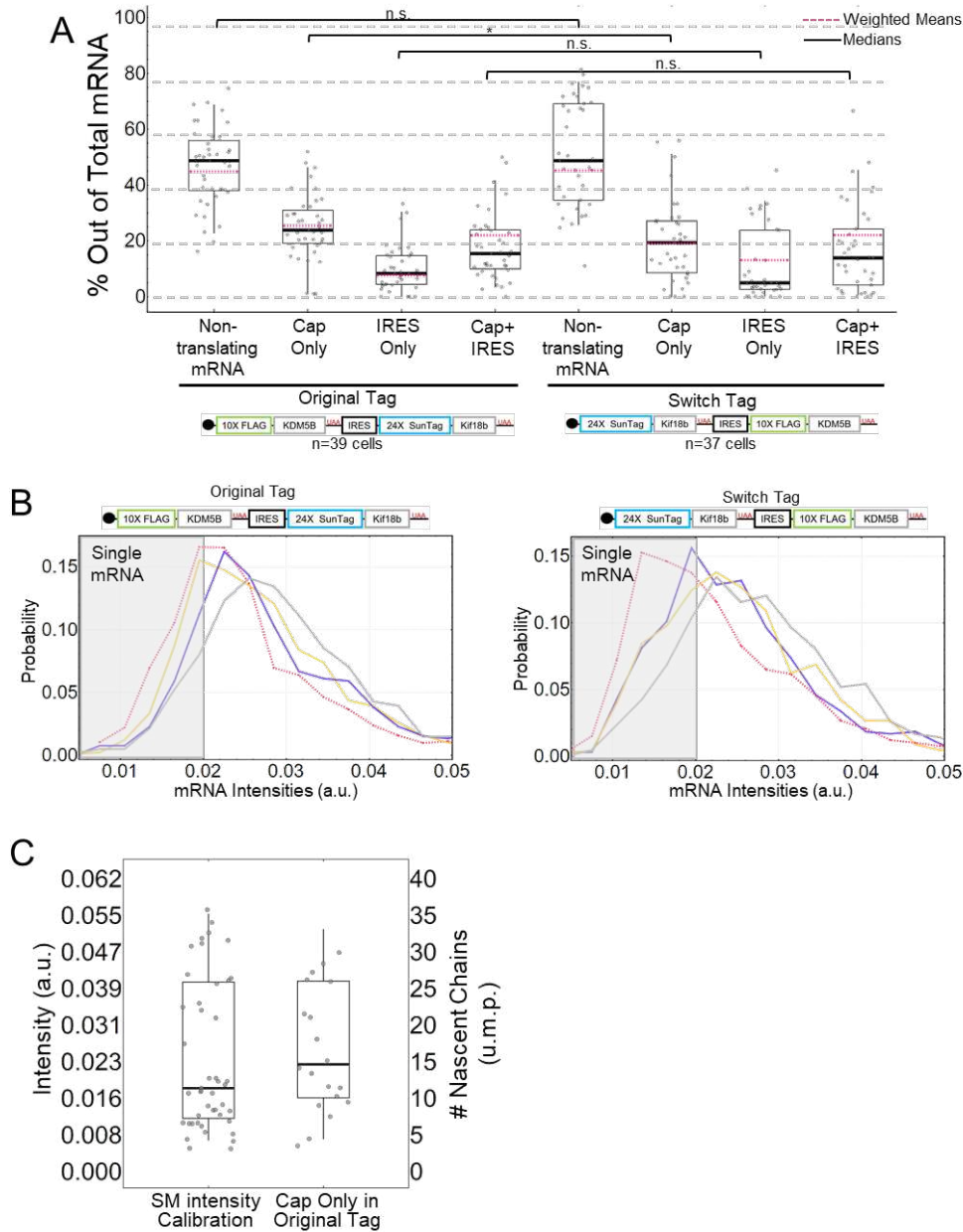
(A) One representative cell expressing the Original Tag (from  $n=39$  cells) (top) or the Switch Tag (from  $n=37$  cells) (bottom) with a Cap+IRES translation spot highlighted by the white square. Scale bars are  $10\mu\text{m}$ . Crops of the representative sites are shown in the middle. The construct schematic with the corresponding crop illustrates how the intensity comparisons between Cap and IRES were conducted. (B) Box and Whisker plots showing the intensity comparisons between Cap and IRES. Left graphs shows intensity comparisons of  $10\times$  FLAG-KDM5B nascent chain signals from Cap in the Original Tag ( $n=302$  spots) and IRES in the Switch Tag ( $n=167$  spots). Right graphs shows intensity comparisons of  $24\times$  SunTag-Kif18b from Cap in the Switch Tag ( $n=262$  spots) and IRES in the Original Tag ( $n=201$  spots). Intensity measurements are in arbitrary units (a.u.). (C) Intensities of Cap in Cap Only (Original Tag) translation sites were compared to Cap in Cap+IRES (Original Tag), IRES in IRES Only (Switch Tag) and IRES in Cap+IRES (Switch Tag) to obtain numbers of ribosomes in units of mature protein (u.m.p.) on all types of translating species. Cap Only sites ( $n=226$  spots) have a median of 14.6 ribosomes, Cap+IRES sites ( $n=76$  spots) have a median of 13.6 Cap-dependent ribosomes and 9.4 IRES-mediated ribosomes. IRES Only sites ( $n=121$  spots) have a median of 5.4 ribosomes. The p-values are based on a two-tailed Mann-Whitney test: \* $p<0.05$ , \*\* $p<0.01$ , \*\*\* $p<0.001$ , \*\*\*\* $p<0.0001$ . For the box and whisker plots, the thick black lines indicate the medians, the boxes indicate the 25-75% range, and the whiskers indicate the 5-95% range.

Reassuringly, when the Switch Tag was expressed in cells, it yielded nearly the same percentages of each type of translation site as the Original Tag (Figure 2.9A), demonstrating the 10×FLAG and 24×SunTag reporters do not interfere with translation dynamics and have similar detection efficiencies. As expected, there were notable differences in the intensities of translation sites. A direct comparison of the intensity of translation sites encoding 10×FLAG-KDM5B initiated in a cap-dependent manner (from the Original Tag) versus an IRES-mediated manner (from the Switch Tag) gave a median intensity ratio of  $2.1 \pm 0.1$  (Figure 2.8B). Similarly, a direct comparison of the intensity of translation sites encoding 24×SunTag-Kif18b initiated in a cap-dependent manner (from the Switch Tag) versus an IRES-mediated manner (from the Original Tag) gave a median intensity ratio of  $2.8 \pm 0.2$  (Figure 2.8B). Here we restricted our analysis to translation sites with relatively dim RNA signals to eliminate complications that could arise from co-translational mRNA clustering (Figure 2.9B) [111]. The similarity of the ratios we measured indicates the presence of between two and three times more ribosomes in cap-dependent translation sites compared to IRES-mediated translation sites. In other words, for every IRES-mediated ribosome on the biosensor, there are two to three cap-dependent ribosomes.

To extend this measurement and obtain absolute ribosome occupancies, we developed a 10×FLAG calibration construct that produces translation sites containing approximately 11 ribosomes (Figure 2.9C). Comparing the intensity of these translation sites to 10×FLAG-KDM5B translation sites in the Original and Switch Tags revealed that Cap-only translation sites have a median of 14.6 ribosomes (Figure 2.8C), while Cap+IRES translation sites have 13.6 cap-dependent ribosomes (p-value = 0.196) and



9.4 IRES-mediated ribosomes, and IRES-only translation sites have just 5.4 ribosomes (p-value =  $5.83E-8$ ) (Figure 2.8C). Thus, cap-dependent translation sites have more ribosomes than IRES-mediated translation sites, consistent with the higher percentage of mRNA translated in a cap-dependent versus IRES-mediated manner. These data demonstrate that cap-dependent translation is overall more efficient than IRES-mediated translation in our biosensor, both at the population level and single-molecule level.



**Figure 2.9: Original Tag comparison to Switch Tag, single mRNA selection, and polysome intensity calibrations.**

(A) Quantification of the percentages of each type of translation sites for the Original Tag (left, n=39 cells) and the Switch Tag (right, n=37 cells). Each point represents a single-cell measurement. (B) Probability histograms showing distributions of mRNA intensities of non-translating mRNA (Red), Cap Only (Yellow), IRES Only (Purple), and Cap+IRES (Gray) translation sites for the Original Tag and the Switch Tag. The gray boxes represent the mRNA intensity threshold used to eliminate multiple mRNAs. Intensities in arbitrary units (a.u.). (C) Translation site calibration measurements. The intensities of Cap in Cap Only translation sites (n=20spots) in the Original Tag were compared to a 10xFlag calibration system (n=47spots) with a known number of ribosomes. These comparisons lead to a calculated number of 14.6 ribosomes in Cap Only translation sites using the Original Tag. For the box and whisker plots, the thick black lines indicate the medians (A and C), and the dashed red line indicate the weighted means (A) the boxes indicate the 25%-75% range, and the whiskers indicate the 5%-95% range. The p-values are based on a two-tailed Mann-Whitney test: \*p<0.05, \*\*p<0.01, \*\*\*p<0.001, \*\*\*\*p<0.0001.

## 2.5 Summary of computational modeling

According to our experiments, the relatively low probability of IRES-mediated translation is due to rate-limiting steps that precede elongation, presumably either ribosome recruitment or initiation. To distinguish these possibilities, we collaborated with Dr. Brian Munsy and his post-doctorate researcher, Dr. Luis Aguilera who developed a set of models with varying levels of complexity (See Appendix A for more information). Models differ in the number of states an mRNA can transition between: Three-state models include an inactive mRNA state (OFF), an active mRNA state that allows cap translation (Cap-ON), and an active mRNA that allows IRES translation (IRES-ON). Four-state models include an additional active mRNA state (Cap+IRES-ON) that allows both cap and IRES-translation (Appendix A).

According to these fitted parameters, the efficiency of IRES translation is not limited by initiation (since  $k_{INIT-I} \sim k_{INIT-C} \sim 1/20 \text{ sec}^{-1}$ ), but rather the IRES spends less time in a translationally active state that can recruit ribosomes. In addition to the above seven parameters (which can be reduced to six by setting the cap and IRES initiation rates equal), one additional parameter was required to fit the data: an enhancement in IRES activation when cap translation is on (i.e.  $k'_{ON-I} > k_{ON-I}$ ). Specifically, in the presence of cap, the IRES refractory period is reduced from 91.3 min to 11 min, leading to a 6.9 fold increase in IRES translation. This enhancement was required to capture the large percentage of Cap+IRES translation sites (which is greater than one would predict if cap and IRES translation were independent) and the larger number of IRES-mediated ribosomes in Cap+IRES translation sites compared to IRES-only translation sites. These

data and the best-fit model therefore provide evidence that translation of an upstream ORF can positively impact translation of a non-overlapping downstream ORF.

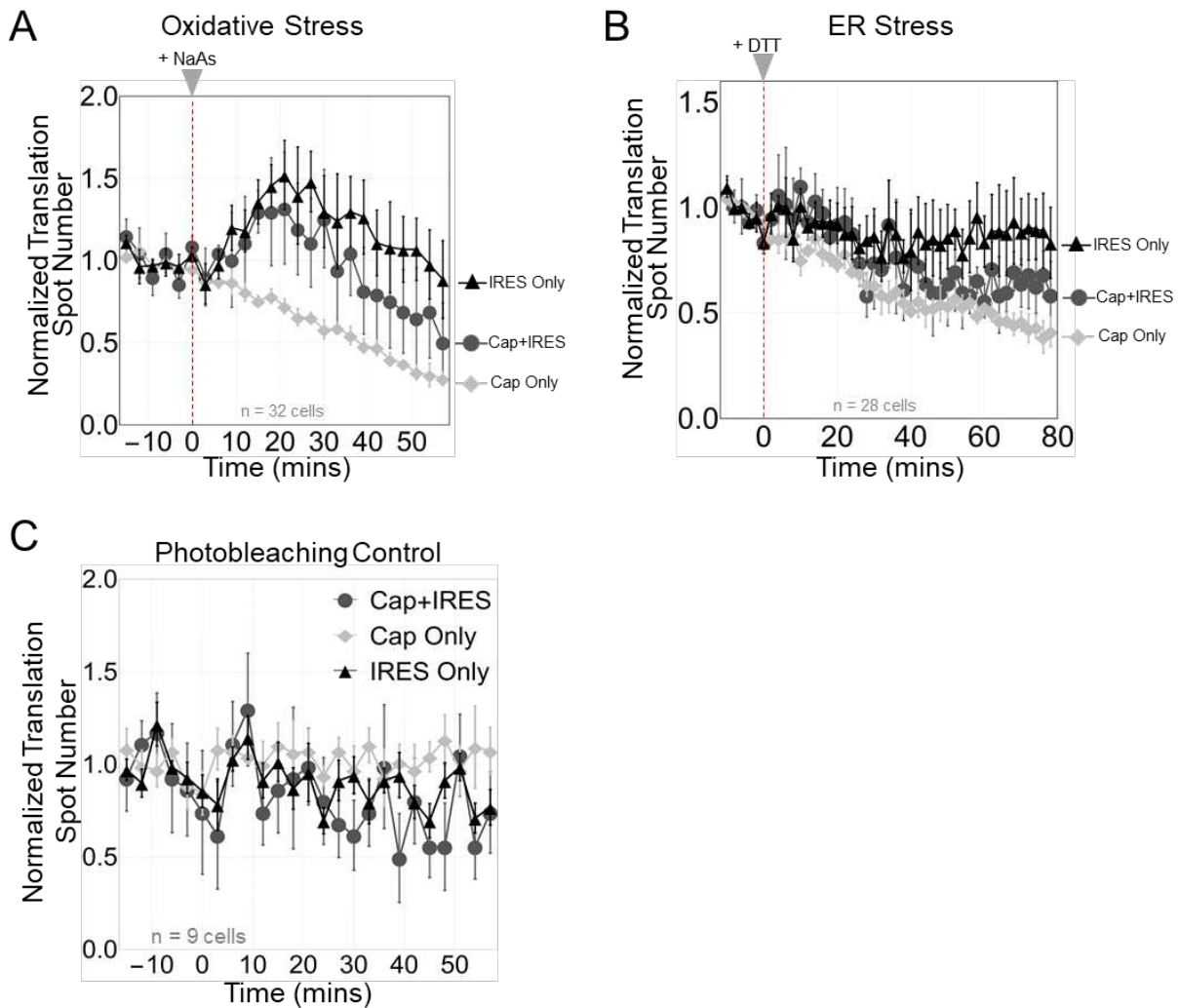
## 2.6 Cap and IRES translation response to specific cellular stresses

Cells encounter a range of stresses that activate an intrinsic network of adaptive changes. Because the process of protein synthesis is energetically costly, cells will divert energy away from protein synthesis and towards adapting to the stress [119]. Therefore, the shutdown of translation initiation is a key feature of cellular stress. Two translation initiation factors, eIF4E and eIF2, are the main targets of translation regulation during stress [120]. As most eukaryotic transcripts rely on these factors for translation initiation, interfering with these two factors has a robust outcome on translation shutdown [121].

It is well known that viral infections cause increased levels of cellular stress. During these stressful conditions, viral RNAs continue to be translated while canonical translation decreases globally [104]. This is thought to occur in part through IRES-mediated ribosomal recruitment to viral RNAs. Extensive studies have shown that both viral and some endogenous mRNAs remain translationally active during certain types of cellular stress in part because IRES sequences do not entirely rely on the action of eIF4E or eIF2 for translation initiation to occur [103,122,123].

To visualize the impact of stress on cap-dependent and IRES-mediated translation at the single-molecule level, we exposed cells expressing our biosensor to sodium arsenite (NaAs) [124], which induces oxidative stress, and DTT [125], which induces ER stress. In both cases, the intensity of Cap-only translation sites decreased significantly

upon stress. In contrast, the intensity of IRES-only translation sites remained steady or increased, while Cap+IRES translation sites displayed an intermediate response (Figure 2.10A-B). This response was not due to photobleaching or photoactivation as our no-drug control did not exhibit the same phenotypes as with drug addition (Figure 2.10C). These data demonstrate that IRES-mediated translation is generally more robust in response to cellular stress than cap-dependent translation, as would be necessary for efficient viral replication in cells during infection.



**Figure 2.10: Cellular stress addition changes the translational landscape.**

(A) Addition of sodium arsenite (NaAs) an oxidative cell stress. Addition of 0.5 mM NaAs was added after 5 imaging time-points (red dotted line). Cells were imaged with a 180 second interval between every 13 frames (one entire cell volume) for 35 total time-points. (B) Addition of 0.75 mM Dithiothreitol (DTT) after 5 imaging time-points (red dotted line). Cells were imaged with a 120 second interval every 13 frames (one entire cell volume) for 35 total time-points. (C) Photobleaching control. Cells were imaged with a 180 second interval between every 13 frames (one cell volume) for 35 total time-points. Nothing was added at time-point 0. Translation site number was quantified in each frame from each cell then totaled. Average spot number was calculated for the first five time-points before drug addition for each type of translation and used as a normalization. Errors bars are the mean and SEM. NaAs: n=32 cells. DTT: n=28 cells. Photobleaching control: n= 9 cells.

## 2.7 Discussion

The ability to track the translation of single mRNA using Nascent Chain Tracking (NCT) makes it possible to directly visualize the heterogeneity of protein expression within cells from one mRNA to another, which would not be possible with standard protein expression assays based on the detection of GFP or luciferase. Thus far, NCT has been used to investigate translation that is initiated in a cap-dependent manner, the predominant form of translation initiation used by eukaryotes. Here, we extend NCT to simultaneously investigate IRES-mediated translation, a mode of translation exploited almost exclusively by viruses to hijack host translation machinery and efficiently replicate in infected cells. By creating a single-molecule bicistronic biosensor that lights up in different colors depending on whether translation is initiated at the cap or the IRES, we quantify precisely when, where, and to what degree IRES sequences hijack ribosomes within living human cells.

One of the hallmarks of IRES-mediated translation is that it depends on only a subset of host translation factors [97,126]. Due to the lax initiation factor requirements of

IRES-mediated translation compared to cap-dependent translation, one would think that translation of the IRES could occur in special microenvironments that, for example, are enriched or depleted in specific translation factors. In contrast to this notion, however, we find little evidence for specialized microenvironments that support IRES translation. Instead, our data suggest IRES-mediated translation sites are biophysically difficult to distinguish from cap-dependent sites, having roughly the same translation initiation and elongation rates, similar mobilities, sizes, and spatial distributions within cells, and similar propensities to cluster near other translation sites. This overall similarity may have evolved to allow the EMCV IRES to effectively compete with the cap for host ribosomes during infection.

According to our results, the EMCV IRES is not as efficient as the cap mainly because it spends less time in a state that is conducive to ribosome recruitment (computational modeling results can be found in Appendix A). Our observation that burst frequencies are modulated to control translational output is reminiscent of common burst control mechanisms of transcription, and this general principle of regulation could be the natural result of sharing a common subset of initiation and elongation factors [127]. Indeed, according to our best-fit model, bursts of IRES translation are both shorter in duration (2.5 min for IRES versus 8.3 min for cap) and separated by longer periods of inactivity (91.3 min for IRES versus 34.5 min for cap) than bursts of cap translation (computational modeling results can be found in Appendix A). Given the complex secondary and tertiary structure of the EMCV IRES [97], which presumably undergoes dynamic conformational changes in living cells, our results suggest the IRES has trouble adopting and maintaining a conformation that can recruit ribosomes and maintain active

translation. In contrast, the cap relies on a larger set of factors, including the cap-binding protein eIF4E and scaffolding protein eIF4G. Presumably, these additional factors work together to better maintain a conformation that is attractive to ribosomes and more amenable to continuous translation.

One of the most interesting observations with our biosensor was that cap translation enhances that of the IRES, but not the other way around (Figure 2.8C). While surprising, this does make sense in hindsight given the subset of factors the IRES requires compared to the cap. In particular, when cap translation is on, all factors necessary for IRES translation are present at high concentrations. The presence of all these factors nearby would enhance the overall probability the IRES gets translated. In contrast, when the IRES is on, although many translation factors are present, not all factors required for cap translation are available, including eIF4E and eIF4G. Without these factors, cap-dependent translation is not significantly enhanced.

The precise molecular mechanisms that govern the enhancement of IRES-mediated translation in the presence of cap-dependent translation remain unclear. An interesting possibility given our live-cell confirmation that ribosomes generally stretch out translation sites is that the stretching somehow alters the accessibility or structure of the IRES. This could impact the IRES in several different ways. For example, the IRES could be stabilized, its folding could become faster, or possibly ribosomes coming off the cap could reinitiate at the IRES with higher probability. According to our simulations, all of these possibilities do indeed improve the overall fits to our data, but faster folding alone was sufficient to improve the fit to near optimum values (computational modeling results can be found in Appendix A). Thus, our data suggest the stretching out of actively



translated transcripts can impact the translation of downstream ORFs, even when those downstream ORFs do not overlap with the upstream ORF.

Despite the lower overall translation efficiency of the EMCV IRES compared to the cap, the upside of relying on a subset of factors is IRES-mediated translation can persist and actually surpass cap-dependent translation during stressful conditions, a situation viruses have evolved and exploited in their ongoing arms race with eukaryotic cells. We see that in NaAs stress, IRES-mediated translation remains strong, presumably because this stress specifically targets eIF4E, one of the cap-binding proteins, which is not required for IRES translation [123,128]. Though IRES translation also remained strong compared to cap in DTT stress, the effect was smaller than with NaAs, presumably because DTT stress impacts a different set of translation factors than NaAs. In the future, it will be interesting to investigate which factors specifically play the biggest roles and also which IRES sequences are most robust to each type of stress. As many viruses have evolved different IRES architectures to capitalize on this distinction, there are likely an abundance of mechanisms involved in each stress, each dependent on a different set of factors with various relative timings.

## Chapter 3

# Investigating the effects of codon non-optimization on translation elongation and initiation

*The work described in Chapter 3 is a collaboration project between Dr. Timothy Stasevich at Colorado State University and Dr. Olivia Rissland at University of Colorado School of Medicine. Listed here are the contributions based on the results described in this chapter: Chloe Barrington-Ham a graduate student in the Rissland lab built the codon optimized/non-optimized constructs and conducted the qPCR and luciferase assays, Amanda L. Koch conducted NCT assays and analyzed single-molecule data, Dr. Rissland and Chloe designed bulk cell experiments and conceptualized molecular mechanisms to test, Dr. Stasevich and Amanda designed and optimized NCT experiments and designed analysis tools. This work is ongoing and is currently unpublished.*

### 3.1 Introduction to codon optimality

As described in Chapter 1, translation elongation is the process in which a ribosome decodes the mRNA codon by codon to make a protein [56]. 61 possible codons specify different amino acids and 3 stop codons signal the ribosome to halt. Because each cell has only 21 possible amino acids (including selenocysteine), genetic code degeneracy where multiple codons code for a single amino acid exists [129]. Interestingly, tRNA concentrations vary between species and even between cell types within an organism [130]. This variability can lead to ribosome elongation differences [131,132]. The relationship between tRNA concentrations and ribosome decoding efficiency is termed codon optimality. Codon optimality can also be defined based on how frequently

certain codons are used within the genome of a particular species termed codon adaptive index. Regardless of how optimality is defined, the optimality of a coding region has implications in both mRNA stability and translation regulation [133].

Up until recently, the focus of mRNA stability has been with regulation pathways associated with the 3' UTR [134]. The 3' UTR has known binding sites for regulatory factors such as RNA binding proteins (RBP) and microRNA (miRNA) that destabilize mRNA and further recruit exonucleases that ultimately degrade the mRNA [135]. It is also of note that the 3' UTRs of human transcripts are much longer than in other eukaryotes, leading researchers to believe that the 3'UTR was the primary source of mRNA stability regulation in humans [136]. However, it has been shown in multiple model organisms and in human cells that open reading frames (ORFs) have an impact on mRNA stability [137]. In 2019, Narula, A. et al. showed that ribosomes must load for codon optimality to affect mRNA stability [138]. These experiments established that mRNA instability is associated with codons that are translated more slowly linking codon optimality to mRNA instability.

It is established that codon non-optimality decreases mRNA stability. However, little is known about the effects of codon optimality on translation dynamics. In collaboration with Dr. Olivia Rissland, we are beginning to answer these questions in bulk (Rissland lab) and at the single-molecule level (Stasevich lab).

### 3.2 Bulk assays reveal a two-fold loss in mRNA and four-fold loss in protein in codon non-optimal transfected cells

To begin investigating the translational effects of codon optimality, Chloe Barrington-Ham, a graduate student in the Rissland lab, developed two constructs capable of both bulk cell and single-molecule experiments. The first construct is a complete codon optimized for humans using the tRNA Adaptive Index (100% optimized) firefly luciferase while the second has a complete codon non-optimized for humans (0% optimized) firefly luciferase. Both constructs are tagged at the N-terminus with a 10×FLAG tag system (SM) and contain 24×MS2 stem-loops in the 3' UTR allowing for Nascent Chain Tracking (NCT) (Figure 3.1A-B).

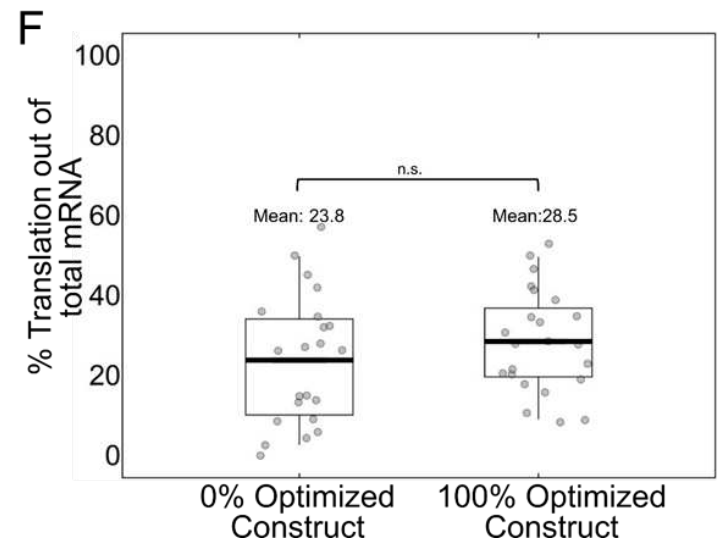
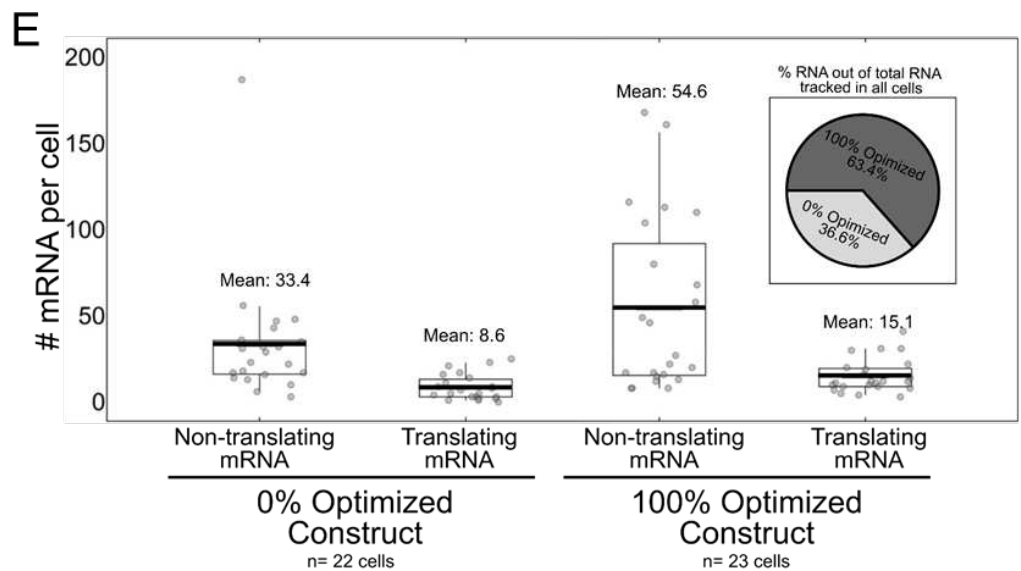
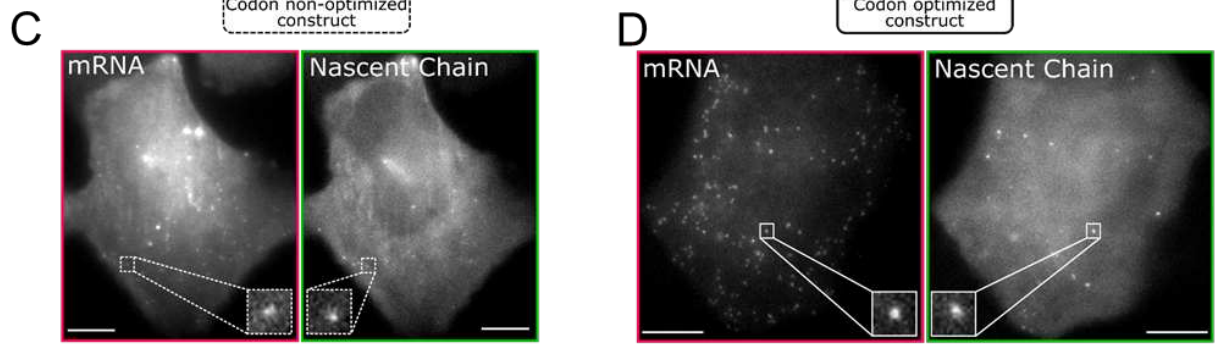
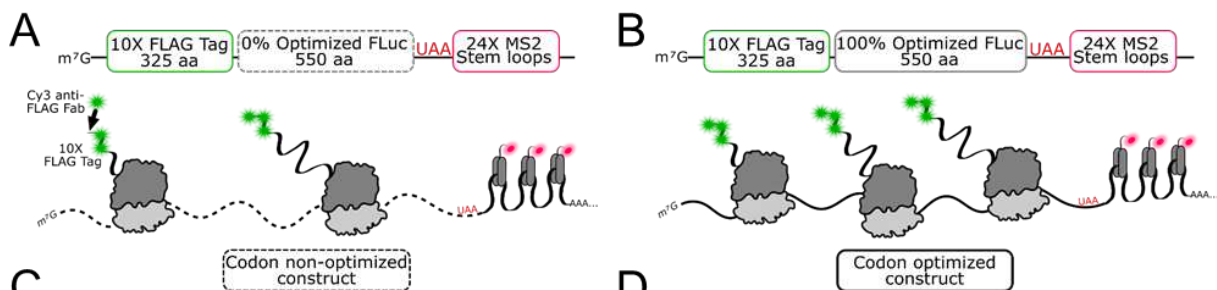
First, we wanted to see what the effect of codon non-optimality was on mRNA and protein levels in bulk cells. To do this, Chloe transfected the optimal and non-optimal constructs into separate dishes of U2OS cells and measured mRNA levels by qPCR. Her results were as expected: the mRNA levels in cells transfected with the non-optimal construct were two times lower than mRNA levels in cells transfected with the optimal construct. Interestingly, when Chloe measured the protein levels using luciferase assays, she saw a four times lower protein level in cells transfected with the non-optimal construct than cells transfected with the optimal construct. If loss of mRNA stability was the only factor dictating protein output, the mRNA difference and protein level difference between the optimal and non-optimal transfected cells should be the same. These results indicate that lower protein output seen in codon non-optimal experiments has a translation repression component along with an increase in mRNA instability.

### 3.3 Using NCT to count the number of translating mRNA per cell

With the NCT assay, we can dissect the translation repression component causing the lower protein output in codon non-optimal conditions. In our assay, we can see translation repression through a few potential ways in codon non-optimal conditions. 1) There are two times less overall mRNA translating. 2) The translating mRNA have fewer ribosomes (dimmer nascent chain signal in translation sites). 3) The ribosomes on codon non-optimal transcripts are elongating slower. 4) There is a combination of all these possibilities.

To further investigate the effects that non-optimal codon usage has on translation dynamics, we imaged either codon optimal and non-optimal constructs in cells. As described in Chapters 1 and 2, to visualize these translation dynamics, we bead-loaded plasmid DNA encoding either the optimal or non-optimal construct along with purified anti-FLAG Cy3-Fab, and HaloTag-MCP into U2OS cells 3-6 hours before imaging. We acquired 2.5-minute movies for cells expressing either the optimal or non-optimal constructs (Figure 3.1C-D). We imaged both constructs on the same day for all experiments to decrease any dish-to-dish variability that may exist. In total, we tracked 924 mRNA 189 of which were translating in the codon non-optimal cells (Figure 3.1E, left). In codon optimal cells, we tracked 1603 mRNA 347 of which were translating (Figure 3.1E, right). An initial observation that we made was that there are around 1.75 times as many total mRNAs in codon optimal loaded cells compared to codon non-optimal cells (63.4% and 36.6%), respectively (Figure 3.1E, inset). These data corroborate Chloe's results for mRNA levels in bulk assays. From these data, we also observe a slight decrease in the number of mRNA that are translating ( $p$ -value = 0.25), though it is not significant and would not account for the overall lower protein levels seen in bulk assays

for codon non-optimal loaded cells (Figure 3.1F). The translating mRNA on codon non-optimal transcripts must then have fewer ribosomes or perhaps those ribosomes are elongating slower.



**Figure 3.1: A nascent chain tracking assay to compare translation dynamics of a codon non-optimal and optimal constructs**

(A&B) Schematic of NCT assay for codon non-optimal (A) and codon optimal (B) constructs. (C&D) Representative cells from imaging each construct. The channels are split for ease of viewing. Channel 1 or the mRNA channel has a red box around it and channel 2 or the nascent chain channel has a green box around it. Crops of mRNA signal and nascent chain signal from translating sites are indicated by a white box in each condition. Scale bars represent 10  $\mu\text{m}$ . (E) Quantification of number of mRNA per cell for both the codon non-optimized (0% Optimized) and codon optimized (100% Optimized) constructs. Each point represents the mRNA count (both non-translating and translating mRNA) in a cell. The inset shows the percent of mRNA (translating and non-translating) for the 100% optimized construct (63.4%) and the 0% optimized construct (36.6%) out of total mRNA tracks in all cells for both constructs (2,527 mRNA). (F) Percent of translating mRNA out of total mRNA. 0% optimized: n= 22 cells. 100% optimized: n= 23 cells. For the box and whisker plots, the thick black lines indicate the means, the boxes indicate the 25-75% range, and the whiskers indicate the 5-95% range.

### 3.4 Quantifying number of ribosomes and translation elongation rates on optimized and non-optimized transcripts

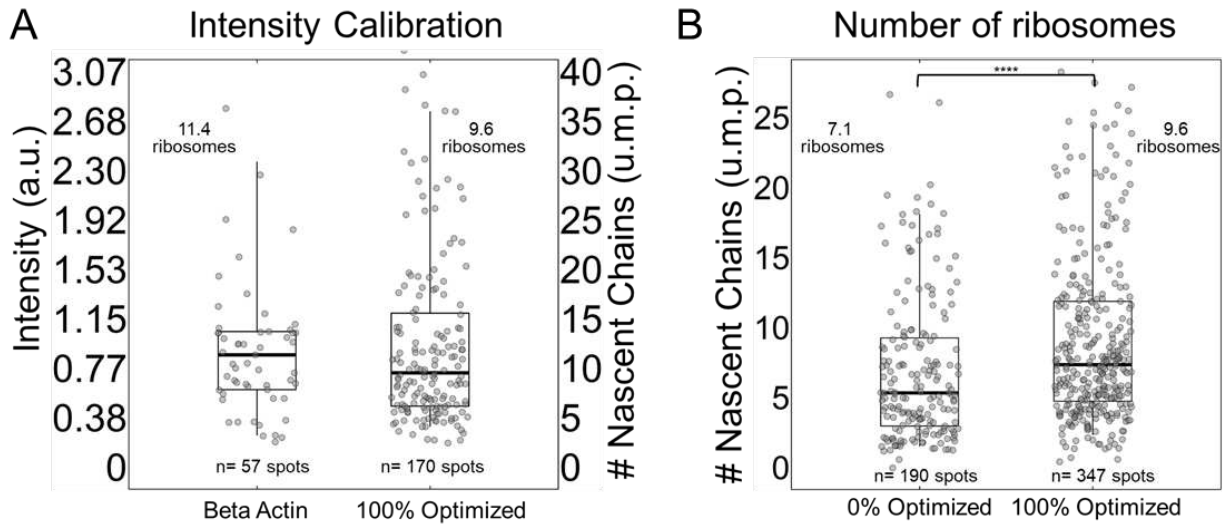
We next measured the number of ribosomes on codon optimal and non-optimal transcripts. Initially, we measured the nascent chain signal intensities which is a gauge of ribosome load on transcripts for both codon optimal and non-optimal transcripts. Here we saw that codon non-optimal translation sites were dimmer than optimal translation sites. To extend our measurements further, we calculated the number of ribosomes present on both optimal and non-optimal transcripts (Figure 3.2). First, we needed to calibrate nascent chain intensities within translation sites to an exact number of ribosomes. For this, we developed a 10 $\times$ FLAG calibration construct that produces translation sites containing approximately 11 ribosomes (Figure 3.2A). Comparing the intensity of these translation sites to optimal (10 $\times$ FLAG-100Luc) and non-optimal (10 $\times$ FLAG-0Luc) translation sites revealed that codon non-optimal transcripts have approximately 7 ribosomes loaded while fully codon optimized transcripts have around 10 ribosomes



loaded (Figure 3.2B). These data suggest that fewer ribosomes can initiate and load onto non-optimal transcripts. However, if slower initiation were the only factor decreasing protein levels seen in bulk assays, we would expect to see two-times not 1.4-times fewer ribosomes on codon non-optimal transcripts. This indicates additional translation mechanism(s) are contributing to the decrease in protein levels in codon non-optimal conditions.

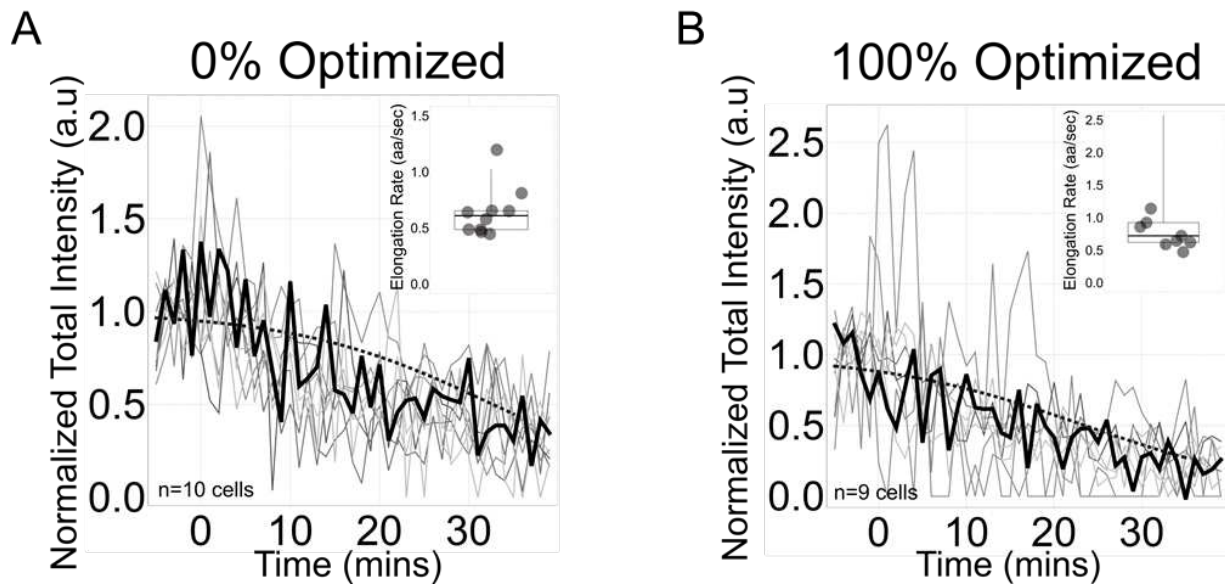
To test if slower elongation is also contributing to lower protein levels in codon non-optimal conditions, we used harringtonine to calculate elongation rates [118]. Consistent with ribosomal run-off, harringtonine led to a steady decay in the nascent chain signals within translation sites for both optimal and non-optimal constructs (Figure 3.3A-B). By eye, it seems as though the elongation rate for ribosomes on non-optimal transcripts is slightly slower than that of optimal transcripts. This can be seen as a more gradual intensity decay overtime for the non-optimal translation run-off curves (Figure 3.3A).

Combined, our data indicate that there are not only fewer ribosomes loading or initiating on codon non-optimal transcripts, but those ribosomes are moving slower meaning fewer overall proteins are made per transcript. Overall, our data corroborate what was seen in bulk cell assays – a two-fold decrease in total mRNA and approaching a four-fold decrease in protein levels due to initiation repression and slower elongation for codon non-optimal conditions.



**Figure 3.2: The non-optimized (0% optimized) transcripts load fewer ribosomes than the fully optimized (100% optimized) transcripts**

(A) Translation site calibration measurements. The intensities of 100% optimized translation sites (n=170) were compared to a 10xFLAG calibration system (n=57 spots) with a known number of ribosomes (11.4 ribosomes). From these comparisons, the number of ribosomes on 100% optimized transcripts are calculated to be 9.6 ribosomes. These cells were imaged on the same day with the same imaging conditions. (B) 100% Optimized translation intensities were then compared to 0% optimized translation intensities to calculate the number of ribosomes on the 0% optimized transcripts to be 7.1 ribosomes. p value(\*\*\*\*) =  $3.08 \times 10^{-6}$ . The p-values are based on a two-tailed Mann-Whitney test: \*p<0.05, \*\*p<0.01, \*\*\*p<0.001, \*\*\*\*p<0.0001. For the box and whisker plots, the thick black lines indicate the means, the boxes indicate the 25-75% range, and the whiskers indicate the 5-95% range.



**Figure 3.3: Calculating the elongation rates on non-optimized and optimized constructs.**

(A-B) Normalized total intensity decay curves of deoptimized and optimized translation, respectively. Each gray line represents a single cell treated with HT. The black line is a representative cell. The dotted black line is the Tanx fit of the representative cell. The inset is the calculated elongation rates of each cell. All cells were imaged for 45 minutes with a 1-minute interval between each capture. Harringtonine was added at time 0. Five pre-treatment images were taken before the addition of HT. Intensity values are in arbitrary units (a.u).

### 3.5 Discussion

Codon usage and optimality have been studied for years. Most investigations contribute lower protein expression of codon nonoptimal transcripts to slower elongation rates [139]. Recently, mRNA stability became a contributing factor to lower protein levels as well [138]. However, the connection between codon optimality and translation regulation has not been well investigated. By combining bulk cell assays and single-molecule translation assays, we were able to begin investigating the translational dynamics on codon optimal and non-optimal transcripts. Preliminarily, in bulk cell assays, Chloe Barrington-Ham in Dr. Rissland's lab found that while mRNA levels decreased two-fold between optimal and non-optimal transfected cells, protein levels decreased four-fold. These data indicated that mRNA instability was not the only factor contributing to lower protein levels in codon non-optimal transfected cells. To further investigate translational dynamics, we turned to our single-molecule translation imaging assay, NCT. With NCT, we were able to quantify the average number of translating mRNA per cell, the number of ribosomes loaded, and the elongation rates on codon optimal and non-optimal transcripts.

We first examined if the number of translating mRNA per cell was different between codon optimal and non-optimal loaded cells. Interestingly, we found that there are around two times fewer non-translating and translating mRNA in codon non-optimal than codon

optimal loaded cells. Because we conduct experiments shortly after loading DNA into cells, the cells are still transcribing plasmid DNA and may not have reached an equilibrium. As a follow-up to this initial analysis, we will count mRNA per cell 24 hours after loading.

As researchers have already established [54], we qualitatively observed that elongation on codon non-optimal transcripts was slower than on codon optimal transcripts. If all else equal, we would then expect nascent chain intensities in codon non-optimal translation sites to be brighter and have more ribosomes. We saw the opposite with nascent chain intensities in codon optimal translation sites brighter and containing ~3 more ribosomes on average than the codon non-optimal sites. These data suggest a reduction in ribosome initiation along with slower rates of elongation on codon non-optimal transcripts.

Preliminarily, bulk cell assays combined with single-molecule analysis of translation on optimal or non-optimal transcripts are painting a clearer picture about the connection between codon optimality and translation regulation. It appears that slow elongation rates are activating a pathway that inhibits translation initiation [140]. Because we do not see a dramatic difference in levels of translating mRNA per cell, it would seem that this is an actively translating dependent mechanism meaning that because a few mRNAs in the cell are being translated slowly does not mean that translation initiation is shutting down globally in the cells. Follow-up studies are being conducted to investigate the connection between slow elongation and initiation inhibition and are discussed further in Chapter 4.

# Chapter 4

## Future Perspectives and Discussion

### 4.1 Future perspectives for IRES-mediated translation and viral translation studies

Overall, our technology to visualize when, where, and to what degree IRES-mediated translation occurs at the single-molecule level in the natural setting of living cells provides a new angle on viral translation that will complement technologies like ribosome profiling [58] and in vitro single-molecule assays [141].

With the translational dynamics of the EMCV IRES element established in our system, it will be interesting to investigate other viral and cellular IRES elements in the same way. The EMCV IRES element is a Type II IRES that requires many host initiation factors to translate [113]. However, Type III and Type IV IRES elements require fewer initiation factors for successful translation to occur [126]. Intuitively, one would think that less initiation factor requirement would lead to increased levels of translation. However, initial unpublished data show that HCV IRES (Type III) translation occurs much less frequently than EMCV IRES translation. Interestingly, translation from Polio IRES (Type I) seems to follow a similar trend. This corroborates what investigators have shown in bulk studies as well [142].

A reasonable explanation for this observation may be that the IRES elements from HCV and poliovirus require certain factors for translation to occur that can only be found in cell types they specifically infect, liver and nerve cells, respectively. To this end, investigating IRES-mediated translation in cells that the corresponding virus infects, can better explain the role IRES elements play in viral replication in a more biological context.

As a next step, it would be interesting to investigate the role cell-specific translational proteins, mRNA binding proteins, or small RNAs have on IRES translation by adding those factors to our system in U2OS cells [143–145]. Experiments such as these can lead to a better understanding of IRES functionality in the presence and absence of certain cellular factors. On the other hand, viruses down-regulate cap-dependent translation not only through stress as described in Chapter 2 but also through the introduction of viral proteins. For example, a protease encoded in the EMCV genome has been shown to cleave eIF4G such that it can no longer bind the cap-binding protein eIF4E but can still be recruited to the IRES and subsequently recruit ribosomes [146].

While our bicistronic biosensor paved the way in understanding IRES-mediated translation in single living cells, our system does not accurately represent true viral infection. Viral mRNAs are not typically bicistronic with a 5' capped ORF upstream of the IRES element. In previous results using the bicistronic construct, we showed that there were not distinguishable differences in subcellular localization, mobility, or size between cap-dependent and IRES-mediated translation. In these particular studies, because of the bicistronic nature of the construct, any effect the cap has on these parameters could be overpowering effects from the IRES. Further, we showed that cap translation increases the likelihood of IRES translation. To fairly study the IRES element in a more viral-like

context, the IRES element and reporter can be *in vitro* transcribed and loaded into cells as mature mRNA. Studying the IRES element outside of the influence of the cap can glean clearer results on how the IRES element alone recruits ribosomes and drives interesting translational phenomena within living cells. This is also one step closer to imaging single-molecule viral life cycle events [147].

## 4.2 Future perspectives for codon optimality studies

We successfully showed that the NCT assay is poised at distinguishing differences in translation dynamics between codon optimal and non-optimal conditions. We are excited by our initial results and will continue to investigate the mechanisms involved in initiation inhibition seen with codon non-optimal transcript translation.

Codon availability is not the only factor slowing ribosomes down. Translation of aberrant mRNA can also cause the ribosome to slow. Slowly moving ribosomes can indicate several potential issues of the mRNA, lack of cellular energy or other needed material, or presence of cell stress [148]. The most common and well-studied type of ribosome pausing is poly-A translation pausing. Here, the paused ribosome is the cue to enact downstream quality control mechanisms that degrade the mRNA and the nascent chain [149]. Over the last few years, groups have teased apart the mechanisms involved in the ribosomal quality control pathway. When ribosomes pause for extended periods of time, the trailing ribosome will collide with the leading paused ribosome. If this collision persists, quality control factors that specifically recognize the interface of these two collided ribosomes will bind and recruit further downstream quality control factors [84]. Ultimately, the stalled ribosome is removed and the nascent protein is degraded. In most

instances, the mRNA is also marked for degradation. This is an extreme case where the leading ribosome is completely stalled, and ribosomal collisions are persistent. However, leading ribosomes that slow during translation of a highly structured part of the transcript for example can transiently collide with lagging ribosomes. This can occur frequently during translation and does not necessarily need the same level of regulation as completely stalled ribosomes. Recently, Juskiewicz, S., et al. showed that an alternative quality control pathway is activated when ribosomes collide transiently [140]. The proposed mechanism is as follows: First, EDF1 recognizes the ribosomal collision which will recruit and stabilize the ZNF598-GIGYF2-4EHP complex. 4EHP inhibits translation initiation by competing off the cap-binding protein, eIF4E. If the collision is incidental, translation elongation will resume on the leading ribosome [140].

The endpoint of this regulation is inhibition of translation initiation. We wondered if the result we are seeing with codon non-optimal transcripts having ~3 fewer ribosomes than codon optimal transcripts is due to activation of this quality control pathway. To test this, the Rissland group has acquired three different strains of HEK 293 cells that have EDF1, GIGYF2, or ZNF598 proteins knocked out. Our working hypothesis is that knocking out either EDF1 or GIGYF2 will rescue the lower ribosome load seen in codon non-optimal translation. If these ribosome collisions are transient and do not persist, knocking out ZNF598 should not affect translation.

Our first main hurdle is to optimize the NCT assay to work in HEK 293 cells. Because we rely on bead loading purified protein for the NCT assay, we generally work with adherent cells such as U2OS, HeLa, or RPE1. HEK 293 cells are somewhat adherent but peel up easily. We are still working through this optimization but have promising initial



results indicating that we can indeed bead load and image HEK 293 cells. Soon, we will focus on quantifying the number of ribosomes on optimal and non-optimal transcripts for each of the knockout HEK 293 cells.

Establishing a regulatory mechanism involved in the translation of codon non-optimal transcripts will further the translation regulatory and codon usage fields helping refine our current definitions of codon optimality. Information that could ultimately give future insight into effective drug and therapeutic development studies such as studies currently taking place at Moderna [150].

### 4.3 Overall discussion

Since its realization in the late 1950s, the quest of molecular biologists has been to tease apart every detail of the central dogma. We now know that both transcription and translation are heavily regulated processes that require the coordinated binding and unbinding of many factors to occur. Misregulation at any step can lead to catastrophic disease.

With the advent of single-molecule imaging technology, studying the when, where, and how of transcription and translation has become possible. Though both imaging technologies are fairly recent, real-time single-molecule translation imaging in living cells was made possible a mere five years ago in 2016. The last five years have provided a wealth of information about translation initiation and elongation dynamics, subcellular localization, translation site structure, and reaction to stress for both canonical and non-canonical translation in living cells.

We anticipate that as NCT tags and probes continue to evolve more detailed questions about translation dynamics and regulation will continue to be unearthed. For example, the creation and evolution of biosensors to study non-canonical translation tactics such as the exploitation of IRES elements will give better insight into both viral and cellular translation mechanisms in healthy and stress conditions. Perhaps these biosensors can be combined with computer-aided high-throughput screens to design and verify new compounds to fight a viral infection in a targeted manner.

Likewise, complimenting classic biochemistry and molecular biology techniques with the NCT assay has proven fruitful at revealing underlying translation regulation pathways in codon optimal and non-optimal conditions. Continued investigation of translation regulatory pathways is essential for not only understanding how translation is switched on or off but how diseases occur when that switch does not work properly. Single-molecule assays reveal details that may otherwise be lost in the noise of other assays. Details of which could lead to a greater understanding of not just translation but the entire central dogma.

# Bibliography for Chapters 1-4

1. Crick FH: **On protein synthesis**. *Symp Soc Exp Biol* 1958, **12**:138–163.
2. Crick F: **Central Dogma of Molecular Biology**. *Nature* 1970, **227**:561–563.
3. **The Nobel Prize in Physiology or Medicine 1974**. *NobelPrize.org* [date unknown],
4. Palade GE: **A small particulate component of the cytoplasm**. *J Biophys Biochem Cytol* 1955, **1**:59–68.
5. Gros F: **The Messenger**. In *Origins of Molecular Biology*. . John Wiley & Sons, Ltd; 2003:143–152.
6. Pardee AB: **NUCLEIC ACID PRECURSORS AND PROTEIN SYNTHESIS**. *Proc Natl Acad Sci U S A* 1954, **40**:263–270.
7. Boivin A, Vendrely R: **Sur le rôle possible des deux acides nucléiques dans la cellule vivante**. *Experientia* 1947, **3**:32–34.
8. Watson JD, Crick FHC: **Genetical Implications of the Structure of Deoxyribonucleic Acid**. *Nature* 1953, **171**:964–967.
9. Hoagland MB, Stephenson ML, Scott JF, Hecht LI, Zamecnik PC: **A soluble ribonucleic acid intermediate in protein synthesis**. *J Biol Chem* 1958, **231**:241–257.
10. Cobb M: **Who discovered messenger RNA?** *Curr Biol* 2015, **25**:R526–R532.
11. Astrachan L, Volkin E: **Properties of ribonucleic acid turnover in T2-infected Escherichia coli**. *Biochim Biophys Acta* 1958, **29**:536–544.
12. Boomsma J (Koos): **Jacobus (Koos) Boomsma**. *Curr Biol* 2015, **25**:R532–R534.
13. Pardee AB, Jacob F, Monod J: **The genetic control and cytoplasmic expression of “Inducibility” in the synthesis of  $\beta$ -galactosidase by E. coli**. *J Mol Biol* 1959, **1**:165–178.
14. Jacob F, Monod J: **Genetic regulatory mechanisms in the synthesis of proteins**. *J*

- Mol Biol* 1961, **3**:318–356.
15. Brenner S, Jacob F, Meselson M: **An Unstable Intermediate Carrying Information from Genes to Ribosomes for Protein Synthesis.** *Nature* 1961, **190**:576–581.
  16. Gros F, Hiatt H, Gilbert W, Kurland CG, Risebrough RW, Watson JD: **Unstable Ribonucleic Acid Revealed by Pulse Labelling of Escherichia Coli.** *Nature* 1961, **190**:581–585.
  17. Furuichi Y. (2015). **Discovery of m(7)G-cap in eukaryotic mRNAs.** *Proceedings of the Japan Academy. Series B, Physical and biological sciences*, **91**(8), 394–409.
  18. Sonenberg N, Gingras A-C: **The mRNA 5' cap-binding protein eIF4E and control of cell growth.** *Curr Opin Cell Biol* 1998, **10**:268–275.
  19. Hinnebusch AG: **Molecular Mechanism of Scanning and Start Codon Selection in Eukaryotes.** *Microbiol Mol Biol Rev MMBR* 2011, **75**:434–467.
  20. Babendure JR, Babendure JL, Ding J-H, Tsien RY: **Control of mammalian translation by mRNA structure near caps.** *RNA* 2006, **12**:851–861.
  21. Höglund M, Säll T, Röhme D: **On the origin of coding sequences from random open reading frames.** *J Mol Evol* 1990, **30**:104–108.
  22. Barrett LW, Fletcher S, Wilton SD: **Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements.** *Cell Mol Life Sci* 2012, **69**:3613–3634.
  23. Pichon X, A. Wilson L, Stoneley M, Bastide A, A King H, Somers J, E Willis A: **RNA Binding Protein/RNA Element Interactions and the Control of Translation.** *Curr Protein Pept Sci* 2012, **13**:294–304.
  24. Edmonds M: **A history of poly A sequences: from formation to factors to function.** *Prog Nucleic Acid Res Mol Biol* 2002, **71**:285–389.
  25. Hinnebusch AG, Lorsch JR: **The Mechanism of Eukaryotic Translation Initiation: New Insights and Challenges.** *Cold Spring Harb Perspect Biol* 2012, **4**.

26. Hershey JWB, Sonenberg N, Mathews MB: **Principles of Translational Control: An Overview.** *Cold Spring Harb Perspect Biol* 2012, **4**.
27. Jackson RJ, Hellen CUT, Pestova TV: **The mechanism of eukaryotic translation initiation and principles of its regulation.** *Nat Rev Mol Cell Biol* 2010, **11**:113–127.
28. Gingras AC, Raught B, Sonenberg N: **eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation.** *Annu Rev Biochem* 1999, **68**:913–963.
29. Sonenberg N, Dever TE: **Eukaryotic translation initiation factors and regulators.** *Curr Opin Struct Biol* 2003, **13**:56–63.
30. Hentze MW: **eIF4G--A Multipurpose Ribosome Adapter?** *Science* 1997, **275**:500–501.
31. Marintchev A, Edmonds KA, Marintcheva B, Hendrickson E, Oberer M, Suzuki C, Herdy B, Sonenberg N, Wagner G: **Topology and Regulation of the Human eIF4A/4G/4H Helicase Complex in Translation Initiation.** *Cell* 2009, **136**:447–460.
32. Poulin F, Sonenberg N: *Mechanism of Translation Initiation in Eukaryotes.* Landes Bioscience; 2013.
33. Hinnebusch AG: **5 Mechanism and Regulation of Initiator Methionyl-tRNA Binding to Ribosomes.** *Cold Spring Harb Monogr Arch* 2000, **39**:185–243.
34. Kimball SR: **Eukaryotic initiation factor eIF2.** *Int J Biochem Cell Biol* 1999, **31**:25–29.
35. LeFebvre AK, Korneeva NL, Trutschl M, Cvek U, Duzan RD, Bradley CA, Hershey JWB, Rhoads RE: **Translation Initiation Factor eIF4G-1 Binds to eIF3 through the eIF3e Subunit\*.** *J Biol Chem* 2006, **281**:22917–22932.
36. Pestova TV, Kolupaeva VG: **The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection.** *Genes Dev* 2002, **16**:2906–2922.
37. Hinnebusch AG: **The Scanning Mechanism of Eukaryotic Translation Initiation.** *Annu Rev Biochem* 2014, **83**:779–812.

38. Kozak M: **The scanning model for translation: an update.** *J Cell Biol* 1989, **108**:229–241.
39. Kozak M: **At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells.** *J Mol Biol* 1987, **196**:947–950.
40. Kozak M: **Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes.** *Cell* 1986, **44**:283–292.
41. Kozak M: **Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes.** *Nucleic Acids Res* 1981, **9**:5233–5252.
42. Kozak M: **Evaluation of the “scanning model” for initiation of protein synthesis in eucaryotes.** *Cell* 1980, **22**:7–8.
43. Kozak M: **Influences of mRNA secondary structure on initiation by eukaryotic ribosomes.** *Proc Natl Acad Sci U S A* 1986, **83**:2850–2854.
44. Kozak M: **Inability of circular mRNA to attach to eukaryotic ribosomes.** *Nature* 1979, **280**:82–85.
45. Shatsky IN, Terenin IM, Smirnova VV, Andreev DE: **Cap-Independent Translation: What’s in a Name?** *Trends Biochem Sci* 2018, **43**:882–895.
46. Merrick WC, Komar AA: **Translation Elongation in Eukaryotes.** In *Encyclopedia of Biological Chemistry*. Edited by Lennarz WJ, Lane MD. Elsevier; 2004:224–229.
47. Gromadski KB, Schümmer T, Strømgaard A, Knudsen CR, Kinzy TG, Rodnina MV: **Kinetics of the Interactions between Yeast Elongation Factors 1A and 1B $\alpha$ , Guanine Nucleotides, and Aminoacyl-tRNA\*.** *J Biol Chem* 2007, **282**:35629–35637.
48. Dever TE, Dinman JD, Green R: **Translation Elongation and Recoding in Eukaryotes.** *Cold Spring Harb Perspect Biol* 2018, **10**.
49. Demeshkina N, Jenner L, Westhof E, Yusupov M, Yusupova G: **A new understanding of the decoding principle on the ribosome.** *Nature* 2012, **484**:256–259.
50. Andersen GR, Valente L, Pedersen L, Kinzy TG, Nyborg J: **Crystal structures of**

- nucleotide exchange intermediates in the eEF1A-eEF1B $\alpha$  complex.** *Nat Struct Biol* 2001, **8**:531–534.
51. Jobe A, Liu Z, Gutierrez-Vargas C, Frank J: **New Insights into Ribosome Structure and Function.** *Cold Spring Harb Perspect Biol* 2019, **11**:a032615.
  52. Frank J, Gao H, Sengupta J, Gao N, Taylor DJ: **The process of mRNA–tRNA translocation.** *Proc Natl Acad Sci* 2007, **104**:19671–19678.
  53. Xie J, de Souza Alves V, von der Haar T, O’Keefe L, Lenchine RV, Jensen KB, Liu R, Coldwell MJ, Wang X, Proud CG: **Regulation of the Elongation Phase of Protein Synthesis Enhances Translation Accuracy and Modulates Lifespan.** *Curr Biol* 2019, **29**:737-749.e5.
  54. Richter JD, Collier J: **Pausing on Polyribosomes: Make Way for Elongation in Translational Control.** *Cell* 2015, **163**:292–300.
  55. Collart MA, Weiss B: **Ribosome pausing, a dangerous necessity for co-translational events.** *Nucleic Acids Res* [date unknown], doi:10.1093/nar/gkz763.
  56. Tahmasebi S, Sonenberg N, Hershey JWB, Mathews MB: **Protein Synthesis and Translational Control: A Historical Perspective.** *Cold Spring Harb Perspect Biol* 2019, **11**.
  57. Ingolia NT: **Ribosome Footprint Profiling of Translation throughout the Genome.** *Cell* 2016, **165**:22–33.
  58. Ingolia NT, Hussmann JA, Weissman JS: **Ribosome Profiling: Global Views of Translation.** *Cold Spring Harb Perspect Biol* 2019, doi:10.1101/cshperspect.a032698.
  59. Morisaki T, Lyon K, DeLuca KF, DeLuca JG, English BP, Zhang Z, Lavis LD, Grimm JB, Viswanathan S, Looger LL, et al.: **Real-time quantification of single RNA translation dynamics in living cells.** *Science* 2016, **352**:1425–1429.
  60. Pichon X, Bastide A, Safieddine A, Chouaib R, Samacoits A, Basyuk E, Peter M, Mueller F, Bertrand E: **Visualization of single endogenous polysomes reveals the dynamics**

- of translation in live human cells.** *J Cell Biol* 2016, **214**:769–781.
61. Wu B, Eliscovich C, Yoon YJ, Singer RH: **Translation dynamics of single mRNAs in live cells and neurons.** *Science* 2016, **352**:1430–1435.
  62. Wang C, Han B, Zhou R, Zhuang X: **Real-Time Imaging of Translation on Single mRNA Transcripts in Live Cells.** *Cell* 2016, **165**:990–1001.
  63. Yan X, Hoek TA, Vale RD, Tanenbaum ME: **Dynamics of Translation of Single mRNA Molecules In Vivo.** *Cell* 2016, **165**:976–989.
  64. Morisaki T, Stasevich TJ: **Quantifying Single mRNA Translation Kinetics in Living Cells.** *Cold Spring Harb Perspect Biol* 2018, **10**.
  65. Hayashi-Takanaka Y, Yamagata K, Wakayama T, Stasevich TJ, Kainuma T, Tsurimoto T, Tachibana M, Shinkai Y, Kurumizaka H, Nozaki N, et al.: **Tracking epigenetic histone modifications in single cells using Fab-based live endogenous modification labeling.** *Nucleic Acids Res* 2011, **39**:6475–6488.
  66. Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD: **A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging.** *Cell* 2014, **159**:635–646.
  67. Zhao N, Kamijo K, Fox PD, Oda H, Morisaki T, Sato Y, Kimura H, Stasevich TJ: **A genetically encoded probe for imaging nascent and mature HA-tagged proteins in vivo.** *Nat Commun* 2019, **10**:2947.
  68. Boersma S, Khuperkar D, Verhagen BMP, Sonneveld S, Grimm JB, Lavis LD, Tanenbaum ME: **Multi-Color Single-Molecule Imaging Uncovers Extensive Heterogeneity in mRNA Decoding.** *Cell* 2019, **178**:458-472.e19.
  69. Cialek CA, Koch AL, Galindo G, Stasevich TJ: **Lighting up single-mRNA translation dynamics in living cells.** *Curr Opin Genet Dev* 2020, **61**:75–82.
  70. Svitkin YV, Evdokimova VM, Brasey A, Pestova TV, Fantus D, Yanagiya A, Imataka H, Skabkin MA, Ovchinnikov LP, Merrick WC, et al.: **General RNA-binding proteins have a**



- function in poly(A)-binding protein-dependent translation.** *EMBO J* 2009, **28**:58–68.
71. Femino AM, Fay FS, Fogarty K, Singer RH: **Visualization of Single RNA Transcripts in Situ.** *Science* 1998, **280**:585–590.
  72. Raj A, Bogaard P van den, Rifkin SA, Oudenaarden A van, Tyagi S: **Imaging individual mRNA molecules using multiple singly labeled probes.** *Nat Methods* 2008, **5**:877–879.
  73. Taliaferro JM: **Classical and emerging techniques to identify and quantify localized RNAs.** *WIREs RNA* 2019, **10**:e1542.
  74. Adivarahan S, Livingston N, Nicholson B, Rahman S, Wu B, Rissland OS, Zenklusen D: **Spatial Organization of Single mRNPs at Different Stages of the Gene Expression Pathway.** *Mol Cell* 2018, **72**:727-738.e5.
  75. Khong A, Parker R: **mRNP architecture in translating and stress conditions reveals an ordered pathway of mRNP compaction.** *J Cell Biol* 2018, **217**:4124–4140.
  76. Vicens Q, Kieft JS, Rissland OS: **Revisiting the closed loop model and the nature of mRNA 5'–3' communication.** *Mol Cell* 2018, **72**:805–812.
  77. Gonatopoulos-Pournatzis T, Cowling VH: **Cap-binding complex (CBC).** *Biochem J* 2014, **457**:231–242.
  78. Hoek TA, Khuperkar D, Lindeboom RGH, Sonneveld S, Verhagen BMP, Boersma S, Vermeulen M, Tanenbaum ME: **Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay.** *Mol Cell* 2019, **75**:324-339.e11.
  79. Fonkeu Y, Kraynyukova N, Hafner A-S, Kochen L, Sartori F, Schuman EM, Tchumatchenko T: **How mRNA Localization and Protein Synthesis Sites Influence Dendritic Protein Distribution and Dynamics.** *Neuron* 2019, **103**:1109-1122.e7.
  80. Cioni J-M, Lin JQ, Holtermann AV, Koppers M, Jakobs MAH, Azizi A, Turner-Bridger B, Shigeoka T, Franze K, Harris WA, et al.: **Late Endosomes Act as mRNA Translation Platforms and Sustain Mitochondria in Axons.** *Cell* 2019, **176**:56-72.e15.

81. Langille JJ, Ginzberg K, Sossin WS: **Polysomes identified by live imaging of nascent peptides are stalled in hippocampal and cortical neurites.** *Learn Mem* 2019, **26**:351–362.
82. Latallo MJ, Livingston NM, Wu B: **Translation imaging of single mRNAs in established cell lines and primary cultured neurons.** *Methods* 2019, **162–163**:12–22.
83. Voigt F, Zhang H, Cui XA, Triebold D, Liu AX, Eglinger J, Lee ES, Chao JA, Palazzo AF: **Single-Molecule Quantification of Translation-Dependent Association of mRNAs with the Endoplasmic Reticulum.** *Cell Rep* 2017, **21**:3740–3753.
84. Juszkievicz S, Chandrasekaran V, Lin Z, Kraatz S, Ramakrishnan V, Hegde RS: **ZNF598 Is a Quality Control Sensor of Collided Ribosomes.** *Mol Cell* 2018, **72**:469–481.e7.
85. Ingolia NT, Lareau LF, Weissman JS: **Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity and Dynamics of Mammalian Proteomes.** *Cell* 2011, **147**:789–802.
86. Lyon K, Aguilera LU, Morisaki T, Munsky B, Stasevich TJ: **Live-Cell Single RNA Imaging Reveals Bursts of Translational Frameshifting.** *Mol Cell* 2019, **75**:172–183.e9.
87. Weingarten-Gabbay S, Elias-Kirma S, Nir R, Gritsenko AA, Stern-Ginossar N, Yakhini Z, Weinberger A, Segal E: **Systematic discovery of cap-independent translation sequences in human and viral genomes.** *Science* 2016, **351**:aad4939.
88. Komar AA, Mazumder B, Merrick WC: **A new framework for understanding IRES-mediated translation.** *Gene* 2012, **502**:75–86.
89. Gebauer F, Hentze MW: **IRES unplugged.** *Science* 2016, **351**:228–228.
90. Pelletier J, Sonenberg N: **Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA.** *Nature* 1988, **334**:320–325.
91. Jang SK, Kräusslich HG, Nicklin MJ, Duke GM, Palmenberg AC, Wimmer E: **A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal**

- entry of ribosomes during in vitro translation. *J Virol* 1988, **62**:2636–2643.**
92. Lee KA, Sonenberg N: **Inactivation of cap-binding proteins accompanies the shut-off of host protein synthesis by poliovirus.** *Proc Natl Acad Sci* 1982, **79**:3447–3451.
93. Balvay L, Rifo RS, Ricci EP, Decimo D, Ohlmann T: **Structural and functional diversity of viral IRESes.** *Biochim Biophys Acta BBA - Gene Regul Mech* 2009, **1789**:542–557.
94. Filbin ME, Kieft JS: **Toward a structural understanding of IRES RNA function.** *Curr Opin Struct Biol* 2009, **19**:267–276.
95. Firth AE, Brierley I: **Non-canonical translation in RNA viruses.** *J Gen Virol* 2012, **93**:1385–1409.
96. Doudna JA, Sarnow P: **Translation Initiation by Viral Internal Ribosome Entry Sites.** 2007,
97. Kieft JS: **Viral IRES RNA structures and ribosome interactions.** *Trends Biochem Sci* 2008, **33**:274–283.
98. Martinez-Salas E, Francisco-Velilla R, Fernandez-Chamorro J, Embarek AM: **Insights into Structural and Mechanistic Features of Viral IRES Elements.** *Front Microbiol* 2018, **8**.
99. Komar AA, Hatzoglou M: **Cellular IRES-mediated translation: The war of ITAFs in pathophysiological states.** *Cell Cycle* 2011, **10**:229–240.
100. Jackson RJ: **The Current Status of Vertebrate Cellular mRNA IRESs.** *Cold Spring Harb Perspect Biol* 2013, **5**:a011569–a011569.
101. Bornes Stéphanie, Prado-Lourenco Leonel, Bastide Amandine, Zanibellato Catherine, Iacovoni Jason S., Lacazette Eric, Prats Anne-Catherine, Touriol Christian, Prats Hervé: **Translational Induction of VEGF Internal Ribosome Entry Site Elements During the Early Response to Ischemic Stress.** *Circ Res* 2007, **100**:305–308.
102. Xue S, Tian S, Fujii K, Kladwang W, Das R, Barna M: **RNA regulons in Hox 5' UTRs confer ribosome specificity to gene regulation.** *Nature* 2015, **517**:33–38.

103. J. P, M. H, Gomaa M, Ye X, Qiu Y, Yang D: **Viral Replication Strategies: Manipulation of ER Stress Response Pathways and Promotion of IRES-Dependent Translation.** In *Viral Replication*. Edited by Rosas-Acosta G. InTech; 2013.
104. Holcik M, Sonenberg N: **Translational control in stress and apoptosis.** *Nat Rev Mol Cell Biol* 2005, **6**:318–327.
105. Hanson PJ, Zhang HM, Hemida MG, Ye X, Qiu Y, Yang D: **IRES-Dependent Translational Control during Virus-Induced Endoplasmic Reticulum Stress and Apoptosis.** *Front Microbiol* 2012, **3**.
106. Thompson SR, Sarnow P: **Regulation of host cell translation by viruses and effects on cell function.** *Curr Opin Microbiol* 2000, **3**:366–370.
107. Koch A, Aguilera L, Morisaki T, Munsky B, Stasevich TJ: **Quantifying the dynamics of IRES and cap translation with single-molecule resolution in live cells.** *Nat Struct Mol Biol* 2020, doi:10.1038/s41594-020-0504-7.
108. Gale M, Tan S-L, Katze MG: **Translational Control of Viral Gene Expression in Eukaryotes.** *Microbiol Mol Biol Rev* 2000, **64**:239–280.
109. Viswanathan S, Williams ME, Bloss EB, Stasevich TJ, Speer CM, Nern A, Pfeiffer BD, Hooks BM, Li W-P, English BP, et al.: **High-performance probes for light and electron microscopy.** *Nat Methods* 2015, **12**:568–576.
110. Pichon X, Robert M-C, Bertrand E, Singer RH, Tutucci E: **New Generations of MS2 Variants and MCP Fusions to Detect Single mRNAs in Living Eukaryotic Cells.** In *RNA Tagging: Methods and Protocols*. Edited by Heinlein M. Springer US; 2020:121–144.
111. Carocci M, Bakkali-Kassimi L: **The encephalomyocarditis virus.** *Virulence* 2012, **3**:351–367.
112. Bochkov YA, Palmenberg AC: **Translational efficiency of EMCV IRES in bicistronic vectors is dependent upon IRES sequence and gene location.** *BioTechniques* 2006,

- 41:283–292.
113. Chamond N, Deforges J, Ulryck N, Sargueil B: **40S recruitment in the absence of eIF4G/4A by EMCV IRES refines the model for translation initiation on the archetype of Type II IRESs.** *Nucleic Acids Res* 2014, **42**:10373–10384.
  114. Pestova TV, Hellen CU, Shatsky IN: **Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry.** *Mol Cell Biol* 1996, **16**:6859–6869.
  115. McNeil PL, Warder E: **Glass beads load macromolecules into living cells.** *J Cell Sci* 1987, **88**:669–678.
  116. Azzam ME, Algranati ID: **Mechanism of Puromycin Action: Fate of Ribosomes after Release of Nascent Protein Chains from Polysomes.** *Proc Natl Acad Sci U S A* 1973, **70**:3866–3869.
  117. Tokunaga M, Imamoto N, Sakata-Sogawa K: **Highly inclined thin illumination enables clear single-molecule imaging in cells.** *Nat Methods* 2008, **5**:159–161.
  118. Fresno M, Jiménez A, Vázquez D: **Inhibition of translation in eukaryotic systems by harringtonine.** *Eur J Biochem* 1977, **72**:323–330.
  119. Liu B, Qian S-B: **Translational reprogramming in stress response.** *Wiley Interdiscip Rev RNA* 2014, **5**:301–305.
  120. Advani VM, Ivanov P: **Translational control under stress: reshaping the translome.** *BioEssays News Rev Mol Cell Dev Biol* 2019, **41**:e1900009.
  121. Lacerda R, Menezes J, Candeias MM: **Alternative Mechanisms of mRNA Translation Initiation in Cellular Stress Response and Cancer.** In *The mRNA Metabolism in Human Disease*. Edited by Romão L. Springer International Publishing; 2019:117–132.
  122. Fernandez J, Yaman I, Sarnow P, Snider MD, Hatzoglou M: **Regulation of Internal Ribosomal Entry Site-mediated Translation by Phosphorylation of the Translation Initiation Factor eIF2 $\alpha$ .** *J Biol Chem* 2002, **277**:19198–19205.

123. Spriggs KA, Stoneley M, Bushell M, Willis AE: **Re-programming of translation following cell stress allows IRES-mediated translation to predominate.** *Biol Cell* 2008, **100**:27–38.
124. Ruiz-Ramos R, Lopez-Carrillo L, Rios-Perez AD, De Vizcaya-Ruíz A, Cebrian ME: **Sodium arsenite induces ROS generation, DNA oxidative damage, HO-1 and c-Myc proteins, NF-κB activation and cell proliferation in human breast cancer MCF-7 cells.** *Mutat Res Toxicol Environ Mutagen* 2009, **674**:109–115.
125. Yorimitsu T, Nair U, Yang Z, Klionsky DJ: **Endoplasmic Reticulum Stress Triggers Autophagy\***. *J Biol Chem* 2006, **281**:30299–30304.
126. De Quinto SL, Lafuente E, Martínez-Salas E: **IRES interaction with translation initiation factors: Functional characterization of novel RNA contacts with eIF3, eIF4B, and eIF4GII.** *RNA* 2001, **7**:1213–1226.
127. Munsky B, Neuert G, Oudenaarden A van: **Using Gene Expression Noise to Understand Gene Regulation.** *Science* 2012, **336**:183–187.
128. Gingras AC, Svitkin Y, Belsham GJ, Pause A, Sonenberg N: **Activation of the translational suppressor 4E-BP1 following infection with encephalomyocarditis virus and poliovirus.** *Proc Natl Acad Sci U S A* 1996, **93**:5578–5583.
129. Hanson G, Collier J: **Codon optimality, bias and usage in translation and mRNA decay.** *Nat Rev Mol Cell Biol* 2018, **19**:20–30.
130. Roth AC: **Decoding properties of tRNA leave a detectable signal in codon usage bias.** *Bioinformatics* 2012, **28**:i340–i348.
131. Sabi R, Tuller T: **Modelling the efficiency of codon-tRNA interactions based on codon usage bias.** *DNA Res Int J Rapid Publ Rep Genes Genomes* 2014, **21**:511–526.
132. Neelagandan N, Lamberti I, Carvalho HJF, Gobet C, Naef F: **What determines eukaryotic translation elongation: recent molecular and quantitative analyses of protein synthesis.** *Open Biol* [date unknown], **10**:200292.

133. Harigaya Y, Parker R: **Codon optimality and mRNA decay**. *Cell Res* 2016, **26**:1269–1270.
134. Mayr C: **Regulation by 3'-Untranslated Regions**. *Annu Rev Genet* 2017, **51**:171–194.
135. Mazumder B, Seshadri V, Fox PL: **Translational control by the 3'-UTR: the ends specify the means**. *Trends Biochem Sci* 2003, **28**:91–98.
136. Mignone F, Pesole G: **mRNA Untranslated Regions (UTRs)**. In *eLS*. . American Cancer Society; 2018:1–6.
137. Presnyak V, Alhusaini N, Chen Y-H, Martin S, Morris N, Kline N, Olson S, Weinberg D, Baker KE, Graveley BR, et al.: **Codon Optimality Is a Major Determinant of mRNA Stability**. *Cell* 2015, **160**:1111–1124.
138. Narula A, Ellis J, Taliaferro JM, Rissland OS: **Coding regions affect mRNA stability in human cells**. *RNA* 2019, **25**:1751–1764.
139. Novoa EM, Ribas de Pouplana L: **Speeding with control: codon usage, tRNAs, and ribosomes**. *Trends Genet* 2012, **28**:574–581.
140. Juszkievicz S, Slodkowicz G, Lin Z, Freire-Pritchett P, Peak-Chew S-Y, Hegde RS: **Ribosome collisions trigger cis-acting feedback inhibition of translation initiation**. *eLife* 2020, **9**:e60038.
141. Johnson AG, Grosely R, Petrov AN, Puglisi JD: **Dynamics of IRES-mediated translation**. *Philos Trans R Soc B Biol Sci* 2017, **372**:20160177.
142. Borman AM, Le Mercier P, Girard M, Kean KM: **Comparison of Picornaviral IRES-Driven Internal Initiation of Translation in Cultured Cells of Different Origins**. *Nucleic Acids Res* 1997, **25**:925–932.
143. Wilson JA, Zhang C, Huys A, Richardson CD: **Human Ago2 Is Required for Efficient MicroRNA 122 Regulation of Hepatitis C Virus RNA Accumulation and Translation**. *J Virol* 2011, **85**:2342–2350.
144. Yang Y, Wang Z: **IRES-mediated cap-independent translation, a path leading to**

- hidden proteome.** *J Mol Cell Biol* 2019, **11**:911–919.
145. Walsh D, Mohr I: **Viral subversion of the host protein synthesis machinery.** *Nat Rev Microbiol* 2011, **9**:860–875.
146. Lee K-M, Chen C-J, Shih S-R: **Regulation Mechanisms of Viral IRES-Driven Translation.** *Trends Microbiol* 2017, **25**:546–561.
147. Boersma S, Rabouw HH, Bruurs LJM, Pavlovič T, van Vliet ALW, Beumer J, Clevers H, van Kuppeveld FJM, Tanenbaum ME: **Translation and Replication Dynamics of Single RNA Viruses.** *Cell* 2020, **183**:1930-1945.e23.
148. Inada T: **Quality control systems for aberrant mRNAs induced by aberrant translation elongation and termination.** *Biochim Biophys Acta BBA - Gene Regul Mech* 2013, **1829**:634–642.
149. Juszkievicz S, Hegde RS: **Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination.** *Mol Cell* 2017, **65**:743-750.e4.
150. Mauger DM, Cabral BJ, Presnyak V, Su SV, Reid DW, Goodman B, Link K, Khatwani N, Reynders J, Moore MJ, et al.: **mRNA structure regulates protein expression through changes in functional half-life.** *Proc Natl Acad Sci* 2019, **116**:24075–24083.



# Appendix A

## IRES-mediated translation computational modeling

*The following is based on the publication Koch, et al. 2020 (Chapter 2). This is the computational work done by Dr. Luis Aguilera from Dr. Brian Munsky's group for Koch, A, et al. publication.*

### **A.1 Computational modeling reveals ribosomal recruitment limits IRES translation**

According to our experiments, the relatively low probability of IRES-mediated translation is due to rate-limiting steps that precede elongation, presumably either ribosome recruitment or initiation. To distinguish these possibilities, we developed a set of models with varying levels of complexity. All models consider the kinetics of individual ribosomes translating along an mRNA, with stochastic initiation and codon-dependent elongation proportional to the prevalence of the associated tRNA in the human genome [1]. Models differ in the number of states an mRNA can transition between: Three-state models include an inactive mRNA state (OFF), an active mRNA state that allows cap translation (Cap-ON), and an active mRNA that allows IRES translation (IRES-ON). Four-state models include an additional active mRNA state (Cap+IRES-ON) that allows both cap and IRES-translation (Figure A.1A).

The stochastic dynamics for all models were simulated over large ranges of potential parameters and automated searches were conducted to identify combinations of mechanisms and parameters that maximize the likelihood of all data, including the

fraction of translating spots (Figure 2.1), harringtonine run-off kinetics (Figure 2.6), and the translation site intensity distributions (Figure 2.7 and A.1B-E). In total, we considered 14 unique models with between 7 and 12 free parameters, some of which included interdependence between cap- and IRES- translation, either in the form of enhanced transition rates between states or via reinitiation of ribosomes from cap to IRES (Equations 1 and 2; See Computational Methods below and Figure A.2A-B). The simplest model that reproduces all data has eight parameters (Figure A.1A and Table 1): (1) a baseline elongation rate of  $= 1.7 \text{ aa/sec}$ , agreeing with our earlier estimate and consistent with previously measured rates [2,3], (2) an initiation rate  $k_{\text{INIT-C}} \sim 1/21 \text{ sec}^{-1}$  for cap-dependent translation; (3) an initiation rate  $k_{\text{INIT-I}} \sim 1/20 \text{ sec}^{-1}$  for IRES-mediated translation; (4) cap activation bursts with refractory periods ( $1/k_{\text{ON-C}}$ ) of 34.5 min and (5) durations of ( $1/k_{\text{OFF-C}}$ ) of 8.3 min, leading to the synthesis of  $k_{\text{INIT-C}}/k_{\text{OFF-C}} = 24$  nascent proteins on average per cap burst; (6) In the absence of cap, the model predicts that typical bursts of IRES translation would have a refractory period ( $1/k_{\text{ON-I}}$ ) of 91.3 min and (7) a duration of ( $1/k_{\text{OFF-I}}$ ) of 2.5 min, leading to the synthesis of 7.5 nascent proteins on average per IRES burst. According to these fitted parameters, the efficiency of IRES translation is not limited by initiation (since  $k_{\text{INIT-I}} \sim k_{\text{INIT-C}} \sim 1/20 \text{ sec}^{-1}$ ), but rather the IRES spends less time in a translationally active state that can recruit ribosomes. In addition to the above seven parameters (which can be reduced to six by setting the cap and IRES initiation rates equal), one additional parameter was required to fit the data: an enhancement in IRES activation when cap translation is on (i.e.  $k'_{\text{ON-I}} > k_{\text{ON-I}}$ ). Specifically, in the presence of cap, the IRES refractory period is reduced from 91.3 min to 11 min, leading to a 6.9 fold increase in IRES translation. This enhancement was required to capture the large percentage of Cap+IRES translation sites (which is greater than one would predict if cap and IRES translation were independent) (Figure 2.1C, left) and the larger number of IRES-mediated ribosomes in Cap+IRES translation sites compared to IRES-only translation sites (Figure 2.7C).

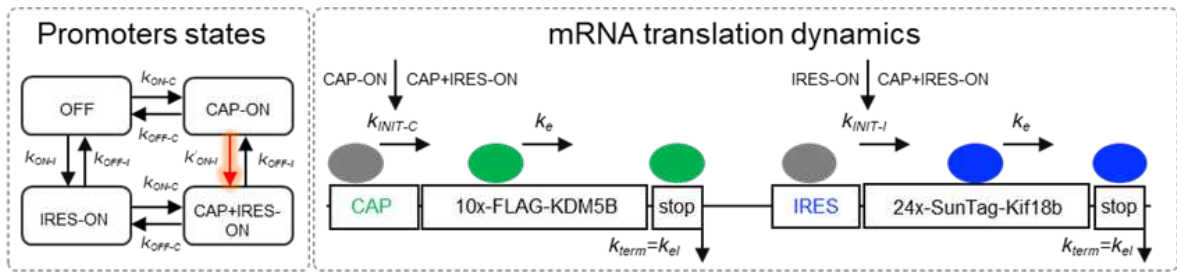
These data and the best-fit model therefore provide evidence that translation of an upstream ORF can positively impact translation of a non-overlapping downstream ORF.

**Table 1. Estimated parameter values for the final selected model.** 4 promoter states and promoter activation for IRES influenced by Cap.

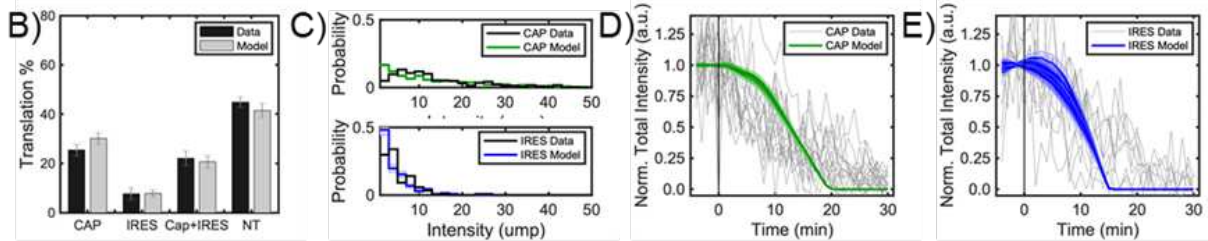
Parameter	Description	Value
$k_{INIT-C}$	Initiation rate for Cap	$4.8 \times 10^{-2} \pm 4.6 \times 10^{-2} \text{ sec}^{-1}$
$k_{INIT-I}$	Initiation rate for IRES	$4.9 \times 10^{-2} \pm 5.3 \times 10^{-2} \text{ sec}^{-1}$
$*k_e$	Average elongation rate	$1.7 \pm 0.18 \text{ aa/sec}$
$k_{ON-C}$	Transition rate from $S_{off}$ to $S_{CAP}$ or $S_{IRES}$ to $S_{CAP-IRES}$	$4.8 \times 10^{-4} \pm 9.7 \times 10^{-5} \text{ sec}^{-1}$
$k_{OFF-C}$	Transition rate from $S_{CAP}$ to $S_{off}$ .	$2 \times 10^{-3} \pm 1.8 \times 10^{-3} \text{ sec}^{-1}$
$k_{ON-I}$	Transition rate from $S_{off}$ to $S_{IRES}$ .	$1.8 \times 10^{-4} \pm 5.5 \times 10^{-5} \text{ sec}^{-1}$
$k_{OFF-I}$	Transition rate from $S_{IRES}$ to $S_{off}$ .	$6.6 \times 10^{-3} \pm 2.7 \times 10^{-3} \text{ sec}^{-1}$
$k'_{ON-I}$	Transition rate from $S_{CAP}$ to $S_{CAP-IRES}$ .	$1.5 \times 10^{-3} \pm 1.0 \times 10^{-3} \text{ sec}^{-1}$

$*k_e$  is the elongation rate calculate as the average elongation of all codons.

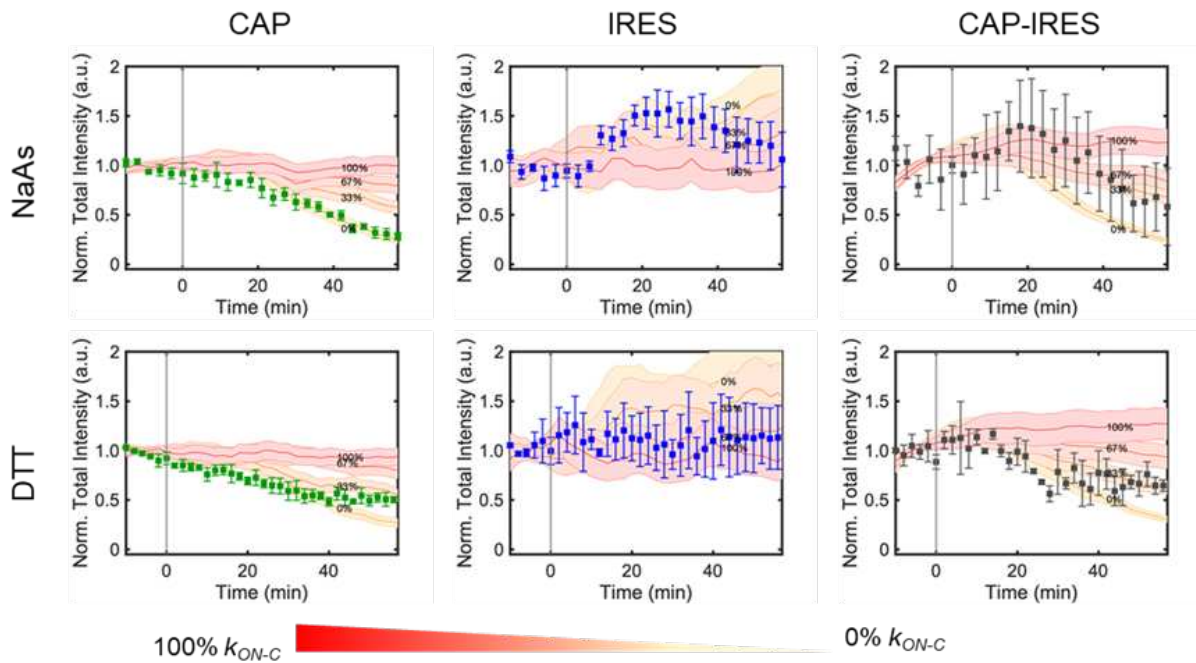
A) Best model that fits all data



Best model fit to all data



F) Stressing cells with NaAs and DTT and best model predictions



**Figure A.1: Modeling bicistronic translation of the multicolor biosensor.**

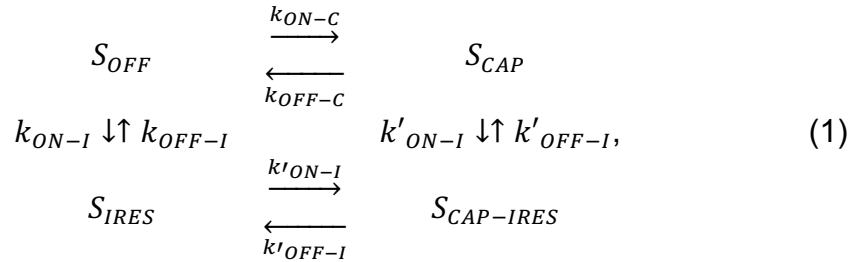
(A) The mathematical model considers four mutually exclusive RNA states: non-translating (OFF), Cap-dependent (CAP-ON), IRES-mediated (IRES-ON), and both Cap and IRES (CAP+IRES-ON), in which initiation can take place from the Cap or IRES as indicated. Elongation and termination processes continue independent of RNA state. To capture interdependence between Cap and IRES states, multiple hypotheses were tested, and the best model was selected after parameter optimization and model reduction (Figure A.2). The selected model considers 4 promoter states, in which the IRES activation rate depends on the Cap-state (model  $4S_{Im2}$ ). (B) Mean values and SD for the model and mean values and SEM for experimental data for the fraction of Cap-dependent (Cap), IRES-mediated (IRES), both Cap-dependent and IRES-mediated (Cap+IRES), and non-translating spots (NT). The prevalence of translation events are shown as the percentage of total RNA. For simulation  $n = 4000$  trajectories were used, SD was calculated using bootstrapping (X50 with sampling of 300 trajectories). For experimental data,  $n=39$  cells. (C) Experimental and model intensity distributions for Cap-dependent and IRES-mediated translation. Distributions consider only those spots that have intensities greater than or equal to one unit of mature protein (u.m.p.). (D) Decrease in intensity after harringtonine application for Cap-mediated translation spots and (E) IRES-mediated translation spots. To denote variability, 10 independent model simulations are plotted in D and E. (F) Experimental data and simulated predictions for translation inhibition by the chemical stressors NaAs and DTT. Chemical stressors were simulated by reducing the Cap activation rates at the RNA state level (i.e. blocking  $k_{ON-C}$ ) for NaAs and by reducing Cap-initiation (i.e. blocking  $k_{INIT-C}$ ) rates for DTT. Experimental data are represented by the square symbols. Errors bars in the experimental data are the mean and SEM and for the simulations mean and SD. NaAs:  $n=32$  cells. DTT:  $n=28$  cells. Simulations:  $n = 4,000$  trajectories. The values given in the figure represents the percentage of inhibition. Cap, Cap-only spots; IRES, IRES-only spots; and Cap-IRES, Cap translation intensity in spots with both Cap and IRES intensities.

## A.2 Computational Details.

A stochastic model was implemented to simulate Cap and IRES activation, ribosome initiation, elongation, termination, and potential ribosome recycling mechanisms for cap-dependent and IRES-mediated genes.

In the mathematical model, initiation events are dictated by the mRNA state.

Specifically, four possible mRNA activation states were proposed ( $S_{OFF}$ ,  $S_{CAP}$ ,  $S_{IRES}$ ,  $S_{CAP-IRES}$ ), where:  $S_{OFF}$  represents a non-permissive initiation state;  $S_{CAP}$  allows for only cap-dependent ribosomal initiation;  $S_{IRES}$  allows for only IRES-mediated initiation; and  $S_{CAP-IRES}$  allows both cap-dependent and IRES-mediated initiation. Eq. 1 represents the transition reactions between mRNA states.



where each  $k_x$  represents a first-order transition rate between two RNA states, and  $k'_x$  is the transition rate conditioned on the activation state of the other construct (e.g.,  $k'_{ON-I}$  is the cap-dependent activation rate of IRES). A simpler three-state model was considered by removing the fourth RNA state (i.e.,  $S_{CAP-IRES}$ ). The parameter estimation section describes how a system with three or four mRNA states was chosen.

When the system is in one of the appropriate mRNA activity states, cap-dependent and IRES-mediated initiation events occur with propensities  $w_{INIT-C}$  and  $w_{INIT-I}$ , respectively, which are defined:

$$w_{INIT-C} = \begin{cases} k_{INIT-C}, & \text{if } S_{CAP} \text{ or } S_{CAP-IRES} , \\ 0, & \text{otherwise,} \end{cases} \quad (2)$$

$$w_{INIT-I} = \begin{cases} k_{INIT-I}, & \text{if } S_{IRES} \text{ or } S_{CAP-IRES} , \\ 0, & \text{otherwise,} \end{cases} \quad (3)$$

where  $k_{INIT-C}$  and  $k_{INIT-I}$  represent the cap and IRES initiation rates, respectively.

To simulate the model under stochastic dynamics, Eqs. (2) and (3) were used to generate a vector of random initiation event times for each gene,  $\tau_{INIT_{IRES}}$  and  $\tau_{INIT_{CAP}}$ .

A codon-dependent model for translation was used, in which the elongation rate for each codon is given by  $\bar{k}_e \left( \frac{u_i}{\bar{u}} \right)$ , where  $u_i$  is the known frequency of the  $i^{th}$  codon in the human genome,  $\bar{u}$  is the average codon usage frequency in the human genome, and  $\bar{k}_e$  is the basal elongation rate (to be estimated from the data). In the models, the final codon termination rates were assumed to be equal to the average elongation rate. For increased computational efficiency, ribosome elongation was approximated using a coarse-grained procedure. For this, sparse ribosome loading was assumed to enable simple calculation of the average time needed by a ribosome to complete gene elongation,  $\tau_{k_e}$ , as follows:

$$\tau_{k_e} = \sum_{i=1}^L \frac{1}{\bar{k}_e \left( \frac{u_i}{\bar{u}} \right)}, \quad (4)$$

where  $L$  represents the gene length in codons. Using the specific gene sequence for the cap-dependent gene and IRES-mediated gene, we calculated the total elongation time  $\tau_{CAP}$  and  $\tau_{IRES}$ , respectively. At any time,  $t$ , such that  $0 < t - \tau_{INIT_{CAP}} < \tau_{CAP}$ , the position of a given cap-translating ribosome was obtained by calculating the proportion of elongated gene as follows:

$$x_{CAP} = j \text{ such that } \sum_{i=1}^j \frac{1}{\bar{k}_e(u_i/\bar{u})} \leq t - \tau_{INIT_{CAP}} < \sum_{i=1}^{j+1} \frac{1}{\bar{k}_e(u_i/\bar{u})}, \quad (5)$$

and for the IRES-mediated gene for  $0 < t - \tau_{INIT_{IRES}} < \tau_{IRES}$ :

$$x_{IRES} = j \text{ such that } \sum_{i=1}^j \frac{1}{\bar{k}_e(u_i/\bar{u})} \leq t - \tau_{INIT_{IRES}} < \sum_{i=1}^{j+1} \frac{1}{\bar{k}_e(u_i/\bar{u})}, \quad (6)$$

where  $\tau_{INIT_{CAP}}$  and  $\tau_{INIT_{IRES}}$  are the times at which the corresponding ribosome initiated translation begins.

To consider potential interaction mechanisms between cap-dependent and IRES-mediated translation, two possible hypotheses were postulated:

A first hypothetical model considers potential ribosome recycling (or crossover) mechanisms, by which a ribosome that completes translation of the cap-dependent gene could immediately re-initiate translation of the IRES-mediated gene. In this context, a new propensity,  $w_{CI}$ , that specifies the probability that a ribosome completing cap will re-initiate at IRES was introduced. The specification of such reactions reflects single-mRNA translation observations by Wang et al., 2016 [4], which suggest ribosome hops between adjacent open reading frames on a single RNA. To test if such recycling mechanisms are necessary to reproduce the experimental data, multiple models with and without nonzero values for the crossover rate  $w_{IC}$  were compared.

In the second hypothetical model, cap and IRES interdependency were tested by assuming that the activation and deactivation of cap or IRES could depend on the activity state for the other sensor (e.g., IRES could activate faster when cap is already active). Including different combinations of these hypothetical mechanisms in the three- and four-state models led us to propose a list of 14 different sub-models, each comprising between 7 and 12 free parameters (Figure A.2). The sub-models test



different hypotheses, including variations of the number of mRNA states (3 or 4 states), dependency on Cap and IRES switching states, and/or the existence of the cross-over mechanism. Cap and IRES dependency are represented in the Figure A.2B by red lines, which denote that the corresponding reaction parameter value has a free value during the optimization process. All models have 3 or 4 mRNA states, denoted by 3S or 4S, respectively. From Figure A.2B left to right, the first seven models lack crossover, while the last seven have cross-over (denoted by subscript 'C', e.g. 3SC). Models can have independent (denoted by subscript 'I') or dependent (denoted by subscript 'D') Cap or IRES activation/deactivation. Models can also have a single dependent activation or deactivation rate (denoted by subscript 'm1' or 'm2').

### ***Converting ribosome elongation times to fluorescence intensity***

To relate the ribosome elongation times to fluorescence intensity, a similar approach as in Aguilera et al.[1] was adopted. Ribosome occupancy is converted to fluorescence intensity by increasing the simulated intensity by one unit after each ribosome moves across the tag-region. For this, a cumulative *probe design vector* was defined that records the number of probe sites upstream from each codon,  $\mathbf{c}_g = [c_1, c_2, \dots, c_L]$ , for the appropriate construct (i.e.,  $g = \text{cap-dependent or IRES-mediated genes, respectively}$ ). Using this, the intensity was calculated as the sum of the product of the position of the ribosome at a given time and  $\mathbf{c}_g$ . For cap-dependent spots, the intensity vector is defined as:

$$I_{CAP}(t) = \sum \mathbf{c}_{CAP}(\mathbf{x}_{CAP}(t)), \quad (8)$$

and for IRES-mediated spots it is:

$$I_{IRES}(t) = \sum \mathbf{c}_{IRES}(\mathbf{x}_{IRES}(t)), \quad (9)$$

where  $c_g(x_g)$  is the intensity of a given ribosome at position  $x_g$ , and the summations are taken over all ribosomes present on the mRNA at time  $t$ . To have consistent units of intensity between model simulations and experimental data, intensity values are reported in units of mature proteins (u.m.p.) as described in detail on the Methods section.

### ***Comparison of experimental data and model***

To reproduce experimental data, the model was simulated using a modified Direct Method [5] for 4000 trajectories representing independent RNA spots. Simulations were run for a burn-in period of 10,000 seconds to approximate steady state. Simulations were processed and used to capture spot intensity for the cap-dependent gene ( $I_{CAP}$ ) and the IRES-mediated gene ( $I_{IRES}$ ). Additionally, simulated spots were classified as cap-dependent with probability  $P_{CAP}$ , IRES-mediated with probability  $P_{IRES}$ ; both with probability  $P_{CAP-IRES}$ , or neither with probability  $P_{None}$ .

### ***Modeling harringtonine experiments***

Harringtonine inhibits new initiation events by directly blocking the 60S subunit in the ribosome, and it has been widely used to perform run-off assays to estimate elongation rates [6]. To mimic the effects of harringtonine in our model, the initiation rate was modified for the first gene as follows:

$$w_{INIT-C} = \begin{cases} k_{INIT-C}, & \text{if } t < t_H, \\ 0, & \text{otherwise,} \end{cases} \quad (10)$$

and the initiation rate for the second gene as follows:

$$w_{INIT-I} = \begin{cases} k_{INIT-I}, & \text{if } t < t_H, \\ 0, & \text{otherwise,} \end{cases} \quad (11)$$

where  $t_H$  is the time of application of harringtonine.

### **Modeling Sodium Arsenite (NaAs) and Dithiothreitol (DTT) experiments**

NaAs and DTT are chemical stresses that have been used to affect cap-dependent initiation in previous single-molecule translation experiments [4]. The mechanism of action for NaAs is not well understood, but it has been suggested to affect ribosome initiation through its action on translation factors, such as eIF2a and eIF4 [7]. To simulate these chemical stresses, two potential mechanisms of action were tested. The first potential mechanism of action involves blocking cap-dependent translation by affecting its RNA state, and was implemented in the model by modifying the cap activation rates,  $k_{ON-C}$  and  $k'_{ON-C}$ , as follows:

$$k_{ON-C} = \begin{cases} k_{ON-C}, & \text{if } t < t_{ST}, \\ k_{ST} \cdot k_{ON-C}, & \text{otherwise,} \end{cases} \quad (12)$$

and

$$k'_{ON-C} = \begin{cases} k'_{ON-C}, & \text{if } t < t_{ST}, \\ k_{ST} \cdot k'_{ON-C}, & \text{otherwise,} \end{cases} \quad (13)$$

where,  $k_{ST}$  is an inhibition constant, where a total inhibition is achieved by  $k_{ST} = 0$ , and a null inhibition is achieved by  $k_{ST} = 1$ . The time  $t_{ST}$  denotes the time of stress application.

In the second mechanism of action, it was hypothesized that the drug directly blocks cap-dependent translation initiation. In the model, this is achieved by modifying  $w_{ini}$  as follows:

$$w_{INIT-C} = \begin{cases} k_{INIT-C}, & \text{if } t < t_{ST}, \\ k_{ST} \cdot k_{INIT-C}, & \text{otherwise.} \end{cases} \quad (14)$$

### **Parameter estimation and optimization routines**

The parameter estimation strategy consists of finding a parameter set ( $\bar{\Lambda}$ ) that statistically reproduces all experimental data, including intensity histograms, fractions of translating spots, and harringtonine ribosomal run-off assays as follows:

#### **Intensity histograms**

To compare experimental and simulated steady-state intensity histograms, the probability to observe the experimentally determined intensities ( $d^{cap}$  or  $d^{IRES}$ ) was estimated given a parameter set ( $\Lambda$ ) in the model implementation. To estimate  $P(d; \Lambda)$ , histograms were collected using  $N_t = 4000$  independent stochastic trajectories per parameter evaluation. The likelihood function was estimated as follows:

$$L_{Dist}(D|M) = \prod_{j=1}^{N_D} P(d_j; \Lambda), \quad (15)$$

and the log-likelihood as:

$$\log L_{Dist}(D|M) = \sum_{j=1}^{N_D} \log P(d_j; \Lambda), \quad (16)$$

where  $D$  represents the data measured in  $N_D$  independent experimental data, and  $M$  corresponds to the model. As the experimental measurements can only detect protein intensities above a threshold of one mature protein, all spots with intensities below 1 u.m.p. were defined as non-translating mRNA. This metric was applied to experimental data consisting of cap-dependent spots (CAP) and IRES-mediated spots (IRES). With this, a total log-likelihood function was calculated as the sum of the functions for cap and IRES spots, that is:

$$\log L_{T_{Dist}}(D|M) = \sum_{j=1}^{N_t} \log P_{CAP}(d_j^{cap}; \Lambda) + \sum_{j=1}^{N_t} \log P_{IRES}(d_j^{IRES}; \Lambda). \quad (17)$$

#### **Fraction of translating spot**

A similar approach was used to compute the likelihood to observe the experimentally determined number of spots classified as Cap-only, IRES-only, Cap+IRES, and non-translating. The likelihood function was computed as follows:

$$L_F(D|M) = \prod_{j=1}^{N_D} P(f_j; \Lambda), \quad (18)$$

and the log-likelihood as:

$$\log L_F(D|M) = \sum_{j=1}^{N_D} \log P(f_j; \Lambda) = \sum_i N_i \log P(f_i; \Lambda), \quad (19)$$

where each  $f_j$  denotes the type (i.e., Cap, IRES, Cap+IRES, or non-translating) of the  $j^{th}$  spot,  $N_D$  is the total number of independent observed spots,  $N_i$  is the number of independent observed spots of the  $i^{th}$  type, and  $P(f_j; \Lambda)$  is the categorical distribution of spots of each type estimated by the model simulations with parameters  $\Lambda$ .

### ***Harringtonine induced ribosomal run-off***

To compare simulated and experimental time course data representing the intensity after harringtonine application, a Gaussian likelihood function was assumed and calculated as follows:

$$L_{HT}(I_D|I_M) = \prod_{i=1}^{N_D} \frac{1}{\sqrt{2\pi\sigma(t_i)^2}} \exp\left(-\frac{(I_D(t_i)-I_M(t_i;\Lambda))^2}{2\sigma(t_i)^2}\right), \quad (20)$$

with a log-likelihood form given by:

$$\log L_{HT}(I_D|I_M) = C_{HT} - \sum_{i=1}^{N_D} \frac{(I_D(t_i)-I_M(t_i;\Lambda))^2}{2\sigma(t_i)^2}, \quad (21)$$

where  $\sigma(t_i)$  is approximated by the measured SEM, and  $N_D$  is the number of time points from the Harringtonine run-off curve. In this log-likelihood formulation,  $C_{HT}$  is a constant that doesn't depend on the parameters.

Experimental data was quantified for the total intensities for cap ( $I_{CAP-D}$ ) and IRES ( $I_{IRES-D}$ ) within all spots (after subtraction of the base level of intensity). These two data sets were collected to compute a total log-likelihood function as follows:

$$\log L_{T_{HT}}(I_D|I_M) = \log L_{HT_{CAP}}(I_{CAP-D}|I_M) + \log L_{HT_{IRES}}(I_{IRES-D}|I_M). \quad (22)$$

Parameter searches consisted of optimization routines based on genetic algorithms (GA) using the function *ga* in MATLAB. The optimization routine was implemented with a population of 100 individuals for 30 generations, and the implementation was run multiple times with random initial conditions. Additionally, the Pattern Search Algorithm[8] was implemented using the function *patternsearch* in MATLAB to ensure convergence. The best parameter values were selected by minimizing a global objective function that considers all data sets, that is:

$$-\log L_{Total}(D|M) = -(\log L_{T_{Dist}}(D|M) + \log L_F(D|M) + \log L_{T_{HT}}(I_D|I_M)). \quad (23)$$

The comparison of the optimization results for all tested models is given in Figure A.2C-D.

### **Assessing how well models predict Sodium Arsenite (NaAs) and Dithiothreitol (DTT) experiments**

After optimizing the models, cross-validation experiments were predicted using the chemical stresses, NaAs and DTT. For this, simulated and experimental time course data representing the total translation spot intensity after NaAs or DTT application were compared. The likelihood function was calculated as follows:

$$L_{ST}(I_D|I_M) = \prod_{i=1}^{N_D} \frac{1}{\sqrt{2\pi\sigma(t_i)^2}} \exp\left(-\frac{(I_D(t_i)-I_M(t_i;\lambda))^2}{2\sigma(t_i)^2}\right), \quad (24)$$

and the log-likelihood function is:

$$\log L_{ST}(I_D|I_M) = C_{ST} - \sum_{i=1}^{N_D} \frac{(I_D(t_i) - I_M(t_i; \Lambda))^2}{2\sigma(t_i)^2}, \quad (25)$$

where  $\sigma(t_i)$  is approximated by the measured SEM, and  $N_D$  is the number of time points measured in the drug-treatment curve, and  $C_{ST}$  is constant that doesn't depend on model parameters.

For chemical stress experiments, three data sets were used representing the intensity for Cap-only spots, IRES-only spots, and green (Cap) intensity in both cap and IRES spots. These three data sets were considered on a total log-likelihood function as follows:

$$-\log L_{TST}(I_D|I_M) = -(\log L_{STCAP}(I_{CAP-D}|I_M) + \log L_{STIRES}(I_{IRES-D}|I_M) + \log L_{STCI}(I_{Cap+IRES-D}|I_M)). \quad (26)$$

### **Uncertainty Quantification**

To quantify uncertainty, the best parameter set from fitting was initially used and 100 runs of 1,000 step Markov Chain Monte Carlo (MCMC) algorithm were run to explore an additional 100,000 possible parameter combinations. At each step, a random perturbation of 10% to the current parameters was proposed, and every proposal for which the log-likelihood for the new parameter set was within a 1% of that found for the best fit was accepted (i.e., all parameters for which  $\log(L(I_D|I_{Best})/L(I_D|I_{New})) < 60$  were accepted). The standard deviation of the resulting 26,650 accepted parameter sets was then used as a measure of parameter uncertainty as shown in Table 1.





**Figure A.2: Model of the bicistronic gene construct.**

(A) The most complete mathematical model considers four mutually exclusive RNA states: non-translating (OFF), Cap-dependent (CAP-ON), IRES-mediated (IRES-ON), and both Cap and IRES (CAP+IRES-ON). All transition rate values between RNA states are free-independent parameters. A cross-over mechanism (CO symbol in the figure), by which a ribosome that completes the translation of the Cap-dependent protein could immediately re-initiate translation of the IRES-mediated protein, is represented by the reaction parameter  $k_{CO}$ .

(B) Comparison of 14 different sub-models. The sub-models test different hypotheses, including variations of the number of mRNA states (3 or 4 states), dependency on Cap and IRES switching states, and/or the existence of the cross-over mechanism. Cap and IRES dependency are represented in the figure by red lines, which denote that the corresponding reaction parameter value has a free value during the optimization process. All models have 3 or 4 mRNA states, denoted by 3S or 4S, respectively. From left to right, the first seven models lack crossover, while the last seven have cross-over (denoted by subscript 'C', e.g. 3S<sub>C</sub>). Models can have independent (denoted by subscript 'I') or dependent (denoted by subscript 'D') Cap or IRES activation/deactivation. Models can also have a single dependent activation or deactivation rate (denoted by subscript 'm1' or 'm2'). The number of free parameters in the sub-models ranges from 7 to 12.

(C) Cross-validation is used to compare two possible mechanisms of translation inhibition under DTT stress. The first mechanism mimics the inhibition of the Cap activation rates at the promoter level ( $L_{NaAs-STATE-CAP}$ ; i.e., block of  $k_{ON-C}$  and  $k'_{ON-C}$ ). The second mechanism considers blocking ribosomal initiation for Cap ( $L_{NaAs-INIT-CAP}$ ; i.e., block of  $k_{INIT-C}$ ).

(D) Optimization process and cross-validation for the DTT stress. The same inhibitory mechanisms described in C are tested for DTT stress. Relative Log-likelihood values for the optimization process are calculated according to Eq. 23 and according to Eq. 26 for the NaAs and DTT cross-validation experiments, respectively. The log-likelihood reported are relative to the minimum value from all models. Relative log-likelihood values over 500 are not plotted.

A selection threshold (dashed red line) was defined by a log-likelihood of 100 worse than the most complex and best fitting model. Models above the selection threshold were discarded (gray background), and their cross-validation log-likelihood values are not shown. The best model shown (green background) was chosen as the model with fewest free parameters below the selection threshold.

(E) Model simulations for the best-fit model 4S<sub>Im2</sub> under NaAs and DTT stresses. The figure shows the effect of blocking ribosomal initiation for Cap. The assumption of blocking initiation results in significantly worse predictions compared to the hypothesis that the drugs block Cap activation (compare to Figure A.1).

# Appendix B

## Experimental material and methods for Chapter 2

### *Plasmid construction*

The Original Tag (SM-KDM5B-EMCV-SunTag-Kif18b-MS2) contains a spaghetti monster (SM) with 10× FLAG epitopes, a SunTag with 24× SunTag epitopes, and an MS2 repeat with 24× MS2 stem loops. The coding region of the SunTag and Kif18b was obtained by polymerase chain reaction (PCR) of a pCMV-SunTag-Kif18b-PP7 template (Addgene #128606), using the following primers: 5'-GCC GAA AGG TTT AAA CGC TAG CTC TGG AGG AGA AGA ACT TTT GAG CAA GAA T-3'; 5'-AGT AAC AGT CCG CCT AGG TCC TTA TCG GAC ACC TTG GT-3'. The PCR product contained arms of homology to the acceptor plasmid (SM-KDM5B-MS2; Addgene #81084). The acceptor plasmid was cut with NheI (New England BioLabs) between the end of KDM5B and the MS2 stem loops. The PCR product and cut acceptor plasmid were assembled via Gibson Assembly (homemade mixture). The resulting plasmid was SM-KDM5B-Nhe1-SunTag-Kif18b-MS2, which was also used as the NoIRES construct. The EMCV IRES sequence was amplified by PCR from EMCV\_IRES\_pcDNA4TO\_H2B\_SunTag24x\_v1 (Addgene #246719) using the following primers: 5'- CCG AAA GGT TTA AAC GCT AGC acg tta ctg gcc gaa -3'; 5'- TTC TTC TCC TCC AGA GCT AGC tat tat cat cgt gtt ttt caa agg aaa -3'. The PCR product contained arms of homology to the acceptor plasmid (SM-KDM5B-Nhe1-SunTag-Kif18b-MS2). The acceptor plasmid was cut with NheI (New England BioLabs) between the end of KDM5B and beginning of SunTag. The PCR

product and cut acceptor plasmid were assembled via Gibson Assembly. The start codon for SunTag-Kif18b is within the EMCV IRES sequence.

For the construction of the Switch Tag (SunTag-Kif18b-EMCV-SM-KDM5B-MS2), the coding region of the SunTag and Kif18b was obtained by PCR of a pCMV-SunTag-Kif18b-PP7 template (Addgene #128606), using the following primers: 5'-TCG CTG TGA TCG TCA CTT GGC GGA cac cat gGA AGA ACT TTT GAG CAA GAAT-3'; 5'-CGT CCT TGT AGT CCA TGG TGG Cgg cgc gcc GTC Tta gaT ATC GGA CAC CTTG-3'. The PCR product contained arms of homology to the acceptor plasmid (SM-KDM5B-MS2 Addgene #81084). The acceptor plasmid was cut with NotI (New England BioLabs) at the beginning of SM. The PCR product and cut plasmid were assembled via Gibson Assembly. The resulting plasmid was SunTag-Kif18b-Nhe1-SM-KDM5B-MS2. The EMCV IRES sequence was amplified by PCR from EMCV\_IRES\_pcDNA4TO\_H2B\_SunTag24x\_v1 (Addgene #246719) using the following primers: 5'- CCA AGG TGT CCG ATA tct aAG ACg gcg tta ctg gcc gaa gcc gct -'3; 5'- CCT TGT AGT CCA TGG TGG Cgg cgc ata tta tca tcg tgt ttt tca aag gaa aac cac- 3'. The PCR product contained arms of homology to the acceptor plasmid (SunTag-Kif18b-Ascl-SM-KDM5B-MS2). The acceptor plasmid was cut with Ascl (New England BioLabs) between the end of Kif18b and beginning of SM. The PCR product and cut acceptor plasmid were assembled via Gibson Assembly. The start codon for SM-KDM5B is within the EMCV IRES sequence.

### ***anti-FLAG Fab generation and dye-conjugation***

Pierce mouse IgG1 preparation kit (Thermo Scientific) was used to generate Fab according to the manufacturer's instruction. Briefly, immobilized ficin in the presence of

25 mM cysteine was used to digest FLAG antibodies (Wako, 012-22384 Anti DYKDDDDK mouse IgG2b monoclonal) to create Fab. Fab were separated from the Fc region using NAb Protein A column. After elution, Fab were concentrated to 1 mg/ml and conjugated to Cy3. Cy3 N-hydroxysuccinimide ester (Invitrogen) was dissolved in DMSO and stored at -20°C. 100 µg of Fab were diluted into 100 µl of 100 mM NaHCO<sub>3</sub> (pH 8.5). 1.33 µl of Cy3 was added to this solution and incubated with end-over-end rotation for 1-2 hours at room temperature. The conjugated Fab were then eluted from a PBS pre-equilibrated PD-mini G-25 desalting column (GE Healthcare) that removed unconjugated dye. Conjugated Fabs were then concentrated using an Ultrafree 0.5 filter (10k-cut off; Millipore) to 1 mg/ml. The Fab:dye ratio was calculated using the absorbance at 280 and 550 nm, and using the extinction coefficient of Fab with the dye correction factor at 280 nm provided by the manufacturers (0.08 for Cy3). The degree of labeling was calculated using the following formula:

$$DOL = \left( \frac{\epsilon_{IgG}}{\epsilon_{dye}} \right) \left( \frac{1}{(A_{rat})^{-1} - CF} \right)$$

### ***MCP and scFv-GFP purification***

His-tagged MCP/scFv-GFP was purified with Ni-NTA-agarose (Qiagen) following the manufacturer's instructions with minor modifications. Briefly, the bacteria were lysed in a PBS-based buffer containing a complete set of protease inhibitors (Roche), binding to the Ni-NTA resin was carried out in the presence of 10 mM imidazole. After washing with 20 and 50 mM imidazole in PBS, the protein was eluted with 300 mM imidazole in PBS, and directly used for experiments. The rest was dialyzed against a HEPES-based buffer (10% glycerol, 25 mM HEPES pH 7.9, 12.5 mM MgCl<sub>2</sub>, 100 mM KCl, 0.1 mM

EDTA, 0.01 % NP-40 detergent, and 1 mM DTT) and stored at -80 °C after snap-freezing by liquid nitrogen.

### ***Cell culture and beadloading***

U2OS cells were grown in DMEM (Thermo Scientific) supplemented with 10% (v/v) FBS, 1 mM L-glutamine and 1% (v/v) Penicillin-streptomycin (DMEM+). One to two days prior to experiments, cells were plated into a 35 mm MatTek chamber at approximately 70-80% confluency. Two to four hours prior to experiments, cells were put in OPTI-MEM (Thermo Scientific) supplemented with 10% (v/v) FBS (OPTI-MEM+). Cells were then bead-loaded with fluorescently labeled Fab, GFP-fused scFv, MCP-HaloTag protein and purified DNA of interest. Briefly, 100 µg/ml of fluorescently labeled Fab, 100 µg/ml of purified GFP-fused scFv, 33 µg/ml of purified MCP HaloTag protein, and 750ng of DNA of interest were prepared in a total volume of 4µl of 1xPBS. After removing OPTI-MEM and FBS, the 4µL solution was pipetted to the top of the cells. Then, ~106 µm glass beads (Sigma Aldrich) were sprinkled evenly over the cells. The chamber was tapped firmly 12 times on the bench, and OPTI-MEM+ was added back to the cells. Two hours after bead-loading, cells were washed twice with phenol-red-free DMEM+ such that all beads were removed. 200 nM of JF646-HaloTag ligand was next added (1µL of 200nM to 1mL of phenol-red-free DMEM+). After 20 minutes of incubation at 37 °C, the cells were washed twice with phenol-red-free DMEM+ to remove excess ligand. 2 mL of phenol-red-free DMEM+ was added back to the cells. Translation experiments were conducted immediately after washing. U2OS cells were purchased from ATCC and were authenticated by STR profiling by ATCC and

morphological assessments. We also confirmed that all cell lines tested negative for mycoplasma contamination.

### ***Single molecule tracking microscopy***

To track single-molecule mRNA translation events, we used a custom-built widefield fluorescence microscope with a highly inclined illumination scheme.[6,9] Briefly, the excitation beams, 488, 561 and 637 nm solid-state lasers (Vortran), were coupled and focused on the back focal plane of the objective (60X, NA 1.49 oil immersion objective, Olympus). The emission signals were split by an imaging grade, ultra-flat dichroic mirror (T660lpxr, Chroma) and detected by two aligned EM-CCD cameras (iXon Ultra 888, Andor) by focusing with 300 mm tube lenses (this lens combination produces 100X images with 130 nm/pixel). Live cells were placed into an incubation chamber (Okolab) at 37 °C and 5% CO<sub>2</sub> on a piezoelectric stage (PZU-2150, Applied Scientific Instrumentation). The focus was maintained with the CRISP Autofocus System (CRISP-890, Applied Scientific Instrumentation). Image acquisition was performed using open source Micro-Manager.[10] With this setting, one camera detected far-red emission signals while the other detected either red or green emission signals.

Far-red signals were excited with the 637 nm laser with a 731/137 nm emission filter (FF01-731/137/25, Semrock). Red and green signals were separated by the combination of the excitation lasers and the emission filters installed in a filter wheel (HS-625 HSFW TTL, Finger Lakes Instrumentation); namely, the 561 nm laser and 593/46 nm emission filter (FF01-593/46-25, Semrock) were used for Cy3 imaging, and the 488 nm laser and 510/ 42 nm emission filter (FF01-510/42-25, Semrock) were used

for sfGFP or A488 imaging. The lasers, filter wheel, cameras, and the piezoelectric stage were synchronized by an Arduino Mega board (Arduino). The exposure time of the cameras was selected as 53.64 msec throughout the experiments. The readout time for the cameras from the combination of imaging size, readout mode, and the vertical shift speed was 23.36 msec, resulting in an imaging rate of 13 Hz (77 msec per image). The excitation laser lines were digitally synched to ensure they only illuminated cells when the camera was exposing to avoid excessive photobleaching. To capture the entire volume of the cytoplasm of U2OS cells, 13 z stacks with a step size of 500 nm (6  $\mu$ m in total) were acquired using the piezoelectric stage. Because one image of Cy3 was captured on one camera and one image of sfGFP/A488 + JF646 was captured on the other camera in the same stack of the cell, the z-position within the cell changed every two images. The position of the filter wheel was changed during the camera readout time. This resulted in a total cellular imaging rate of 0.5 Hz (2 s per volume for 3-colors). Note that all colors described in the text and that are shown in the figures are based on the color of the excitation laser: RNA in red (JF646) and protein in green (Cy3) or blue (sfGFP).

#### ***Cell imaging conditions with no drugs added for all constructs***

Cells beadloaded with SM-KDM5B-EMCV-SunTag-Kif18b-MS2 (Original Tag), SunTag-Kif18b-EMCV-SM-KDM5B-MS2 (Switch Tag), or SM-KDM5B-SunTag-Kif18b-MS2 (NoIRES Tag), Cy3 labeled anti-FLAG Fab, Halo-MCP protein (labeled with JF646-HaloTag ligand), and anti-SunTag scFv-GFP were imaged with a 6 second interval between every 13 captures (one entire cell volume) for 25-50 total time-points.

Laser powers for all images were: 15mW for 637nm, 9mW for 488nm, and 5mW for 561nm with an ND10 neutral density filter at the beam expander.

### ***Particle tracking***

Collected images were first pre-processed with Fiji.[11] Briefly, the 3D images were projected to 2D images by a maximum intensity projection and background subtracted. Post-processed images were then analyzed by a custom-written *Mathematica* (Wolfram Research) routine to detect and track particles in the RNA channel (red color). Specifically, particles were emphasized with a band-pass filter such that the positions could be detected using the built-in *Mathematica* routine `ComponentMeasurements` “IntensityCentroid.” Detected particles were linked through time by allowing a maximum displacement of 5 pixels between consecutive frames. Particle tracks lasting at least 5 frames were selected. To properly account for the offset between the two cameras, a geometric transformation function (see method below) was applied to the coordinates of the center of mRNAs. For each frame of each track, 15x15 (pixels x pixels) crops centered on the registered mRNA coordinate were made and averaged through time. Using *Mathematica*’s bandpass filter and `ComponentMeasurements` described above, the time-averaged crops corresponding to each track were categorized based on the presence of detectable signals in the green and blue nascent chain channels: Red – mRNA not translating, Red + Green = Yellow – mRNA translating in Cap Only for Original Tag or IRES Only for Switch Tag, Red + Blue = Purple – mRNA translating in IRES Only for Original Tag and Cap Only for the Switch Tag, Red + Green + Blue = White – mRNA translating in both cap and IRES manner.



Once the spots were categorized in this automated fashion, all spots were again hand-checked to minimize error.

Finally, the original 2D maximum intensity projected images corresponding to each hand-checked track were fit to find their precise coordinates and intensities (using the built-in *Mathematica* routine NonlinearModelFit) to a 2D Gaussians of the following form:

$$I(X, Y) = I_{BG} + I e^{-\frac{(x-x_0)^2}{2\sigma_x^2} - \frac{(y-y_0)^2}{2\sigma_y^2}}$$

where  $I_{BG}$  is the background intensity,  $I$  the particle peak intensity,  $(\sigma_x, \sigma_y)$  the spread of the particle, and  $(x_0, y_0)$  the particle location. From these data, the intensity, position through time, and number of spots over time in each track were quantified for downstream analysis.

### ***Fast imaging for mean square displacement analysis***

For fast particle tracking to accurately quantify the mean square displacements, single planes of cells loaded with the Original Tag construct, Cy3 anti-FLAG Fab, anti-GCN4 scFv-GFP, and Halo-MCP were imaged with an imaging rate of 77 msec.

### ***Geometric transformation function***

The offset between the two cameras was registered using the built-in *Mathematica* routine FindGeometricTransform. To find the transformation function that best aligned the fitted positions, 100 nm diameter Tetraspeck beads evenly spread out across the image field-of-view were imaged on the same day experiments were taken. Only the fitted particle positions were registered to avoid introducing any distortion into images. Therefore, a slight offset can be observed between the red and the green or

blue particles even though they are within a diffraction limited spot according to our registration.

### ***Calibrating translation site intensity***

We wanted to quantify the units of mature protein (i.e. number of nascent chains or active ribosomes) at a translation site using its intensity signal. For this purpose, we imaged two calibration constructs.[6] The two calibration constructs were equal in length, one containing the spaghetti monster 10×FLAG tag (SM-BetaActin) which contains 10 repeats of the FLAG epitope, the other containing just a single FLAG epitope (1×FLAG-filler-BetaActin). Note the spaghetti monster 10×FLAG tag (SM) is the same tag utilized in the Original Tag (SM-KDM5B-IRES-ST-Kif18b) and the Switch Tag (ST-Kif18b-IRES-SM-KDM5B). The 1×FLAG-filler-BetaActin tag was used to measure the number of ribosomes translating in a translation site. With the 1×FLAG-filler-BetaActin calibration construct, each nascent chain in a translation site contains just one FLAG epitope labeled by a single Fab conjugated (on average) to a single Cy3 fluorophore. By imaging this 1×FLAG construct at high laser powers such that individual translation sites and single Cy3 fluorophores (confirmed by single-step photobleaching, see below) can both be visualized, the ratio of total Cy3 signal in translation sites to single Cy3 fluorophore intensity signals approximates the number of nascent chains (or ribosomes) per translation site. To quantify this, we imaged cells beadloaded with 1xFLAG-filler-BetaActin and Cy3 anti-FLAG Fab in a single optical plane at high laser powers (50 mW for 561nm and 15mW for 637nm laser). A short movie was acquired, after which cells were continually illuminated to photobleach them to the point at which single probe fluorescence could easily be detected by single-step photobleaching. At

this point, a second short 250-frame movie was acquired. The intensity of polysomes (verified by the presence of an RNA signal intensity) from the first frame of the first movie was then measured (as described in the 'Particle tracking' section above) and compared to the plateau intensity of a single Cy3 just prior to single-step photobleaching. The ratio of polysome to Cy3 intensities gives us an estimate for the number of ribosomes on a single RNA. By averaging all measured RNA together, we obtain a mean value for the ribosomal occupancy of a single, translating, 1×FLAG-filler-BetaActin mRNA. From these measurements,  $11.4 \pm 2.0$  ribosomes were estimated to be translating the 1×FLAG-filler-BetaActin calibration construct.

Since the 1×FLAG-filler-BetaActin calibration construct and the SM-BetaActin calibration construct (with 10×FLAG) are the same length with the same promoters and 3' and 5' UTRs, their translation sites should contain roughly the same number of ribosomes ( $11.4 \pm 2.0$ ). With a known number of ribosomes on SM-BetaActin, we wished quantify the number of ribosomes on the cap ORF in the Original Tag (SM-KDM5B-IRES-ST-Kif18b) and the IRES ORF in the Switch Tag (ST-Kif18b-IRES-SM-Kif18b). To do this, we imaged cells in two different chambers. In one chamber, we beadloaded cells with anti-FLAG Fab (Cy3) and SM-BetaActin. In the second chamber, we beadloaded cells with anti-FLAG (Cy3) and the Original Tag (SM-KDM5B-IRES-ST-Kif18b). Both chambers were imaged on the same day with the same imaging conditions (50 mW for 561nm and 15mW for 637nm laser). By taking the intensity ratio of SM-BetaActin containing  $11.4 \pm 2.0$  ribosomes and Cap Only translation sites in the Original Tag (SM-KDM5B-IRES-ST-Kif18b), we measured the ribosomal occupancy of the cap ORF to be  $14.6 \pm 5.6$  ribosomes. Taking the intensity ratio of these translation

sites and all other SM translation sites in the Original Tag (Cap in Cap+IRES translation sites) and Switch Tag (ST-Kif18b-IRES-SM-KDM5B) (IRES in IRES Only and Cap+IRES translation sites) gave the number of ribosomes translating in all possible translation sites, as shown in Fig. 4c.

***Ribosome run-off experiments using harringtonine treatment and elongation estimates***

To measure average elongation rates, cells beadloaded with the Original Tag (SM-KDM5B-EMCV-SunTag-Kif18b-MS2), Cy3 labeled anti-FLAG Fab, Halo-MCP protein (labeled with JF646-HaloTag ligand), and anti-SunTag scFv-GFP were imaged with a 60 second interval between every 13 frames (one entire cell volume) for 50 total time-points. Laser powers were the same as previously described for general imaging. After acquiring 5 time-points of pre-treated images, cells were treated with a final concentration of 3  $\mu\text{g/mL}$  of harringtonine (Cayman Chemical). After treatment, cells were imaged for the remaining 45 time-points as described. As a photobleaching control, cells were imaged at the exact same imaging conditions described previously however no drug was added.

To generate ribosomal run-off curves, images were analyzed with the particle tracker as previously described. In each frame of each cell image, nascent chain signal intensities from all translation sites were totaled resulting in an intensity decay curve over time for each individual cell. Each decay curve was normalized to the average value from the first five frames (preceding drug addition). Each individual curve was fit to the following phenomenological equation:

$$\frac{1 - \text{Tanh}(ax - b)}{2}$$

where the linear part of the fit (the slope at  $ax=b$ ) provides a good estimate of the elongation rates in amino acids over time in seconds.[12] From this, elongation rates for each cell were calculated.

### ***Puromycin treatment***

Cells beadloaded with the Original Tag (SM-KDM5B-EMCV-SunTag-Kif18b-MS2) or the Switch Tag (SunTag-Kif18b-EMCV-SM-KDM5B-MS2), Cy3 labeled anti-FLAG Fab, Halo-MCP protein (labeled with JF646-HaloTag ligand), and scFv-GFP were imaged with a 60 second interval between every 13 frames (one entire cell volume) for 15 total time-points. After acquiring 5 time-points of pre-treated images, cells were treated with a final concentration of 50 $\mu$ g/mL of puromycin. After treatment, cells were imaged for the remaining 10 time-points as described previously. As a photo-bleaching control, cells were imaged at the exact same imaging conditions described previously with no drug added. Three biological replicates were taken.

### ***Sodium Arsenite (NaAs) and Dithiothreitol (DTT) treatment***

Cells beadloaded with the Original Tag (SM-KDM5B-EMCV-SunTag-Kif18b-MS2) or Switch Tag (SunTag-Kif18b-EMCV-SM-KDM5B-MS2), Cy3 labeled anti-FLAG Fab, Halo-MCP protein (labeled with JF646-HaloTag ligand), and anti-GCN4 scFv-GFP were imaged with a 180 second interval for NaAs and 120 for DTT between every 13 frames (one entire cell volume) for 35 total time-points. After acquiring 5 time-points of pre-treated images, cells were treated with a final concentration of 0.5 mM of NaAs or 0.75 mM of DTT. After treatment, cells were imaged for the remaining time-points. As a photo-bleaching control, cells were imaged at the exact same imaging conditions described previously with no drug added. Four biological replicates were taken.

## Appendix C

### Experimental material and methods for Chapter 3

#### ***anti-FLAG Fab generation and dye-conjugation***

Pierce mouse IgG1 preparation kit (Thermo Scientific) was used to generate Fab according to the manufacturer's instruction. Briefly, immobilized ficin in the presence of 25 mM cysteine was used to digest FLAG antibodies (Wako, 012-22384 Anti DYKDDDDK mouse IgG2b monoclonal) to create Fab. Fab were separated from the Fc region using NAb Protein A column. After elution, Fab were concentrated to 1 mg/ml and conjugated to Cy3. Cy3 N-hydroxysuccinimide ester (Invitrogen) was dissolved in DMSO and stored at -20°C. 100 µg of Fab were diluted into 100 µl of 100 mM NaHCO<sub>3</sub> (pH 8.5). 1.33 µl of Cy3 was added to this solution and incubated with end-over-end rotation for 1-2 hours at room temperature. The conjugated Fab were then eluted from a PBS pre-equilibrated PD-mini G-25 desalting column (GE Healthcare) that removed unconjugated dye. Conjugated Fabs were then concentrated using an Ultrafree 0.5 filter (10k-cut off; Millipore) to 1 mg/ml. The Fab:dye ratio was calculated using the absorbance at 280 and 550 nm, and using the extinction coefficient of Fab with the dye correction factor at 280 nm provided by the manufacturers (0.08 for Cy3). The degree of labeling was calculated using the following formula:

$$DOL = \left( \frac{\epsilon_{IgG}}{\epsilon_{dye}} \right) \left( \frac{1}{(A_{rat})^{-1} - CF} \right)$$

#### ***MCP purification***

His-tagged MCP was purified with Ni-NTA-agarose (Qiagen) following the manufacturer's instructions with minor modifications. Briefly, the bacteria were lysed in a PBS-based buffer containing a complete set of protease inhibitors (Roche), binding to the Ni-NTA resin was carried out in the presence of 10 mM imidazole. After washing with 20 and 50 mM imidazole in PBS, the protein was eluted with 300 mM imidazole in PBS, and directly used for experiments. The rest was dialyzed against a HEPES-based buffer (10% glycerol, 25 mM HEPES pH 7.9, 12.5 mM MgCl<sub>2</sub>, 100 mM KCl, 0.1 mM EDTA, 0.01 % NP-40 detergent, and 1 mM DTT) and stored at -80 °C after snap-freezing by liquid nitrogen.

### ***Cell culture and beadloading***

U2OS cells were grown in DMEM (Thermo Scientific) supplemented with 10% (v/v) FBS, 1 mM L-glutamine and 1% (v/v) Penicillin-streptomycin (DMEM+). One to two days prior to experiments, cells were plated into a 35 mm MatTek chamber at approximately 70-80% confluency. Two to four hours prior to experiments, cells were put in OPTI-MEM (Thermo Scientific) supplemented with 10% (v/v) FBS (OPTI-MEM+). Cells were then bead-loaded with fluorescently labeled Fab, GFP-fused scFv, MCP-HaloTag protein and purified DNA of interest. Briefly, 100 µg/ml of fluorescently labeled Fab, 33 µg/ml of purified MCP HaloTag protein, and 750ng of DNA of interest were prepared in a total volume of 4µl of 1xPBS. After removing OPTI-MEM and FBS, the 4µL solution was pipetted to the top of the cells. Then, ~106 µm glass beads (Sigma Aldrich) were sprinkled evenly over the cells. The chamber was tapped firmly 12 times on the bench, and OPTI-MEM+ was added back to the cells. Two hours after bead-loading, cells were washed twice with phenol-red-free DMEM+ such that all beads were

removed. 200 nM of JF646-HaloTag ligand was next added (1 $\mu$ L of 200nM to 1mL of phenol-red-free DMEM+). After 20 minutes of incubation at 37 °C, the cells were washed twice with phenol-red-free DMEM+ to remove excess ligand. 2 mL of phenol-red-free DMEM+ was added back to the cells. Translation experiments were conducted immediately after washing. U2OS cells were purchased from ATCC and were authenticated by STR profiling by ATCC and morphological assessments. We also confirmed that all cell lines tested negative for mycoplasma contamination.

### ***Single molecule tracking microscopy (Chapters 2 and 3)***

To track single-molecule mRNA translation events, we used a custom-built widefield fluorescence microscope with a highly inclined illumination scheme.[6,9] Briefly, the excitation beams, 488, 561 and 637 nm solid-state lasers (Vortran), were coupled and focused on the back focal plane of the objective (60X, NA 1.49 oil immersion objective, Olympus). The emission signals were split by an imaging grade, ultra-flat dichroic mirror (T660lpxr, Chroma) and detected by two aligned EM-CCD cameras (iXon Ultra 888, Andor) by focusing with 300 mm tube lenses (this lens combination produces 100X images with 130 nm/pixel). Live cells were placed into an incubation chamber (Okolab) at 37 °C and 5% CO<sub>2</sub> on a piezoelectric stage (PZU-2150, Applied Scientific Instrumentation). The focus was maintained with the CRISP Autofocus System (CRISP-890, Applied Scientific Instrumentation). Image acquisition was performed using open source Micro-Manager [10]. With this setting, one camera detected far-red emission signals while the other detected either red or green emission signals.



Far-red signals were excited with the 637 nm laser with a 731/137 nm emission filter (FF01-731/137/25, Semrock). Red and green signals were separated by the combination of the excitation lasers and the emission filters installed in a filter wheel (HS-625 HSFW TTL, Finger Lakes Instrumentation); namely, the 561 nm laser and 593/46 nm emission filter (FF01-593/46-25, Semrock) were used for Cy3 imaging, and the 488 nm laser and 510/ 42 nm emission filter (FF01-510/42-25, Semrock) were used for sfGFP or A488 imaging. The lasers, filter wheel, cameras, and the piezoelectric stage were synchronized by an Arduino Mega board (Arduino). The exposure time of the cameras was selected as 53.64 msec throughout the experiments. The readout time for the cameras from the combination of imaging size, readout mode, and the vertical shift speed was 23.36 msec, resulting in an imaging rate of 13 Hz (77 msec per image). The excitation laser lines were digitally synched to ensure they only illuminated cells when the camera was exposing to avoid excessive photobleaching. To capture the entire volume of the cytoplasm of U2OS cells, 13 z stacks with a step size of 500 nm (6  $\mu$ m in total) were acquired using the piezoelectric stage. Note that all colors described in the text and that are shown in the figures are based on the color of the excitation laser: RNA in red (JF646) and protein in green (Cy3).

### ***Cell imaging conditions with no drugs added for all constructs***

Cells beadloaded with 100% optimized construct (SM-100Luc-MS2) or 0% optimized construct (SM-0Luc-MS2), Cy3 labeled anti-FLAG Fab, and Halo-MCP protein (labeled with JF646-HaloTag ligand) were imaged with a 6 second interval between every 13 captures (one entire cell volume) for 25-50 total time-points. Laser

powers for all images were: 20mW for 637nm, and 5mW for 561nm with an ND10 neutral density filter at the beam expander.

### ***Particle tracking***

Collected images were first pre-processed with Fiji.[11] Briefly, the 3D images were projected to 2D images by a maximum intensity projection and background subtracted. Post-processed images were then analyzed by a custom-written *Mathematica* (Wolfram Research) routine to detect and track particles in the RNA channel (red color). Specifically, particles were emphasized with a band-pass filter such that the positions could be detected using the built-in *Mathematica* routine `ComponentMeasurements` “IntensityCentroid.” Detected particles were linked through time by allowing a maximum displacement of 5 pixels between consecutive frames. Particle tracks lasting at least 5 frames were selected. To properly account for the offset between the two cameras, a geometric transformation function (see method below) was applied to the coordinates of the center of mRNAs. For each frame of each track, 15x15 (pixels x pixels) crops centered on the registered mRNA coordinate were made and averaged through time. Using *Mathematica*’s bandpass filter and `ComponentMeasurements` described above, the time-averaged crops corresponding to each track were categorized based on the presence of detectable signals in the green nascent chain channel: Red – mRNA not translating, Red + Green = Yellow – mRNA translating. Once the spots were categorized in this automated fashion, all spots were again hand-checked to minimize error.

Finally, the original 2D maximum intensity projected images corresponding to each hand-checked track were fit to find their precise coordinates and intensities (using

the built-in *Mathematica* routine NonlinearModelFit) to a 2D Gaussians of the following form:

$$I(X, Y) = I_{BG} + I e^{-\frac{(x-x_0)^2}{2\sigma_x^2} - \frac{(y-y_0)^2}{2\sigma_y^2}}$$

where  $I_{BG}$  is the background intensity,  $I$  the particle peak intensity,  $(\sigma_x, \sigma_y)$  the spread of the particle, and  $(x_0, y_0)$  the particle location. From these data, the intensity, position through time, and number of spots over time in each track were quantified for downstream analysis.

### ***Geometric transformation function***

The offset between the two cameras was registered using the built-in *Mathematica* routine FindGeometricTransform. To find the transformation function that best aligned the fitted positions, 100 nm diameter Tetraspeck beads evenly spread out across the image field-of-view were imaged on the same day experiments were taken. Only the fitted particle positions were registered to avoid introducing any distortion into images. Therefore, a slight offset can be observed between the red and the green or blue particles even though they are within a diffraction limited spot according to our registration.

### ***Calibrating translation site intensity***

We wanted to quantify the units of mature protein (i.e. number of nascent chains or active ribosomes) at a translation site using its intensity signal. For this purpose, we imaged two calibration constructs [6]. The two calibration constructs were equal in length, one containing the spaghetti monster 10×FLAG tag (SM-BetaActin) which contains 10 repeats of the FLAG epitope, the other containing just a single FLAG epitope (1×FLAG-filler-BetaActin). Note the spaghetti monster 10×FLAG tag (SM) is the

same tag utilized in the 100% optimized (SM-100Luc-MS2) and the 0% optimized (SM-0Luc-MS2). The 1×FLAG-filler-BetaActin tag was used to measure the number of ribosomes translating in a translation site. With the 1×FLAG-filler-BetaActin calibration construct, each nascent chain in a translation site contains just one FLAG epitope labeled by a single Fab conjugated (on average) to a single Cy3 fluorophore. By imaging this 1×FLAG construct at high laser powers such that individual translation sites and single Cy3 fluorophores (confirmed by single-step photobleaching, see below) can both be visualized, the ratio of total Cy3 signal in translation sites to single Cy3 fluorophore intensity signals approximates the number of nascent chains (or ribosomes) per translation site. To quantify this, we imaged cells beadloaded with 1×FLAG-filler-BetaActin and Cy3 anti-FLAG Fab in a single optical plane at high laser powers (50 mW for 561nm and 15mW for 637nm laser). A short movie was acquired, after which cells were continually illuminated to photobleach them to the point at which single probe fluorescence could easily be detected by single-step photobleaching. At this point, a second short 250-frame movie was acquired. The intensity of polysomes (verified by the presence of an RNA signal intensity) from the first frame of the first movie was then measured (as described in the 'Particle tracking' section above) and compared to the plateau intensity of a single Cy3 just prior to single-step photobleaching. The ratio of polysome to Cy3 intensities gives us an estimate for the number of ribosomes on a single RNA. By averaging all measured RNA together, we obtain a mean value for the ribosomal occupancy of a single, translating, 1×FLAG-filler-BetaActin mRNA. From these measurements,  $11.4 \pm 2.0$  ribosomes were estimated to be translating the 1×FLAG-filler-BetaActin calibration construct.

Since the 1×FLAG-filler-BetaActin calibration construct and the SM-BetaActin calibration construct (with 10×FLAG) are the same length with the same promoters and 3' and 5' UTRs, their translation sites should contain roughly the same number of ribosomes ( $11.4 \pm 2.0$ ). With a known number of ribosomes on SM-BetaActin, we wished quantify the number of ribosomes on the 100% and 0% Optimized ORF. To do this, we imaged cells in two different chambers. In one chamber, we beadloaded cells with anti-FLAG Fab (Cy3) and SM-BetaActin. In the second chamber, we beadloaded cells with anti-FLAG (Cy3) and the 100% Optimized construct. Both chambers were imaged on the same day with the same imaging conditions (50 mW for 561nm and 15mW for 637nm laser). By taking the intensity ratio of SM-BetaActin containing  $11.4 \pm 2.0$  ribosomes and 100% Optimized translation sites, we measured the ribosomal occupancy of the cap ORF to be 9.6 ribosomes. Taking the intensity ratio of these translation sites and all other SM translation sites in the 0% optimized construct gave the number of ribosomes translating in all possible translation sites, as shown in Figure 3.2.

### ***Ribosome run-off experiments using harringtonine treatment and elongation estimates***

To measure average elongation rates, cells in two separate chambers beadloaded with the 100% Optimized and 0% Optimized constructs, Cy3 labeled anti-FLAG Fab and Halo-MCP protein (labeled with JF646-HaloTag ligand) were imaged with a 60 second interval between every 13 frames (one entire cell volume) for 50 total time-points. Laser powers were the same as previously described for general imaging. After acquiring 5 time-points of pre-treated images, cells were treated with a final

concentration of 3  $\mu\text{g}/\text{mL}$  of harringtonine (Cayman Chemical). After treatment, cells were imaged for the remaining 45 time-points as described.

To generate ribosomal run-off curves, images were analyzed with the particle tracker as previously described. In each frame of each cell image, nascent chain signal intensities from all translation sites were totaled resulting in an intensity decay curve over time for each individual cell. Each decay curve was normalized to the average value from the first five frames (preceding drug addition).

## Permission of use

D.1 Chapter 3: Investigating the effects of codon non-optimization on translation elongation and initiation

05/7/2021

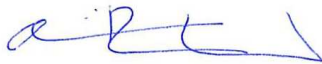
Amanda Lynn Koch has my permission to include unpublished material of which I am collaborator and author in her doctoral dissertation.

---



Chloe Barrington-Ham

---



Dr. Olivia Rissland

## Bibliography for Appendices

1. Aguilera LU, Raymond W, Fox ZR, May M, Djokic E, Morisaki T, Stasevich TJ, Munsky B: **Computational design and interpretation of single-RNA translation experiments.** *PLoS Comput Biol* 2019, **15**:e1007425.
2. Munsky B, Neuert G, Oudenaarden A van: **Using Gene Expression Noise to Understand Gene Regulation.** *Science* 2012, **336**:183–187.
3. Ingolia NT, Hussmann JA, Weissman JS: **Ribosome Profiling: Global Views of Translation.** *Cold Spring Harb Perspect Biol* 2019, doi:10.1101/cshperspect.a032698.
4. Wang C, Han B, Zhou R, Zhuang X: **Real-Time Imaging of Translation on Single mRNA Transcripts in Live Cells.** *Cell* 2016, **165**:990–1001.
5. Gillespie DT: **A general method for numerically simulating the stochastic time evolution of coupled chemical reactions.** *J Comput Phys* 1976, **22**:403–434.
6. Morisaki T, Lyon K, DeLuca KF, DeLuca JG, English BP, Zhang Z, Lavis LD, Grimm JB, Viswanathan S, Looger LL, et al.: **Real-time quantification of single RNA translation dynamics in living cells.** *Science* 2016, **352**:1425–1429.
7. Holcik M, Sonenberg N: **Translational control in stress and apoptosis.** *Nat Rev Mol Cell Biol* 2005, **6**:318–327.
8. HookeRobert, A J: **“Direct Search” Solution of Numerical and Statistical Problems.** *J ACM JACM* 1961,
9. Tokunaga M, Imamoto N, Sakata-Sogawa K: **Highly inclined thin illumination enables clear single-molecule imaging in cells.** *Nat Methods* 2008, **5**:159–161.



10. Edelstein AD, Tsuchida MA, Amodaj N, Pinkard H, Vale RD, Stuurman N: **Advanced methods of microscope control using  $\mu$ Manager software.** *J Biol Methods* 2014, **1**.
11. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al.: **Fiji: an open-source platform for biological-image analysis.** *Nat Methods* 2012, **9**:676–682.
12. Morisaki T, Stasevich TJ: **Quantifying Single mRNA Translation Kinetics in Living Cells.** *Cold Spring Harb Perspect Biol* 2018, **10**:a032078.